

## Presentazione

Dalla mitologia degli strumenti inglesi di valutazione, all'outcome movement americano, alle esperienze valutative nei servizi italiani di salute mentale di comunità

Nella mitologia romana *Honos* era un Dio particolarmente adorato dai soldati, in quanto identificato come il Dio dell'onore militare e della moralità e per questo spesso accompagnato dalla Dea *Virtus*. Era raffigurato come un giovane che aveva armato la sua mano destra con una lancia (ma in altre raffigurazioni ha un ramo d'ulivo o uno scettro) e che portava in quella sinistra una cornucopia. Il suo culto fiorì nella Roma repubblicana. In suo onore furono eretti due templi e la sua immagine riprodotta su alcune monete (tra cui un sesterzio coniato da Traiano) e anche su di un pannello dell'arco di Traiano (114 d.C.). Era dunque il Dio dell'onore e della virtù.

Sappiamo che HoNOS sta anche per *Health of the Nation Outcome Scales*, una scala di valutazione, messa a punto da John Wing, che appartiene alla "mitologia" inglese della creazione di strumenti di misura per la psichiatria epidemiologica e per la ricerca clinica e valutativa in salute mentale. Tra gli altri strumenti "mitici" che si devono allo stesso autore, ricordiamo l'intervista PSE (*Present State Examination*), passata attraverso nove edizioni e denominata poi, nella decima, *Schedules for Clinical Assessment in Neuropsychiatry* (SCAN). Ricordo che lo SCAN (Manuale e Glossario) è stato tradotto in italiano e pubblicato in forma di tre volumi, su iniziativa del Pensiero Scientifico Editore (WHO, 1996a e 1996b). Fare valutazione, in un settore difficile come quello dei disturbi mentali, richiede impegno, per "combattere" e contrastare i numerosi fattori di confondimento e superare gli ostacoli che rendono difficile questo compito, vale a dire una dose notevole di virtù metodologiche per portare a termine con "onore" questo compito.

John Wing, l'autore di HoNOS, professore di psichiatria sociale, era uno dei due "principali" allievi di Sir Aubrey Lewis, il fondatore dell'Institute of Psychiatry di Londra e della psichiatria sociale europea. L'altro era Michael Shepherd, professore di psichiatria epidemiologica.

Secondo la tradizione inglese, i professori, quando lasciano il servizio, lasciano anche la loro stanza e molte delle attività accademiche svolte fino ad allora. Il *retirement* può essere più o meno “attivo” a seconda delle circostanze, ma non è mai il proseguimento, in forma non ufficiale, di ciò che si faceva prima. Ciò sarebbe considerato cosa poco elegante, anche perché possibile intralcio alle attività del successore. Molti professori inglesi *retired* continuano però a fare ricerca e presentare *application* per ricevere *grant*, a insegnare, a viaggiare per partecipare a congressi e conferenze. Ad esempio, Sir Aubrey Lewis, dopo il suo *retirement* nel 1966, lasciò la direzione e la sua segreteria, al primo piano dell’Institute of Psychiatry, e andò a occupare una stanza nella Biblioteca dell’Istituto. Ebbi occasione di vederlo lì, nell’autunno del 1969. Per alcuni anni continuò a studiare e a dare suggerimenti a coloro che avevano il coraggio di chiederglieli. Non era, infatti, persona facile da avvicinare, per i comuni mortali. Michael Shepherd, quando venne il suo turno, ventidue anni dopo, nel 1988, dopo il *retirement* andò a dividere per alcuni anni una stanza con i giovani *registrar* dell’Istituto. Ai suoi tempi i professori in pensione erano, infatti, diventati più numerosi e non era possibile ospitarli in Biblioteca. Egli disse poi che quell’esperienza gli aveva consentito di capire molte più cose dell’Istituto di quelle che aveva potuto percepire sedendo per anni nella direzione della General Practice Research Unit.

John Wing, invece, ha avuto un *retirement* diverso: andato in pensione nel 1989, lasciò la direzione della MRC Social Psychiatry Unit ed anche l’Istituto, ma fu nominato Direttore della neo-creata Royal College of Psychiatrists’ Research Unit (CRU). Fu un *retirement* in qualche modo anomalo e molto produttivo. Nei sette anni in cui rimase alla direzione di quella Unità tra le altre cose sviluppò l’HoNOS ed ebbe anche un grande ruolo sullo sviluppo della prima *Clinical Practice Guideline*, dedicata alla gestione della violenza in *setting* psichiatrici. Faccio notare che essa è stata pubblicata due anni prima della creazione del National Institute for Health and Clinical Excellence (NICE), che tanta influenza ha avuto ed ha nelle revisioni critiche della letteratura per la scelta di trattamenti razionali in salute mentale.

L’HoNOS costituì invero la risposta a un’azione strategica del Governo inglese, che nel 1992 aveva pubblicato il programma denominato “*Health of the Nation*”. Esso, per la parte dedicata alla salute mentale, aveva identificato tre obiettivi principali: ridurre del 33% i suicidi delle persone con disturbi mentali; ridurre del 15% il numero totale di suicidi, prima

dell'anno 2000, e “*migliorare in modo significativo la salute e il funzionamento sociale di coloro che soffrono di disturbi mentali*”. Mentre i primi due obiettivi erano stati ben identificati e quantificati, il terzo appariva vago. C'era bisogno, dunque, di una metodologia e di uno strumento che permettesse di misurare, su larga scala, salute e funzionamento sociale dei pazienti e che potessero anche valutare i cambiamenti, nel tempo, del loro stato. Il governo inglese affidò l'incarico di sviluppare tale strumento, nel 1993, all'Unità di ricerca diretta dal professor Wing.

Una prima versione dell'HoNOS è stata pubblicata da Wing e collaboratori nel 1996. I primi risultati su validità e attendibilità della scala apparvero nel 1998 (Wing et al., 1998), mentre il numero di Maggio 1999 del *British Journal of Psychiatry* (BJP) fu quasi esclusivamente dedicato a fare il punto sull'uso di questa scala e delle versioni parallele per bambini e adolescenti e per anziani uscite nel frattempo. L'HoNOS è la scala più usata nei servizi di salute mentale inglesi per la valutazione dell'esito ed è stata tradotta in numerose lingue.

I quattordici articoli del Fascicolo del BJP del 1999 ricordato sopra costituiscono qualcosa di simile al volume della serie PEQuOD che esce ora in Italia, grazie all'impegno di Arcadio Erlicher, Antonio Lora, Emiliano Monzani e Alberto Parabiaghi, ai quali deve andare una forte riconoscenza da parte di tutti gli psichiatri italiani.

Nel presentare questo volume desidero, però, continuare a utilizzare una prospettiva storica per discutere alcuni aspetti della questione della valutazione dell'esito in psichiatria (Tansella, 2003).

L'inizio del cosiddetto *outcome movement* può essere fatto risalire al 1988, anno nel quale, negli Stati Uniti, Paul Ellwood, nella sua *Shattuck lecture*, fece per la prima volta la proposta di archiviare, in *database* appositamente creati, grandi quantità di dati sull'esito dei trattamenti sanitari. I dati andavano raccolti in modo sistematico e su larga scala, nella pratica clinica di routine. Si sarebbe così potuta costituire una nuova risorsa per la ricerca medica, e in particolare per la ricerca sui servizi sanitari (Ellwood, 1988).

Sapevamo tutti, sin d'allora, che la metodologia da seguire per dimostrare in modo scientificamente corretto l'efficacia delle cure e dei trattamenti, sia a livello individuale (i pazienti) sia a livello locale (i servizi) era quella dello studio clinico controllato e randomizzato (*Randomised Controlled Trial* o RCT). Ho avuto modo, in altra sede, di ricordare il ruolo di Sir Bradford Hill nella messa a punto di quella metodologia ed anche di sottoli-

neare come essa abbia consentito lo sviluppo della valutazione “scientifica” delle cure, sviluppo iniziato non più di 60 anni fa (Tansella, 2002 e 2009). La proposta di Ellwood, accolta subito con grande interesse, non suggeriva di sostituire *tout court* ai RCT le ricerche di *outcome*. La cosa non sarebbe stata sostenibile. Essa voleva estendere l’attenzione, negli studi valutativi, dagli aspetti quantitativi del monitoraggio delle attività dei servizi (i dati e le misure di processo) agli aspetti qualitativi della valutazione propriamente detta (i dati e le misure di esito). Quella proposta aveva, a monte, anche altre motivazioni che conviene ricordare: diminuire i costi della ricerca basata sugli studi clinici controllati, costi che continuavano ad aumentare, mettendo in difficoltà gli Enti finanziatori e in particolare le industrie farmaceutiche; favorire l’introduzione nella ricerca clinica delle nuove, promettenti tecnologie (erano gli anni del forte sviluppo dell’informatica); poter documentare e, se possibile, correggere, la disomogeneità tra Stati e aree diverse degli USA nella quantità e nella qualità dei servizi sanitari. Ma doveva e deve anche oggi intendersi come una proposta integrativa, non sostitutiva.

L’interesse per l’*outcome movement* dette luogo l’anno dopo, nel 1989, alla creazione della Agency for Health Care Policy and Research (AHCPR), ora denominata Agency for Healthcare Research and Quality (AHRQ), agenzia fondata per condurre ricerche su larga scala sull’esito delle malattie più comuni. Nacquero poi i Patient Outcome Research Teams (PORT), con lo scopo di monitorare e confrontare tra loro i diversi programmi sanitari americani. L’Agenzia ricevette dal Governo americano, nel primo anno, uno stanziamento di 6 milioni di dollari; somma che salì a 63 milioni nel 1991. Quello che il Governo voleva era utilizzare i dati raccolti per determinare, in modo abbastanza attendibile (ma, ripeto, a mio avviso non sufficientemente rigoroso), “gli esiti, l’efficacia e l’appropriatezza dei trattamenti” (Anderson, 1994; Gilbody et al., 2002). Il Congresso americano, nel 1992, a evitare equivoci e per sottolineare l’esigenza di risparmiare sui costi della ricerca, precisò poi che le ricerche condotte dai PORT non avrebbe dovuto utilizzare la metodologia degli studi clinici controllati, ma quella, meno complessa e costosa, degli studi osservazionali (Gilbody et al., 2002). Ma le persone più attente e informate capirono presto che le ricerche dei PORT non avrebbero comunque potuto essere considerate esaustive.

Era nata così ufficialmente l’*outcome research* e insieme ad essa i “*database analysts*”, che adottarono il motto “*Happiness is a humongous database*” (Ano-

nymous, 1989 e 1992; Smith, 1997). Nella sua versione all'*americaine* quella ricerca si caratterizzava non solo per la vastità delle risorse economiche disponibili, ma anche per l'uso di misure di esito rappresentate, nella maggior parte dei casi, da dati amministrativi raccolti nella pratica clinica di routine, più qualche altra misura aggiuntiva.

Già a metà degli anni '90, tuttavia, l'entusiasmo che, al di là dell'Atlantico, aveva caratterizzato la nascita della AHCPR si stemperava nella seguente dichiarazione: "*Contrariamente alle attese espresse nella legge istitutiva della AHCPR... i database amministrativi non si sono generalmente mostrati utili a rispondere alle domande relative al confronto, in termini di effectiveness, fra trattamenti medici alternativi*" (Office for Health Technology Assessment, US Congress, 1994). Possiamo dire che tutto ciò era prevedibile.

Anche i servizi di salute mentale, ancora una volta con qualche ritardo rispetto agli altri servizi sanitari, sono stati coinvolti nell'*outcome movement*, soprattutto dopo lo sviluppo dei servizi di comunità. Ciò è avvenuto grazie alle pressioni di alcuni amministratori i quali, con la riduzione delle risorse, sentivano sempre più l'esigenza di confrontare tra loro, in termini di rapporti costi/efficacia, programmi diversi, per scegliere ovviamente i meno costosi, a parità di esito: il confronto, quindi, doveva riguardare sia i volumi di attività, sia l'esito (Sederer et al., 1995; Becker et al., 2002). Con il progredire di queste esperienze è diventato chiaro che, anche nel caso dei servizi di salute mentale, queste valutazioni sistematiche e di routine, effettuate utilizzando dati raccolti dagli stessi operatori che forniscono le cure e l'assistenza, avevano limiti che non è possibile ignorare. Il problema non è solo quello di superare le difficoltà "pragmatiche" nella raccolta e nell'organizzazione dei dati (le motivazioni e la disponibilità di tempo degli operatori; il costo del training; l'accuratezza e la completezza), difficoltà sulle quali i fautori dell'*outcome movement* avevano subito concentrato la loro attenzione. Bisognava considerare anche gli errori sistematici e le distorsioni insite in questo tipo di raccolta, errori e distorsioni che possono, tra l'altro, configurarsi a volte come veri e propri conflitti di interesse, visto che sarebbero gli stessi operatori a valutare gli esiti del loro lavoro (Speer, 1998). Dopo Bradford Hill, ripetiamo, è chiaro come in ogni processo di raccolta e interpretazione dei dati, in assenza di adeguati controlli e precauzioni metodologiche, si nascondano *bias* che possono ridurne o inficiarne il valore (Tansella, 2003). Nel caso dell'autovalutazione dei servizi sanitari essi appartengono a tre categorie: *selection bias* (gli operatori tendono a

raccogliere dati *baseline* dai pazienti che hanno maggiori possibilità di beneficiare del programma oggetto di valutazione); *attrition bias* (gli operatori tendono a raccogliere dati di follow-up dai pazienti che hanno avuto un esito favorevole); *detection bias* (gli operatori tendono ad attribuire punteggi più elevati ai pazienti che hanno completato il programma) (Feinstein, 1985).

Una revisione sistematica degli studi relativi all'uso di routine di misure di *outcome* (Routine Outcome Assessment o ROA) nei servizi di salute mentale ha sollevato molti dubbi sull'utilità di questo approccio (Gilbody et al., 2002). Il dibattito in questi anni è stato molto vasto. Rimandiamo il lettore interessato ad alcune ottime revisioni della letteratura (Salvador Carulla, 1999; Slade, 2002a e 2002b; Lloyd Evans et al., 2007; Thornicroft & Slade, 2009). Slade et al. (2006), in uno studio clinico controllato, hanno dimostrato che l'impiego di routine di misure dedicate a valutare bisogni, qualità della vita, gravità dei problemi di salute mentale e alleanza terapeutica non aveva migliorato la percezione soggettiva dell'esito, anche se era risultato associato ad un minore uso del ricovero ospedaliero. Rimane il problema di identificare gli ingredienti necessari e sufficienti a dare significato e spessore, da un punto di vista clinico, all'uso di queste metodiche di *assessment* e alle pratiche di ROA.

È necessario però fare anche un altro ordine di considerazioni. Innanzitutto bisogna tener presente che anche gli studi clinici controllati (RCT), che pure rappresentano, come si è detto, il *golden standard* per la valutazione dell'efficacia dei trattamenti e delle cure, hanno dei limiti:

1. sono condotti in gruppi di pazienti selezionati, in condizioni sperimentali che appaiono "artificiali" rispetto alla pratica clinica quotidiana (validità "interna" alta, validità "esterna" bassa);
2. valutando l'efficacia di trattamenti e interventi specifici, spesso richiedono la sospensione di ogni altro trattamento, il che contrasta con quanto accade nella comune pratica clinica quotidiana;
3. raramente valutano l'efficacia di decisioni pratiche di management (ad esempio ricoverare o no un paziente), nonostante l'importanza che queste decisioni hanno nella pratica clinica;
4. valutano, in genere, solo effetti a breve-medio termine, mentre i clinici e gli amministratori hanno bisogno di avere anche informazioni su quelli a lungo termine;
5. sono poco pratici per valutare eventi rari (ad esempio il suicidio);
6. non tengono conto di variabili di contesto;

7. nei RCT opinioni “politiche” o considerazioni e convincimenti dei medici e del personale sanitario possono ostacolare una corretta randomizzazione;
8. i clinici e il personale coinvolto nei RCT sono spesso “atipici” ed hanno un interesse particolare nel problema oggetto dello studio;
9. molti trial sono condotti in “*Rolls Royce Services*”, che godono di fondi e di risorse aggiuntive e non sono pertanto confrontabili con la maggior parte dei servizi “normali”;
10. non sempre il trattamento di controllo è scelto in modo ineccepibile (ad esempio il dosaggio del farmaco di confronto non è sempre adeguato e, nelle ricerche sponsorizzate dall’industria, la sua scelta è talora influenzata dagli obiettivi che la ricerca si pone; e ancora, le caratteristiche dell’*usual care*, ricevuta dal gruppo di controllo, a volte sono di non elevata qualità, il che avvantaggia il gruppo sperimentale).

In secondo luogo dobbiamo distinguere gli studi messi in campo per giudicare l’efficacia di un (nuovo) trattamento, confrontandolo con quelli già in uso, con gli studi che hanno come *focus* i servizi sanitari e le loro attività, in termini di identificazione dei possibili benefici per i pazienti. Per gli studi del secondo tipo è oggi possibile condurre una *outcome research* di qualità, a patto di osservare alcune regole (Rosenheck et al., 1999). Le principali sono:

1. includere nella ricerca sull’esito un elevato numero di soggetti, che siano però rappresentativi della totalità dei pazienti trattati nell’area o nella regione nella quale si conduce la ricerca;
2. usare strumenti di misura standardizzati, scelti accuratamente, in relazione alle domande alle quali la ricerca deve rispondere;
3. misurare l’esito in modo multidimensionale (usando anche valutazioni fatte dagli stessi pazienti e dai familiari) e multifasico;
4. monitorare anche un buon numero di variabili socio-demografiche (per facilitare la *matching*);
5. garantire tassi elevati di partecipazione al follow-up, nel quale bisogna includere, se possibile, un campione rappresentativo dei pazienti che, nel frattempo, hanno perso contatto con i servizi monitorati. Infine,
6. garantire, almeno per alcune misure, una valutazione “indipendente”, fatta da operatori non direttamente coinvolti nell’assistenza ai pazienti (Tansella, 2003).

Si tratta dunque di associare, in alcuni servizi, all'attività assistenziale un'attività di ricerca relativamente "leggera", vale a dire non particolarmente impegnativa e costosa, e al tempo stesso rigorosa, che utilizzi strumenti semplici, clinicamente significativi e di provata validità (Tansella, 2003). Seguendo questa strada un vantaggio aggiunto, certamente non trascurabile, è quello di consolidare e far progredire, anche nella pratica clinica quotidiana, la cultura della valutazione, il che potrà consentire di far attecchire poi, in quei *setting*, altri, più sofisticati, approcci valutativi. La scala HoNOS è uno degli strumenti che meglio soddisfa le suddette esigenze di promozione della cultura della valutazione.

La proposta lanciata entusiasticamente negli USA, nel 1988, con l'*outcome movement* era altra cosa, essendo essenzialmente finalizzata a un'utilizzazione "secondaria", a fini di ricerca, di dati primariamente raccolti per scopi amministrativi.

Considero, per concludere, la pubblicazione di questo volume una tappa importante di un percorso di promozione della ricerca nei servizi di salute mentale italiani, che va oltre l'*outcome movement all'americaine*. Ritengo che ce ne sia, oggi, un forte bisogno.

## Bibliografia

Anonymous (1989). Databases for healthcare outcomes. *Lancet* 396, 195-196.

Anonymous (1992). Cross design synthesis: a new strategy for studying medical outcomes? *Lancet* 340, 944-946.

Anderson C. (1994). Measuring what works in health care. *Science* 263, 1080, 1082.

Becker T., Hulsmann S., Knudsen H.C., Martiny K., Amaddeo F., Herran A., Knapp M., Schene A.H., Tansella M., Thornicroft G. & Vazquez-Barquero J.L. (2002). Provision of services for people with schizophrenia in five European regions. *Social Psychiatry and Psychiatric Epidemiology* 37, 465-474.

Ellwood P.M. (1988). Shattuck lecture--outcomes management. A technology of patient experience. *New England Journal of Medicine* 318, 1549-1556.

Feinstein A. (1985). *Clinical epidemiology: The architecture of clinical research*. Saunders: Philadelphia.

Gilbody S.M., House A.O. & Sheldon T.A. (2002). Outcomes research in mental health. Systematic review. *British Journal of Psychiatry* 181, 8-16.

Lloyd-Evans B., Johnson S. & Slade M. (2007). Assessing the content of mental health services: a review of measures. *Social Psychiatry and Psychiatric Epidemiology* 42, 673-682.

Office for Health Technology Assessment (US Congress) (1994). *Identifying health strategies that work: Searching for evidence*. US Government Printing Office: Washington DC.

Rosenheck R.A., Fontana A. & Stolar M. (1999). Assessing quality of care: administrative indicators and clinical outcomes in post traumatic stress disorder. *Medical Care* 37, 180-188.



- Salvador-Carulla L. (1999). Routine outcome assessment in mental health research. *Current Opinion in Psychiatry* 12, 207-210.
- Sederer L.I., Hermann R. & Dickey B. (1995). The imperative of outcome assessment in psychiatry. *American Journal of Medical Quality* 10, 127-132.
- Slade M. (2002a). Routine outcome assessment in mental health services. *Psychological Medicine* 32, 1339-1343.
- Slade M. (2002b). What outcomes to measure in routine mental health services, and how to assess them: a systematic review. *Australian and New Zealand Journal of Psychiatry* 36, 743-753.
- Slade M., McCrone P., Kuipers E., Leese M., Cahill S., Parabiaghi A., Priebe S. & Thornicroft G. (2006). Use of standardised outcome measures in adult mental health services. *British Journal of Psychiatry* 189, 330-336.
- Smith D.M. (1997). Database research: is happiness a humongous database? *Annals of Internal Medicine* 127, 725-756.
- Speer D. (1998). *Mental health outcome evaluation*. Academic Press: San Diego.
- Tansella M. (2002). The scientific evaluation of mental health treatments: an historical perspective. *Evidence Based Mental Health* 5, 4-5.
- Tansella M. (2003). La valutazione dell'esito delle cure e dei trattamenti nella pratica psichiatrica di routine. In *L'Esito della schizofrenia. Gli studi internazionali e i dati italiani* (di A. Lasalvia e M. Ruggeri), pp. XI-XX. Centro Scientifico Editore: Torino.
- Tansella M. (2009). La valutazione scientifica dei trattamenti in medicina ed il ritardo della psichiatria. In *Filosofia della medicina. Modelli, metodo, cura ed errori* (a cura di P. Giaretta, A. Moretto, G.F. Gensini e M. Trabucchi), pp. 125-149. Il Mulino: Bologna.
- Thornicroft G. & Slade M. (2009). Are routine outcome measures feasible in mental health? *Quality in Health Care* 9, 84.
- Wing J.K., Curtis R.H. & Beevor A.S. (1996). *HoNOS: Health of the Nation Outcome Scales: Report on Research and Development July 1993-December 1995*. London: Royal College of Psychiatrists.
- Wing J.K., Beevor A.S., Curtis R.H., Park S.B.G., Hadden S. & Burns A. (1998). Health of the Nation Outcome Scales (HoNOS): Research and development. *British Journal of Psychiatry* 172, 11-18.
- World Health Organization (1996a). *SCAN, Schedules for Clinical Assessments in Neuropsychiatry. Manuale*. Edizione italiana a cura di M. Tansella e M. Nardini (traduzione di S. Nicolau, L. Pietrangeli e I. d'Errico, con la revisione di G. de Girolamo), pp. 1-325. Il Pensiero Scientifico Editore: Roma.
- World Health Organization (1996b). *SCAN, Schedules for Clinical Assessment in Neuropsychiatry. Glossario e opuscolo per le codifiche (2 volumi)*. Edizione italiana a cura di M. Tansella e M. Nardini (traduzione di L. Pietrangeli e S. Nicolau, con la revisione di A. Giannini), pp. 1-108 e pp. 1-243. Il Pensiero Scientifico Editore: Roma.

## Michele Tansella

Professore ordinario di Psichiatria,  
 Direttore del Dipartimento Interaziendale di Salute Mentale  
 e del WHO Collaborating Centre for Research and Training  
 in Mental Health and Service Evaluation,  
 Università di Verona