

Medicina basata sulle prove di efficacia e pratica clinica quotidiana: il caso degli antidepressivi

*Evidence-based medicine and routine clinical practice:
the case of new antidepressant drugs*

CORRADO BARBUI

*Dipartimento di Medicina e Sanità Pubblica, Sezione di Psichiatria
Università di Verona, Verona*

CONTESTO DI RIFERIMENTO

In medicina, ormai da diversi decenni, si ritiene che ogni intervento terapeutico, prima di essere utilizzato nella pratica clinica quotidiana, debba essere studiato in condizioni sperimentali, al fine di valutarne il profilo di tollerabilità ed efficacia. Questo è importante per non introdurre nella pratica clinica dei comportamenti terapeutici inefficaci o, addirittura, dannosi. Ad oggi, il più raffinato strumento che il mondo della ricerca ha a propria disposizione per misurare efficacia e tollerabilità degli interventi sanitari è lo studio clinico controllato (RCT), detto anche *trial*. Nei *trial* si seleziona un campione di soggetti 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 di sintomatologia e di contesto assistenziale in cui i trattamenti sono stati somministrati.

Naturalmente, in ogni ambito della medicina, per ogni specifico problema clinico, possono essere a disposizione nessuno, pochi o molti RCT. Questi RCT costi-

tuiscono il bagaglio di conoscenze, dette anche evidenze o prove di efficacia, disponibili per ciascun argomento, da utilizzarsi come guida nella pratica clinica quotidiana. 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 denominato revisione sistematica (SR) della letteratura. In alcuni casi le SR della letteratura 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 SR. 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. Si tratta di un procedimento piuttosto complesso, 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 di SR della letteratura. La *Cochrane Collaboration* è il più autorevole di questi *network* di ricercatori (<http://hiru.mcmaster.ca/cochrane/>).

Naturalmente, tanto maggiore è il numero di RCT disponibili per un certo argomento, tanto più rilevante è compiere una SR di quegli studi ed, eventualmente, una metanalisi. Il caso degli antidepressivi (AD) è emblematico a questo riguardo. Si tratta di un settore caratterizzato dalla presenza di diverse centinaia di *trial*

Indirizzo per la corrispondenza: Dr. C. Barbui, Dipartimento di Medicina e Sanità Pubblica, Sezione di Psichiatria, Università di Verona, Ospedale Policlinico, 37134 Verona.

Fax: +39-045-585.871

E-mail: corrado.barbui@univr.it

che hanno studiato il profilo di tollerabilità ed efficacia di questi farmaci nei soggetti con depressione. Ciascuno di questi studi, mediamente, è stato condotto arruolando piccoli campioni di soggetti. Queste due caratteristiche – elevato numero di *trial* e pochi soggetti inclusi in ciascun *trial* – rendono l'approccio metanalitico particolarmente appropriato. Questo per due motivi. Il primo si riferisce alla necessità di un approccio sistematico. Laddove vi sono molti *trial*, infatti, capita che le indicazioni per la pratica si compiano a partire dall'analisi solo di alcuni studi (quelli che si conoscono, oppure quelli che si ritiene, soggettivamente, essere i migliori). Questo è naturalmente un errore dal punto di vista metodologico, in quanto la selezione arbitraria dei *trial* introduce una distorsione nella valutazione delle evidenze. Ecco dunque che la sistematicità dell'approccio metanalitico, in cui i criteri di inclusione degli studi vengono dichiarati a priori, è garanzia di assenza di distorsioni, per lo meno a questo livello. Il secondo motivo si riferisce al fatto che, laddove i *trial* sono stati condotti arruolando pochi pazienti, è possibile che tali studi, presi singolarmente, non fossero in grado di riconoscere differenze significative tra i trattamenti in competizione. E qui l'approccio metanalitico è pure decisivo, in quanto la nuova analisi statistica viene compiuta includendo i soggetti arruolati in tutti i *trial* contemporaneamente, con l'effetto ultimo di aumentare il potere statistico dell'analisi.

Il gruppo della *Cochrane Collaboration* che si occupa di trattamenti nella depressione è denominato CC-DAN, acronimo che sta per *Cochrane Collaboration Depression, Anxiety & Neurosis Group* (<http://www.ccdan.auckland.ac.nz/>). Il CCDAN, negli ultimi anni, ha portato a termine alcune SR sulla efficacia e tollerabilità degli AD di nuova generazione, gli inibitori della ricaptazione della serotonina (SSRI), rispetto ai vecchi AD triciclici (TCA). Di seguito vengono presentati alcuni risultati di queste metanalisi, al fine di esemplificare come queste ultime siano in grado di sintetizzare le evidenze della letteratura in modo sistematico. La disponibilità di tali evidenze, tuttavia, non si traduce automaticamente in un utilizzo degli AD basato sulle prove di efficacia. Esiste, infatti, un problema di trasferibilità dell'informazione dalla ricerca alla pratica. Per questo motivo, in questo articolo i risultati delle metanalisi degli AD verranno messe a confronto con l'epidemiologia degli AD, al fine di evidenziare la discrepanza esistente tra evidenze sperimentali e pratica clinica. Infine, alcuni motivi che possono spiegare tale discrepanza, e alcune strategie per tentare di superarla, verranno presentate e discusse.

ANTIDEPRESSIVI DI VECCHIA E NUOVA GENERAZIONE: LE PROVE DI EFFICACIA

Una delle più note e influenti metanalisi nel campo degli AD è sicuramente quella portata a termine, quasi 10 anni fa, dal gruppo di Song *et al.* (1993). Questo lavoro, contraddicendo le impressioni cliniche dei prescrittori di SSRI, evidenziava un'efficacia clinica di questa classe di farmaci molto simile a quella dei TCA. Tuttavia, gli ampi intervalli di confidenza, calcolati sui punteggi medi, mettevano in evidenza il basso potere statistico di questi studi e, dunque, la bassa capacità di evidenziare eventuali differenze. Gli autori non trovavano differenze quantitative nemmeno per quanto riguarda la frequenza degli effetti collaterali, per lo meno giudicata come tasso di soggetti che interrompono il trattamento (*dropout*). Questi risultati sono stati successivamente messi in discussione da una nuova metanalisi che escludeva dal confronto SSRI verso TCA gli AD atipici, inclusi invece nella precedente metanalisi di Song. La nuova analisi evidenziava ancora una volta una analoga efficacia, ma in questo caso metteva anche in evidenza una differenza statisticamente significativa nei tassi di *dropout* per effetti collaterali a favore degli SSRI (Montgomery & Kasper, 1995). In questo lavoro, tuttavia, gli autori non avevano preso in considerazione il tasso totale di *dropout* (ossia il numero di coloro che interrompono il trattamento per inefficacia più il numero di coloro che interrompono il trattamento a causa degli effetti collaterali). Anderson & Tomenson (1995), al fine di produrre questo dato, hanno pubblicato nel 1995 una nuova metanalisi di 62 studi clinici controllati che analizzavano circa 6000 pazienti affetti da depressione maggiore unipolare). La differenza calcolata dagli autori nel tasso totale di *dropout* risultava di piccola entità: 30.8% per gli SSRI e 34.4% per i TCA. Lo stesso gruppo di ricerca ha recentemente confermato questi tassi in un campione più numeroso di studi (Anderson, 2000).

Certamente, un limite di queste metanalisi è legato all'eterogeneità delle molecole raggruppate e classificate come TCA. Hotopf *et al.* (1996) hanno affrontato questo problema osservando come l'amitriptilina e l'imipramina, considerati i TCA di riferimento, siano le molecole con il profilo beneficio-rischio meno favorevole. Questi autori hanno quindi analizzato i tassi di *dropout* degli SSRI rispetto ai "vecchi triciclici" (amitriptilina e imipramina), "nuovi triciclici" (dotie-pina, nortriptilina, desipramina, clomipramina, doxepina) e rispetto agli AD eterociclici (bupropione, mianserina, trazodone, maprotilina, amineptina, nomifensi-

na). Questa metanalisi dimostrava che i tassi di *dropout* degli SSRI erano inferiori ai vecchi triciclici; tuttavia, quando gli SSRI venivano comparati ai nuovi triciclici e agli AD eterociclici non si evidenziavano differenze statisticamente significative.

Con il passare degli anni, ovviamente, nuovi RCT che confrontavano gli SSRI con i TCA aggiungevano prove di efficacia a quelle già esistenti, rendendo necessario il continuo aggiornamento delle metanalisi. Nel 2000, il CCDAN pubblicava sulla *Cochrane Library* l'aggiornamento del lavoro di Hotopf *et al.*, che arrivava ad includere 136 RCT di confronto tra AD di vecchia e nuova generazione (Barbui *et al.*, 2000). Questo lavoro, oltre a calcolare il tasso totale di *dropout*, ha calcolato anche il tasso di *dropout* causati dagli effetti collaterali e dalla mancata risposta terapeutica. In sintesi, il lavoro indica che se un medico tratta 100 pazienti con un TCA di vecchia generazione, circa 30 pazienti interromperanno il trattamento nel corso delle prime 6-8 settimane; viceversa, se un medico tratta 100 pazienti con un SSRI, circa 27 pazienti interromperanno il trattamento nelle fasi iniziali. Vi è dunque una differenza di 3 soggetti ogni 100 trattati, a favore dei nuovi farmaci. Una informazione aggiuntiva proveniente da questo lavoro è che, stratificando i *dropout* per ragione della interruzione, si evidenziava un vantaggio statisticamente significativo per gli SSRI in termini di interruzioni per effetti collaterali (*odds ratio* 1.48, intervallo di confidenza 1.32, 1.66), ma anche un vantaggio statisticamente significativo per i vecchi TCA in termini di *dropout* per inefficacia (*odds ratio* 0.83, intervallo di confidenza 0.71, 0.98). Questo dato è importante perché da un lato conferma le impressioni dei medici prescrittori di SSRI, i quali ritengono i nuovi farmaci più tollerabili dei vecchi, dall'altro suggerisce che, per lo meno in termini di efficacia antidepressiva, i vecchi farmaci TCA possano rimanere i composti di riferimento.

L'ipotesi della maggiore efficacia dei vecchi farmaci rispetto agli SSRI è stata testata in un ulteriore lavoro metanalitico che confrontava il più utilizzato e noto AD di vecchia generazione, l'amitriptilina, con tutti gli altri AD, triciclici e SSRI (Barbui & Hotopf, 2001a). La metanalisi è stata condotta su di un campione totale di 146 RCT pubblicati nel corso di oltre 40 anni, dai primi anni '60 ai giorni nostri. L'esito veniva misurato mediante quattro indicatori: *responder*, ossia proporzione di soggetti che al termine dello studio erano in remissione sintomatologica (variabile dicotomica); punteggi medi alla scala di Hamilton al termine dello studio (variabile continua); proporzione di

soggetti che interrompono lo studio (variabile dicotomica); proporzione di soggetti che lamentano effetti collaterali (variabile dicotomica). Un aspetto di metodo importante in questo lavoro è che l'analisi dell'efficacia escludeva i soggetti che non avevano terminato lo studio (analisi endpoint) nel confronto dei punteggi alla scala di Hamilton (giocoforza, poiché tale punteggio era disponibile solo per i soggetti ai quali lo strumento era stato somministrato, e cioè a quelli che terminavano lo studio). Tuttavia, l'analisi dei *responder* includeva tutti i soggetti, anche i *dropout*, e li considerava come fallimenti terapeutici (analisi *intention-to-treat*). La metanalisi dimostrava un piccolo, ma statisticamente significativo, vantaggio per l'amitriptilina in termini di *responder* e di punteggi medi alla scala di Hamilton rispetto agli AD di confronto; al contrario, in termini di tollerabilità, l'amitriptilina evidenziava un profilo peggiore rispetto agli altri farmaci, sia in termini di tassi di *dropout* che in termini di pazienti che lamentavano effetti collaterali.

EPIDEMIOLOGIA DEGLI ANTIDEPRESSIVI NELLA PRATICA CLINICA

La medicina basata sulle prove di efficacia presuppone che le evidenze generate in contesti sperimentali vengano tradotte in pratica. Per quanto riguarda gli AD, le prove di efficacia suggeriscono che gli SSRI sono farmaci gravati da minore rischio di effetti collaterali, tuttavia le differenze quantitative con i vecchi TCA sono modeste. In termini di efficacia i farmaci di riferimento rimangono i TCA, disponibili già da molti anni; ad ogni modo, il diverso profilo di tollerabilità degli SSRI rende questi composti potenzialmente idonei per essere utilizzati in determinate popolazioni di pazienti, per esempio, quelle in cui gli effetti collaterali dei TCA potrebbero essere dannosi.

Le evidenze provenienti dal mondo sperimentale si scontrano tuttavia con la pratica clinica, dove gli AD vengono utilizzati in modo difforme da come le evidenze stesse suggerirebbero. Un primo dato interessante è la progressiva espansione del mercato di questi farmaci: in Italia, negli ultimi 15 anni le vendite sono aumentate di ben oltre il 50%; questo aumento, grosso modo, corrisponde al volume di vendite degli SSRI. Non si tratta cioè di un processo di sostituzione (si prescrivono gli SSRI nei casi in cui, prima, venivano impiegati i TCA), quanto piuttosto di un processo additivo (si utilizzano gli SSRI nei soggetti

che un tempo non ricevevano alcun trattamento). Tale trend si è mantenuto anche negli ultimi anni e, ormai, il volume di vendite degli SSRI ha superato quello dei TCA (Barbui *et al.*, 2001). Diviene allora interessante caratterizzare le popolazioni di soggetti che ricevono questi farmaci. A questo proposito, recentemente è stata portata a termine una indagine di farmacoepidemiologia che utilizzava l'archivio delle ricette di AD di 6 farmacie territoriali piemontesi (Pietraru *et al.*, 2001). Le farmacie territoriali archiviano tutte le ricette dei farmaci rimborsati dal Servizio Sanitario Nazionale (SSN) e, dunque, costituiscono un osservatorio privilegiato per questo tipo di indagini. In un periodo indice di 6 mesi sono state identificate tutte le ricette di AD compilate in un territorio di oltre un milione di abitanti. Si trattava di ricette compilate da medici di medicina generale, psichiatri, neurologi e altri specialisti, pubblici o privati. Tutte le ricette erano nell'archivio, purché si riferissero a farmaci prescritti in regime extraospedaliero e rimborsati dal SSN. Dal totale delle ricette si è quindi risaliti al totale di soggetti che, nel periodo indice, avevano ricevuto una o più prescrizioni di AD. Si è così costituito un campione di oltre 22000 soggetti, per una prevalenza d'uso del 21 per 1000 circa. In altre parole, circa il 2% della popolazione studiata aveva ricevuto almeno una prescrizione di AD nei 6 mesi indice. La distribuzione dei tassi per sesso ed età evidenziava due dati interessanti, il primo atteso, il secondo inatteso. Il dato atteso si riferisce al fatto che i tassi di prevalenza d'uso degli AD erano circa doppi nel sesso femminile rispetto a quello maschile. Tale dato è comprensibile alla luce della maggiore prevalenza delle sindromi affettive nel sesso femminile. Il dato inaspettato, viceversa, è che i tassi di prevalenza d'uso degli AD aumentavano progressivamente all'aumentare dell'età ed erano massimi negli anziani e nei grandi anziani. Questo è meno comprensibile, se si considera che la prevalenza delle sindromi depressive è sostanzialmente costante al crescere dell'età. Se la prevalenza di depressione è costante al crescere dell'età, mentre la prevalenza d'uso di AD aumenta progressivamente, si può ipotizzare che la prescrizione di AD negli anziani avvenga per problemi e sintomi non immediatamente riconducibili a forme depressive, quanto piuttosto a forme di disagio psicologico non ben specificate che i medici di medicina generale affrontano prescrivendo farmaci AD. Questa ipotesi è suffragata da analoghi studi condotti in altri contesti assistenziali che mostrano un utilizzo di AD che non si limita alle

indicazioni approvate per questi farmaci. Ohayon *et al.* (1998) hanno evidenziato che circa un quarto degli utilizzatori di AD nel Regno Unito assumeva questi farmaci per dormire. Rouillon *et al.* (1996), in una coorte di popolazione generale francese, hanno dimostrato che gli AD venivano prescritti seguendo le indicazioni approvate solo nel 65% dei casi. In Italia il 34% circa degli utilizzatori di AD nei *setting* specialistici assume tali composti per colpire sindromi ansiose (Munizza *et al.*, 1995).

In aggiunta al problema dell'indicazione all'uso degli AD, l'analisi del database piemontese evidenziava anche quello relativo alla durata dei trattamenti (Pietraru *et al.*, 2001). La distribuzione dei pazienti per numero di prescrizioni di AD ricevute nel corso dei 6 mesi mostrava che circa il 50% del campione aveva ricevuto solo una o due prescrizioni di farmaco. Se, grossolanamente, assumiamo che una scatola di AD corrisponde ad un mese di terapia, si ottiene che la metà circa dei soggetti non ha protratto l'assunzione del farmaco oltre i 2 mesi. Viceversa, i soggetti con un numero di prescrizioni tali da coprire tutti i 6 mesi corrispondevano a meno di un quarto del campione. Anche questo è un dato interessante, che ulteriormente ribadisce la distanza tra la pratica clinica e le raccomandazioni provenienti dal mondo sperimentale, che suggeriscono di prolungare il trattamento per almeno 6-8 mesi.

Per quanto riguarda il tipo di AD prescritto, emergeva il primato della paroxetina. Il dato interessante, tuttavia, è che la distribuzione dei 5 AD più utilizzati per fasce di età metteva in risalto dei *pattern* di utilizzo difforni dalle indicazioni provenienti dal mondo sperimentale. I dati sperimentali suggeriscono cautela nell'utilizzo dei TCA nei soggetti anziani a causa del rischio di effetti collaterali; gli SSRI, scevri da rischi cardiovascolari e anticolinergici, sono viceversa, per lo meno teoricamente, adatti in queste fasce di età. I dati hanno, invece, mostrato che al crescere dell'età diminuiva l'uso degli SSRI e aumentava progressivamente quello dei vecchi TCA e di altri composti come il trazodone. Addirittura, nei soggetti ultrasettantacinquenni il trazodone era il farmaco più prescritto, seguito dall'amitriptilina. Questo andamento, che evidentemente non si spiega tenendo in considerazione le evidenze della letteratura, fornisce ulteriore testimonianza della discrepanza esistente tra mondo sperimentale e mondo reale.

LE RAGIONI DELLA DISCREPANZA TRA RICERCA E PRATICA

Tra le ragioni della difficoltà a tradurre in pratica le prove di efficacia provenienti dal mondo sperimentale vi è certamente la bassa qualità di queste ultime. Recentemente è stata portata a termine un'analisi qualitativa di oltre 300 *trial* sugli AD pubblicati a partire dagli anni '60 fino alla fine degli anni '90 (Barbui & Hotopf, 2001b). Per quanto riguarda la numerosità campionaria è sorprendente sottolineare che, nonostante il progressivo aumento nel corso degli anni del numero di studi che seguivano oltre 100 pazienti, negli anni '90 il 50% circa dei *trial* continuava ad arruolare meno di 100 soggetti. Anche la durata di questi studi è aumentata davvero poco nel corso degli anni: nell'ultima decade la maggior parte dei *trial* ha seguito i pazienti non oltre la sesta settimana. Ancora, l'utilizzo di criteri diagnostici è cambiato in maniera radicale: si è passati dall'uso di criteri impliciti (per esempio, venivano reclutati i soggetti che, a giudizio del medico, necessitavano di trattamento con AD), a criteri espliciti, in cui i soggetti reclutati dovevano soddisfare i criteri diagnostici del DSM o ICD ed avere una severità sintomatologica grave attestata dai punteggi ad una scala di valutazione. Chiaramente, se da un lato l'impiego di strumenti diagnostici sofisticati ha certamente aumentato la precisione nel selezionare i soggetti con depressione, aumentando la validità interna degli studi, ha purtroppo anche parallelamente contribuito a diminuirne la validità esterna, rendendo meno generalizzabili le informazioni: nella pratica clinica quotidiana, infatti, non si somministrano AD solamente a chi ha una diagnosi di depressione certificata dal DSM o dall'ICD e un punteggio elevato ad una scala per la depressione. Zimmerman *et al.* (2002), a questo proposito, hanno dimostrato che i soggetti inclusi nei RCT rappresentano solo una piccola frazione dei soggetti trattati con AD tutti i giorni. I dati dei *trial*, dunque, non si riferiscono alle popolazioni "reali" di pazienti. Ecco allora la difficoltà a tradurli in pratica. Lo stesso discorso è valido per gli indicatori di esito, il cui numero è aumentato progressivamente nel corso degli ultimi 40 anni: se da un lato tale aumento può essere interpretato in senso positivo, in quanto segnale di una cresciuta precisione nel definire il miglioramento sintomatologico, dall'altro esso pone, di nuovo, un problema di trasferibilità del dato, poiché non è vero che nella pratica clinica quotidiana si stabilisce se un

trattamento con AD è stato efficace mediante la somministrazione di strumenti di esito.

Queste considerazioni non sono nuove. Già molti anni fa Keller & Lavori (1988), in un articolo rimasto esemplare, notavano che i RCT studiano campioni di soggetti non rappresentativi dei pazienti che quotidianamente affollano gli ambulatori dei medici, valutano trattamenti che non vengono somministrati nella realtà (poiché nella realtà il farmaco AD è solo un frammento di un percorso terapeutico che comprende una miriade di interventi concomitanti) e misurano gli esiti tralasciando di utilizzare gli indicatori che normalmente vengono impiegati per affermare che un paziente "sta meglio". I *trial*, secondo Keller e Lavori, studiano "not my patients, not my treatments, not what I try to do".

MEDICINE-BASED EVIDENCE, A PREREQUISITE FOR EVIDENCE-BASED MEDICINE

In tutta la medicina, e non solo in ambito psichiatrico, il dibattito è aperto. Gli ultimi anni hanno visto affermarsi prepotentemente la metodologia delle SR e metanalisi della letteratura come strumento per ribadire che l'esercizio della medicina deve basarsi su prove di efficacia. Si parla addirittura di paradigma della medicina basata sulle evidenze. Allo stesso tempo si riconosce che la qualità delle cosiddette evidenze spesso è discutibile e che gli studi clinici controllati, e le SR di questi ultimi, devono incrementare il proprio standard qualitativo per produrre informazioni trasferibili alla pratica clinica quotidiana. In psichiatria, come sopra accennato, il cambiamento nel corso degli anni ha prodotto una sofisticazione sempre maggiore dei *trial* che, verosimilmente, ha essa stessa contribuito allo scollamento tra ricerca e pratica.

Esistono, tuttavia, alcuni segnali di cambiamento. Uno di questi è rappresentato dal razionale che sottende lo studio condotto negli Stati Uniti da Greg Simon e collaboratori (Simon *et al.*, 1996; 1999). Il gruppo di Seattle ha portato a termine un RCT in cui un campione di soggetti con depressione, reclutati nel *setting* della medicina generale, veniva randomizzato a ricevere un trattamento farmacologico con fluoxetina, imipramina o desipramina. Lo studio si proponeva di confrontare il profilo di tollerabilità, efficacia e costo di questi tre farmaci. Il *trial* non aveva fini registrativi, intendeva "solamente" capire se fosse più conveniente iniziare un trattamento con SSRI o TCA. A questo

fine, gli autori hanno adottato un disegno di studio assolutamente antitetico all'ortodossia attuale dei RCT, per i seguenti motivi:

- lo studio non era in doppio cieco;
- lo studio non selezionava i soggetti depressi utilizzando strumenti diagnostici validati e non identificava i casi più gravi per mezzo di scale di gravità;
- l'unico criterio di inclusione era il bisogno, a giudizio del medico di medicina generale, di un trattamento AD;
- lo studio non imponeva ai medici la somministrazione dei farmaci AD a una posologia prestabilita dal protocollo di studio; addirittura, i medici erano liberi di aumentare, diminuire, sospendere o cambiare farmaco se ve ne fosse stato bisogno da un punto di vista clinico.

Il senso di questa metodologia era quello di produrre evidenze facilmente trasferibili in pratica. A questo fine, le condizioni sperimentali hanno riprodotto quelle in cui, ogni giorno, i medici si vengono a trovare, ossia condizioni di relativa incertezza diagnostica, senza la possibilità di selezionare i casi in base ai punteggi ottenuti mediante la somministrazione di scale di gravità, con la necessità di modificare la posologia e la durata della terapia (Simon *et al.*, 1995).

Lo studio di Simon e collaboratori esemplifica chiaramente la posizione di molti ricercatori che teorizzano la necessità di ripartire dall'esercizio della medicina per sviluppare una nuova generazione di studi dal disegno innovativo. Lo slogan di questa posizione, parafrasando la terminologia "medicina basata sulle evidenze", è "evidenze basate sulla medicina" (figura 1). Il titolo di questo paragrafo conclusivo, tratto da un editoriale apparso sul *British Medical Journal* qualche anno fa (Knottnerus & Dinant, 1997), esprime con un gioco di parole questo approccio: la produzione di prove di efficacia deve tenere in maggiore considerazione ciò che avviene nella pratica clinica quotidiana, ossia la produzione di evidenze scientifiche deve trarre ispirazione dall'esercizio della medicina (figura 1). L'esercizio quotidiano della medicina, a sua volta, non può prescindere dalle prove di efficacia a nostra disposizione per ciascun argomento. Si viene a realizzare in questo modo un *continuum* tra ricerca e pratica virtuoso per entrambi i settori: per la ricerca, finalmente calata nella realtà, e per la pratica, finalmente basata sulle prove di efficacia.

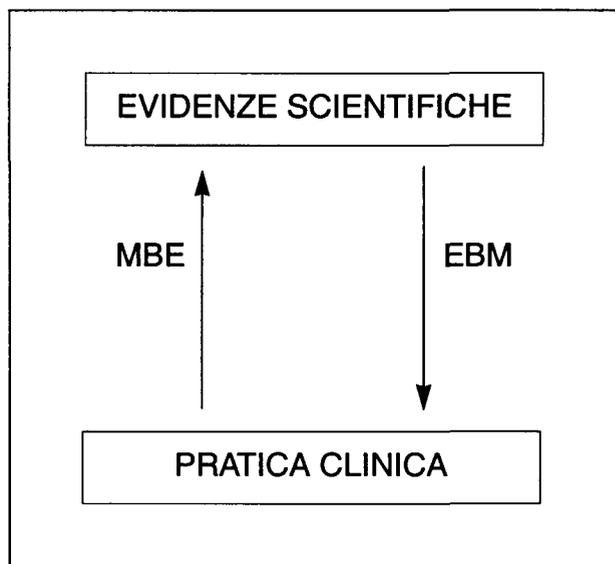


Figura 1. - Relazione tra evidenze scientifiche e pratica clinica quotidiana. Nella pratica clinica quotidiana è necessario basare l'esercizio della medicina sulle prove di efficacia (Evidence-Based Medicine, EBM). Viceversa, la produzione di prove di efficacia o evidenze deve avvenire a partire dalla pratica clinica quotidiana, cioè a partire dall'esercizio della medicina (Medicine-Based Evidence, MBE).

Si tratta di trovare un po' di coraggio per promuovere tali studi. I ricercatori sanno infatti che vi sono criteri di supposta qualità che debbono essere rispettati per poter vedere il proprio lavoro pubblicato sulle riviste di maggiore impatto (per esempio, i criteri diagnostici devono essere quelli del DSM-IV, vi deve essere la cecità, le diagnosi devono essere validate) e che la mancata osservanza di tali criteri espone alle critiche. Siamo, quindi, ad un bivio: o seguiamo la corrente e produciamo studi copia dei vecchi, di media qualità, che sicuramente verranno pubblicati su buone riviste internazionali e che verosimilmente avranno scarso impatto sulla pratica clinica quotidiana, oppure proviamo a percorrere strade nuove e testiamo ipotesi originali conducendo studi dal disegno innovativo. Questi studi, probabilmente, verranno criticati per il fatto di essere anomali, ma con il passare del tempo faranno cultura e, forse, permetteranno di generare evidenze scientifiche nuove, finalmente utili per migliorare la gestione farmacologica dei soggetti che quotidianamente seguiamo e trattiamo nei nostri servizi.

BIBLIOGRAFIA

- Anderson I.M. (2000). Selective serotonin reuptake inhibitors versus tricyclic antidepressants: a meta-analysis of efficacy and tolerability. *Journal of Affective Disorders* 58, 19-36.
- Anderson I.M. & Tomenson B.M. (1995). Treatment discontinuation with selective serotonin reuptake inhibitors compared with tricyclic antidepressants: a meta-analysis. *British Medical Journal* 310, 1433-1438.
- Barbui C. & Hotopf M. (2001a). Amitriptyline versus the rest: still the leading antidepressant after 40 years of randomised controlled trials. *British Journal of Psychiatry* 178, 129-144.
- Barbui C. & Hotopf M. (2001b). 40 years of antidepressant drug trials. *Acta Psychiatrica Scandinavica* 104, 92-95.
- Barbui C., Hotopf M., Freemantle N., Boyton J., Churchill R., Eccles M.P., Geddes J.R., Hardy R., Lewis G. & Mason J.M. (2000). Selective serotonin reuptake inhibitors versus tricyclic and heterocyclic antidepressants: comparison of drug adherence. In *The Cochrane Library*, Issue 4. Update Software: Oxford.
- Barbui C., Campomori A., Mezzalana L., Da Cas R. & Garattini S. (2001). Psychotropic drug use in Italy 1984-1999: impact of regulatory changes. *International Clinical Psychopharmacology* 16, 227-233.
- Hotopf M., Hardy R. & Lewis G. (1996). Discontinuation rates of SSRIs and tricyclic antidepressants: a meta-analysis and investigation of heterogeneity. *British Journal of Psychiatry* 170, 120-127.
- Keller M.B. & Lavori P.W. (1988). The adequacy of treating depression. *Journal of Nervous and Mental Disease* 176, 471-474.
- Knottnerus A. & Dinant G.J. (1997). Medicine based evidence, a prerequisite for evidence based medicine. *British Medical Journal* 315, 1109-1110.
- Montgomery S.A. & Kasper S. (1995). Comparison of compliance between serotonin reuptake inhibitors and tricyclic antidepressants: a meta-analysis. *International Clinical Psychopharmacology* 9, 33-40.
- Munizza C., Tibaldi F. & Bollini P. (1995). Prescription pattern of antidepressants in out-patient psychiatric practice. *Psychological Medicine* 25, 771-778.
- Ohayon M.M., Caulet M., Priest R.G. & Guilleminault C. (1998). Psychotropic medication consumption patterns in the UK general population. *Journal of Clinical Epidemiology* 150, 571-577.
- Pietraru C., Barbui C., Poggio L. & Tognoni G. (2001). Antidepressant drug prescribing in Italy, 2000: analysis of a general practice database. *European Journal of Clinical Pharmacology* 57, 605-609.
- Rouillon F., Blachier C., Dreyfus J.P., Bouhassira M. & Allicar M.P. (1996). Pharmaco-epidemiology study of the use of antidepressant drugs in the general population. *Encephale* 1, 39-48.
- Simon G.E., Wagner E.H. & VonKorff M. (1995). Cost-effectiveness comparisons using "real-world" randomised trials: the case of new antidepressant drugs. *Journal of Clinical Epidemiology* 48, 363-373.
- Simon G.E., VonKorff M., Heiligenstein J.H., Revicki D.A., Grothaus L., Katon W. & Wagner E.H. (1996). Initial antidepressant choice in primary care. Effectiveness and cost of fluoxetine vs tricyclic antidepressants. *Journal of the American Medical Association* 275, 1897-1902.
- Simon G.E., Heiligenstein J.H., Revicki D.A., VonKorff M., Katon W., Ludman E., Grothaus L. & Wagner E.H. (1999). Long-term outcomes of initial antidepressant drug choice in a "real-world" randomised trial. *Archives of Family Medicine* 8, 319-325.
- Song F., Freemantle N., Sheldon T.A., House A., Watson P., Long A., & Mason J. (1993). Selective serotonin reuptake inhibitors: meta-analysis of efficacy and acceptability. *British Medical Journal* 306, 683-687.
- Zimmerman M., Mattia J.I. & Posternack M.A. (2002). Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *American Journal of Psychiatry* 159, 469-473.