

1 • UN DECENNIO DI EBM: UN BILANCIO NON PROPRIO IMPARZIALE

Alessandro Liberati

INTRODUZIONE

Il 1992 è l'anno del battesimo ufficiale del termine evidence-based medicine (EBM) e dell'apparizione del primo di una serie di articoli pubblicati sul *Journal of the American Medical Association (JAMA)*¹ dalla quale ha avuto origine una lunga serie di contributi successivamente raccolti in testi divenuti *best seller*.^{2,3} Da allora la medicina ha attraversato anni molto critici per il proprio sviluppo e la propria credibilità, sociale prima ancora che scientifica. Le polarizzazioni pro o contro l'EBM hanno occupato, all'estero⁴⁻⁸ e in Italia,⁹⁻¹¹ la letteratura scientifica e non. È quindi tempo di tentare una lettura critica:

1. di quanto l'EBM abbia effettivamente mantenuto le molte aspettative che aveva generato;
2. se e quanto almeno parte di esse fossero praticamente, ed epistemologicamente, legittime;
3. come possa essere oggi proposta una visione «aggiornata» dell'EBM, che ne sappia preservare la carica innovativa attrezzandola alle sfide che, in generale, si sono rivelate assai più complesse di quanto alcuni dei suoi iniziatori avessero previsto.

Se questa è, in estrema sintesi, l'agenda di quanto vorrei considerare in questo capitolo, mi auguro che il lettore comprenda la necessità di affrontare il dibattito da diverse angolazioni. È infatti mia convinzione profonda che i limiti e i pregi di tutto quel fiorire di idee e spunti critici, e tutte le discussioni, talora anche molto aspre, nate attorno al ter-

mine «EBM» altro non siano che aspetti diversi e compositi delle tante reazioni che la crisi di credibilità della medicina e della professione medica, nei confronti della società e dei cittadini, e la crescente domanda di salute e partecipazione hanno posto all'*establishment* medico-scientifico.

Per affrontare tutte queste responsabilità l'EBM aveva «gambe metodologiche», e soprattutto culturali e filosofiche, troppo fragili. È per questo che si è trovata, ancor troppo giovane, ad essere gettata in un dibattito più grande di lei, al quale poteva sì portare qualche contributo, ma di cui certo non era la soluzione. E così, cedendo alla lusinga di soddisfare i propri tifosi, l'EBM non ha capito che – al di là delle importantissime questioni di metodo e culturali che aveva contribuito a sollevare – la posta in gioco era ben più alta e si dovevano fare i conti con questioni di grande rilievo. Quali lo sbilanciamento strutturale nei meccanismi di produzione delle conoscenze, la storica resistenza a politiche di effettivo *empowerment* dei cittadini e dei pazienti, la forte resistenza dell'*establishment* medico sia al cambiamento sia a sottoporre il proprio operato a verifica esterna, il forte ritardo culturale, oltretutto pratico, dei sistemi sanitari nel cogliere l'importanza della ricerca, dell'innovazione e del coinvolgimento attivo e partecipato dei cittadini. Tutto ciò ha reso in parte vani – o ha comunque fortemente mitigato – gli sforzi per migliorare la qualità delle prove sull'efficacia degli interventi sanitari in quanto troppo basati su un'ingenuamente positivistica fiducia nel fatto che il miglioramento degli standard metodologici fosse requisito sufficiente per una crescita della capacità di dare risposte rilevanti.

L'eccesso di aspettative create da molti sostenitori acritici e un po' *naive*, l'interesse non sempre disinteressato di chi ne ha colto un utile veicolo per il controllo dei costi in sanità, il sacrosanto entusiasmo di chi ne ha visto le potenzialità quale strumento di *empowerment* nei confronti di una medicina intrinsecamente autoritaria e paternalistica sono alcuni dei fattori che hanno fatto sì che l'EBM in questi primi dieci anni di vita sia stata allevata come creatura nata per diventare star, con genitori e familiari troppo ansiosi di successo per potersi permettere analisi più complesse, e che dovevano gioco forza proporre uno scenario epistemologico semplificato e, come tale, più facilmente vendibile.

Ma non è mai troppo tardi per diventare adulti. A patto però che l'EBM accetti l'idea che per avere una sana, e soprattutto utile, vita adulta sono necessarie alcune robuste iniezioni di complessità, indispensabili a garantire un'ordinata transizione e soprattutto una possibilità di dia-

logo con le altre discipline che contribuiranno a determinare il futuro dei sistemi sanitari.

Questo obiettivo pare non solo necessario, ma anche maturo, proprio ora che la fase delle promesse e delle semplificazioni è anche per l'EBM, per fortuna, terminata, e proprio ora che ci si comincia ad accorgere dei pericoli di queste eccessive semplificazioni. I segnali ci sono, a partire da una maggior complessità di analisi pur nella consapevolezza e, credo legittima, soddisfazione per aver constatato che molte delle reazioni che il movimento EBM ha suscitato derivano dall'aver saputo toccare le corde sensibili della relazione medico-paziente e del difficile equilibrio accettabile tra approccio scientifico alla pratica della medicina ed esperienza personale.

A pochi movimenti come a quello EBM è stata attribuita una tale molteplicità ed eterogeneità di obiettivi. Alcuni hanno infatti definito l'EBM come tentativo (estremo?) del mondo clinico di mantenere sotto il proprio controllo i processi decisionali per resistere alla crescente intrusione dei *policy maker* in un mondo, una volta, caratterizzato da sostanziale autonomia decisionale. Altri ne hanno criticato la prevalente attenzione verso i quesiti di tipo terapeutico (arrivando a sostenere che si dovrebbe piuttosto parlare di «evidence-based therapy» piuttosto che di «evidence-based medicine») con la conseguenza di eleggere a metodo standard di «disegno di ricerca» il trial clinico randomizzato. Altri ancora hanno visto nello sviluppo delle idee dell'EBM nient'altro che un ulteriore stadio del confronto tra scienza e progresso da un lato e mito e reazione dall'altro. Insomma un bel mix di questioni di vario spessore e portata!

Per tutti questi motivi nelle pagine che seguono si discuterà se l'attuale dibattito sull'EBM ha davvero toccato le questioni rilevanti o non si è invece un po' perso nel cercare di affrontare, tutte insieme, criticità come:

- la comprensione epistemologica di termini come «prova» ed «evidenza»;
- le resistenze ai necessari cambiamenti emerse dall'interno dell'*establishment* medico;
- le critiche strumentali formulate da certi «EBM-scettici»;
- i danni – magari involontari, ma non per questo meno gravi – causati dagli «EBM-entusiasti» creatori di ricette, acronimi e slogan ritenuti utili per un'EBM che doveva soprattutto far parlare di sé per affermarsi senza capire che quasi sempre l'eccesso di promesse, soprattutto se difficili o impossibili da mantenere, si paga.

COME È NATA L'EBM

Il termine EBM, come viene utilizzato oggi, è stato introdotto nel 1992 dallo stesso gruppo di clinici e ricercatori che anni prima aveva avviato la disciplina chiamata epidemiologia clinica (EC).^{1 12} L'EC nasceva sostanzialmente dall'idea di adattare ed espandere i metodi e le applicazioni dell'epidemiologia all'analisi dei processi decisionali sia nel campo dell'assistenza al paziente individuale sia in quello degli interventi mirati alle popolazioni. Da qui la definizione di EC come «lo studio delle relazioni di occorrenza delle decisioni mediche in relazione ai fattori che tali decisioni determinano».¹²

L'EC si è subito conquistata una grande popolarità grazie alla sua capacità di innovare le modalità di valutazione critica della qualità della letteratura medica e di apprendimento facendo sì che il «critical appraisal» si affermasse come una delle competenze essenziali per un medico moderno oltre a quelle di tipo diagnostico-terapeutico e di inter-relazione personale con il paziente.

Un primo importante by-prodotto dei concetti e dei metodi del critical appraisal è stata la documentazione – applicabile in grado variabile alle diverse aree della pratica clinica – che la qualità metodologica e la trasferibilità della letteratura medica sulla valutazione degli interventi sanitari era generalmente scadente.¹³

Questo ha certamente avuto il merito di esercitare un forte richiamo alla necessità di migliorare la qualità delle basi scientifiche della medicina moderna nella quale venivano troppo spesso utilizzati interventi di efficacia non documentata o non utilizzati invece interventi potenzialmente efficaci.¹³ Fu questo in estrema sintesi il retroterra attorno a cui nacque il primo articolo della serie del *JAMA* dove, per la prima volta appariva il termine «evidence-based medicine».¹

Se la qualità della maggior parte della ricerca non resiste a un rigoroso vaglio critico ne discende che il principio secondo cui «all medical action of diagnosis, prognosis and therapy should rely on solid quantitative evidence based on the best of clinical epidemiological research»^{1 2} sollecitava un atteggiamento ancor più critico, e di sano scetticismo, nei confronti di quegli interventi sostenuti dalla sola esperienza clinica (del singolo medico o tramandata dalla esperienza collettiva di esperti). Come puntualmente scandiva lo stesso articolo del *JAMA* aggiungendo: «we should be cautious about actions that are only based on experience or extrapolation from basic science».¹

Certo non si trattava di un concetto o di un affermazione nuova, così come il recente interesse per l'approfondimento delle origini del movimento EBM nella storia della medicina ha dimostrato.³ Vandenbroucke ha recentemente discusso alcuni importanti precedenti storici che ricordano molto alcuni slogan di fondo dell'EBM. Tra gli altri, Vanderbroucke cita Alexandre Louis che nel 1830 in Francia fu a capo della *Medecine d'Observation*.¹⁴ Louis incontrò, non sorprendentemente, forti resistenze tra i suoi contemporanei quando affermò che «i medici non dovevano basarsi su teorie e ipotesi circa la causa delle malattie, e neppure su esperienze derivate da casi singoli, ma piuttosto dovevano contribuire ad accumulare informazioni (esperienze collettive) basate su casistiche ampie, raccolte con criteri metodologici espliciti e definiti; solo così – concludeva Louis – si potrà capire quali trattamenti sono realmente efficaci e quali no».¹⁴

È interessante notare similarità e differenze tra l'epoca di Louis e quella, attuale, dell'EBM. All'inizio del XIX secolo i sostenitori della *Medicine d'Observation* reagivano nei confronti di una medicina che derivava le sue teorie da molti «fatti» che, con gli standard scientifici di oggi, noi considereremmo puro «non senso». Oggi l'EBM opera in un contesto assai diverso, caratterizzato da scienze mediche di base dotate di forte background sperimentale. La *Medicine d'Observation* scomparve poco dopo la sua nascita a causa di una forte reazione negativa dell'*establishment* medico-scientifico di allora e dell'assenza di condizioni sociali e culturali di contesto che potessero garantire quanto meno la permanenza di una dialettica così radicale.

Ciò che dobbiamo chiederci è se l'EBM farà, prima o poi, la stessa fine – per colpe proprie o per fattori esterni certamente più complessi di quelli dei tempi di Louis – nonostante operi in un mondo medico maggiormente orientato al metodo scientifico. Non sembri peregrino porsi questa domanda visti i crescenti segnali di pericolo e insofferenza che si addensano attorno all'EBM sia a causa delle risposte che essa non è stata in grado di dare (dopo essersi magari un po' incautamente sovraesposta) sia a motivo della capacità, già largamente fatta intravedere da parte degli interessi commerciali, di metabolizzare il patrimonio innovatore dell'EBM se questo non viene costantemente arricchito e tenuto all'altezza delle aspettative che ha generato.

Non c'è dubbio che una delle ragioni delle reazioni negative all'EBM sia stato l'averla considerata un «cambio di paradigma»^{19 14} nel senso attribuito a questo termine dal filosofo della scienza Thomas Khun.

Questa definizione, anche da me un po' superficialmente accettata alcuni anni fa,⁹ comporterebbe che solo l'EBM significa medicina scientifica e che, di conseguenza, quella praticata in epoca pre-EBM non fosse tale *tout court*. Tutto ciò non è solamente semplicistico, ma è sostanzialmente sbagliato e tradisce un certo entusiasmo che non pochi all'inizio dell'EBM hanno avuto (anche per la necessità di farsi sentire da orecchie spesso molto resistenti al messaggio) e che – come recentemente anche Brian Haynes, uno dei pionieri del moderno movimento EBM, ha ammesso – tradisce una scarsa preoccupazione da parte dei fautori dell'EBM di costruire una base filosofica alla propria disciplina.¹⁵

In realtà, la differenza che deve essere sottolineata tra l'era pre- e post-EBM non sta nel fatto che nella prima gli operatori non usavano le prove empiriche («evidence»). La differenza sta invece nel fatto che in epoca pre-EBM non esisteva una cornice concettuale e metodologica condivisa per usare le prove empiriche in modo esplicito e sistematico. Potremmo dire, in altre parole, che l'eccesso di aspettativa sull'EBM è nato perché si è scambiato uno strumento utile per la produzione e l'uso delle informazioni in medicina per una nuova definizione dello statuto epistemologico della medicina stessa.

Vista in questa prospettiva, l'attuale controversia sulla natura dell'EBM potrebbe passare dal terreno dei principi a uno maggiormente pragmatico nel quale l'attenzione è orientata al miglior uso delle informazioni scientifiche disponibili (riconoscendone a fondo anche tutti i limiti!). Alla ricerca disperata di un nuovo acronimo potremmo proporre di passare dalla EBM alla BUEM («better use of evidence in medicine»). Questo avrebbe l'enorme vantaggio di indicare che il vero, e sostanziale, cambiamento da introdurre è nel modo e nelle regole con cui usiamo e integriamo le informazioni scientifiche nella pratica della medicina.

A differenza del buon senso basato sull'esperienza, insistere sulla necessità di utilizzare al meglio le prove empiriche significa sottolineare la convinzione profonda che l'intuizione o il razionale fisiopatologico, da soli, rappresentano un terreno insufficiente per l'esercizio di un processo decisionale soddisfacente. Al contrario, una pratica moderna della medicina si alimenta dell'applicazione, intelligente ma rigorosa, di regole esplicite mirate a interpretare – in un modo che sia il meno possibile influenzato dai propri bias personali o professionali – le informazioni derivanti dalla ricerca clinico-epidemiologica. Queste regole esplicite devono integrare e non sostituire, nella forma-

zione del medico, l'accumulazione dell'esperienza personale che tuttavia, da sola, non è più eticamente accettabile come riferimento per la pratica professionale.

Impegnarsi per una migliore applicazione delle prove di efficacia in medicina ha anche altri importanti vantaggi. Rende più facile superare l'autoritarismo insito in molti stili di pratica ancora assai diffusi e rende assai più cogente il principio del diritto del paziente di partecipare alle scelte che riguardano la propria salute/malattia. Aiuta poi a meglio cogliere lo sbilanciamento selettivo che esiste nella «epidemiologia delle conoscenze» del tutto squilibrata a favore degli interventi riparativi semplici e a sfavore di quelli preventivi o riabilitativi complessi. Acquisendo quel bagaglio di «critical appraisal skills» che, lo ricordo, rappresenta la più importante innovazione del movimento EBM credo si finisca per acquisire quella «immunizzazione metodologica» contro le false certezze e le facili novità che è, alla fine, il bene formativo più importante per un operatore sanitario.

LE MOLTE FACCE DELL'EVIDENZA PROVE, CAUSALITÀ E INCERTEZZA

Dopo aver brevemente rivisitato il quadro di riferimento vorrei ora dare un pò di spazio alla discussione del concetto di «prova» per cercare di stabilire un confine tra medicina scientifica e tutto quello che, in un *continuum* difficilmente categorizzabile, porta a una riduzione del *tasso di scientificità* fino ad arrivare, a fine corsa, alla ciarlataneria. Una delle consapevolezze più importanti del movimento EBM è che il ricorso al «razionale fisiopatologico», per spiegare e giustificare l'effetto di un intervento sanitario, ha fallito ripetutamente nella storia della medicina. La diffusa pratica della flebotomia nel XVIII secolo aveva una qualche base fisiopatologica, ma certo nessuna efficacia. I più recenti esempi della controversia sull'utilità dei farmaci antiaritmici nell'infarto miocardico¹⁶ e della terapia ormonale sostitutiva¹⁷ sono lì a indicare come la valutazione empirica degli interventi possa, anche se con notevole fatica, dimostrare che interventi ritenuti «logicamente» utili o efficaci possano essere in realtà inutili quando non dannosi. La definizione, concettuale e operativa, di che cosa accettiamo come «prova» è un problema di regole e trasparenza, sia per la ricerca sia per la pratica clinica.

Sfortunatamente, la presenza/assenza di prove non è mai così chiaramente definibile, per diverse ragioni:

- per molte pratiche cliniche, anche laddove sono disponibili i risultati di studi clinici controllati, spesso abbiamo a disposizione informazioni da «filtrare» visto che i risultati dei diversi studi non sempre concordano e dobbiamo pertanto valutare attentamente il valore delle differenze in termini di validità del disegno, rilevanza e trasferibilità dei risultati (box 1.1);
- in molti casi le informazioni di cui avremmo bisogno non sono disponibili poiché gli studi necessari non sono stati condotti oppure perché non era eticamente e praticamente possibile condurre gli studi che sarebbe stato necessario fare;
- in altri casi la qualità dei randomised controlled trial (RCT) è scadente e anche le eventuali revisioni sistematiche o metanalisi non sono facilmente interpretabili.

Per quanto siano molti i casi nei quali la realizzazione di revisioni sistematiche e metanalisi ha dato importanti risposte circa l'efficacia di specifici interventi, resta il fatto che, anche in era EBM, per molti interventi non abbiamo solide ed esauritive basi scientifiche che ci permettano di valutarne l'efficacia. Per di più, le informazioni sull'efficacia teorica degli interventi dovrebbero essere integrate con le preferenze dei pazienti, i vincoli di natura economica e finanziaria, la compatibilità con gli assetti organizzativi dei servizi, e altro ancora.

In poche parole l'obiettivo che ci si dovrebbe porre con la produzione delle linee-guida clinico-organizzative; tuttavia questo strumento ancora oggi viene difficilmente o raramente sviluppato con il rigore e la completezza necessaria,¹⁸⁻²¹ anche perché le informazioni sull'efficacia clinica sono una condizione necessaria ma tutt'altro che sufficiente per la sua implementazione.

Un modello utile per conciliare gli elementi contrastanti che devono convergere nella sfida tesa a integrare informazioni scientifiche e scelte di pratica clinica e politica sanitaria è quello suggerito dal filosofo John Mackie. Egli sostiene che la dimostrazione di causalità non può essere ridotta all'identificazione di una singola causa necessaria e sufficiente ma deve piuttosto essere concettualizzata e descritta in termini di più elementi che chiama, con acronimo inglese, INUS (*Insufficient non-redundant component of an unnecessary sufficient complex*).

BOX 1.1 - VALUTARE LE INFORMAZIONI PER LA PRATICA CLINICA: I LIMITI DELLE TRADIZIONALI "SCALE DI EVIDENZA" E GLI SVILUPPI DELLA RICERCA METODOLOGICA

A partire dagli anni Ottanta, lo sviluppo della metodologia EBM e delle linee-guida ha fatto crescere la necessità di definire regole per la valutazione "aggregata" delle informazioni prodotte su un determinato argomento. Il concetto è semplice e importante in quanto si tratta di cercare di standardizzare e fornire agli operatori sanitari regole cogenti per valutare criticamente gli studi pubblicati, definirne la validità e sintetizzarne l'utilità per la pratica clinica.²³ Il primo tentativo di produrre delle «classificazioni» della bontà e credibilità delle informazioni disponibili a favore o contro un determinato intervento risale alla Canadian Task Force on Preventive Health Care (CTFPHC),²⁴ un'organizzazione dedicata all'analisi sistematica e alla produzione di raccomandazioni relativamente all'implementazione di interventi di prevenzione. Nel 1979 la CTFPHC mise a punto la prima classificazione relativa alla qualità intrinseca delle informazioni («livelli di evidenza») e alla convinzione/impegno con cui una o più raccomandazioni devono essere implementate nella pratica clinica («forza delle raccomandazioni»). Tale schema era essenzialmente basato su una «gerarchizzazione» della qualità delle prove secondo la quale le informazioni derivate da studi controllati e randomizzati (RCT) avevano il massimo di validità interna, seguite da quelle derivate da studi controllati non randomizzati, da studi osservazionali e, al fondo, dalle informazioni derivate dai database o sostenute esclusivamente dalla opinione di esperti. La classificazione della «forza delle raccomandazioni» discendeva direttamente da questa sistematizzazione e veniva legata in modo lineare alla classificazione dei livelli di prova per cui una raccomandazione era «forte» (indicata con la lettera «A» nella classificazione canadese) quando derivava da un RCT. La maggiore attrattiva di questo sistema era la semplicità che ne rappresentava tuttavia anche il limite maggiore.²⁵ Proprio a causa della sua semplicità, infatti, l'applicazione di questo metodo richiedeva molte assunzioni implicite relative a come comportarsi nella classificazione in caso di risultati conflittuali di più RCT o in caso di studi non sperimentali molto ben condotti e assenza di RCT. Specificamente, l'applicazione rigida di questo metodo suscitava domande del tipo: a) i risultati di RCT piccoli e mal disegnati devono sempre essere considerati «top evidence»? b) se vi sono importanti incoerenze nei risultati dei diversi RCT (eterogeneità), possiamo ancora considerare le informazioni prodotte «top evidence»? c) i risultati di studi osservazionali di buona qualità devono invece sempre essere considerati prove di qualità e credibilità inferiore? Il sistema peraltro aveva un importante tallone di Achille nel legame di dipendenza unidirezionale creato tra «livelli di prova» e «forza delle raccomandazioni». Nella definizione della «forza delle raccomandazioni» ci si basava meccanicisticamente sul livello di prova senza considerare la reale rilevanza clinico-epidemiologica delle informazioni, l'effettiva trasferibilità e applicabilità, il trade-off tra benefici e rischi, e così via. Dopo l'esperienza canadese sono stati sviluppati diversi sistemi di classificazione, analoghi come idea di fondo ma più o meno articolati e flessibili per quanto riguarda la definizione dei livelli e il legame tra livelli di prova e forza delle raccomandazioni.²⁶ Nel 2001 si è formato un gruppo di lavoro internazionale denominato GRADE (Grade of Recommendation Assessment, Development and Evaluation) che si è posto l'obiettivo di rivalutare in modo critico i vari sistemi sin qui prodotti da diversi gruppi e organizzazioni internazionali²⁷ e di mettere a punto un nuovo approccio – denominato appunto GRADE – che fosse capace di superare i limiti di quelli esistenti. Alla fine

del 2003 questa prima fase del lavoro del gruppo è terminata e la proposta di un nuovo sistema di grading unificato è ora disponibile.²⁸

Le novità introdotte da questo nuovo sistema riguardano, soprattutto: a) la chiara separazione concettuale del costruito «livelli di evidenza/prova» da quello della «forza delle raccomandazioni»; b) l'attribuzione ai due costrutti delle dimensioni specifiche; c) la definizione di un set di regole esplicite per definire il giudizio complessivo di qualità in funzione di specifiche considerazioni legate alla natura del quesito da affrontare e alle caratteristiche del setting nel quale le raccomandazioni dovranno poi essere applicate.

Molti sono i problemi ancora aperti, visto che il nuovo sistema è piuttosto complesso da applicare, deve ancora essere finalizzato per quanto riguarda la considerazione delle implicazioni economiche, di equità e applicabilità in diversi contesti e deve essere associato ad approcci semplici e facilmente comprensibili di rappresentazione visiva/grafica dei risultati. Elementi che ad oggi non sono ancora disponibili.²⁹

Mackie illustra solitamente la propria teoria con l'esempio del perché «una casa brucia». Il «complesso causale» è formato dall'associazione tra la presenza del fuoco nel caminetto, di un forte vento, da un difetto nel sistema di allarme antincendio della casa e dal fatto che la casa è fatta di legno. Se analizziamo ciascuna singola componente, nessuna di esse basta al divampare dell'incendio e soltanto la loro coesistenza dà origine a un «complesso» sufficiente. E tuttavia il complesso non è «necessario» poiché la casa può bruciare in tanti modo differenti (ad esempio, perché qualcuno decide deliberatamente di incendiarla). Sebbene nessuno dei singoli elementi che compongono il complesso sia «sufficiente» almeno uno è però necessario (se nell'esempio di cui sopra eliminiamo il fuoco acceso nel caminetto l'intero «complesso» diviene non più sufficiente).

Proviamo a fare un paragone con una decisione clinica. Ogni clinico deve integrare differenti elementi per prendere una decisione. Pensiamo, ad esempio, di dover decidere se sottoporre o meno a intervento di ovariectomia una giovane donna (di età inferiore ai 50 anni) alla quale è stato diagnosticato un tumore al seno. Se ci basiamo sui risultati della revisione sistematica della Cochrane Library sull'argomento troviamo 12 studi²² che, in maggioranza, suggeriscono un vantaggio della ovariectomia in questo gruppo di pazienti, anche se nessuno degli studi singolarmente raggiunge la significatività statistica. La relativa metanalisi indica un odds ratio, statisticamente significativo, pari a 0,72, indicativo di una riduzione del 18% nella mortalità dopo questo trattamento. Cosa dovrebbe decidere l'oncologo? Dovrebbe decidere, sulla base dei dati della metanalisi, di prescrivere il trattamento pur se la base biologica che

ne spiega l'esistenza è piuttosto debole? Il punto è che l'esempio dell'ovariectomia, lungi dal rappresentare un'eccezione, rappresenta un caso piuttosto frequente: solamente in una minoranza di casi uno studio, o una metanalisi di più studi condotti su uno stesso argomento, o anche una spiegazione fisiopatologica, possono essere considerati conclusivi. In altre parole dobbiamo aspettarci di trovarci spesso in presenza di «aree grigie». Da una parte l'oncologo potrebbe decidere che, sulla base della metanalisi Cochrane, le prove sono sufficientemente robuste in quanto esiste una buona base fisiopatologica a priori relativamente alla sensibilità del tumore del seno alla risposta ormonale nelle donne giovani. Così facendo, egli interpreterebbe i risultati degli studi empirici nel contesto di una plausibilità a priori, decidendo di ricorrere all'ovariectomia. E tuttavia un altro medico porrebbe fare un ragionamento esattamente opposto: dal momento che il vantaggio è modesto (riduzione della mortalità del 18%) soprattutto nelle donne che sono già state trattate con chemioterapia, e dal momento che vi possono essere effetti indesiderati importanti, compresi i problemi di limitazione della funzione riproduttiva, il valore da attribuire alle prove è limitato tanto da portare a decidere di non proporre il trattamento alla paziente.

In questo contesto è molto importante comprendere a fondo un'implicazione della definizione di INUS proposta da Mackie: il fatto che almeno una delle componenti del complesso deve essere necessaria (non-ridondante). Credo che si possa dire che questa componente «necessaria» sia proprio l'esistenza di prove empiriche: in assenza di prove empiriche è difficile sostenere che vi sia stata una buona decisione clinica. Se accettiamo che le prove empiriche siano componenti necessarie, resta comunque il fatto che il peso da dare a queste prove dipende dalla definizione stessa che diamo al termine «prova». È il concetto stesso di efficacia (o meglio di efficacia nella pratica, cioè l'*effectiveness*) che diventa un concetto «sfuggente/elusivo». Le prove empiriche non sono quasi mai chiare e definite in senso assoluto, cioè definite sulla base di un'unica proprietà, e tendono quindi ad essere «elusive». Il concetto di efficacia, in particolare, non può essere definito sulla base di un'unica proprietà (nel caso prima discusso, capacità di ridurre la mortalità) ma dell'insieme di diverse proprietà che tendono a sovrapporsi in casi specifici. Per alcuni, infatti, il concetto di efficacia è essenzialmente soggettivo mentre per altri è essenzialmente oggettivo, ma nessuna definizione specifica sembra in grado di soddisfare tutti. Questo significa che i risultati dei trial (o delle loro revisioni sistematiche e metanalisi) non possono

rappresentare l'unica fonte di informazione per prendere decisioni in campo di assistenza sanitaria: sebbene le evidenze/prove empiriche siano una componente necessaria è comunque compito dell'operatore sanitario integrare i diversi tipi di conoscenza e verificare che esistano anche i presupposti di plausibilità e fattibilità senza i quali nessun intervento teoricamente efficace è trasferibile nella pratica clinica. La breve discussione sull'esistenza o meno della «evidence-based homeopathy» bene illustra l'ambiguità di definizioni eccessivamente riduttive di evidence-based medicine (box 1.2).

BOX 1.2 - PUÒ ESISTERE L'EVIDENCE-BASED HOMEOPATHY?

L'omeopatia è certamente tra le forme della cosiddetta «medicina complementare» quella che più vigorosamente sta cercando di guadagnarsi uno statuto di scientificità. I dati attestano una sua crescente diffusione rispetto alla medicina convenzionale in virtù di una maggiore «naturalità» delle medicine omeopatiche e per reazione a una spersonalizzazione del rapporto medico-paziente nella medicina ufficiale.

In questa marcia verso il riconoscimento, parte del movimento dell'omeopatia sta cercando di incorporare i principi dell'EBM enfatizzando soprattutto il principio – stressato da quest'ultima – secondo cui i risultati di ricerche «patient-oriented» sono *necessari e sufficienti* per stabilire se un trattamento funziona, anche in assenza di un razionale scientifico e biologico. Rovesciando dialetticamente uno degli aspetti dell'EBM che più ha fatto discutere in questi anni, la parte «più scientifica» dell'omeopatia si appella al principio secondo cui la non-esistenza di un razionale fisiopatologico non è un argomento sufficiente a demolire la possibile efficacia di un intervento. Se così è – procede questa corrente di pensiero interna all'omeopatia – se vi sono RCT che sostengono l'efficacia dei rimedi omeopatici non vale più l'obiezione della mancanza di presupposto scientifico.

Tralasciando il piccolo particolare che, alla prova dei fatti, la maggior parte degli omeopati rifiuta di sottoporre i propri rimedi ai test di rigorosi RCT giudicando il metodo eccessivamente riduzionistico e inadatto alla «complessità dei loro interventi», il vero problema è se il risultato di uno o più RCT – in assenza di qualunque ipotesi biologica o fisiopatologia – possa rappresentare una sufficiente prova della efficacia dell'omeopatia. Rimandando alla diretta consultazione dei diversi documenti presenti nel sito web della Società Italiana di Medicina Omeopatica (www.omeomed.org) – conviene riprendere ed espandere brevemente quanto discusso relativamente alla definizione di prova, causalità e incertezza nelle pagine precedenti di questo capitolo in quanto questo può essere utile a chiarire alcuni equivoci ingenerati da una certa interpretazione riduzionistica dell'EBM.

In un recente articolo Victora et al.⁷¹ discutono la rete di informazioni necessarie per stabilire in modo convincente un'inferenza causale. Secondo la loro concettualizzazione – espressa ancor più compiutamente in un loro precedente articolo⁷² – esistono due assi principali nella formulazione dell'inferenza causale sull'efficacia di un

intervento. Il primo riguarda il tipo di quesito e il secondo il grado di certezza che si ritiene necessario, in funzione dell'obiettivo della valutazione.

Relativamente a quest'ultimo il gradiente di certezza/incertezza prevede tre dimensioni denominate rispettivamente: a) adeguatezza, b) plausibilità, c) probabilità. La prima si riferisce alla documentazione che i presupposti perché l'intervento potesse esercitare la propria efficacia si siano verificati. La seconda, la plausibilità, al fatto che si sia ragionevolmente esclusa l'esistenza di fattori estranei («confondenti») in grado di determinare l'effetto osservato. Infine, la terza, la probabilità, è il momento nel quale si valuta e quantifica formalmente la possibilità che l'effetto osservato sia dovuto al confondimento, al caso o ad una reale differenza di efficacia tra gli interventi posti a confronto. Il gold standard per quest'ultima dimensione è l'RCT, mentre per gli altri due sono sufficienti valutazioni meno complesse (osservazionale o quasi sperimentale).

Il punto che qui si vuole sollevare riguarda quindi proprio il grado di «autosufficienza inferenziale» dell'RCT. La valutazione che si ottiene con un RCT, infatti, fornisce informazioni con alta validità interna ma povere sul piano della contestualizzazione della generalizzabilità («validità esterna»). Da questo punto di vista l'applicazione del metodo RCT alla valutazione di efficacia è utile e importante come step finale nella sequenza dell'inferenza causale, ma non permette – proprio per la standardizzazione imposta nel setting della valutazione controllata – la raccolta di informazioni che sono necessarie a valutare la sussistenza dei presupposti di plausibilità dell'effetto dell'intervento.

È quindi pericoloso alimentare equivoci e generare ambiguità attorno a uno dei messaggi-chiave dell'EBM: la dimostrazione di un effetto importante di un trattamento sulla salute deve essere documentato assieme all'esistenza di una razionale («efficacia potenziale»). Così come il razionale fisiopatologico da solo non basta, egualmente, la sola dimostrazione di un effetto senza che esistano presupposti per poterne spiegare l'occorrenza non può costituire per l'EBM elemento sufficiente per un'inferenza causale. In altre parole, non si può saltare la fase dell'adeguatezza e della plausibilità arrivando subito a quella della probabilità per affermare che un intervento funziona. E questo deve valere, in una visione matura, sia per l'omeopatia sia per qualunque altro tipo di intervento sanitario. Tra l'altro, solamente rendendo esplicite le assunzioni che guidano le formulazioni dell'inferenza in tutte e tre le sue fasi si potrà porre un argine al meccanicismo e alla mancanza di rilevanza che caratterizza oggi la continua proposta di pseudo-innovazioni nel campo dell'assistenza sanitaria. Insomma, la via dell'«evidence-based homeopathy» è ancora lunga e deve passare attraverso vagli ben più corposi di qualche singolo piccolo trial favorevole (a fronte peraltro della presumibile grande quantità di studi negativi non pubblicati).⁷³

NEMICI INTERNI ED ESTERNI

Ma le difficoltà che impediscono un utilizzo prudente e sistematico delle prove non derivano soltanto dai limiti intrinseci che le caratterizzano e dalle resistenze al cambiamento espresse dall'*establishment* medico-scientifico. Esistono anche quelli che, con una definizione forse un po'

dura, possono essere definiti i «nemici interni», e che da qui in avanti chiamerò «gli EBM-entusiasti». Mi riferisco a due tipologie molto diverse, per caratteristiche e motivazioni, di persone. Un primo gruppo è composto da quelli che, gattopardescamente, si sono fatti un nuovo guardaroba «vestendosi all'EBM» per continuare a fare «business as usual». Il secondo è fatto da coloro che, per eccesso di semplificazione, hanno contribuito a diffondere irrealistiche aspettative circa la capacità dell'EBM di incidere sulla produzione delle conoscenze e sulla pratica clinica.

Pur nella loro diversità, e nel differente giudizio personale sui componenti dei due gruppi, resta vero che il danno che entrambi provocano al movimento EBM è sostanziale. Nelle considerazioni che seguono discuterò appunto gli errori e le conseguenze più che i punti di partenza soggettivi. Nella speranza che quanto discuterò, sia pur ingeneroso verso alcuni singolarmente, serva comunque ad eliminare una volta per tutte quelle critiche che invece che all'EBM dovrebbero essere indirizzate a chi ne strumentalizza l'uso o ne dimentica i limiti.

Ricordiamoci anzitutto dello sbilanciamento strutturale che caratterizza l'agenda di ricerca in campo bio-medico. La crescente, e non contrastata, prevalenza degli interessi commerciali nel campo della salute (si veda il capitolo 6) ha prodotto sin qui, e appare destinata sempre di più a produrre, una distorsione profonda sulla tipologia e qualità dei programmi della ricerca.³⁰ Questo è particolarmente evidente nel campo degli studi sui farmaci dove ormai la predominanza degli interessi commerciali dell'industria è palese continuando a promuovere una ricerca centrata sulla messa a punto di nuovi farmaci «me-too», poca o nulla comparazione diretta tra farmaci, e un ricorso esasperato agli studi di «equivalenza» invece che di «superiorità» accompagnata dall'uso di end-point surrogati di dubbia o nulla rilevanza clinica per il paziente.³¹ Per contro si continua a investire poco sulla ricerca non farmacologica e su quella mirata ad affrontare quesiti di «care» (box 1.3).

Solo di recente vi è stata, in alcuni paesi, una certa attenzione allo sviluppo di un filone di ricerca legato maggiormente ai bisogni informativi dei servizi sanitari; questo filone si è focalizzato sull'impatto reale sulla salute e sugli aspetti economici e organizzativi della sanità tanto delle tecnologie quanto degli interventi.³³ Le risorse a disposizione per questo tipo di ricerca sono tuttavia ancora troppo limitate (in Italia non esiste un programma nazionale) e deve ancora essere sviluppata una cultura relativa all'importanza strategica di programmi di ricerca e sviluppo per la sostenibilità e sopravvivenza dei servizi sanitari nazionali.

BOX 1.3 - SBILANCIAMENTO STRUTTURALE NEI QUESITI DI RICERCA E REVISIONI SISTEMATICHE DI LETTERATURA

La Cochrane Collaboration (CC) è nata 10 anni fa con l'obiettivo di preparare, disseminare e mantenere aggiornati i risultati di revisioni sistematiche sull'efficacia dell'assistenza sanitaria. L'agenda di lavoro e le priorità dei diversi gruppi di revisione della CC sono largamente autodefinitive e riflettono sia la disponibilità di fondi sia la disponibilità di studi primari sui singoli argomenti. Esiste, in altre parole, il pericolo che l'attività di *systematic reviewing* si concentri sulle aree per le quali esiste maggior quantità di informazioni disponibili piuttosto che su quelle teoricamente importanti e prioritarie ma per le quali non esistono dati.

Per verificare quanto questo sia vero è stata condotta un'analisi a partire dall'Issue 2 del 2002 della Cochrane Library (CL) andando a verificare la tipologia delle revisioni sistematiche disponibili in differenti aree cliniche quali: Asma nei bambini, Mal di schiena, Tumori (seno, colon-retto e prostata), Gravidanza e parto, Demenze, Depressione, Malattie cardiovascolari, Protesi d'anca, Parkinson, Polmoniti, Schizofrenia.³² Dall'analisi è emerso che: a) nella CL vi erano oltre 400 revisioni sistematiche su questi argomenti che complessivamente analizzavano circa 700 tipi di intervento; b) il 73% degli interventi riguardava interventi terapeutici, il 20% interventi preventivi e il 7% interventi di «care»; c) questo sbilanciamento a favore degli interventi terapeutici era presente in tutte le aree con un range compreso tra il 45% per l'area cardiovascolare al 95% per l'asma pediatrico; d) il 53% degli interventi studiati riguardava la valutazione di efficacia di uno o più farmaci.

La stragrande maggioranza degli interventi testati riguardava quindi la sfera della decisione terapeutica individuale (93%) e solamente il 7% interventi di tipo organizzativo. Un simile squilibrio si verificava anche per la valutazione degli esiti, che riguardava la sfera clinica nell'87% e considerava invece gli assetti organizzativi solamente nel 13% dei casi.

Le conclusioni da trarre da questa parziale rilettura critica delle revisioni Cochrane in relazione alla loro rilevanza per la sfera delle decisioni di politica sanitaria, si possono così schematicamente riassumere: a) vi è una scarsa attenzione agli interventi di «care» rispetto a quelli di natura terapeutica e di fase acuta; b) l'attenzione agli esiti non clinici è limitata; c) si rischia di dare troppa importanza agli studi randomizzati a scapito di studi che per specifici quesiti e condizioni possono anche essere più appropriati degli RCT.

Ricordiamoci anche del fatto che il punto di vista e l'input degli utenti nell'agenda di ricerca è sostanzialmente assente e le stesse associazioni dei pazienti e consumatori rischiano di giocare ruoli quanto meno ambigui finendo il più delle volte per svolgere un'attività di lobby per una particolare malattia piuttosto che un'azione chiara e trasparente a favore di una «prioritarizzazione» dei bisogni sanitari.³⁴

La mancanza di indipendenza e trasparenza dell'informazione in campo sanitario, soprattutto per quanto riguarda quella sui farmaci, è un

altro dei fattori chiave che limitano l'impatto dell'EBM. Lo sbilanciamento tra informazione commerciale e informazione indipendente (si veda in proposito il capitolo 11) sta diventando così evidente e marcato da rappresentare una minaccia mortale per la sopravvivenza di un sistema credibile di produzione e disseminazione delle informazioni scientifiche. È ipocrita pensare che la ricerca di buona qualità prodotta secondo i principi dell'EBM possa sopravvivere in assenza di investimenti strutturali nel campo della ricerca, della formazione degli operatori condotta secondo approcci didattici e contenuti moderni e senza una politica e un programma di informazione indipendente. Non è un caso che in questo libro manchi un capitolo proprio sulla «formazione evidence-based», a testimonianza della sostanziale impermeabilità della nostra accademia a un approccio che metterebbe profondamente in crisi equilibri, consuetudini e prassi consolidati e obbligherebbe a valorizzare l'etica dell'incertezza; questa più di tutto metterebbe in crisi l'autoritarismo che ancora caratterizza il modo di insegnare la medicina.

Il recente esempio delle reazioni suscitate da un grande studio internazionale di confronto dell'efficacia di diversi farmaci antipertensivi sintetizzato nel box 1.4 e discusso in un editoriale apparso nel 2003 sul *British Medical Journal*³⁵ illustra quale sia la posta in gioco. Farmaci antipertensivi costosi sono stati promossi e spinti sul mercato per molti anni senza che ci fossero prove credibili di una loro maggiore efficacia rispetto ai più vecchi e molto meno costosi diuretici. Grazie a un grande studio internazionale finanziato dai National Institutes of Health americani è stato osservato che i «vecchi» diuretici possono svolgere il loro compito altrettanto bene dei nuovi farmaci.³⁶ Ma l'industria farmaceutica non poteva rimanere indifferente e si è così scatenata una massiccia campagna di «controinformazione» dove «opinion leader» più o meno prestigiosi si sono impegnati mettendo in gioco la personale credibilità.

BOX 1.4 - IL CASO DELL'ASSOCIAZIONE EUROPEA DELL'INDUSTRIA FARMACEUTICA (EFPIA) E DELLO STUDIO ALLHAT

Il 31 marzo 2003 il *British Medical Journal* ha dedicato un numero speciale al tema dello spinoso rapporto tra classe medica e industria farmaceutica dal titolo emblematico *Time to untangle doctors from drug companies*. Un editoriale³⁵ presente in quel numero del *BMJ* ha esaminato specificamente il problema del ruolo degli opinion leader nell'interpretazione e disseminazione delle informazioni derivate dalla ricerca citando due esempi che sono stati al centro del dibattito scientifico, alla fine

del 2002. Il primo è un documento dell'Associazione europea dell'industria farmaceutica (EFPIA), che attaccava frontalmente i sistemi sanitari pubblici a base universalistica giudicandoli colpevoli di un sostanziale razionamento che priverebbe larghi strati della popolazione di interventi farmacologici di provata efficacia. Il documento – come rileva l'editoriale – vestito ambiguamente da testo scientifico – e firmato da un docente universitario tedesco di *health policy* – in oltre 100 fitte pagine di testo denuncia come in almeno 16 aree terapeutiche vengano ignorate solide prove di efficacia a favore di interventi che non sono «offerta» ai cittadini di diversi paesi. Il documento EFPIA rappresenta un classico *prodotto di parte* che dovrebbe essere usato come materiale di insegnamento nelle scuole di medicina tanto è parziale nell'uso delle citazioni, nella sistematica omissione del riferimento a revisioni sistematiche di letteratura che sconfesserebbero molte delle affermazioni contenute nel testo basate invece solo sul riferimento a singoli studi positivi, nella citazione selettiva di studi che dimostrerebbero il sottoutilizzo di specifici interventi e nella sistematica omissione di studi – pur disponibili – che documenterebbero invece un eccesso di utilizzo.

L'altro esempio discusso nell'editoriale è quello dello studio ALLHAT, un trial di grandi dimensioni condotto negli USA e pubblicato sul *JAMA* alla fine del 2002³⁶ che confrontava l'efficacia di diuretici, calcio-antagonisti, e ACE-inibitori nel trattamento dell'ipertensione. Lo studio si è concluso dimostrando essenzialmente un'equivalenza del «vecchio diuretico» rispetto ai più costosi nuovi farmaci e – come detto in precedenza – ha determinato una forte reazione di «controinformazione» da parte delle ditte produttrici di questi ultimi farmaci che temevano di vedere messi fortemente in discussione i propri profitti se i risultati dello studio si fossero trasformati in una massiccia adozione dell'alternativa di pari efficacia e minor costo.³⁷ Come l'editoriale del *BMJ* prima citato faceva notare, allo studio ALLHAT sono state mosse molte critiche che hanno trovato un'ampia eco soprattutto su riviste italiane e all'interno di congressi e corsi di aggiornamento largamente finanziati dall'industria e con un massiccio reclutamento di opinion leader nostrani.

Queste polemiche e discussioni sono state solo in parte attenuate dalla pubblicazione di una documentata metanalisi³⁸ che ha sostanzialmente confermato come i risultati dell'ALLHAT fossero in realtà coerenti con quelli della maggioranza degli studi sulla terapia della ipertensione: l'insieme degli studi disponibili indica infatti che i diuretici tiazidici a basse dosi – fatte salve specifiche eccezioni in cui il paziente iperteso non può assumere diuretici per un'allergia o un effetto avverso non tollerato – appaiono più efficaci per i casi non complicati, rispetto ad altre terapie antipertensive, nel ridurre gli eventi cardiovascolari. Come peraltro discusso in un recente articolo del *Bollettino d'informazione sui farmaci*,³⁹ anche i due principali argomenti usati contro lo studio ALLHAT – e cioè il timore di effetti negativi sulla glicemia e la scarsa tollerabilità del diuretico a basse dosi rispetto ad altri farmaci – non sono più sostenibili. E tuttavia la massiccia opera di controinformazione nei confronti dello studio ALLHAT sembra aver funzionato se osserviamo i dati di prescrizione italiani a sei mesi di distanza dalla pubblicazione della metanalisi del *JAMA* e ad oltre un anno dalla pubblicazione dell'ALLHAT stesso. Non si sono infatti verificati quei cambiamenti prescrittivi che sarebbe stato auspicabile attendersi;⁴⁰ il consumo di clortalidone (diuretico) non associato ad altri principi attivi è stato solo dello 0,3% delle Dosi Definite Giornaliere (DDD) consumate.³⁹

L'assenza di indipendenza e il monopolio della informazione scientifica si manifesta anche attraverso una crescente medicalizzazione dei problemi di tutti i giorni che passa anche attraverso la creazione di «nuove malattie» (si veda il capitolo 6) come strumento per aumentare il ritorno economico del mercato sanitario.⁴¹ In questo scenario è molto pericoloso classificare come «evidence-based» quelle procedure per le quali esiste «qualche studio più o meno a supporto» e invece come «non evidence-based» quelle per le quali non ci sono studi disponibili. Ignorare la distorsione da «mancata ricerca» è uno dei più clamorosi errori dei cosiddetti «troppo entusiasti», che finiscono per far diventare EBM una parola d'ordine confondente, ad alto rischio di quella strumentalizzazione che invece è nata per evitare.⁴²

Il paternalismo che sottende l'idea secondo cui «gli esperti sono la miglior fonte di informazione disponibile» è un'altra grave minaccia all'EBM. Da questo deriva un'altra idea largamente diffusa in medicina e cioè che la crescente complessità tecnologica corrisponda necessariamente a una migliore assistenza. Sempre più la medicina viene frammentata in sub-specialità nelle quali i singoli hanno una conoscenza sempre più approfondita e dettagliata di uno spettro di problemi sempre più ristretto. Questa conoscenza specialistica – e le relative competenze tecniche necessarie per poter applicare in modo adeguato una specifica tecnologia – porta a una sovrastima del grado di applicabilità e, in definitiva, dell'effettiva resa complessiva di quella particolare tecnologia sullo stato di salute di una popolazione, sopravvalutandone i benefici e sottostimandone i rischi e i costi.⁴³ A questo si lega l'inestricabile conflitto di interessi fra il prestigio sociale e professionale degli «esperti» in un determinato settore (con tutto il correlato di accessibilità a fondi e opportunità di ricerca) e il successo, la diffusione e affermazione di quella particolare tecnologia di cui un esperto si fa «campione».⁴⁴ Anche in questo caso una visione miope e riduttiva dell'EBM può fare più danni che benefici se non si va al di là della pura e semplice esistenza di «un qualche studio» e non se ne analizza a fondo e criticamente l'importanza e rilevanza clinica. Non vi è dubbio che ciò sia accaduto in questo primo decennio di EBM. Basta pensare alla fioritura di corsi e di esperti dell'ultima ora, ansiosi – secondo le regole del buon marketing italiano un po' «pataccaro» – di vendere con entusiasmo il «tool kit» dell'«ebiemmistà militante» che – con un po' di competenze Medline e un'infarinatura sui livelli di evidenza e la forza delle raccomandazioni – si proponeva come facile risolutore di problemi di variabilità e inappropriatazza della pratica clinica.

La mancanza di consapevolezza dei problemi che ho solo sommariamente tratteggiato può provocare gravi danni alla credibilità e integrità dell'EBM. Assumere che: a) tutte le informazioni rilevanti sono disponibili nella letteratura pubblicata, e che basta solo saperle cercare e filtrare; b) il rigore e la correttezza formale degli studi è la sola vera dimensione che distingue ciò che è buono da ciò che non lo è per l'EBM – senza preoccuparsi troppo della rilevanza clinico-epidemiologica e del profilo costo-efficacia; c) le scelte di politica sanitaria debbano essere definite (piuttosto che più sobriamente «informate») dalle prove di efficacia clinica, è il maggiore limite intrinseco che deve essere seriamente considerato e adeguatamente contrastato se si vuole garantire un futuro positivo e utile all'EBM.

COSA ABBIAMO IMPARATO SIN QUI GRAZIE ALL'EBM?

Dopo aver accennato ai limiti epistemologici, strutturali e pratici dell'EBM mi pare opportuno chiedersi se comunque essa è stata utile e quali sono le criticità da risolvere per lo sviluppo e l'applicazione futura degli strumenti che essa ha messo a disposizione. A scanso di equivoci, e prima di entrare nella discussione specifica, dirò subito che la mia risposta è positiva: l'EBM è stata utile. Poiché tuttavia ritengo necessaria una visione e una pratica dell'EBM più matura e analitica proverò a proporre alcune riflessioni conclusive orientate a far procedere in modo costruttivo il dibattito su vantaggi e limiti dell'EBM.

Per ragioni dovute più all'esigenza di schematizzazione che di reale separazione concettuale ho cercato di distinguere tra questioni che hanno a che fare con i problemi legati alla pratica clinica (per cui rimando al capitolo 3 e al capitolo 8), alla ricerca e alla politica sanitaria (rimandando al capitolo 9).

Pratica clinica

Non vi è dubbio che l'enfasi posta dall'EBM sulla necessità di promuovere l'utilizzo degli interventi di documentata efficacia ha creato un'attenzione specifica alla verifica del grado di effettivo trasferimento delle conoscenze dalla ricerca alla pratica clinica. La ricerca in questo

campo ha coinciso con la progressiva acquisizione della consapevolezza che il modello semplicistico del trasferimento delle conoscenze⁴⁵ (figura 1.1) non funziona in assenza di uno sforzo sistematico e organizzato per ridurre la distanza tra il momento della produzione e quello dell'applicazione delle conoscenze sull'efficacia degli interventi.⁴⁶

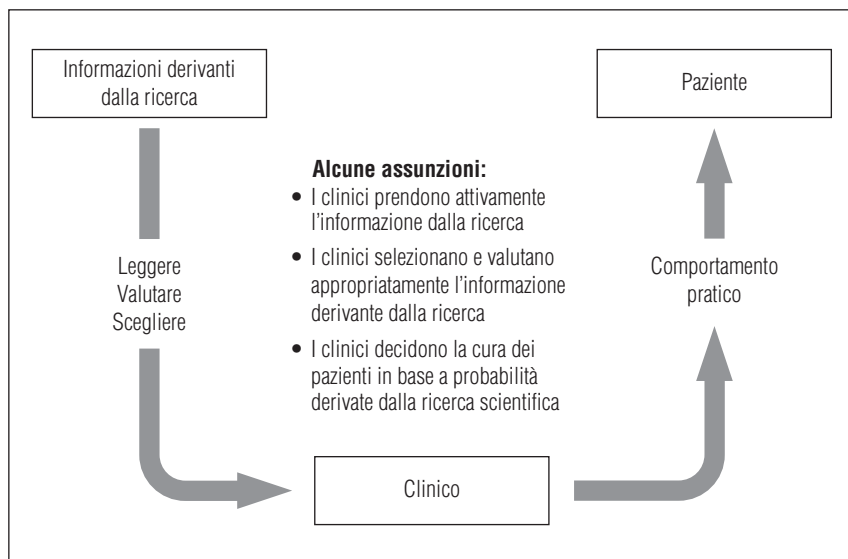


Figura 1.1 • Il modello semplicistico di divulgazione dell'informazione e cambiamento del comportamento «Come vorremmo che fosse»

L'ampia letteratura sulla «epidemiologia della qualità dell'assistenza» ha ben documentato che la pratica clinica è piena di interventi di dubbia o nulla efficacia mentre esistono spesso ritardi inspiegabili nel percorso di adozione di interventi di efficacia documentata.^{47 48} Come mai questo avviene è un quesito di notevole complessità al punto che si è andato delineando un campo specifico di ricerca per studi che analizzano i determinanti delle dinamiche di accettazione degli avanzamenti della ricerca nella pratica clinica.⁴⁹ Quello che la ricerca in questo campo ha contribuito a chiarire è che: a) la semplice diffusione passiva dell'informazione produce effetti nulli o molto modesti; b) le caratteristiche specifiche del messaggio (ad esempio, modifiche della pratica clinica) possono influenzare il risultato dell'intervento migliorativo; c) la

messa in atto di un intervento di implementazione richiede sforzo e impegno assai maggiore di quello necessario per la diffusione e la disseminazione e necessita di un'attenta analisi dei fattori favorevoli e di ostacolo al cambiamento (vedi il capitolo 8); d) insegnare le capacità critiche di interpretazione della letteratura scientifica («critical appraisal skill») ai singoli operatori sanitari non basta nemmeno a colmare quella «asimmetria informativa» che caratterizza le relazioni tra i «fornitori» e gli «utilizzatori» delle informazioni scientifiche.⁵⁰ Si è passati cioè dalla fiducia nel modello semplicistico della diffusione dell'informazione (figura 1.1) a concettualizzazioni più complesse del nesso tra disponibilità dell'informazione e suo impatto sulla qualità dell'assistenza e la salute dei pazienti (riassunte schematicamente nelle figure 1.2 e 1.3). Per tutte queste ragioni le informazioni che oggi abbiamo disponibili dagli studi sulla variabilità della pratica clinica^{51 52} e sugli insuccessi degli interventi di disseminazione passiva ci dicono che i sistemi sanitari hanno una specifica responsabilità nell'organizzare canali e modalità di accesso a un'informazione il più possibile indipendente e di buona qualità (si veda il capitolo 11). Attività che deve, ovviamente, essere considerata come condizione necessaria ma assolutamente non sufficiente e che deve soprattutto puntare a creare quella «cultura critica di fondo» in assenza della quale ogni altro sforzo di implementazione di politiche per l'efficacia e l'appropriatezza clinica sarebbe destinato al fallimento.

È necessario un profondo cambiamento culturale, prima ancora che scientifico, nel modo con il quale viene concepito il processo di accumulazione delle informazioni scientifiche. Uno dei più importanti messaggi dell'EBM è stato, in questo senso, il richiamo a considerare la totalità delle informazioni disponibili quando si tratta di capire se, nella pratica, un intervento funziona o meno. L'affinamento della metodologia delle revisioni sistematiche di letteratura e delle metanalisi – senza dubbio uno dei maggiori contributi che l'EBM ha dato alla metodologia clinico-epidemiologica in questo decennio – insieme con l'attività della Cochrane Collaboration⁵³ sono state di importanza fondamentale nel dimostrare i pericoli legati alla mancata sistematicità nella produzione di sintesi della letteratura. Inoltre, la ricerca metodologica condotta sui determinanti del bias di pubblicazione⁵⁴ ha permesso di meglio comprendere e dimensionare le distorsioni che possono essere indotte nelle sintesi di letteratura dai vari trucchi legati ai meccanismi di pubblicazione.

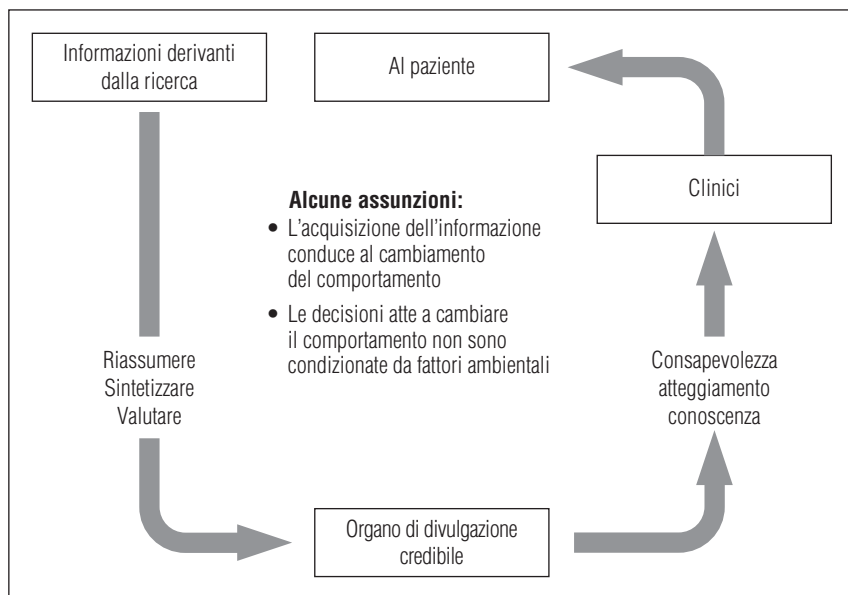


Figura 1.2 • Prima revisione del modello semplicistico: come veicolare il passaggio delle informazioni, ovvero «La nascita delle linee-guida»

Ricerca clinico-epidemiologica

La valutazione critica e la sintesi delle informazioni scientifiche di buona qualità, che nell'ultimo decennio hanno ricevuto una forte spinta culturale dal movimento EBM e dall'attività della Cochrane Collaboration⁵³ e delle diverse agenzie internazionali di Technology Assessment, hanno documentato in modo chiaro lo sbilanciamento strutturale nella agenda di ricerca e la mancanza di studi su aree molto rilevanti, soprattutto in campo di prevenzione e di assistenza nella fasi post-acute e di riabilitazione (box 1.2). Lo scarso supporto pubblico alla ricerca e il quasi sostanziale monopolio dell'industria farmaceutica per quanto riguarda la ricerca sui farmaci stanno creando una gap crescente che mette a forte rischio di irrilevanza e di dipendenza l'intero movimento EBM. Perché l'EBM possa mantenere le promesse e le aspettative che ha generato è necessario che le sue istanze siano accompagnate da un forte impegno pubblico nel sostegno alla ricerca nelle aree orfane.³⁰ Oltre a questo sbilanciamento strutturale vi è un'importante e crescente minac-

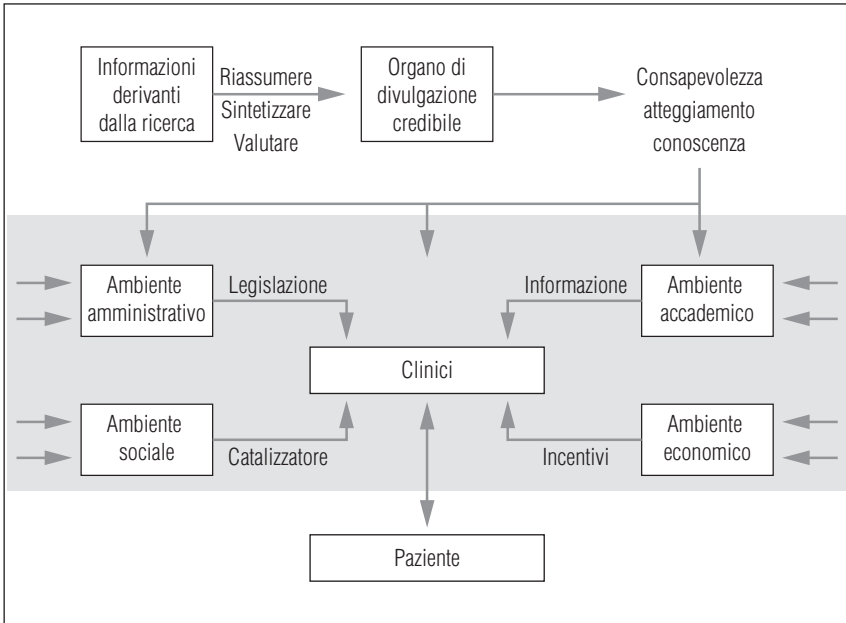


Figura 1.3 • Seconda revisione del modello semplificato
«Guardando dentro la scatola nera»

cia alla credibilità della ricerca scientifica e quindi all'EBM: l'emergere della rete di conflitti di interessi e la progressiva documentazione delle sue conseguenze sia per la ricerca sia per la pratica clinica. In modo particolare negli ultimi 5-6 anni il tema dei conflitti di interessi in campo medico sanitario è emerso con forza e si sono moltiplicati gli studi che hanno esplorato il potenziale ruolo distorsivo dell'influenza delle sponsorizzazioni specialmente per quel che riguarda la ricerca farmacologica.⁵⁵⁻⁵⁹ Ma la cosa più allarmante, da questo punto di vista, è che gli strumenti a disposizione dell'EBM si presentano oggi come «armi spuntate» ai fini di un effettivo esercizio di vigilanza e controllo apriori perché la vera soluzione sarebbe poter impedire che si moltiplichino e diffondano studi irrilevanti e produttori di risultati fuorvianti. Il problema, infatti, non è più quello di cogliere la presenza di errori nella metodologia di disegno e analisi dei trial quanto piuttosto quello di saper valutare se le ricerche sono state apriori disegnate in modo da far «avverare le aspettative» (dello sponsor in questo caso) «adottando trucchi» quali la scel-

ta di un gruppo di controllo inappropriato, utilizzando il placebo quando invece il nuovo farmaco dovrebbe essere testato contro il farmaco più attivo già disponibile, ritardando o sopprimendo le pubblicazioni quando lo studio non dia i risultati attesi o infine promuovendo soprattutto studi di «equivalenza» invece che di «superiorità» (si veda il capitolo 2).

Ancora una volta, per mantenere la credibilità e integrità della ricerca scientifica e di chi, come il movimento EBM, difende le regole di rigore e trasparenza, è necessario impegnarsi per un'azione di contrasto impostata su: a) la definizione di precise regole del gioco alle quali gli sponsor devono attenersi quando sostengono e finanziano la ricerca; b) la creazione di meccanismi internazionali di trasparenza stimolando la messa in opera di registri internazionali delle sperimentazioni cliniche* e impegnando le società scientifiche e il mondo professionale a un impegno etico e di trasparenza; c) la definizione condivisa di regole di dichiarazione («disclosure») e soglie di incompatibilità per la partecipazione ad attività di ricerca (si veda in proposito il sito web del Coordinamento per la integrità della ricerca biomedica, CIRB, www.cirb.it).

Proprio esempi come quello del CIRB in Italia, e altri ben più forti a livello internazionale, indicano che in questi ultimi dieci anni, grazie al contributo dell'EBM, si sono fatti passi avanti nella consapevolezza e nella salvaguardia di alcuni di questi obiettivi.

La strada è tuttavia in salita e le difficoltà crescono in assenza di una presa di coscienza della necessità di fermare la deriva commerciale e di mercato che la ricerca sta prendendo, soprattutto perché manca una controparte di ricerca pubblica e indipendente guidata dai bisogni operativi del servizio sanitario. L'esempio del potenziale innovativo offerto dalla Cochrane Collaboration e le difficoltà che questa iniziativa trova nel guadagnarsi il necessario supporto pubblico (tranne che in singoli paesi come l'Inghilterra, il Canada e l'Australia) non sono altro che un chiaro indicatore negativo di quanto sia vero ciò che ho appena affermato.

Tra gli operatori sanitari, gli amministratori e i pazienti bisogna far crescere la consapevolezza dell'importanza della ricerca come strumento di miglioramento continuo della qualità e appropriatezza delle prestazioni. È necessario uno sforzo massiccio di informazione per far crescere l'esigenza della ricerca come strumento per ridurre l'incertezza pre-

* Importantissima, a questo proposito, la presa di posizione degli editor delle maggiori riviste apparsa su JAMA del 15 settembre 2004, nel volume 292 alle pagine 1363-64.

sente in molte aree della pratica clinica e che – pur essendo in linea di principio ineliminabile – deve essere esplicitamente discussa e affrontata attraverso strategie di ricerca e sviluppo⁶⁰⁻⁶² mirate anche a impedire che certi interventi di non provata efficacia vengano comunque utilizzati perché non esiste alcuna credibile alternativa.⁶³ La partecipazione dei pazienti deve essere attivamente ricercata e stimolata al fine di aumentare la rilevanza e l'applicabilità dei risultati della ricerca e per facilitare una partecipazione realmente informata.

Cominciano infatti a essere disponibili anche prove empiriche che pazienti e cittadini, se coinvolti in modo appropriato quando si devono identificare le priorità, sono in grado di fornire stimoli utili per l'identificazione dei quesiti di ricerca.^{34 64} Si tratta evidentemente di un processo complesso che richiede una capacità di indirizzo e governo per evitare che le buone intenzioni si traducano in un'ulteriore frammentazione delle attività di ricerca e che chi ha interessi commerciali non usi questa nuova leva per spingere a favore della sperimentazione e della adozione precoce di tecnologie di non provata efficacia.⁶⁵ Anche in Italia qualche piccolo segnale in questa direzione comincia a vedersi con l'attivazione di progetti e iniziative (si veda in proposito il sito www.partecipasalute.it).

Politiche sanitarie

I risultati della ricerca e le prove di efficacia dovrebbero «informare», ma sono spesso inadeguate per orientare in modo specifico le scelte di politica sanitaria. Su questo il panorama del «mondo EBM» è piuttosto composito. Se è vero che vi è una parte tutt'altro che trascurabile (penso qui agli «EBM entusiasti») che non sembra avere coscienza di questo vi sono tuttavia numerosi segnali sul piano metodologico e della riflessione epistemologica che portano a ridimensionare le ambizioni eccessive di cui l'EBM ha certamente sofferto nella sua fase infantile. È quindi salutare che da settori estranei, almeno ufficialmente, ai «ranghi EBM» arrivino richiami forti al fatto che il tipo di informazioni prodotte dalla ricerca tradizionale, nel suo paradigma del RCT, hanno valore troppo limitato e mancano di informazioni ed elementi contestuali essenziali per il *policy making*.⁷⁰⁻⁷³

Un'altra importante consapevolezza che l'EBM ha contribuito a far crescere è quella dell'inefficienza intrinseca conseguente all'inazione di

fronte al permanere e diffondersi nella pratica di interventi inefficaci o ai tentativi di implementare l'innovazione con strategie inefficaci.

Il Technology Assessment – termine che riassume quell'insieme di attività che cercano di integrare gli elementi di *efficacy*, sicurezza, accettabilità e fattibilità di una tecnologia – è stato, nella maggior parte dei casi, un esercizio «retrospettivo» e il trasferimento delle conoscenze dalla ricerca alla pratica è stato considerato una sorta di processo evolutivo naturale derivante dall'emergere di novità da parte della ricerca di base. Questa ambiguità è ancora oggi perpetuata dalla confusione che circonda il termine «ricerca traslazionale», termine che di per sé perpetua e rinforza l'idea di una dipendenza strutturale dei sistemi sanitari dalla conoscenza «naturalmente» prodotta dal mondo scientifico, visto come entità separata. Non deve quindi sorprendere che voci autorevoli lancino l'allarme circa la relativa inefficienza di un sistema che produce conoscenze che in larga parte vanno poi perse nella fase di trasferimento,⁴⁶ non divenendo mai reali innovazioni. Per modificare questo stato di cose sembrano necessari alcuni cambiamenti di rotta sostanziali: a) una maggiore interazione, nelle fasi precoci di sviluppo delle tecnologie, tra produttori e potenziali utilizzatori intermedi e finali; b) lo sviluppo di una capacità di osservazione e predizione da parte delle strutture tecniche dei sistemi sanitari che devono, anche, saper essere interlocutori attivi nella concettualizzazione e nel disegno di programmi di valutazione controllata dell'introduzione di queste nuove potenziali innovazioni; c) la sperimentazione delle nuove tecnologie nella loro fase ancora immatura non deve più essere lasciata alla spontanea creatività dei clinici (sempre meno libera e sempre più condizionata dalla logica del ritorno commerciale)⁷⁴ ma deve essere guidata da competenze multidisciplinari e interprofessionali; d) la sperimentazione controllata delle nuove tecnologie non deve più essere limitata alla valutazione di efficacia ma deve includere, in modo sistematico e precoce, la valutazione delle implicazioni organizzative, sociali, economiche e finanziarie dell'innovazione.

CONCLUSIONI

Non c'è dubbio che l'illusione, forse coltivata da alcuni, che l'EBM fosse la risposta principale e autosufficiente ai principali problemi statuari ed epistemologici, oltreché pratici, della medicina debba essere definitivamente archiviata. D'altro canto è assolutamente necessario tentare

di definire aspettative credibili per l'EBM evitando il rischio di perdere una grossa occasione di innovazione culturale per il mondo scientifico e per le aspirazioni legittimamente crescenti di partecipazione dei pazienti alle scelte in campo sanitario.

L'EBM, se decide di caratterizzarsi soprattutto come un corpo di regole metodologiche ed etiche, può essere molto utile al processo di produzione di informazioni valide e rilevanti in campo sanitario. Culturalmente l'EBM è in grado di esprimere una spinta anti-autoritaria salutare per la promozione dell'interdisciplinarietà nell'approccio ai problemi clinici e assistenziali. La rilettura di questi 10 anni di grande popolarità dell'EBM ci dice che (anche per sua responsabilità «interna») il suo potenziale non è stato sfruttato appieno e che anzi si è perso tempo in dibattiti e polemiche che avrebbero, forse, potuto essere evitati se alcuni «EBM-entusiasti» non avessero avuto troppo spazio grazie anche a un corteggiamento tutt'altro che disinteressato da parte di chi pensava di piegare lo spirito dell'EBM alle proprie convenienze immediate.

Un limite chiaro di questi 10 anni è stata una concentrazione eccessiva sull'approfondimento monotematico della metodologia degli studi di efficacia intorno al modello degli RCT, particolarmente per lo studio dei farmaci. Non deve stupire quindi che oggi critici attenti ed esigenti sottolineino che, più che l'EBM, la metodologia ha in questi anni definito approcci alla EB-therapy con strumenti oggi formalmente robusti soprattutto per valutare interventi relativamente semplici, come sono per la maggior parte dei casi quelli farmacologici.

Il fatto che finora l'EBM abbia sviluppato strumenti limitati per valutare l'efficacia di interventi complessi deve essere visto non solo come segno dell'intrinseca maggiore complessità di queste aree ma anche come conseguenza di un minore investimento culturale e scientifico. Risvegliarsi bruscamente da queste crisi di certezza può essere salutare ma anche pericoloso se ci si limita a discutere sui limiti delle informazioni scientifiche dividendosi in dispute tra chi pensa di essere portatore della «risposta giusta» autoproclamandosi espressione della metodologia più rigorosa. La recente controversia sull'efficacia o meno dello screening mammografico esplosa dentro e fuori la Cochrane Collaboration ne è un esempio evidente (box 1.5).

Malgrado le molte limitazioni discusse, l'EBM ha già avuto effetti positivi in alcune aree della medicina nel migliorare la qualità della ricerca e nell'aumentare il grado di consapevolezza e di «sano scetticismo» degli operatori e dei *policy maker* nei confronti del mercato delle

false innovazioni. Molto, come detto nella parte finale di questo capitolo, rimane da fare per creare un ambiente culturale che: a) valorizzi, invece che negare, l'etica dell'ignoranza; b) rifletta a fondo sullo statuto epistemologico e sulla parzialità ed evolutività dei concetti di prova, evidenza, incertezza; c) si accorga della necessità di correggere lo sbilanciamento strutturale negli investimenti per la ricerca; d) sviluppi una visione più organica e integrata del rapporto tra ricerca, innovazione e pratica clinica; e) riconosca come priorità lo sviluppo di interventi mirati a dare maggior ruolo decisionale al paziente. Detto ciò, sarebbe interessante capire davvero quanto di tutto questo debba entrare a far parte dei buoni propositi che l'EBM deve formulare per se stessa o non vada ben al di là dipendendo, in ultima analisi, da quali saranno le politiche sanitarie e da cosa deciderà di fare la politica con la «P» maiuscola per il futuro della tutela della salute e della difesa di sistemi sanitari pubblici e universalistici.

BOX 1.5 - LA CONTROVERSIA SULL'EFFICACIA DELLO SCREENING MAMMOGRAFICO: CRISI DI NERVI O LIMITE STRUTTURALE DELL'EBM?

Verso la fine del 2001 è esploso nuovamente il dibattito sulla reale efficacia dello screening mammografico. Sulla qualità e sufficienza delle prove a favore o contro l'utilità dello screening nel ridurre la mortalità si era già discusso diverse volte nel recente passato. La novità questa volta è sembrata essere non solo e non tanto la natura di un verdetto scientifico cui una gran parte della popolazione femminile è ovviamente interessata, quanto piuttosto la criticità del dibattito stesso per il futuro e la credibilità dell'EBM.⁶⁶ Il dibattito attuale è stato innescato da una nuova analisi dei dati compiuta da due ricercatori danesi che hanno pubblicato il loro lavoro in tre forme: sulla Cochrane Library, sotto forma di un breve sommario sul *Lancet* in versione cartacea e come rapporto esteso nella versione elettronica della stessa rivista.

La novità introdotta dai due danesi riguarda il «livello soglia» utilizzato nella valutazione della qualità metodologica degli otto studi randomizzati sull'efficacia dello screening. A causa di irregolarità più o meno gravi in fase di reclutamento, follow-up e analisi, solamente due studi potevano – secondo gli autori danesi – essere giudicati di qualità accettabile e considerati per una metanalisi. Il verdetto fu negativo sulla efficacia dello screening mammografico. Stimolati da questa analisi, altri organismi e agenzie internazionali hanno rianalizzato i dati, accettando in alcuni casi ma rifiutando in altri la reale importanza dei rilievi posti da Gotzsche e Olsen e comunque non concordando sulla opportunità di limitare la metanalisi a due soli studi. Perché questo è sembrato a molti un punto di crisi dell'EBM?

A parte la malignità di alcuni rilievi che l'Editor del *Lancet*, Richard Horton, sollevò nel suo editoriale⁶⁷ contro i membri del gruppo editoriale Cochrane Breast Cancer,⁶⁸

per giustificare la sua decisione di ospitare la controversa analisi di Gotzsche e Olsen sulla sua rivista, resta il fatto che questa vicenda ha rappresentato un vero e proprio «caso editoriale». Assunto fondamentale dell'EBM è proprio che la dimostrazione scientifica derivata da studi randomizzati e studi controllati di buona qualità debba avere la precedenza rispetto all'opinione degli esperti. Questa controversia sembra tuttavia far rientrare dalla finestra il vecchio «parere dell'esperto» che questa volta si traveste da metodologo per decidere quali sono gli studi buoni e quali quelli cattivi. Tutto ciò appare poco differente dalle tradizionali controversie tra esperti, usuali in epoca pre-EBM, e che proprio i fautori della novità e superiorità dell'EBM avevano aspramente criticato. Certo, si può sempre ribattere che è proprio grazie al fatto che l'EBM impone criteri espliciti che un dibattito di questo tipo può avvenire. Il problema vero è che la radice sostanziale del disaccordo sta nell'importanza che si dà alla valutazione della qualità metodologica degli RCT e soprattutto nel come utilizzare poi questa valutazione di qualità in fase di analisi dei risultati.

I due autori danesi e la US Preventive Task Force, per esempio, arrivano a valutazioni abbastanza simili circa la qualità dei trial ma a decisioni opposte sul come trattare (escludendo o includendo i dati in fase di analisi, rispettivamente). Entrambi sostenendo che l'approccio da loro scelto è preferibile perché porta a conclusioni più vicine al vero. Il punto critico in questo caso sembra riguardare il fondamento epistemologico di una visione eccessivamente riduzionistica e meccanicistica dell'EBM secondo la quale essa presume di poter trovare una base sicura per l'affermazione della verità scientifica solamente facendo ricorso alla osservazione empirica («evidence»).⁶⁹ Questa convinzione è destinata a fallire in quanto assume che le osservazioni siano indipendenti da una teoria sottostante relativa a cosa rende l'osservazione stessa affidabile. Nello specifico dell'EBM, quando le osservazioni vengono fatte attraverso un RCT (che, ricordiamo, è considerato lo standard) la teoria deve dirci che cosa rende un RCT «sufficientemente valido». Ed è proprio a questa domanda che l'EBM oggi non sa ancora rispondere. Non deve stupire, a questo punto, che se non vi è consenso sulla teoria non vi può essere consenso sui fatti osservati.

Resta solo da dire, a conclusione di questa vicenda, che in una situazione di incertezza dovrebbero essere coloro che vogliono escludere una parte delle informazioni disponibili a convincerci della giustezza della loro posizione, piuttosto che il contrario. E tuttavia bisogna essere coscienti che controversie simili a quella esplosa attorno allo screening mammografico saranno sempre più frequenti in futuro, man mano che ci si cimenterà in modo critico nella valutazione di diverse tecnologie e interventi sanitari. Il risultato di queste valutazioni ci farà spostare sempre più da aree di certezza ad aree di relativa incertezza ed è per questo che una delle priorità dello sviluppo epistemologico e metodologico della EBM è lo sviluppo di approcci maggiormente espliciti e complessi proprio al trattamento dell'incertezza.

BIBLIOGRAFIA

1. Evidence-based Medicine Working Group. Evidence-based medicine: a new approach to the teaching of medicine. *JAMA* 1992; 268: 2420-5.
2. Sackett DL, Richardson WS, Rosenberg W et al. Evidence based medicine: how to practice and teach EBM. New York, Edimburgh: Churchill Livingstone, 1997.
3. Guyatt G, Rennie D. Users' Guides to the medical literature. Chicago: AMA Press, 2002.
4. Gupta M. A critical appraisal of evidence based medicine: some ethical considerations. *J Evaluat Clin Practice* 2003; 9(2): 111-21.
5. Culpepper L, Gilbert TT. Evidence and ethics. *Lancet* 1999; 353: 829-31.
6. Kerridge I, Lowe M, Henry D. Ethics and evidence based medicine. *BMJ* 1998; 316: 1151-3.
7. Sehon SR, Stanley DE. A philosophical analysis of the evidence-based medicine debate. *BMC Health Services Research* 2003; 3: 14.
8. Djulbegovic B, Morris L, Lyman G. Evidentiary challenges to evidence based medicine. *J Evaluat Clin Practice* 2000; 6(2): 99-109.
9. Liberati A, ed. La medicina delle prove di efficacia: potenzialità e limiti della EBM. Roma: Il Pensiero Scientifico Editore, 1997.
10. Cavicchi I. Evidence. Keiron 2000 (Internet). Disponibile all'indirizzo: www.farindustria.it/farindustria/documenti/004/apertura.pdf (ultimo accesso: settembre 2004).
11. Pagliaro L. La medicina basata sull'evidenza: l'uso della letteratura scientifica nella medicina clinica. *Bollettino di Informazione sui Farmaci* 2003: 127-37.
12. Spitzer WO. Clinical epidemiology. *J Chronic Disease* 1986; 39(6): 412-5.
13. Liberati A. Valore, importanza e limiti dell'informazione scientifica. In: Grilli R, Penna A, Liberati A, eds. *Migliorare la pratica clinica: come promuovere ed implementare le linee guida*. Roma: Il Pensiero Scientifico Editore, 1995; 2-20.
14. Vandembroucke IP. Evidence-based medicine and «Medicine d'Observation». *J Clin Epidem* 1996; 49: 1335-8.
15. Haynes B. What kind of evidence is it that evidence based medicine advocates want health care providers and consumers to pay attention

- to? BMC Health Services Research 2002; 2: 3. (Internet). Disponibile all'indirizzo: www.biomecentral.com.
16. The Cardiac Arrhythmia Suppression Trial II Investigators. Effects of the antiarrhythmic agent moricizine on survival after myocardial infarction. *N Engl J Med* 1992; 327: 227-33.
 17. Grodstein F, Clarkson TB, Manson JA. Understanding the divergent data on postmenopausal hormone therapy. *N Engl J Med* 2003; 348(7): 645-50.
 18. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of changes in patients' care. *Lancet* 2003; 362: 1225-30.
 19. Grilli R, Lomas J. Evaluating the message: the relationship between compliance rate and the subject of a practice guideline. *Med Care* 1994; 32: 202-13.
 20. Cabana M, Rand C, Power N et al. Why don't physicians follow clinical practice guidelines? *JAMA* 1999; 282: 1458-65.
 21. Nonino F, Liberati A. Essential requirements for practice guidelines at national and local levels. *Neurol Sci* 2004; 25(1) 260-4.
 22. Breast Cancer Trialists Collaborative Review Group. Ovarian ablation for early breast cancer (Cochrane review). In: *The Cochrane Library*, Issue 1; 2004. Chichester, UK: John Wiley & Sons.
 23. Liberati A, Buzzetti R, Grilli R et al. Which guidelines can we trust? Assessing strength of evidence behind recommendations for clinical practice. *West J Med* 2001; 174: 262-5.
 24. Canadian Task Force on Preventive Health Care. History and methods (Internet). Disponibile all'indirizzo: www.ctfpchc.org (ultimo accesso: aprile 2004).
 25. Glasziou P, Vandenbroucke I, Chalmers I. Assessing the quality of research. *BMJ* 2004; 328: 39-41.
 26. Upshur RE. Are all evidence-based practice alike? Problems in the ranking of evidence. *CMAJ* 2003; 169: 672-4.
 27. The GRADE Working Group. A critical appraisal of systems for grading evidence and formulating recommendations. *BMC Health Services Research* (submitted for publication).
 28. The GRADE Working Group. Grading the quality of evidence and the strength of recommendations. *BMJ* 2004; 328: 1490.
 29. Schunemann HJ, Best D, Vist G on behalf of GRADE. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. *CMAJ* 2003; 169: 677-80.

30. Garattini S, Liberati A. Bias by omitted research. *BMJ* 2000; 321: 845-6.
31. Bobbio M. Dove è finita la stella polare? *Bollettino di Informazione sui Farmaci* 2004; XI(1): 1-2.
32. Grilli R, Liberati A, Buzzetti R, et al. Is information from Cochrane systematic reviews useful for health policy? Paper presented at the XI Cochrane Colloquium, Barcelona (Spain), October 2003.
33. Liberati A, Sheldon TA, Banta HD. EUR-Assess project subgroup report on methodology. *Int J Technol Assess Health Care* 1997; 13(2): 186-219.
34. Liberati A. Consumer participation in research and health care. *BMJ* 1997; 315: 499.
35. Liberati A, Magrini N. Information from drug companies and opinion leaders: double standards in information from medical journals and practitioners should go. *BMJ* 2003; 326: 1156-7.
36. ALLHAT Collaborative Research Group. Major outcomes in high risk hypertensive patients randomised to angiotensin-converting enzyme inhibitors or calcium channel blockers vs. diuretics (ALLHAT). *JAMA* 2002; 288: 2981-97.
37. Appel L. The verdict from the ALLHAT: thiazide diuretics are the preferred initial therapy for hypertension. *JAMA* 2002; 288: 3039-41.
38. Psaty BM, Lumley T, Furberg CD et al. Health outcome associated with various antihypertensive therapies used as first line agents. *JAMA* 2003; 289: 2534-44.
39. Terapie antipertensive a confronto: la realtà italiana dopo lo studio ALLHAT. *Bollettino di Informazione sui Farmaci* 2004; XI(1). Disponibile all'indirizzo: www.ministerosalute.it/imgs/c_17_bif_bollettino_23_file_itemName_5_filePdf.pdf (ultimo accesso: novembre 2004).
40. Naylor CD. The complex world of prescribing behaviour. *JAMA* 2004; 291: 104-6.
41. Freemantle N, Hill S. Medicalisation, limits to medicine or never enough money to go around. *BMJ* 2002; 4: 864-5.
42. Charlton B. The rise and fall of EBM. *QJMed* 1998; 91: 371-4.
43. Sackett DL. The arrogance of preventive medicine. *CMAJ* 2002; 167(4): 363-4.
44. Bekelman JE, Li Y, Gross CP. Scope and impact of financial conflicts of interests in biomedical research. *JAMA* 2003; 289: 454-65.

45. Lomas J. Words without action? The production, dissemination and impact of consensus recommendations. *Ann Rev Public Health* 1991; 12: 41-65.
46. Lenfant C. Clinical research to clinical practice. Lost in translation? *New Engl J Med* 2003; 349: 868-74.
47. Institute of Medicine. Crossing the quality chasm: a new health system for the 21st century. Washington DC: National Academy Press, 2001.
48. Schuster M, McGlinn B, Brook R. How good is quality of care in the United States? *Milbank Q* 1998; 76: 517-63.
49. Wennberg JE. Understanding geographic variations in health care delivery. *New Engl J Med* 1999; 340: 52-3.
50. Guyatt G, Meade M, Jaenske R et al. Practitioners of evidence based care. *BMJ* 2000; 320: 954-5.
51. Wennberg JE, Gittleston A. Small area variations in health care delivery. *Science* 1973; 182: 1102-8.
52. Kaufman DW, Kelly JP, Rosemberg L et al. Recent patterns of medication use in the ambulatory adult population in the United States: the Sloan survey. *JAMA* 2002; 287: 337-44.
53. Ault G. Climbing a medical Everest: the Cochrane Collaboration. *Science* 2003; 300: 2024-5.
54. Song F, Eastwood AJ, Gillbody S et al. Publication and related biases. *Health Technology Assessment* 2000; 4: 10.
55. Korn D. Conflicts of interest in biomedical research. *JAMA* 2000; 284(17): 2234-36.
56. Angell M. Is Academic Medicine for sale? *New Engl J Med* 2000; 342(20): 1516-8.
57. Bero L, Rennie D. Influences on the quality of published drug studies. *Int J Technol Assess Health Care* 1996; 12: 209-37.
58. Lexchin J, Bero LA, Djulbegovic B et al. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ* 2003; 326: 1167-70.
59. Rochon PA, Gurwitz GM, Simms RW et al. A study of manufacturer-sponsored trials of non steroidal anti-inflammatory drugs in the treatment of arthritis. *Arch Int Med* 1994; 154: 157-63.
60. Chalmers I. Well informed uncertainties about the effects of treatments. *BMJ* 2004, 328: 475-6.
61. Chalmers I. Its official. Evaluative research must become part of routine care in the NHS. *J R Soc Med* 2000; 93: 555-6.

62. Woods K. Filling the evidence gaps: identifying topics for research in the NHS Health Technology Assessment Programme. *Journal of Clinical Excellence* 2002; 4: 355-6.
63. Doust J, Del Mar C. Why do doctors use treatments that do not work? *BMJ* 2004; 328: 474-5.
64. Hanley B. Involvement works: the 2nd report of the Standing Group on Consumer in NHS research. NHS Executive 1999 (Internet). Disponibile all'indirizzo: www.conres.co.uk/pdfs/involvement_works.pdf (ultima consultazione: novembre 2004).
65. Herxheimer A. Relationship between the pharmaceutical companies industry and patient's organisation. *BMJ* 2003; 326: 1208-10.
66. Goodman SN. The mammography dilemma: a crisis for evidence based medicine? *Ann Int Med* 2002; 137: 363-5.
67. Horton R. Screening mammography: an overview revisited. *Lancet* 2001; 358: 1284-5.
68. The Cochrane Breast Cancer Review Group. Screening mammography: setting the record straight. *Lancet* 2002; 359: 439-40.
69. Langhorne P. Screening mammography: setting the record straight. *Lancet* 2002; 359: 440-2.
69. Liberati A. Forzature e schematizzazioni nel dibattito sulla efficacia dello screening mammografico. *Epid & Prev* 2002; 26(2): 53-4.
70. Maynard A. Evidence based medicine: an incomplete method to inform treatment choice. *Lancet* 1997; 349: 126-8.
71. Victora CG, Habicht JP, Bryce J. Evidence based public health: moving beyond randomised trials. *Am J Publ Health* 2004; 94: 400-5.
72. Habicht JP, Victora CG, Vaughan JP. Evaluation design for adequacy, plausibility and probability of public health programme performance and impact. *Int J Epidem* 1999; 28: 10-8.
73. Miller FG, Emanuel AJ, Rosenstein DL et al. Ethical issues concerning research in complementary and alternative medicine. *JAMA* 2004; 291: 599-604.
74. Gotzsche P, Liberati A, Torri V et al. Beware of surrogate outcome measures. *Int J Technol Assess Health Care* 1996;12(2): 238-46.