



L'UTILIZZO DI LINEE GUIDA NELLA PRATICA CLINICA

Mario Cruciani

Sezione Screening HIV - Centro di Medicina Preventiva - Azienda ULSS 20 Verona

NELL'ERA DELLA MEDICINA BASATA SULL'EVIDENZA

Negli ultimi decenni si è progressivamente assistito ad un cambiamento dei criteri utilizzati nel processo decisionale in campo medico. In un passato, più o meno remoto, il processo decisionale medico si svolgeva in un ambiente sociale e culturale ampiamente dominato dall'opinione di esperti. Il pensiero medico, l'insegnamento e la pratica clinica derivavano in massima parte da conoscenze trasmesse da esperti, in un contesto definito da alcuni di "medicina autoritaria". Il livello di evidenza clinica si limitava in genere alla descrizione del caso, o di una serie di casi, sulla base di informazioni retrospettive o osservazionali. Pur riconoscendo che le osservazioni valide sono basilari in tutti i campi scientifici, è anche vero che le stesse sono sostanzialmente volte a generare ipotesi, piuttosto che a confermarle. Il primo studio clinico randomizzato, sull'efficacia della streptomina nella terapia della tubercolosi, risale al 1948; da allora la qualità dell'evidenza, sempre di più basata su osservazioni raccolte in maniera sistematica e quantitativa, è andata gradualmente migliorando, in un contesto questa volta definibile di "medicina autorevole", come per l'appunto quella basata sull'evidenza. Davidoff riporta un esempio specifico che può aiutare a capire tali concetti (vedi tabella 1).

Nell'esempio fornito, i rischi sono ampiamente inferiori ai potenziali benefici, e mi trovo pertanto nelle condizioni di poter prendere una decisione terapeutica sulla base dell'evidenza clinica consistente sia qualitativamente che quantitativamente. In effetti, tra le finalità della medicina basata sull'evidenza rientrano la valutazione della qualità dell'evidenza disponibile, e l'esplicitazione di quanto il processo decisionale sia basato su dati piuttosto che su opinioni. Ciò non significa che la medicina basata sull'evidenza miri a negare trattamenti per condizioni per le quali vi sia mancanza di evidenza; al contrario sottolinea l'importanza di rilevare la mancanza di evidenza, indicando allo stesso tempo nuove linee di ricerca.

LE LINEE GUIDA NELLA PRATICA CLINICA

È difficile stabilire chi abbia prodotto le prime linee-guida. Secondo alcuni Mosè con le tavole dei 10 comandamenti (Esodo 20: 1-17), secondo altri il Re di Babilonia Hammurabi, con il famoso codice del 2500 AC (16). Di fatto, in campo sanitario si assiste ad un'esplosione di interesse per le linee-guida, al punto che qualcuno ha coniato l'espressione "industria delle linee-guida".

Una definizione ampiamente accettata di linee guida è quella dell'Institute of Medicine (USA), secondo la quale le linee-guida sono raccomandazioni di comportamento clinico, prodotte attraverso un processo sistematico, allo scopo di assistere medici e pazienti nel decidere quali siano le modalità assistenziali più appropriate in determinate situazioni cliniche. Le linee guida tendono principalmente a migliorare la qualità, l'appropriatezza ed il costo-efficacia degli interventi sanitari, nonché a fornire strumenti educativi.

Le linee guida dovrebbero essere efficaci, attendibili, riproducibili, applicabili dal punto di vista clinico e flessibili; dovrebbero inoltre essere scritte in maniera comprensibile ed accessibile ad operatori sanitari di vario tipo, ed essere revisionate ed aggiornate periodicamente.

È importante sottolineare che le linee guida non possono sostituire il giudizio clinico, dal momento che non riescono ad essere talmente specifiche da poter essere applicate a qualsivoglia situazione clinica: la discrezione clinica rimane pertanto di fondamentale importanza nell'applicare le linee-guida.

Le linee-guida, così come i processi che le alimentano, hanno subito un progressiva evoluzione, di pari passo con l'estendersi dei concetti legati alla medicina basata sull'evidenza. È importante che il metodo con il quale si raccoglie l'evidenza scientifica sia chiaramente specificato. In effetti, il più delle volte l'operatore sanitario si trova nella necessità di prendere decisioni in situazioni per le quali vi è una limitata quantità di informazione o, per contro, una sovrabbondanza di informazioni, spesso contraddittorie. Storicamente le linee-guida venivano prodotte mediante consensi informali: un pannello di esperti, a conoscenza della letteratura, sviluppa raccomandazioni basate sulla documentazione che si ritiene adeguata, al di fuori di qualsiasi regola formale di evidenza ed in assenza di descrizione dei metodi utilizzati. Tali raccomandazioni venivano in genere preparate in 1-2 giorni. Negli anni 70' questa prima generazione di linee guida, autoritaria ma non autorevole, ha cominciato ad essere soppiantata da una seconda generazione, volta a stabilire perlomeno procedure formali per raggiungere il consenso. Ancora una volta però tali metodi si basavano largamente sull'opinione di esperti, con scarsa documentazione dei metodi utilizzati e di conseguenza ampio adito a critiche relative ad errori, scarso rigore metodologico e conflitti di interesse. A partire dagli anni 80', cominciano a prendere piede concetti attinenti alla medicina basata sull'evidenza, che vengono estesi anche allo sviluppo di linee guida. In effetti, le decisioni basate sull'evidenza non necessariamente concordano con quelle degli esperti: quest'ultime sono infatti suscettibili di idiosincrasia, e talora possono essere francamente errate.

LINEE GUIDA BASATE SULL'EVIDENZA

Le linee-guida basate sull'evidenza vengono prodotte attraverso varie fasi, in un processo che richiede in genere 1-2 anni (tab. 2). Nella prima fase si seleziona il panel di esperti (in genere da 6 a 10) ed un leader, sufficientemente rappresentativi da poter analizzare l'argomento in questione nei dettagli. In una seconda parte vengono definiti in maniera più precisa gli scopi delle linee-guida, il tipo di intervento, e relativamente a quali pazienti ed a quali esiti (outcome), è indirizzata la verifica. Si dovrà inoltre specificare il processo che verrà utilizzato.

In una fase successiva si svolge una revisione sistematica della letteratura rilevante disponibile, specificando la strategia di ricerca utilizzata per l'identificazione degli studi. Questa fase è la stessa che sta alla base della meta-analisi.

La validità della letteratura relativa ad efficacia e convenienza di un intervento non è



regolarmente di alto livello, per cui l'esperienza e l'opinione degli esperti è spesso richiesta come sostituto dell'evidenza scientifica. È importante che in tali occasioni venga chiaramente espresso che le raccomandazioni si basano, in assenza o carenza di evidenza di altro tipo, sull'opinione. Come accennato, le procedure per determinare l'opinione degli esperti possono variare da un'informale tavola rotonda a più strutturate e formali metodiche quantitative.

Tab. 1 : Esempio di decisione clinica basata sull'evidenza disponibile

PRESENTAZIONE CLINICA ED IPOTESI
<p>Una donna di 77 anni vi si presenta, per la prima volta, in quanto ha iniziato ad accusare disturbi legati ad uno stato di insufficienza cardiaca di grado lieve, ed in presenza di fibrillazione atriale. Anamnesticamente vi è solo una storia di lieve ipertensione, ben controllata; l'ecocardiogramma mostra lieve insufficienza ventricolare sinistra. La terapia con digitale e diuretici risolve la situazione, ma non la fibrillazione. Questo rappresenta il problema clinico da risolvere. La fibrillazione rende infatti la paziente a rischio per ictus embolici, ed in tal caso una terapia anticoagulante con warfarin sarebbe appropriata. Per contro tale terapia, soprattutto per periodi prolungati, non è scevra di rischi emorragici, che vanno attentamente pesati.</p>
DECISIONI
<p>Quale decisione devo pertanto prendere, e come arrivarci. Ecco alcune possibilità: i) ascolto i consigli di un collega fidato; ii) applico gli standard terapeutici in uso localmente; iii) mi rivolgo a un farmacologo; iv) seguo quello che ho letto nell'ultima brochure sul farmaco; v) mi fido delle mie precedenti esperienze con pazienti simili; vi) ascolto le preferenze della paziente; vii) chiamo telefonicamente un illustrissimo collega (Eugene Braunwald nell'esempio di Davidoff), viii) metto in atto la medicina basata sull'evidenza.</p>
SE SCELGO L'ULTIMA, DEVO PROCEDERE COME SEGUE
<ol style="list-style-type: none"> 1. Formulare il problema in maniera corretta. Gli interrogativi sono due: a) qual è il rischio di embolizzazione in una paziente di 77 anni con fibrillazione atriale in assenza di terapia? b) come posso quantificare la riduzione del rischio ischemico indotta dal warfarin e bilanciarla con gli effetti indesiderati gravi (emorragici) indotti dallo stesso farmaco? 2. Ricercare le informazioni desiderate. Mi collego ad internet, ed eseguo una MEDLINE utilizzando le parole chiave appropriate. Della moltitudine di articoli estratti ne scelgo 7 (2 sulla prognosi, 4 sulla terapia, ed una meta-analisi). Il tutto mi costa 15 minuti di tempo e il prezzo di una telefonata. 3. Valutazione critica dell'evidenza. La letteratura in mio possesso mi fornisce i seguenti dati: a) il rischio embolico, per l'età ed in assenza di terapia, è del 18 % l'anno (= CER, control event rate); b) la terapia anticoagulante riduce il rischio di circa il 70 % (= RRR, riduzione relative del rischio) ; c) la riduzione assoluta del rischio (AAR) è pertanto del 13 %, dal 18 % al 5 % (= EER, experimental event rate). d) la percentuale di complicanze emorragiche maggiori in pazienti in warfarin è dell'1 % /anno. A questo punto posso anche calcolare il numero di pazienti che devono essere trattati per prevenire un evento (NNT), un indicatore utile in campo clinico in quanto offre al medico la possibilità di ragionare in termini di efficacia (o inefficacia) e di utilità (3). A tal scopo applico la formula: $NNT = 1/ARR$, dove $AAR = CER - EER$; <p>Pertanto, nel caso del rischio embolico, $NNT = 1/18 - 5 = 7,6$; in altri termini, per ogni 7,7 pazienti trattati con warfarin per un anno avrò prevenuto un episodio di ictus. Applicando le stesse formule per gli eventi emorragici, trovo un NNT di 100 (cioè 1 evento indesiderato grave ogni 100 pazienti/anno).</p>
CONCLUSIONE
DECIDO DI UTILIZZARE LA TERAPIA CON WARFARIN

Tab. 2: Le varie fasi nello sviluppo di linee-guida

PASSAGGI	DESCRIZIONE	TEMPO RICHIESTO
1	Selezione del panel di esperti	2-4 settimane
2	Organizzazione interna del gruppo	1-2 mesi
3	Definizione precisa degli scopi, del target, dei destinatari	Concomitante a punto 2
4	Definizioni delle metodiche per la ricerca dell'evidenza (ad es. MEDLINE, consensus, ecc.)	Concomitante a punto 2
5	Revisione sistemica della letteratura	2-3 mesi
6	Quantificazione della forza e qualità dell'evidenza	Concomitante a punto 5
7	Revisione esterna delle raccomandazioni	Dopo 9-10 mesi dall'inizio
8	Modifiche sulla base della revisione esterna	1-2 mesi
9	Revisioni ed aggiornamenti periodici	Ogni 2 anni

Tab. 3: Scala per la determinazione della forza e della qualità delle raccomandazioni

CATEGORIA, GRADO	DEFINIZIONE
FORZA DELLA RACCOMANDAZIONE	
A	Buona evidenza a supporto della raccomandazione di utilizzare l'intervento nella pratica clinica
B	Moderata evidenza a supporto della raccomandazione di utilizzare l'intervento nella pratica clinica
C	Scarsa evidenza a supporto della raccomandazione
D	Moderata evidenza a sfavore della raccomandazione di utilizzare l'intervento nella pratica clinica
E	Buona evidenza a sfavore della raccomandazione di utilizzare l'intervento nella pratica clinica
QUALITÀ DELL'EVIDENZA	
I	Evidenza da ≥ 1 studio clinico adeguatamente randomizzato
II	Evidenza da ≥ 1 studio clinico non randomizzato ma con disegno adeguato; da studi analitici di coorte o caso-controllo (preferibilmente multicentrici); risultati straordinari da studi non controllati
III	Evidenza basata sull'opinione di autorevoli esperti, basate sull'esperienza clinica, studi descrittivi, o report di comitati di esperti

A questo punto il panel di esperti potrebbe già redigere delle raccomandazioni o, meglio



ancora, definire una più approfondita linea di condotta, considerando altri aspetti quali il costo-efficacia, la disponibilità di tecnologie, l'accettabilità degli interventi da parte dei pazienti e della società, senza escludere osservazioni di tipo etico e legale.

È importante che le raccomandazioni contengano chiare indicazioni sulla forza e la qualità dell'evidenza che le supporta (o contrasta). La tabella 3 riporta un sistema in scala, ampiamente applicato per la valutazione delle raccomandazioni.

Le linee guida sono a questo punto pronte per essere revisionate esternamente da "peer reviewers", e su questa base modificate ed inviate per la pubblicazione.

LINEE GUIDA BASATE SULL'OPINIONE

Come accennato, in molte circostanze manca l'evidenza di elevata qualità (ad esempio studi clinici randomizzati) a supporto di una raccomandazioni. In tal caso i membri del panel devono ricorrere all'opinione. I limiti di tale approccio sono facilmente comprensibili e ben conosciuti: l'opinione degli esperti, per quanto illuminata, non è esente da errori metodologici di valutazione, preconcetti e conflitti di interesse. Esistono comunque metodi che permettono di incorporare le opinioni degli esperti nelle linee guida, attraverso la ricerca del consenso. Il consenso può derivare semplicemente dalla discussione collettiva delle opzioni, con l'accordo finale che viene fornito sulla base della percezione che il leader del gruppo ha della situazione. Per contro, sono state messe a punto tecniche del tipo "Delphi metodo" e successive modifiche (approccio Rand), finalizzate ad ottenere, per uno specifico problema, un giudizio da un gruppo di esperti che, nel corso di ripetuti momenti di valutazione, esprimano la propria opinione (19, 20). I membri del gruppo quantificano l'adeguatezza della procedura secondo una scala da 1 a 9, dove 9 indica l'assoluta correttezza, ed 1 l'estrema inappropriatazza.

Un classico, ed anche isolato, esempio di come tale tecnica possa permettere di incorporare l'opinione degli esperti con l'evidenza scientifica e' fornito dallo studio dell'American Society of Hematology sul trattamento della porpora trombocitopenica idiopatica -ITP (21). Riconosciuta la carenza di letteratura a proposito, il panel ITP affinò le metodiche per quantificare il livello di consenso tra esperti. Si distribuirono questionari simulanti numerosi (~ 2000) e differenti scenari clinici, nei quali si richiedeva di quantificare in una scala da 1 a 9 l'appropriatezza e la necessità dell'impiego di test diagnostici e terapie. L'indipendenza della valutazione era assicurata dal fatto che la votazione si svolgeva in condizioni di isolamento dei singoli membri. Il panel definì come appropriato o necessario un intervento che avesse ottenuto un punteggio medio da 7 a 9, inappropriato o non necessario un punteggio da 1 a 3 (tab. 4).

Per punteggi da 3.01 a 6.99 non venivano espresse raccomandazioni. Si valutò inoltre il livello di consenso per le raccomandazioni espresse dal panel (deviazioni dalla media) secondo una scala di 5 punti, dalla A (completa unanimità) alla E (forte disaccordo) (tabella 4). Secondo tali scale, un punteggio 1,5-A significa pertanto "completo accordo sulla non necessità o inadeguatezza" dell'intervento in esame, mentre un 7,5-D significa "a favore dell'intervento, ma con ampie variazioni nelle opinioni individuali".

Le raccomandazioni fornite dal panel ITP non risultarono particolarmente ampie, dal momento che molti degli interventi analizzati confluivano nel punteggio da 3.01 a 6.99, insufficiente per esprimere consenso. In genere molte delle conclusioni erano a sfavore dei test diagnostici (punteggio da 1 a 3). La valutazione della variabilità nel punteggio assegnato fornì altrettante importanti informazioni sulla natura del consenso. Ad esempio, la semplice presa visione di un punteggio medio da 3-01 a 6.99 non permetteva di per sé di evincere se il panel era concorde sulla non necessità o adeguatezza dell'intervento (la maggior parte dei votanti esprimeva un punteggio in tale range), o se tale punteggio fosse il risultato di un voto bimodale, con parte del panel favorevole (punteggio da 7 a 9) e parte sfavorevole (da 1 a 3), ma con un risultato in termini di media molto simile al precedente.

Una tale definizione e quantificazione sulla natura del consenso ottenuto per la stesura di linee guida è comunque rara se non eccezionale. Un'altra importante serie di informazioni che

scaturì dal panel ITP (e da altri studi) era relativa alla variazione delle conclusioni ottenibili in base alla composizione uni- o multi-disciplinare del panel stesso. Nel caso ITP il panel era costituito equamente da ematologi che lavoravano come liberi professionisti o come ricercatori universitari. La mancanza di consenso ed i differenti patterns di voto erano appunto dovuta all'eterogeneità dei partecipanti. Se il gruppo fosse stato composto da soli ricercatori, o da soli libero professionisti, le conclusioni sarebbero state definitive, anche se con i due gruppi in contrasto tra loro; sarebbero però state, allo stesso tempo, meno generalizzabili ed estrapolabili alla pratica clinica.

Tab. 4: Scala per la valutazione del livello di consenso tra esperti del panel ITP (porpora trombocitopenica idiopatica).

PUNTEGGIO	DEFINIZIONE
APPROPRIATEZZA O NECESSITÀ	
1.0 – 3.0	Non appropriato, non necessario (in base al quesito iniziale)
3.01 – 6.99	Dubbia l'appropriatezza o la necessità
7.0 – 9.0	Appropriato o necessario
CODICI DI CONSENSO	
A	Completa o quasi unanimità (varianza del panel > 2 deviazioni standard sotto la varianza media)
B	Forte unanimità (varianza del panel 1- 2 deviazioni standard sotto la varianza media)
C	Modesta unanimità (varianza del panel < 1 deviazione standard sotto la varianza media)
D	Modesto disaccordo (varianza del panel < 1 deviazione standard sopra la varianza media)
E	Forte disaccordo (varianza del panel > 1 deviazione standard sopra la varianza media)

LA VALUTAZIONE DELLA QUALITÀ DELL'EVIDENZA IN BASE ALLA METODOLOGIA DEGLI STUDI CLINICI

Le sperimentazioni cliniche, purché condotte appropriatamente, forniscono uno strumento insostituibile per valutare efficacia e sicurezza di nuovi interventi (terapeutici, profilattici, diagnostici, etc.), in genere allo scopo di migliorare la qualità degli interventi sanitari e di controllarne i costi.

I principi che regolano le sperimentazioni cliniche sono evoluti nel corso degli anni, ed alcuni di questi, riguardanti più le procedure d'esecuzione e documentazione della sperimentazione che la metodologia della stessa, sono attualmente regolati da uno standard di base che va sotto il nome di Good Clinical Practice (GCP). Le norme per la buona prassi clinica sono riportate nella dichiarazione di Helsinki del World Medical Association (settembre 1989), dalla Gazzetta Ufficiale delle Comunità Europea (direttiva CEE luglio 1991) e dalla Gazzetta Ufficiale della Repubblica Italiana (aprile 1992).

Ogni studio clinico ben disegnato richiede la stesura di un protocollo, che deve contenere adeguate informazioni sulla natura dello studio stesso. Qui di seguito sono schematizzati alcuni dei punti che dovrebbero essere presi in considerazione in ogni sperimentazione clinica.



Background, scopi generali ed obiettivi specifici

Il protocollo dovrebbe contenere una descrizione del background motivazionale dello studio, ed un'adeguata definizione non solo dell'obiettivo generale dello studio, ma anche della domanda specifica alla quale ci si prefigge di rispondere, di eventuali domande secondarie o sottogruppi di ipotesi da testare.

Criteri di selezione e di valutazione dei pazienti

Dal momento che i pazienti che entrano nello studio dovrebbero essere rappresentativi di qualche futura classe di pazienti che possa beneficiare dei dati originati dallo studio stesso, è essenziale che siano ben documentati i criteri di selezione, includendo quelli specifici, la fonte di reclutamento (ad esempio se pazienti ospedalizzati o meno) e le condizioni patologiche oggetto di studio.

È inoltre essenziale fornire un'adeguata descrizione sulle modalità di trattamento. Nel caso si sperimentino farmaci si dovrà specificare formulazioni, vie di somministrazione, dosi e ritmi, durata della terapia, effetti collaterali, modificazioni di dosaggio e sospensioni, aderenza del paziente al trattamento, trattamenti ancillari.

La metodica per la valutazione dei pazienti deve essere accurata ed oggettiva, e deve comprendere una valutazione iniziale (baseline), la definizione dei principali criteri di risposta al trattamento (ad esempio tempo di sopravvivenza, parziale o completa guarigione, durata della risposta, etc.), la definizione dei criteri sussidiari (ad es. effetti collaterali) ed altri oggetti del monitoraggio. Idealmente le misure dovrebbero essere precise e riproducibili, e non dipendenti dall'osservatore. Ciò non è però possibile per parametri non valutabili quantitativamente; in tali casi si può ricorrere a metodi specifici di valutazione (ad es. "rating scale"), all'esperienza di qualcuno degli sperimentatori o a "consensus" tra gli sperimentatori. Se è poi richiesta l'opinione del paziente, con le eventuali interferenze dello sperimentatore che la richiede, aumentano le possibilità di bias. Molte di queste problematiche sono però ovviate dalla cecità dello studio.

Disegno dello studio

Il disegno dello studio merita poi una più dettagliata descrizione. Si distinguono studi non controllati e controllati (vedi tabella 5).

Tab. 5: Principali tipi di sperimentazione clinica (da Bollettino d'informazione sui farmaci, Ministero della Sanità, modificato)

MODELLO DI SPERIMENTAZIONE	NOTE
1. Non controllata	Il trattamento sperimentale viene assegnato a tutti i pazienti eleggibili consecutivamente osservati. Non c'è confronto diretto con altri gruppi di pazienti, e gli effetti del trattamento sperimentale sono valutati in base al confronto con il decorso della malattia trattata con terapie standard, che si ritiene ben noto.
2. Controllata, non randomizzata: - controlli paralleli - controlli storici - controlli da banche dati	Il trattamento sperimentale viene assegnato a tutti o ad una parte dei pazienti eleggibili consecutivamente osservati. C'è un gruppo di pazienti trattati in altro modo, arruolati con procedure diverse, che servono come controlli. Rimane incerta la comparabilità fra pazienti che ricevono in trattamento sperimentale ed i controlli.

3. Controllata e randomizzata	L'assegnazione dei trattamenti, sperimentale e di controllo, è fatta mediante un sistema di sorteggio (randomizzazione) che favorisce la comparabilità fra i gruppi.
4. Meta-analisi	È una sintesi quantitativa, che permette di assemblare sperimentazioni multiple di uno stesso trattamento (in genere randomizzate).

I primi sono in grado di generare una visione distorta di un intervento. Possono comunque essere utili, come nel caso di alcuni tipi di sperimentazione di fase II. Tra gli studi controllati occorre distinguere quelli randomizzati da quelli con controlli d'altro tipo. Negli studi non randomizzati rimane incerta la comparabilità fra pazienti che ricevono differenti trattamenti. Gli studi con controlli storici presentano limiti sostanziali, legati soprattutto alla bilanciatura del confronto. Le principali incompatibilità derivano dalla selezione dei pazienti e dall'ambiente sperimentale. In genere i criteri di scelta nel gruppo storico sono meno specificati che nel gruppo osservazionale, e la qualità dei dati storici è inferiore. Nel tempo, inoltre, possono variare il tipo di paziente ed il criterio di risposta. Sono in ogni modo preferibili controlli storici nella stessa organizzazione che dai dati della letteratura.

Tra gli studi prospettici controllati occorre distinguere quelli non randomizzati con controlli concomitanti da quelli randomizzati. Nel caso degli studi non randomizzati la scelta di un trattamento non è affidata al caso, ma si utilizzano metodi predeterminati sistematici o criteri di valutazione eseguiti dallo sperimentatore e/o paziente. In tal modo, soprattutto nel secondo caso (in genere non accettato), si generano bias selettivi, che spesso a loro volta generano ulteriori bias. Si ricorderà che per bias si intende la deviazione dei risultati o delle inferenze dalla verità; si tratta in generale di errori sistematici, e non stocastici, che possono generarsi a vari livelli della sperimentazione (ad es. bias di selezione, di rilevazione, di valutazione, etc.) Nel caso dei metodi predeterminati, si utilizzano ad esempio date di nascita o di presentazione (giorni pari/dispari - trattamento nuovo/vecchio); l'investigatore conosce facilmente il farmaco che andrà somministrato, e questo per influenzare la sua decisione. Facilmente si osservano sbilanciature nelle dimensioni dei gruppi, ad indicare preferenze. Ad esempio, in uno studio che confrontava eparina versus con trattamento nell'infarto del miocardio, i pazienti ricoverati nei giorni pari ricevevano eparina, mentre quelli ricoverati nei giorni dispari non ricevevano niente. Alla fine dello studio in 68% dei pazienti era stato ricoverato in giorni pari, ed aveva perciò ricevuto eparina, mentre solo il 32% non aveva ricevuto trattamento.

Si è inoltre evidenziato che in circa il 60% delle sperimentazioni non randomizzate la distribuzione dei fattori prognostici è sbilanciata, e che nella maggior parte dei casi il trattamento sperimentale è assegnato a pazienti con prognosi migliore. È pertanto evidente che il rischio di risultati falsi positivi è elevato. Ad esempio si è visto che i dati pubblicati di sperimentazioni non controllate in campo psichiatrico fornivano risultati positivi nel 85% dei casi, mentre negli studi controllati tale percentuale scendeva al 25%. Al di là di questi limiti, le sperimentazioni non randomizzate e con controlli d'altro tipo possono avere un ruolo in alcune situazioni, quali: la malattia è talmente rara da non consentire un trial controllato; la malattia presenta decorso prevedibile ed esito invariabilmente fatale; l'efficacia terapeutica del nuovo trattamento è drammaticamente evidente; non esistono trattamenti alternativi impiegabili per confronto.



Registrazione e randomizzazione dei pazienti

Il controllo randomizzato è ritenuto l'unico strumento valido. L'uso appropriato della randomizzazione aiuta a ridurre il rischio di differenze nell'ambiente sperimentale e a produrre gruppi comparabili. Garantisce inoltre che non vi sono bias nella selezione dei pazienti nei gruppi, anche se altri tipi di bias possono originarsi in altre fasi della sperimentazione. I metodi di randomizzazione sono vari e più o meno complessi, e possono andare dalla monetina a metodiche più formali e meglio definite (tavole di randomizzazione). Esistono poi randomizzazioni stratificate, a blocchi permutati, non eguali (ad esempio randomizzazioni 2:1, 3:1), per le quali si rimanda ad apposite pubblicazioni.

Va ricordato che, prima della randomizzazione, è necessario verificare che il paziente sia eleggibile nello studio (vedi punto 3). È inoltre essenziale che lo sperimentatore sia d'accordo nell'accettare il trattamento random prima che il paziente entri nello studio. Legato a questo è il problema dell'eticità della randomizzazione. Ogni studio clinico dovrebbe nascere dal presupposto che vi è incertezza nel potenziale beneficio di un nuovo intervento. Nel caso un investigatore ritenga che -per qualsiasi motivo- il nuovo intervento sia più benefico o più pericoloso dovrebbe non partecipare allo studio. Per contro, se lo sperimentatore ha dubbi sulla superiorità di un trattamento verso l'altro, troverà etico partecipare ad uno studio randomizzato.

A questo punto, dopo l'ottenimento del consenso da parte del paziente, lo stesso può entrare formalmente nello studio. Lo sperimentatore potrà registrarlo e, solo a questo punto, venire a conoscenza del trattamento assegnato, con modalità differenti in base al tipo di studio e all'organizzazione dello stesso. La lista potrà ad esempio, negli studi monocentrici, essere in possesso di una singola persona, coinvolta unicamente nelle fasi di registrazione e randomizzazione; o la lista di randomizzazione verrà trasferita mediante busta chiusa che contiene il tipo di trattamento da somministrare al paziente registrato; od ancora vi è una registrazione centralizzata, negli studi multicentrici, e il trattamento viene comunicato per via telefonica. Negli studi randomizzati in doppio-cieco è in genere il farmacista che prepara i farmaci e fornisce le confezioni da somministrare, assolutamente non distinguibili l'una dall'altra.

È necessario a questo punto introdurre i concetti di cecità e placebo. Alcune delle distorsioni nel confronto tra due trattamenti possono originare dal fatto che lo sperimentatore e/o il paziente sono a conoscenza del farmaco che somministra e/o riceve. Il paziente può essere favorevolmente impressionato dal fatto che riceve un nuovo trattamento rispetto ad un trattamento standard o ad un non trattamento, e reagire con maggior aderenza allo studio con eventuali vantaggi psicologici. Il contrario può capitare nei gruppi di controllo. A loro volta anche il team di sperimentatori può essere influenzato dalla conoscenza del tipo di trattamento. Ad esempio un certo entusiasmo dello sperimentatore può essere trasferito al paziente, o lo sperimentatore può seguire più strettamente ed attentamente i pazienti trattati con il nuovo farmaco rispetto a quelli che ricevono il farmaco standard. Gli stessi problemi interessano chi deve valutare comparativamente i differenti trattamenti (vedi anche punto 5). Questi problemi possono essere parzialmente risolti rendendo cieco lo studio, che deve chiaramente essere randomizzato. In uno studio in doppio cieco, sia il paziente sia il team di sperimentatori (compreso chi valuta) non è a conoscenza di quale trattamento sia in corso. Nel caso per la specifica condizione morbosa non esista un trattamento standard efficace, può essere appropriato lasciare il gruppo di controllo non trattato. Gli studi in doppio cieco, con o senza gruppo placebo, devono rispondere ad alcuni requisiti di eticità e praticabilità. Innanzitutto non vi deve essere alcun rischio addizionale per il paziente. In altre circostanze occorre valutare il reale peso del bias in assenza di cecità. Talora è sufficiente che solo il paziente non sia a conoscenza del tipo di trattamento che sta ricevendo (single-blind), soprattutto se è lui stesso a dover valutare gli effetti di una terapia. Anche in questo caso è essenziale che l'investigatore non eserciti alcuna influenza, più o meno diretta, sul paziente. Altre volte è sufficiente che la sola valutazione sia cieca.

Il rationale per l'impiego del placebo è sostanzialmente quello di valutare se, nel gruppo trattato, la risposta osservata è realmente legata alla molecola farmacologicamente attiva o al solo fatto di ricevere qualcosa. Tale aspetto è ben esemplificato nel trial del clofibrato riportato

dal comitato dei principali investigatori. In tale studio la mortalità a 5 anni per infarto del miocardio risultava simile nel gruppo clofibrato (18%) e placebo (19%). Nello studio era stata valutata anche la "compliance" dei pazienti al clofibrato: i pazienti che assumevano almeno l'80% delle prescrizioni presentavano una mortalità a 5 anni del 15%, mentre in quelli con compliance inferiore al 80% la mortalità saliva al 25% ($p < 0.0001$). Ad una prima osservazione si potrebbe pertanto, erroneamente, affermare che i pazienti che assumono regolarmente la terapia con clofibrato hanno una mortalità inferiore rispetto a quelli che ricevono placebo (19 vs 15%). In effetti nello studio era stata valutata anche la compliance al placebo, e si era visto che, allo stesso modo che per il gruppo clofibrato, la mortalità era significativamente inferiore ($p < 0.0001$) nei pazienti che assumevano regolarmente placebo rispetto a quelli che lo assumevano con minor regolarità (15 versus 27%). In ultima analisi, l'aderenza al trattamento, e non il tipo di trattamento, era significativamente correlato con riduzione della mortalità. Osservazioni di questo tipo sono state fatte in altri studi randomizzati con gruppo di controllo costituito da placebo.

Un discorso a parte merita l'analisi dei dati ricavati da molteplici studi clinici condotti sullo stesso argomento, i quali hanno fornito in genere risultati contrastanti. Questa metodica è stata particolarmente perfezionata negli ultimi anni, e viene generalmente definita come meta-analisi. Una delle principali difficoltà nell'esecuzione di studi clinici è data dal fatto, sopra accennato, che per evidenziare con affidabilità differenze reali tra trattamenti, è spesso richiesto un numero elevato di pazienti da randomizzare. Lo scopo principale della meta-analisi è di fornire un unico dato a partire dalle molte sperimentazioni condotte sull'argomento in questione. Per la metodologia ed il rationale delle indagini meta-analitiche si rimanda ad esaurienti rassegne.

Consenso informato

Le problematiche etiche relative alla sperimentazione clinica riguardano, in ultima analisi, la contrapposizione tra azioni dirette nell'interesse del singolo paziente e quelle dirette al progresso in campo medico. I requisiti etici essenziali della ricerca clinica sono elencati nella dichiarazione di Helsinki del 1960 e nei successivi adattamenti. Ad essa fanno riferimento le disposizioni nazionali e comunitarie. In tali documenti si fa specifico riferimento all'ottenimento del consenso (preferibilmente scritto) alla partecipazione allo studio da parte del paziente, previa adeguata informazione sulle caratteristiche della sperimentazione in oggetto.

La costituzione di comitati etici ad hoc, come da decreto ministeriale del 1992, dovrebbe facilitare il difficile aspetto dell'applicabilità delle linee guida etiche ai singoli studi, e del rapporto tra industria farmaceutica e sperimentatore.

Dimensioni del campione richieste dallo studio

Accanto a problemi pratici ed etici, esistono considerazioni scientifiche che devono essere necessariamente prese in considerazione nel disegno di un protocollo di studio clinico. Una di queste è la definizione delle dimensioni del campione richiesto allo scopo di soddisfare gli obiettivi scientifici che ci si prefigge, in modo da poter evidenziare differenze nei gruppi in oggetto considerate di interesse clinico (il cosiddetto calcolo della potenza dello studio).

Analisi dei dati e deviazioni dal protocollo

Generalmente le deviazioni dal protocollo si verificano per singoli pazienti in quanto il paziente stesso, od il suo trattamento e/o la valutazione non rientrano nei criteri prestabiliti nel protocollo. Nel caso i criteri di eleggibilità siano ben precisati e seguiti dagli investigatori, il rischio di osservare deviazioni dal protocollo, che comunque persistono in altre fasi dello studio, si ridimensiona di molto. Esiste però la possibilità che il criterio di ineleggibilità emerga più avanti nello studio, dopo la randomizzazione (ad es. a seguito del risultato di un esame



istologico). Alcuni problemi possono nascere dalla non-compliance del paziente e dalla valutazione incompleta. Nel caso si verificano deviazioni dal protocollo, è importante evidenziarle, nel tentativo di trovare delle spiegazioni ed eventualmente di prevenire ulteriori deviazioni. Se ad esempio si verificano ripetutamente violazioni maggiori del protocollo, la causa può essere nella natura stessa dello studio, mal amministrato e con scarsa cooperazione tra investigatori e/o pazienti. In alcuni casi è possibile modificare, sulla base di tali osservazioni, in senso migliorativo il disegno e l'esecuzione dello studio.

Nel caso della non-compliance, occorre distinguere tra reale scarsa collaborazione del paziente e le modificazioni od interruzioni della terapia causate da effetti collaterali o progressione della malattia, che devono necessariamente rientrare nella valutazione clinica.

L'uscita dallo studio può essere voluta dal paziente stesso o dallo sperimentatore; in quest'ultimo caso la sospensione del trattamento deve essere basata su adeguata documentazione clinica e nell'interesse del paziente (ad esempio per progressione della malattia, o per eccesso di tossicità), altrimenti sarà da ritenere una violazione del protocollo da parte dell'investigatore.

Uno dei punti di notevole interesse ed attualità nelle sperimentazioni cliniche è quello relativo a chi includere nella valutazione. Nell'approccio "Intention to treat" (approccio pragmatico) vengono inclusi nell'analisi tutti i pazienti eleggibili e randomizzati, indipendentemente dalla compliance al protocollo. Questa tendenza è da preferirsi in quanto fornisce una più completa valutazione del trattamento e della sua praticabilità. Nell'approccio "by treatment" o "explanatory" vengono invece inclusi nell'analisi unicamente i pazienti con completa aderenza al protocollo e alla terapia; in questo caso sono però possibili distorsioni nella valutazione comparativa dei trattamenti, in quanto si valuta l'efficacia assoluta, che può discordare da quella raggiungibile nella pratica. Un esempio è mostrato dal seguente studio, che confrontava la terapia medica e quella chirurgica nella stenosi carotidea bilaterale. L'analisi "by treatment", escludendo cioè i pazienti che decedevano o sviluppavano ictus durante il ricovero, ed includendo pertanto solo pazienti "ideali", mostrava una riduzione del rischio di progressione della patologia (TIA, ictus o decesso) nel gruppo trattato con terapia chirurgica pari al 20 % (54 versus 74 %); tale riduzione era ($p=0.02$) statisticamente significativo. Per contro, l'analisi "intention to treat", che includeva questa volta tutti i pazienti arruolati, mostrava una minor riduzione del rischio nel gruppo chirurgico, 62 versus 74 %, che non raggiungeva ($p=0.9$) la significatività statistica. Nell'analisi by treatment erano stati esclusi 16 pazienti (1 solo dei quali in terapia medica) che erano morti o andati incontro ad ictus quando essere ancora ospedalizzati; i 15 pazienti randomizzati nel gruppo terapia chirurgica avrebbero per contro potuto ricevere una terapia medica, con eventuali risultati positivi. L'esclusione di tali pazienti, nell'analisi "by treatment", tende pertanto a sopravvalutare l'efficacia della terapia chirurgica, mentre nell'analisi "intention to treat", più vicina all'attuabilità, terapia medica e chirurgica non differiscono sostanzialmente.

Analisi statistica

Quando sono disponibili i risultati della sperimentazione, ad esempio un confronto fra l'efficacia di due farmaci (e la sperimentazione è stata condotta in maniera metodologicamente corretta, in assenza di bias o errori sistematici), ci si deve porre la domanda se le eventuali differenze tra trattamenti sono reali (legate alla terapia in quanto tale) o dovute al caso. A tal scopo occorre ricorrere a tests statistici, quali i tests di significatività e i limiti di confidenza. Senza voler entrare nella complessità della materia, per la quale si rimanda a testi specializzati (Armitage, 1971; Fisher et al, 1974), si prenderanno qui di seguito in considerazione alcuni punti, con i quali il clinico dovrebbe avere una certa familiarità.

Va sottolineato innanzitutto che il criterio di significatività statistica non equivale ad importanza clinica, esprimendo unicamente una misura probabilistica della realtà o causalità delle differenze osservate; in altre parole i test di significatività danno l'idea della forza dell'evidenza della superiorità di un trattamento rispetto all'altro. I limiti di confidenza stimano

invece la grandezza delle differenze osservate, e pertanto la rilevanza clinica delle stesse; pur non determinando la significatività statistica, sono correlati con i test di significatività.

I limiti di confidenza specificano il range di valori della differenza reale tra i due trattamenti che avrebbero potuto produrre con relativa facilità la differenza osservata nello studio. Normalmente il criterio per stabilire l'intervallo fiduciale segue lo stesso ragionamento su cui si basa il test di significatività. Viene cioè stabilito il tasso di errore accettato, e si fissano i limiti dell'intervallo di confidenza, in maniera che la frequenza con cui esso non comprende il valore reale corrisponda appunto a questo tasso di errore. I limiti di confidenza più usati sono quelli al 95%, che corrispondono alla classica percentuale di errore del 5%. Quanto maggiore è il numero di pazienti in studio, tanto più è improbabile che la differenza osservata sia molto diversa da quella reale, e questo si rifletterà in un intervallo di confidenza più ristretto intorno al valore osservato.

GLI ALGORITMI CLINICI

Le linee-guida hanno la finalità pratica di trasferire le conoscenze disponibili in decisioni. Per far questo devono utilizzare un linguaggio chiaro, non ambiguo, e devono anche utilizzare modelli di presentazione logici e facili da seguire. Gli algoritmi clinici rappresentano un valido modello per rappresentare delle raccomandazioni. Rispetto ad un testo discorsivo, l'algoritmo presenta alcuni vantaggi, quali la possibilità di apprendimento più rapido, maggior ricordo, maggiore aderenza con gli standard pratici stabiliti.

La realizzazione di un algoritmo presuppone che vengano identificati dei sottogruppi di pazienti che possano beneficiare (o non beneficiare) di un intervento. L'efficacia degli algoritmi si fonda per l'appunto sulla possibilità di sottoclassificare i singoli casi clinici in categorie più omogenee.

La figura 1 evidenzia un esempio di algoritmo per il management dell'epatite da virus C in uso presso il nostro centro. Un cliente si presenta per lo screening, esegue il primo prelievo per la ricerca di anticorpi anti - HCV in ELISA. Vi sono 2 possibilità: il test è negativo, in una o più determinazioni (a distanza di 3 e 6 mesi dall'evento a rischio), per cui si chiude il monitoraggio.

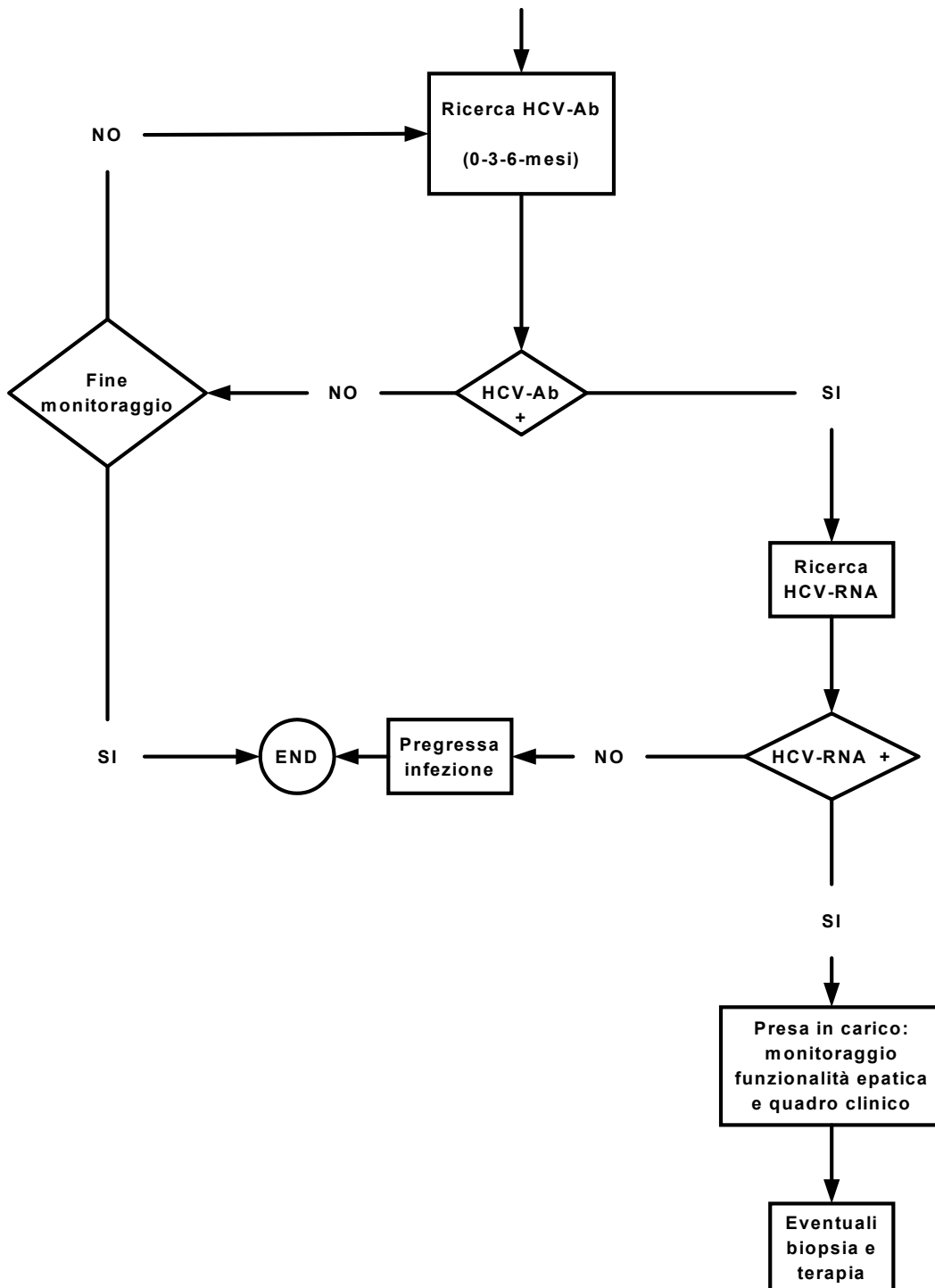
Se il test è positivo, occorre ampliare le indagini diagnostiche per differenziare se si tratta di una pregressa infezione risolta o cronicizzata. La presenza di anticorpi, non protettivi, non è infatti sufficiente allo scopo. Eseguo allora la ricerca del RNA virale. Se negativa, chiudo il processo, considerando che il soggetto ha avuto un'infezione da virus C che si è risolta. Se invece nel sangue è presente l'RNA del virus, devo andare avanti con le indagini. Prendo a questo punto in carico il paziente, eseguo un monitoraggio clinico e della funzionalità epatica più approfondito, e così via sino all'eventuale biopsia prima dell'inizio della terapia.

Rispetto al paragrafo di prosa qui sopra, riteniamo che l'approccio algoritmico alle linee guida apporti sostanziali vantaggi, permettendo ad esempio di cristallizzare le raccomandazioni in un formato più diretto rispetto alla prosa, ed una precisa definizione dei termini e dei dettagli clinici adeguati alle necessità, così come deducibile dalla letteratura.

Nonostante gli apparenti vantaggi, molti operatori sanitari mostrano riluttanza nell'applicare gli algoritmi clinici alla pratica quotidiana. Questo atteggiamento può trovare varie spiegazioni, ma forse la principale è che l'algoritmo impone un approccio rigido, e può essere vissuto come una limitazione alla libertà decisionale. Al contrario, l'algoritmo dovrebbe essere considerato come un'ossatura a supporto della riflessione.



Fig. 1: Flow-chart per il management dell'infezione da HCV



Implementazione e verifica

Il successo delle linee-guida è in relazione a vari fattori che includono il contesto clinico in cui utilizzarle ed i metodi utilizzati per svilupparle, diffonderle ed implementarle. Difficoltà nell'implementazione di linee-guida sono talora osservabili in varie specialità mediche. In una valutazione sistematica della letteratura relativa all'effetto delle linee-guida sulla pratica clinica, si è notato che nella maggior parte dei casi (in 55 di 59 studi analizzati) era possibile evidenziare un generale miglioramento nella pratica (25). Il grado di miglioramento era però ampio. I livelli maggiori di miglioramento si osservavano per gli studi nei quali sviluppo, diffusione e implementazione delle raccomandazioni erano attuate in maniera adeguata. Si è visto ad esempio che una strategia interna di sviluppo delle linee-guida piuttosto che una strategia esterna o nazionale, si accompagnava a maggiore efficacia. Pertanto, può essere utile che operatori sanitari in determinate occasioni modifichino delle linee guida esistenti sulla base di esigenze locali. Allo stesso modo, si è visto che gli interventi educativi specifici costituiscono la più efficace strategia per la diffusione delle raccomandazioni.

LE LINEE GUIDA NEL CAMPO DELLE DIPENDENZE

Una esaustiva rassegna delle linee guida disponibili per il problema delle tossicodipendenze è disponibile nel sito www.dronet.org. Il sito fornisce una raccolta di linee guida raggruppate per categorie: diagnosi, organizzazione e gestione, prevenzione primaria, prevenzione secondaria, trattamento, valutazione. Il sito continuamente aggiornato, fornisce dati descrittivi e collegamenti a documenti (oltre 100 all'agosto 2001) in formato pdf, con la possibilità di scaricarli. Tra i vari links, ricorderemo quelli con l'Organizzazione Mondiale della Sanità, la Presidenza del consiglio dei Ministri, il Ministero della Sanità, organizzazioni sanitarie ed accademiche di vari paesi europei ed extra-europei, e con la Comunità europea.

Le linee-guida rintracciabili nel sito riguardano problemi, tematiche generali sull'abuso di sostanze e sui disturbi mentali, e specifiche legate all'uso di alcool, benzodiazepine, cocaina, oppio ed altre sostanze illecite.

CONCLUSIONI

Le linee-guida hanno la finalità pratica di trasferire le conoscenze disponibili in decisioni. Questo non significa che necessariamente le linee guida possano indicare con certezza quello che occorre fare; in molte occasioni esse potranno solo informarci su quello che è disponibile, creando eventualmente indicazioni per future indagini. Per poter risultare utili, particolare attenzione deve essere posta alle fasi di sviluppo, diffusione ed implementazione delle linee-guida. Dal momento che l'evidenza scientifica su cui le linee-guida si basano non è costantemente elevata, è importante che vengano applicati e migliorati metodi formali per combinare le direttive derivate dall'opinione di esperti con l'evidenza disponibile.

Ciascun clinico dovrebbe essere in grado di valutare la qualità dell'evidenza, sia che si tratti di sperimentazioni cliniche, reviews sistematiche e meta-analisi, ed anche linee-guida pratiche. Le linee guida non possono sostituire il giudizio clinico, ma possono fornire una cornice che può facilitare il processo decisionale e migliorare la qualità dell'assistenza fornita. Né tali raccomandazioni possono essere indicate, per lo meno da sole, per risolvere questioni legali legate alla professione medica.

BIBLIOGRAFIA

- 1) Davidoff F. In the teeth of the evidence. *Mount Sinai J Med.* 1999, 66: 75-83.
- 2) Medical research council. Streptomycin treatment of pulmonary tuberculosis. *BMJ.* 1948, ii: 769-82.



- 3) Ministero della Sanità, Direzione generale della valutazione dei medicinali e della farmacovigilanza. Bollettino d'informazione sui farmaci. N. 2, 2001.
- 4) Woolf SH, Gorge JN. Evidence-based medicine. Interpreting studies and setting policy. *Hem Onc Clin North Am.* 2000, 14: 761-84.
- 5) Pocock SJ. *Clinical trials. A practical approach.* J. Wiley and Sons, Chichester, 1983.
- 6) Altman DG. Statistics and ethics in medical research: I. Misuse of statistics is unethical. *BMJ*, 281: 1182-4, 1980.
- 7) Altman DG. Statistics in medical journals. *Stat Med*, 1: 59-71, 1982
- 8) Gazzetta Ufficiale della Repubblica Italiana, Supplemento ordinario alla Gazzetta Ufficiale n. 139 del 15 giugno 1992, serie generale. Ministero della Sanità, decreto ministeriale 27 aprile 1992. Disposizioni sulla documentazione tecniche da presentare a corredo delle domande di autorizzazione all'immissione in commercio di specialità medicinali per uso umano, anche in attuazione della direttiva n. 91/507/CEE.
- 9) Foulds GA. Clinical research in psychiatry. *J Ment Sci.* 1958, 104: 259-65.
- 10) Committee of principal Investigators. WHO cooperative trial on primary prevention of ischaemic heart disease using clofibrate to lower serum cholesterol: mortality follow-up. *Lancet*, ii: 379-85.
- 11) Fields WS et al. Joint study of extracranial arterial occlusion. *JAMA.* 1970. 211: 1993-2003.
- 12) Messori A, Rampazzo R, Scuffi C. La metanalisi. Area SIFO di formazione e ricerca. Area di metodologie informatiche in farmacocinetica, metanalisi e statistica. Società Italiana di farmacia Ospedaliera.
- 13) Armitage P. *Statistical Methods in Medical Research.* Blackwell, Oxford, 1971.
- 14) Fisher RA, Yates F. *Statistical tables for biological, agricultural and medical research.* 6th ed. Hafner Publishing Company, inc., New York, 1963.
- 15) Ministero della Sanità. Bollettino d'informazione sui farmaci. N. 1-6, 1999
- 16) Gross PA. Practice guidelines for infectious diseases: rationale for a work in progress. *Clin Infect Dis.* 1998, 26 : 1037-41.
- 17) Cook D, Giacomini M. The trials and tribulations of clinical practice guidelines (Editorial). *JAMA.* 1999, 281: 1950-1
- 18) Kish MA. Guide to development of practice guidelines. *Clin Infect Dis.* 2001, 32 : 851-4.
- 19) Park RE, Fink A, Brook RH et al. Physician ratings of appropriate indications for three procedures: theoretical indications vs indications used in practice. *Am J Pub Health.* 1989, 79: 445.
- 20) Grilli R, Penna A, Liberati A. *Migliorare la pratica clinica. Come produrre ed implementare linee-guida.* Il Pensiero Scientifico Editore, Roma. 1995.
- 21) American Society of Hematology ITP Practice Guidelines panel. Diagnosis and treatment of idiopathic thrombocytopenic purpura: recommendations from the American Society of Hematology. *Ann Intern Med.* 1997. 126: 319-26.
- 22) Hadorn DC, McCormick K, Diokno A. An annotated algorithm approach to clinical guideline development. *JAMA.* 1992, 267: 3311-4.
- 23) Sox HC, Sox CH, Tompkins RK. The training of physician's assistants: the use of a clinical algorithm system for patient care audit of performance and education. *N Engl J Med.* 1973. 288: 818-24.
- 24) Donabedian A. *The criterion and standards of quality.* Ann Arbor, Mich: Health Administration Press, 1982.
- 25) Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice. A systematic review of rigorous evaluations. *Lancet.* 1993, 342: 1317-22.

