

Revisioni sistematiche e decisioni cliniche in psichiatria

Systematic reviews and clinical decisions in psychiatry

CORRADO BARBUI and MATTHEW HOTOPIF

INTRODUZIONE

In psichiatria, ormai da alcuni anni, accanto alle revisioni classiche della letteratura, le cosiddette revisioni narrative, gli articoli cioè in cui gli esperti di ciascun settore passano in rassegna le conoscenze riguardanti un determinato argomento per trarne conclusioni utili alla pratica clinica, si è progressivamente affermato un nuovo tipo di rassegna della letteratura, denominato revisione sistematica. Questo è avvenuto a seguito dell'altrettanto forte affermazione dell'idea secondo cui le decisioni in medicina devono basarsi sulle evidenze scientifiche a nostra disposizione. Le evidenze scientifiche sono meglio dette prove di efficacia, e il termine che descrive questo movimento è medicina basata sulle prove di efficacia (Barbui *et al.*, 2002). Secondo il paradigma della medicina basata sulle prove di efficacia (o *evidence-based medicine*, da cui la sigla EBM) le revisioni sistematiche rappresentano lo strumento più idoneo per riassumere le prove di efficacia esistenti per ogni argomento. Scopo di questo editoriale è chiarire che cosa sono le prove di efficacia in medicina ed esemplificare come queste ultime vengono correntemente utilizzate nella preparazione e nel mantenimento delle revisioni sistematiche. Obiettivo ultimo è discutere il possibile ruolo delle revisioni sistematiche nel supportare il processo di acquisizione di competenze che permettano di prendere decisioni in ambito psichiatrico.

Address for correspondence: Dr. C. Barbui, Dipartimento di Medicina e Sanità Pubblica, Sezione di Psichiatria, Università di Verona, Policlinico G.B. Rossi, P.le L.A. Scuro 10, 37134 Verona.

Fax: +39 045 585871

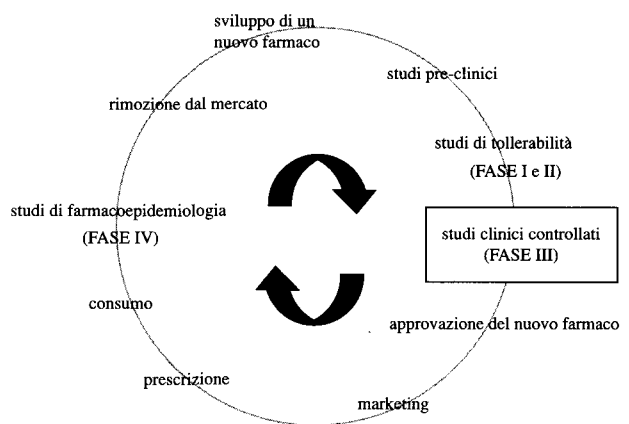
E-mail: corrado.barbui@univr.it

Declaration of Interest: none.

IL CONCETTO DI PROVA DI EFFICACIA

Il termine prova di efficacia o evidenza scientifica non si riferisce a qualsiasi osservazione scientifica reperibile in letteratura. Questo è un primo punto, cruciale, che caratterizza le revisioni sistematiche rispetto a quelle narrative. Per esempio, in ambito farmacologico esistono alcune fasi che caratterizzano lo sviluppo e lo studio di un nuovo composto potenzialmente utile nel trattamento di una determinata malattia (figura 1).

Figura 1. – Il ciclo vitale dei farmaci. Le revisioni sistematiche della letteratura si occupano di analizzare le informazioni contenute negli studi clinici controllati (studi di fase III).



Vi sono innanzitutto i cosiddetti studi pre-clinici, condotti in laboratorio, che consentono di tracciare il profilo farmacodinamico del nuovo farmaco chiarendone il meccanismo d'azione; questi studi sono evidentemente evidenze scientifiche, ma non sono prove di efficacia clinica del farmaco. Vi sono poi gli studi clinici di fase I e II, condotti per saggiare la tollerabilità del composto e stabilirne il *range* di dosaggio utile in clinica, e quelli di fase

III successivamente. Gli studi di fase III sono gli studi clinici controllati, detti anche *trial*. Nei *trial* si seleziona un campione di soggetti affetti che, potenzialmente, potrebbe trarre beneficio dal nuovo intervento e, attraverso un procedimento di assegnazione casuale, si suddivide il campione in due gruppi, il primo trattato mediante intervento standard, il secondo mediante intervento sperimentale. Al termine dello studio, eventuali differenze di esito vengono attribuite proprio ai trattamenti ricevuti, poiché questi ultimi rappresentano l'unico elemento di differenza tra i due gruppi di soggetti, per il resto identici fra loro in ogni caratteristica sociodemografica, clinica, di storia di malattia, di intensità sintomatologica e di contesto assistenziale in cui i trattamenti sono stati somministrati. Gli studi di fase III, i *trial*, sono le evidenze scientifiche che consentono di confermare o meno l'efficacia dei farmaci e dei trattamenti in generale, sono cioè le prove di efficacia che vengono analizzate nelle revisioni sistematiche. La presenza di *trial* positivi consente l'approvazione e l'immissione nel mercato delle novità terapeutiche da parte di enti regolatori come la *Food and Drug Administration* (FDA) negli Stati Uniti e la *European Medicine Evaluation Agency* (EMA) in Europa. Una volta introdotti nel mercato e, quindi, nell'uso clinico quotidiano, gli studi di fase IV permettono di tracciare il profilo epidemiologico delle popolazioni di soggetti che ricevono tali farmaci. Questi studi, tuttavia, non sono prove di efficacia, non rispondono, cioè, ad un quesito di efficacia in termini assoluti e, dunque, non vengono presi in considerazione nell'analisi delle prove di efficacia che si compie all'interno delle revisioni sistematiche. Per fare un esempio pratico, è possibile reperire numerose revisioni narrative della letteratura che supportano in maniera molto forte l'utilizzo degli antidepressivi di nuova generazione negli anziani. Questa indicazione all'uso viene però motivata da ragioni farmacodinamiche, ossia ricordando che gli antidepressivi di nuova generazione non hanno effetto adrenolitico e, dunque, non provocano ipotensione posturale come i vecchi triciclici, dotati invece di una indubbia azione adrenolitica. Cioè si induce un utilizzo clinico di un farmaco sulla base di studi pre-clinici, condotti in laboratorio, che mostrano l'assenza di una determinata azione farmacologica. Ma gli studi pre-clinici non sono le prove di efficacia dei nuovi antidepressivi negli anziani, ovvero il fatto che il meccanismo d'azione sia potenzialmente più vantaggioso per gli anziani costituisce una ottima ipotesi di ricerca clinica, ma non una evidenza di efficacia. Tale ipotesi dovrebbe essere confermata o meno da appositi *trial* condotti negli anziani, di confronto tra antidepressivi di vecchia e nuova generazione. E' solamente attraverso il confronto all'in-

terno dello studio clinico controllato che si potranno valutare eventuali differenze clinicamente rilevanti e stabilire se gli antidepressivi di nuova generazione, rispetto ai vecchi, conferiscano vantaggi negli anziani (per esempio diminuendo la frequenza di cadute accidentali come conseguenza di un minore effetto ipotensivo).

Proprio per evitare questo tipo di ambiguità, le revisioni sistematiche della letteratura si occupano di analizzare esclusivamente le informazioni contenute negli studi clinici controllati. Naturalmente, in ogni ambito della medicina, per ogni specifico problema clinico, possono essere a disposizione nessuno, pochi o molti *trial*. Il lavoro di raccogliere, argomento per argomento, tutte le prove di efficacia pubblicate nella letteratura internazionale e di analizzarle in modo sistematico al fine di presentare sinteticamente ciò che è noto per ciascun argomento, viene a costituire la revisione sistematica della letteratura. In alcuni casi, le revisioni sistematiche si preoccupano di estrarre da ogni studio incluso le informazioni più rilevanti, di solito quelle che hanno a che vedere con l'efficacia e la tollerabilità dell'intervento oggetto del *trial*, e di metterle tutte assieme, a costituire un unico ipotetico studio molto grande, sommatoria di tutti i *trial* inclusi nella revisione. Questo permette di compiere una nuova analisi statistica dei dati, denominata metanalisi. I risultati delle metanalisi vengono in questo modo a rappresentare il sunto delle prove di efficacia disponibili per ciascun argomento.

LA REALIZZAZIONE DELLE REVISIONI SISTEMATICHE

Le revisioni sistematiche devono essere considerate, a tutti gli effetti, dei veri e propri progetti di ricerca. Si tratta di progetti piuttosto complessi, sia nella fase di raccolta sistematica delle evidenze, che in quella di estrazione e rielaborazione statistica dei dati. Per far fronte a questa complessità ormai da alcuni anni si sono sviluppate reti di ricercatori distribuiti in tutti i paesi del mondo che collaborano, ognuno nel proprio ambito di interesse, alla realizzazione e al mantenimento delle revisioni. La *Cochrane Collaboration* è il più autorevole di questi network di ricercatori. Le tappe da percorrere per la produzione e il mantenimento di una revisione sistematica condotta nell'ambito della *Cochrane Collaboration* sono riassunte nella tabella I.

Come in ogni progetto di ricerca, anche lo sviluppo delle revisioni sistematiche inizia con la definizione di una ipotesi da testare. Ovviamente, l'ipotesi deve tenere in considerazione la tipologia degli studi presenti in lette-

Tabella I. - Le tappe per la produzione e il mantenimento di una revisione sistematica nell'ambito della *Cochrane Collaboration* (per dettagli consultare il sito <http://cochrane.org>).

-
- ✓ Ipotesi di ricerca
 - ✓ Titolo della revisione sistematica (registrazione)
 - ✓ Protocollo della revisione (pubblicato se ritenuto valido nella *Cochrane Library*)
 - ✓ Realizzazione dello studio:
 - Identificazione degli studi che soddisfano i criteri di inclusione
 - Reperimento degli studi inclusi
 - Preparazione di una scheda per l'estrazione dei dati
 - Estrazione dei dati (effettuata in modo indipendente da almeno due ricercatori)
 - Inserimento delle informazioni in un database elettronico
 - Analisi statistica (= metanalisi)
 - ✓ Discussione dei risultati e divulgazione:
 - letteratura scientifica
 - *Cochrane Library*
 - ✓ Mantenimento della revisione nel tempo
-

ratura poiché il lavoro metanalitico è possibile solo se siano a disposizione molti studi ricchi di informazioni. Nell'ambito della *Cochrane Collaboration* i titoli di possibili revisioni sistematiche vengono registrati in modo da evitare che più gruppi di lavoro si occupino dello stesso argomento contemporaneamente. A questo punto si procede alla stesura di un protocollo di studio. Questa è una tappa essenziale di tutto il progetto di ricerca. Nel protocollo si traccia una sintesi delle conoscenze acquisite e si formula, nel dettaglio, l'ipotesi di ricerca. I metodi che verranno utilizzati per rispondere a tale ipotesi vengono quindi esplicitati con grande precisione, dal tipo di ricerca bibliografica che permetterà di individuare gli studi ai criteri di inclusione e esclusione degli studi stessi, da come i dati verranno estratti da ogni studio (quante persone estrarranno i dati? Verrà utilizzato un apposito questionario per l'estrazione dei dati?) al tipo di analisi statistica che verrà condotta. Quest'ultimo punto è cruciale. La definizione a priori dell'analisi statistica è un importantissimo criterio di qualità del lavoro, poiché permette di controllare facilmente se effettivamente gli autori analizzano i dati secondo quanto dichiarato a priori. Per questo motivo i protocolli delle revisioni sistematiche vengono resi accessibili mediante pubblicazione sulla *Cochrane Library*, la rivista elettronica che raccoglie tutte le revisioni sistematiche della letteratura e i protocolli stesi secondo i criteri di qualità della *Cochrane Collaboration*.

A questo punto si procede alla realizzazione di quanto dichiarato nel protocollo secondo lo schema della tabella I. In particolare vale la pena sottolineare che, tra i requisiti di qualità richiesti, vi è anche quello secondo cui l'estrazione delle informazioni da ogni studio incluso deve essere compiuta da almeno due ricercatori in modo indipendente, al fine di evitare errori o interpretazioni erronee. Tutte le informazioni estratte vengono inserite in un database elettronico ed analizzate con appositi software statistici. Il prodotto finale viene divulgato mediante il

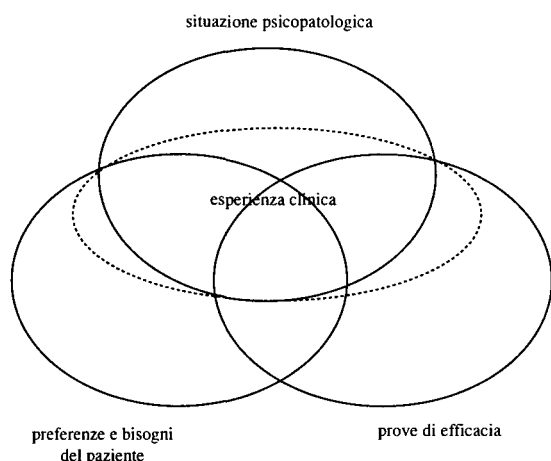
canale della letteratura scientifica (ossia pubblicazioni sulle riviste scientifiche) e mediante pubblicazione sulla *Cochrane Library* della revisione e dei successivi eventuali aggiornamenti della stessa (qualora nuovi studi vengano pubblicati nel frattempo).

IL PROCESSO DECISIONALE IN PSICHIATRIA

È naturalmente importante interrogarsi circa il possibile uso delle revisioni sistematiche nel favorire le decisioni che, continuamente, gli operatori della salute mentale sono costretti a prendere nella gestione quotidiana dei propri pazienti. Come è noto, il processo decisionale in medicina e in psichiatria è frutto di una serie di fattori che si combinano e intersecano secondo il modello presentato in tabella II. Tra questi fattori, che comprendono le circostanze cliniche in cui ci si trova ad operare, le caratteristiche psicopatologiche e di storia di malattia, i bisogni e le preferenze dei pazienti e molte altre variabili, vi sono anche le conoscenze scientifiche in nostro possesso, ovvero le prove di efficacia accumulate nel corso degli anni.

In questo quadro generale, le revisioni sistematiche della letteratura si offrono come strumento decisivo nel supportare il processo decisionale, proprio in quanto rappresentano un sunto critico e sistematico delle conoscenze scientifiche. Le revisioni sistematiche delle prove di efficacia, tuttavia, rappresentano un ausilio al processo decisionale solamente per quanto riguarda l'insieme collocato in basso a destra della figura 2, quello che si riferisce, appunto, alle prove di efficacia. In altre parole, come più volte sottolineato dai sostenitori del movimento della *evidence-based medicine*: "evidence does not take decisions, people do", ossia le revisioni sistematiche non prendono decisioni, sono gli operatori sanitari che le prendono, tenendo in considerazione, fra l'altro, le evidenze scientifiche a disposizione (Haynes *et al.*, 2002).

Figura 2. - Il processo decisionale in psichiatria. Ogni decisione clinica di tipo terapeutico ed assistenziale dovrebbe tenere in considerazione e derivare dall'analisi degli elementi riportati in figura. Le revisioni sistematiche delle prove di efficacia rappresentano un ausilio al processo decisionale solamente per quanto riguarda l'insieme collocato in basso a destra, quello che si riferisce, appunto, alle prove di efficacia.



UN ESEMPIO PRATICO: ANTIDEPRESSIVI A BASSE DOSI NEL TRATTAMENTO DELLA DEPRESSIONE

Il campo dei trattamenti antidepressivi (AD) nella gestione farmacologica della depressione è un buon esempio del possibile ruolo delle revisioni sistematiche nel supportare il processo decisionale del medico. Si tratta di un campo caratterizzato dalla pubblicazione di molti studi clinici controllati, ciascuno con piccoli numeri di pazienti, dai risultati contrastanti. L'approccio metanalitico è dunque particolarmente indicato per sviscerare specifici problemi di ordine clinico. Un problema rimasto irrisolto si riferisce ai dosaggi degli AD triciclici (TCA) nel trattamento dei pazienti depressi. Le evidenze scientifiche sono infatti molto scarse. Le linee guida esistenti raccomandano di prescrivere dosaggi superiori a 100 mg/die o superiori a 125 mg/die, ammonendo che l'utilizzo di TCA a dosaggi inferiori a 125 mg/die corrisponderebbe alla prescrizione di un placebo gravato da effetti collaterali (Hirschfeld *et al.*, 1997). Queste considerazioni, tuttavia, non si basano su dati forti di inefficacia dei TCA a dosaggi inferiori. È possibile, quindi, mettere in dubbio l'idea, ampiamente condivisa in letteratura, che la depressione sia attualmente trattata a dosaggi troppo bassi in ambito di medicina generale.

Per affrontare in modo sistematico questo problema, recentemente è stata condotta una revisione sistematica

degli studi che confrontavano le basse dosi di TCA con il placebo e con dosaggi standard di TCA (Furukawa *et al.*, 2002). La metanalisi definiva basso un dosaggio di TCA inferiore a 100 mg/die e alto un dosaggio superiore a 100 mg/die. Di oltre 2400 articoli che in qualche modo affrontavano il problema dei dosaggi di AD nella depressione, solo 43 erano *trial* clinici che studiavano davvero il problema oggetto dello studio. Di questi, 39 rispondevano ai criteri di inclusione e qualità definiti a priori dagli autori. In 35 studi il confronto era tra le basse dosi di TCA e il placebo, in 6 casi tra le basse dosi di TCA e i dosaggi standard. L'estrazione dei dati di efficacia e tollerabilità da ciascuno studio ha quindi permesso di compiere una nuova analisi che aveva il vantaggio di un maggiore potere statistico rispetto ad ogni studio considerato singolarmente.

Rispetto al placebo, i bassi dosaggi di TCA (tra 75 e 100 mg/die) aumentavano del 65% la probabilità di remissione sintomatologica alla quarta settimana di trattamento, e di oltre il 100% alla sesta settimana. In media, il 45% dei pazienti trattati con bassi dosaggi ha risposto a 4 settimane, il 59% a sei settimane e il 53% a 12 mesi. Per quanto riguarda la tollerabilità delle basse dosi di TCA, misurata come proporzione di soggetti che hanno interrotto prematuramente il trattamento (*drop-out*), il confronto con placebo non mostrava alcuna differenza in termini di tasso totale di *drop-out*. Suddividendo però i soggetti che interrompevano il trattamento a seconda del motivo dell'interruzione, si è evidenziato che i soggetti in trattamento con basse dosi di TCA interrompevano più frequentemente il trattamento, a causa dell'insorgenza di effetti collaterali.

Rispetto alle dosi standard di TCA, i bassi dosaggi (tra 75 e 100 mg/die) non sono risultati meno efficaci nell'indurre l'effetto terapeutico alla quarta e all'ottava settimana di trattamento. La figura 3 traccia in maniera schematica i dettagli di questa analisi. La prima colonna riporta il nome dello studio, la seconda e la terza il numero di soggetti in remissione, rispetto al totale, nel gruppo assegnato alle basse dosi e in quello assegnato alle dosi standard, rispettivamente. La parte grafica della figura riporta la stima dell'effetto per ciascuno studio, sotto forma di quadrato nero. La linea verticale rappresenta l'assenza di differenze; i quadrati neri collocati sulla linea verticale esprimono quindi una uguale efficacia tra dosi basse e quelle standard. I quadrati neri collocati a sinistra della linea verticale esprimono un effetto a favore delle dosi standard, mentre quelli collocati a destra della linea verticale esprimono un effetto a favore delle basse dosi. Ogni quadrato nero è attraversato da una barra orizzontale che esprime la variabilità della stima, ossia l'intervallo

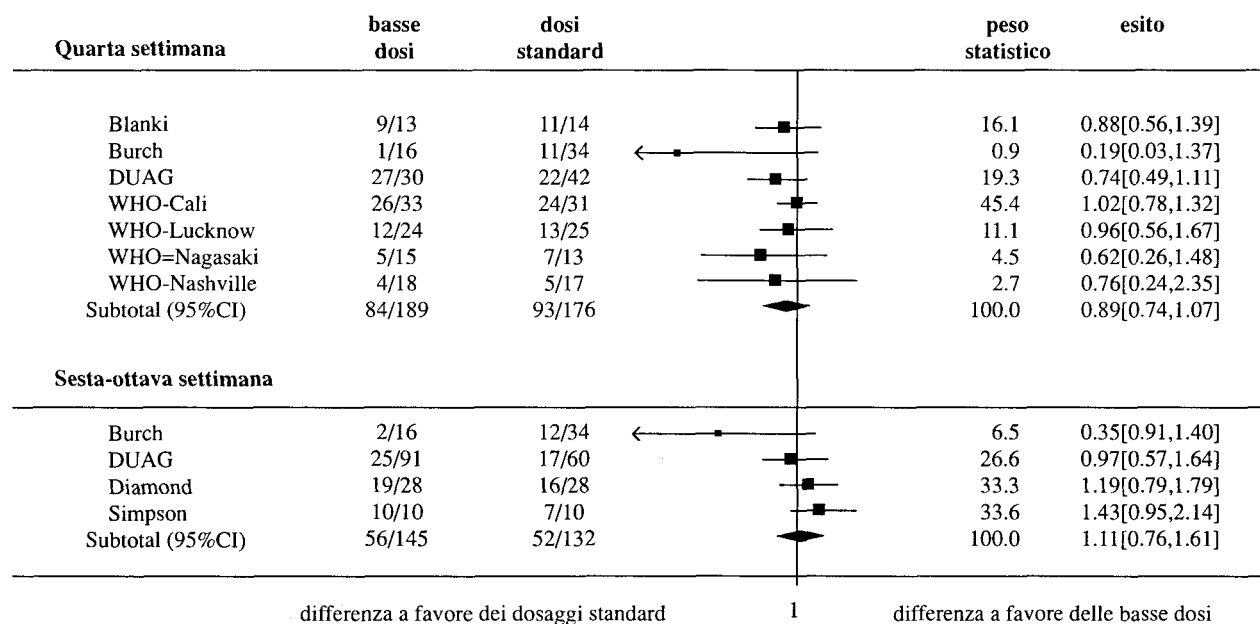
di confidenza. Se l'intervallo di confidenza tocca la linea verticale, significa che la variabilità della stima include anche la possibilità di assenza di differenze e, dunque, il risultato non è statisticamente significativo. La colonna 'esito' esprime in numeri le stesse informazioni presenti nella parte grafica. La stima globale fornita dalla metanalisi è rappresentata dalla figura a forma di diamante riportata sotto le stime dei singoli studi (ve ne sono due, una per gli studi che durano 4 settimane ed una per quelli che durano 6-8 settimane). Come riportato nella Figura 3, le differenze di efficacia tra bassi dosaggi e dosaggi standard non sono statisticamente significative, in quanto l'estremità destra e sinistra delle figure a diamante (che rappresentano gli estremi dell'intervallo di confidenza) incrociano la linea verticale, quella che sancisce l'assenza di differenze. Per quanto riguarda la tollerabilità delle basse dosi di TCA, misurata come proporzione di soggetti che interrompono prematuramente il trattamento, il confronto con i dosaggi standard non ha evidenziato alcuna differenza in termini di tasso totale di *drop-out*. Suddividendo però i soggetti che interrompevano il trattamento a seconda del motivo della interruzione, si è evidenziato che i pazienti in trattamento con basse dosi di TCA interrompevano meno frequentemente il trattamento a causa dell'insorgenza di effetti collaterali rispetto ai soggetti trattati a dosaggio pieno.

DALLE REVISIONI ALLE DECISIONI

Questo esempio indica il possibile uso, nella pratica clinica quotidiana, delle revisioni sistematiche della letteratura. Per molti anni gli esperti nel settore dei trattamenti antidepressivi hanno criticato la pratica, frequente in medicina generale, di prescrivere "basse" dosi di TCA. La revisione sistematica delle prove di efficacia disponibili in letteratura ha però dimostrato che tale atteggiamento critico non è supportato dai dati. Dosaggi di antidepressivo compresi tra 75 e 100 mg/die non sono meno efficaci di dosaggi superiori a 100 mg/die e provocano meno interruzioni del trattamento per effetti collaterali. I medici di medicina generale che prescrivono gli antidepressivi a tali dosaggi hanno così un supporto scientifico a sostegno della propria pratica.

Questo è, dunque, il ruolo che le revisioni sistematiche dovrebbero avere in psichiatria e, più in generale, in medicina. Esse dovrebbero diventare uno strumento che permetta l'acquisizione di competenze utili nel prendere decisioni. Dovrebbero essere quindi uno strumento che, proprio attraverso l'acquisizione di competenze, consenta e stimoli una maggiore autonomia culturale di giudizio. Autonomia di giudizio vuole significare capacità di esprimere una opinione formata attraverso l'analisi critica dei dati scientifici. L'assenza di questa autonomia, pur-

Figura 3. - Efficacia degli antidepressivi triciclici a basse dosi nel trattamento della depressione (vedi il testo per il commento alla figura).



troppo, è inevitabilmente destinata a lasciare gli operatori della salute mentale in balia degli esperti, delle mode, delle pressioni del mercato e delle vere – o presunte – novità terapeutiche che ogni giorno ci vengono proposte.

BIBLIOGRAFIA

- Barbui C. (2002). Medicina basata sulle prove di efficacia e pratica clinica quotidiana: il caso degli antidepressivi. *Epidemiologia e Psichiatria Sociale* 11, 152-158.
- Furukawa T.A., McGuire H. & Barbui C. (2002). Meta-analysis of effects and side-effects of low-dosage tricyclic antidepressants in depression: systematic review. *British Medical Journal* 325, 991-999.
- Haynes R.B., Devereaux P.J. & Guyatt G.H. (2002). Physicians' and patients' choices in evidence based practice. Evidence does not take decisions, people do. *British Medical Journal* 324, 1350.
- Hirschfeld R.M., Keller M.B., Panico S., Arons B.S., Barlow D. & Davidoff F. (1997). The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of depression. *Journal of the American Medical Association* 277, 333-340.