

## Variabilità nelle misure di efficacia per la malattia di Alzheimer

Massimo MUSICCO (a), Carla PETTENATI (b) e Carlo CALTAGIRONE (c)

(a) Istituto di Tecnologie Biomediche, Consiglio Nazionale delle Ricerche, Segrate, Milano

(b) Ospedale di Passirana di Rho, Centro Regionale Alzheimer, Rho, Milano

(c) Università degli Studi "Tor Vergata", Fondazione S. Lucia, Roma

**Riassunto.** - L'efficacia degli interventi è definibile, in medicina, come la loro capacità di modificare positivamente la storia naturale delle malattie. La storia naturale delle demenze è contrassegnata da eventi conseguenti dal declino cognitivo e funzionale che caratterizza queste malattie. Questi eventi sono poco prevedibili nel singolo paziente dato che la loro occorrenza, o entità percepita attraverso le scale di misurazione, è largamente variabile da paziente a paziente. Ne consegue che l'efficacia dei trattamenti, modesta per quelli oggi disponibili, è difficilmente apprezzabile nella pratica clinica e persino negli studi osservazionali. Negli studi clinici randomizzati l'effetto di tale variabilità è infine analogo a quello della misclassificazione in studi epidemiologici caso-controllo o di coorte e causa una sottostima dell'effettiva efficacia dei trattamenti.

*Parole chiave:* demenza, scale cliniche, studi randomizzati e controllati.

**Summary** (*Variability of efficacy measures in Alzheimer's disease*). - The efficacy of medical interventions is their capacity of inducing positive modifications of the natural history of diseases. The natural history of dementia is marked by specific events related to the cognitive and functional decline, but their occurrence is poorly predictable in individual patients being highly variable from patient to patient. For this reason it is difficult that the modest efficacy of available interventions for dementia, or their entity measured by clinical scales, may be perceived in clinical practice or in observational studies. Moreover in randomized clinical studies, the effect of this variability, in analogy to misclassification of exposition and/or disease in case control or cohort epidemiological studies, is that of an underestimation of the true efficacy of interventions.

*Key words:* dementia, clinical scales, randomized controlled trials.

### Premessa

Per affrontare il tema della variabilità delle misure di efficacia impiegate per lo studio e la valutazione dei pazienti con malattia di Alzheimer (AD) è necessario innanzitutto definire il concetto di efficacia in generale e calarlo quindi nella specifica realtà di malattie complesse quali le demenze.

Efficacia in medicina può essere definita come la capacità di un certo intervento di modificare positivamente la storia naturale di malattia. Implicita quindi, nel concetto di efficacia, è l'esistenza di una relazione di causa-effetto in quanto l'intervento in questione causa una modificazione positiva della storia naturale di malattia.

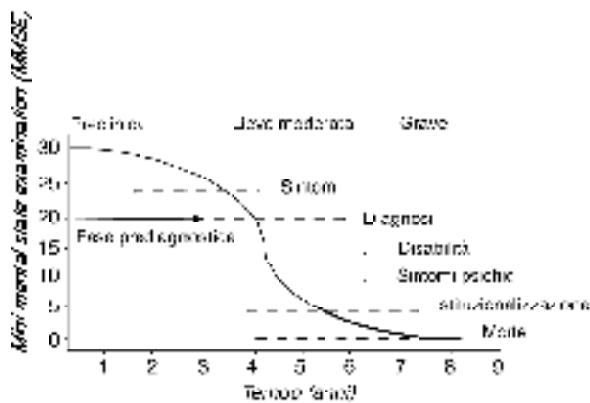
La storia naturale di una malattia è scandita da "eventi" (esiti, *outcome*) ed in particolare, quando si deve decidere se un intervento medico è o meno efficace, ci si rivolge ad eventi rilevanti per l'ammalato. L'efficacia di un intervento può quindi, in modo più circostanziato, essere definita come la sua capacità di favorire l'occorrenza di eventi positivi (guarigione, sopravvivenza, ecc) o la sua capacità di contrastare la comparsa di eventi considerati negativi (morte, disabilità, malattie intercorrenti). Nel caso di malattie progressive, ed in assenza di interventi che possono guarire la malattia stessa o arrestarne la progressione, l'efficacia di un intervento viene valutata in termini di sua capacità di rallentare la velocità di occorrenza di esiti negativi.

**La valutazione di efficacia nella malattia di Alzheimer**

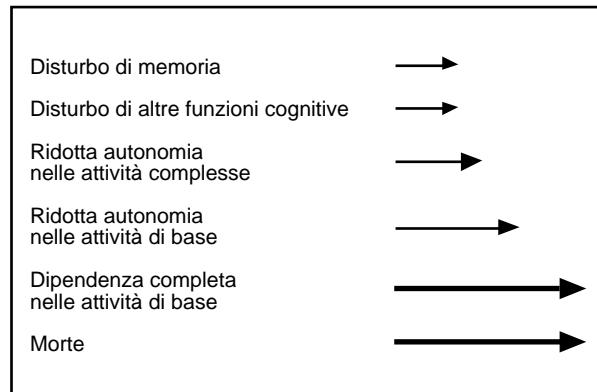
Gli studi di efficacia degli interventi nella AD fanno sempre riferimento all'asse temporale accettando implicitamente l'idea che nessuno dei trattamenti attualmente disponibili è in grado di arrestare l'inevitabile progressione della malattia.

La Fig. 1 rappresenta la storia naturale della AD in funzione del tempo e dell'andamento del declino cognitivo misurato con il test di più comune utilizzo il *mini mental state examination* (MMSE) di Folstein [1]. Dallo schema presentato è possibile apprezzare come gli eventi rilevanti nella storia della malattia siano la comparsa dei sintomi, il momento della diagnosi, l'insorgenza della disabilità e cioè delle limitazioni nella capacità di svolgere le attività del vivere quotidiano, il comparire della sintomatologia comportamentale, l'istituzionalizzazione o più in generale l'incapacità assoluta a svolgere le azioni del vivere quotidiano ed infine la morte. Tutti gli studi di storia naturale hanno mostrato come la comparsa di questi eventi rilevanti sia poco prevedibile in base alla compromissione cognitiva o alla durata della malattia.

Gli eventi caratterizzanti la storia naturale della AD non sono tutti egualmente rilevanti per il paziente. È evidente, ad esempio, che il disturbo di memoria per sé, pur essendo il sintomo principale della AD, può essere poco rilevante quando non determina una riduzione di autonomia del paziente. Ai fini però della valutazione di efficacia nella AD è largamente accettata l'idea che il processo patologico determini primariamente una compromissione della memoria seguita da una compromissione di altre funzioni cognitive che determinerebbe a sua volta l'insorgere di una disabilità via via crescente fino al punto in cui la sua gravità è tale che il paziente è incapace di svolgere autonomamente qualunque attività del vivere quotidiano [2]. Il processo patologico infine e la progressiva disabilità causano poi la morte del paziente. In conclusione la riduzione delle capacità



**Fig. 1.** - Rappresentazione schematica della storia naturale della malattia di Alzheimer.



**Fig. 2.** - Eventi che caratterizzano la storia naturale della malattia di Alzheimer. Lo spessore e la lunghezza delle frecce sono proporzionali alla rilevanza clinica dell'evento.

cognitive, la disabilità e la sopravvivenza costituiscono gli indicatori principali utilizzati nella valutazione di efficacia degli interventi e di cui si dovrà valutare la variabilità.

**La variabilità biologica e la variabilità degli strumenti di misura**

La variabilità in medicina ed in generale nei fenomeni biologici è la risultante di tutti quei fattori che causano modificazioni del fenomeno di interesse. Queste modificazioni possono essere variazioni nell'occorrenza del fenomeno o variazioni nelle sue dimensioni. Per restare all'ambito della AD è variabile la sopravvivenza dei malati come pure ad esempio il livello di compromissione cognitiva al momento della diagnosi. La variabilità è però non solo una caratteristica inerente il fenomeno che si vuole indagare ma dipende anche da fattori che possono influenzare la nostra percezione o misura del fenomeno stesso. In altre parole la variabilità dipende anche dagli strumenti di rilevazione e di misura del fenomeno che vengono utilizzati. La variabilità inerente al fenomeno può essere ridotta utilizzando modelli di predizione. Per restare agli esempi sopra riportati, la variabilità osservata nella sopravvivenza può essere spiegata, e quindi ridotta, considerando le diverse età e il sesso degli ammalati. All'interno di classi omogenee di età e sesso le durate di sopravvivenza sono in genere meno variabili che nell'intera popolazione esaminata. Analogamente la variabilità della compromissione cognitiva al momento della diagnosi può essere meno marcata quando esaminata all'interno di classi omogenee di durata di malattia. Diverso è il discorso qualora si consideri la variabilità risultante dagli strumenti di misura in questo caso la variabilità può essere controllata solo migliorando gli strumenti di misura stessi.

### Variabilità degli esiti e delle scale per la valutazione dell'efficacia nella malattia di Alzheimer

Se consideriamo i diversi indicatori utilizzati per la valutazione di efficacia nella AD possiamo innanzitutto rivolgere la nostra attenzione alla mortalità dove la variabilità osservata è in tutto attribuibile a caratteristiche intrinseche al fenomeno. In generale la letteratura è sostanzialmente concorde nel descrivere una sopravvivenza mediana dal momento della diagnosi in pazienti con AD di 4-8 anni [3-5]. La mortalità è fortemente dipendente dal sesso, età, gravità della compromissione cognitiva e funzionale del paziente, comorbidità e condizioni di vita del paziente. Questi differenti fattori permettono di contraddistinguere pazienti con differente prognosi ma spiegano solo una piccola parte della variabilità osservata nella sopravvivenza dei pazienti. Nella Tab. 1 è riportato il risultato di una analisi multivariata sulla sopravvivenza di 870 pazienti con AD. È evidente come numerose variabili siano significativamente associate con la sopravvivenza dei pazienti ma come la variabilità spiegata da questi fattori sia solo il 20% ( $R^2 = 0,20$ ).

La disabilità viene generalmente misurata con scale che misurano le capacità del paziente nelle attività del vivere quotidiano. Le attività del vivere quotidiano vengono distinte in attività di base e attività strumentali che sono indagate rispettivamente con scale denominate BADL (*basic activities of daily living*) e IADL (*instrumental activities of daily living*) o EADL (*extended activities of daily living*) [6].

La riproducibilità delle scale di disabilità è risultata generalmente buona e quindi si può accettare l'idea che la valutazione della disabilità sia poco influenzata dai difetti intrinseci degli strumenti di misura [7-11].

Ciononostante la variabilità della disabilità osservata sia trasversalmente che longitudinalmente tra i pazienti con AD è sempre molto grande e la variabilità "spiegata" dalle condizioni e caratteristiche del paziente, quando è stata studiata, è risultata generalmente molto limitata [12].

La possibilità di apprezzare l'efficacia di un intervento medico su pazienti con demenza in termini di riduzione della disabilità o di aumento di sopravvivenza è limitata dai tempi necessari per apprezzare questi eventi. Abbiamo già visto come la sopravvivenza dopo la diagnosi sia mediamente superiore ai 4-6 anni e la stessa progressione di disabilità richiede tempi di 2-4 anni perché si osservi nella metà dei pazienti una progressione significativa [2]. È per questo che negli studi clinici controllati e nella pratica clinica si utilizzano scale per la valutazione delle capacità cognitive come indicatori di esito degli interventi. L'assunto è ovviamente che il processo patologico neu-

**Tabella 1.** - Variabili predittive della sopravvivenza di 870 pazienti con malattia di Alzheimer visti consecutivamente fra il 1987 e il 2000 presso il Centro Alzheimer dell'Ospedale di Passirana di Rho (Milano)

Variabili considerate	p
MMSE (*)	0,137
Comorbidità	0,006
Scolarità	0,003
Età	0,099
Disabilità	0,004
Global deterioration score	0,014
Nucleo familiare	0,339
Durata di malattia	0,000
Sesso	0,000

(\*) MMSE: *mini mental state examination*.  $R^2 = 0,20$ .

rodegenerativo causi il deficit cognitivo che a sua volta causa disabilità e che questa limiti la sopravvivenza del paziente. Questo assunto è stato in larga misura verificato anche se non esiste una relazione lineare fra deficit cognitivo e disabilità [2, 13-18].

La performance cognitiva di soggetti sani e dementi viene valutata tramite scale tra cui il MMSE di Folstein [1] e l'ADAS Cog (*Alzheimer disease assessment scale - cognitive section*) [19] sono le più utilizzate. Il MMSE è uno strumento di screening che è stato largamente utilizzato anche in studi epidemiologici mentre l'ADAS è una scala prevalentemente utilizzata per i clinical trial. A differenza della valutazione di mortalità e della stessa disabilità queste scale di valutazione cognitiva presentano difetti di riproducibilità. Quindi quando queste scale sono utilizzate per valutare l'efficacia degli interventi medici alla variabilità inerente la malattia si aggiunge una variabilità dovuta agli strumenti di misura. In particolare uno studio sul MMSE ha concluso che "... a change score from 20 to 21 could reflect a decrease to 19 or an increase to 23" [20]. La variabilità della scala ADAS Cog non è risultata minore o comunque trascurabile quando valutata con studi formali (*Int J Geriatr Psychiatry* 2004; 19:41-50, *Alzheimer disease and Associated disorders* vol 15, No 4:174-183) [21].

Queste osservazioni portano a concludere che la possibilità di prevedere la sopravvivenza, la progressione della disabilità o del deficit cognitivo in singoli pazienti è estremamente limitata dalla variabilità intrinseca di questi eventi e dalla variabilità delle scale di rilevazione. L'implicazione pratica è che difficilmente il clinico potrà apprezzare sul singolo paziente l'efficacia dei suoi interventi basandosi sugli *score* delle scale di valutazione. Ciò non significa che altri strumenti, il giudizio soggettivo dei pazienti o dei *caregiver* e lo stesso intuito clinico non siano in grado

di apprezzare effettivi cambiamenti indotti dai trattamenti nel paziente.

Se quanto sopra affermato è vero ne consegue che l'efficacia degli interventi attualmente disponibili, dai quali ci si può ottimisticamente attendere solo moderate/modeste modificazioni della storia naturale, può essere apprezzata solo con studi adeguatamente controllati. Questa affermazione appare poi particolarmente appropriata per le demenze e la AD che rappresentano le malattie la cui occorrenza presenta il più importante trend di aumento in virtù dell'invecchiamento della popolazione e della crescente attenzione diagnostica. In particolare poi solo gli studi randomizzati e controllati, che garantiscono la comparabilità delle popolazioni (trattati e non trattati) messe a confronto, possono produrre evidenze scientifiche valide di efficacia o inefficacia. In quest'ottica gli studi osservazionali nelle demenze e nella AD sono strumenti assolutamente inadeguati per lo scopo della dimostrazione di efficacia. Per questo motivo sarebbe opportuno introdurre la pratica della randomizzazione e dello studio controllato anche quando si desidera valutare l'efficacia in pratica nell'ambito di iniziative a livello del servizio sanitario pubblico. A livello europeo non mancano gli esempi della praticabilità di una tale strategia come documenta un recente studio inglese [22].

#### **Effetto della variabilità sulle valutazioni di efficacia**

Da ultimo ci si deve chiedere quale sia l'effetto della variabilità delle scale di misura e degli *outcome* sulla valutazione di efficacia. La variabilità intrinseca al paziente e quella delle scale di misura, in studi controllati, assumono le caratteristiche di un errore casuale che è egualmente distribuito fra soggetti trattati e soggetti di controllo in virtù della randomizzazione. È stato dimostrato che questo errore casuale ha l'effetto di ridurre le eventuali differenze fra trattati e non trattati (Appendice) e quindi si può concludere che l'efficacia dimostrata dai trattamenti oggi disponibili per la demenza rappresenta una sottostima del loro reale effetto. Questa sottostima può rendere ragione della discrepanza nel giudizio di efficacia che spesso si osserva fra medici e parenti che spesso riportano un'efficacia che il clinico non è in grado di documentare o obiettivare.

#### **BIBLIOGRAFIA**

1. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975 Nov;12(3):189-98.
2. Roe H, Domonick W, Albert, Steven M *et al.* The rate of cognitive decline and risk of reaching clinical milestones in Alzheimer disease. *Arch Neurol* 2003;60:1137-42.
3. Bonsignore M, Heun R. Mortality in Alzheimer's disease. *Dement Geriatr Cognitve Disorder* 2003;15:231-6.
4. Larson EB, Shadlen MF, Wang L, McCormick WC, Bowen JD, Teri L, Kukull WA. Survival after initial diagnosis of Alzheimer disease. *Ann Intern Med* 2004;140(7):501-9.
5. Brookmeyer R, Corrada MM, Curriero FC, Kawas C. Survival following a diagnosis of Alzheimer disease. *Arch Neurol* 2002;59(11):1764-7.
6. Lindeboom R, Vermeulen M, Holman R, De Haan RJ. Activities of daily living instruments: optimizing scales for neurologic assessments. *Neurology* 2003;60(5):738-42.
7. Persson M, Nilsson S, Iwarsson S. Development of multi-disciplinary team I-ADL assessment in community health care: an interrater reliability study of the measure of instrumental daily activity. *Arch Gerontol Geriatr* 1999;29(2):149-63.
8. Reynolds T, Thornicroft G, Abas M, Woods B, Hoe J, Leese M, Orrell M. Camberwell Assessment of Need for the Elderly (CANE). Development, validity and reliability. *Br J Psychiatr* 2000;176:444-52.
9. Gosman-Hedstrom G, Svensson E. Parallel reliability of the functional independence measure and the Barthel ADL index. *Disabil Rehabil* 2000;22(16):702-15.
10. Hokoishi K, Ikeda M, Maki N, Nomura M, Torikawa S, Fujimoto N, Fukuhara R, Komori K, Tanabe H. Interrater reliability of the Physical Self-Maintenance Scale and the Instrumental Activities of Daily Living Scale in a variety of health professional representatives. *Aging Ment Health* 2001;5(1):38-40.
11. Avlund K, Schultz-Larsen K, Kreiner S. The measurement of instrumental ADL: content validity and construct validity. *Aging (Milano)* 1993;5(5):371-83.
12. Lechowski L, Dieudonné B, Tortar D *et al.* Role of behavioral disturbance in the loss of autonomy for activities of daily living in Alzheimer patients. *Int J Geriatr Psychiatry* 2003;18:977-82.
13. Schultz SK, Ellingrod VL, Moser DJ, Kutschner E, Turvey C, Arndt S. The influence of cognitive impairment and psychiatric symptoms on daily functioning in nursing facilities: a longitudinal study. *Ann Clin Psychiatry* 2002;14(4):209-13.
14. Aguero-Torres H, Thomas VS, Winblad B, Fratiglioni L. The impact of somatic and cognitive disorders on the functional status of the elderly. *J Clin Epidemiol* 2002;55(10):1007-12.

15. Caro J, Ward A, Ishak K, Migliaccio-Walle K, Getsios D, Papadopoulos G, Torfs K. To what degree does cognitive impairment in Alzheimer's disease predict dependence of patients on caregivers? *BMC Neurol* 2002;2(1):6.

16. Sauvaget C, Yamada M, Fujiwara S, Sasaki H, Mimori Y. Dementia as a predictor of functional disability: a four-year follow-up study. *Gerontology* 2002;48(4):226-33.

17. Espiritu DA, Rashid H, Mast BT, Fitzgerald J, Steinberg J, Lichtenberg PA. Depression, cognitive impairment and function in Alzheimer's disease. *Int J Geriatr Psychiatry* 2001;16(11):1098-103.

18. Artero S, Touchon J, Ritchie K. Disability and mild cognitive impairment: a longitudinal population-based study. *Int J Geriatr Psychiatry* 2001;16(11):1092-7.

19. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry* 1984;141:1356-64.

20. Peter Bowie, Tim Branton, John Holmes. Should the mini mental state examination be used to monitor dementia treatments? *Lancet* 1999;354:1527-8.

21. Verhey FR, Houx P, Van Lang N *et al.* Cross-national comparison and validation of the Alzheimer's Disease Assessment Scale: results from the European Harmonization Project for Instruments in Dementia (EURO-HARPID). *Int J Geriatr Psychiatry* 2004;19(1):41-50.

22. Courtney C, Farrell D, Gray R, Hills R, Lynch L, Sellwood E, Edwards S, Hardyman W, Raftery J, Crome P, Lendon C, Shaw H, Bentham P; AD2000 Collaborative Group. Long-term donepezil treatment in 565 patients with Alzheimer's disease (AD2000): randomised double-blind trial. *Lancet* 2004;363(9427):2105-15.

**Appendice.** - Effetti della variabilità sulla stima di efficacia

Esito positivo	Trattamento	
	Si	No
No	A	B
Si	C	-

La tabella rappresenta la situazione reale in cui il trattamento causa un numero di esiti positivi pari a C mentre il non trattamento non induce nessun esito positivo. Per comodità assumiamo un eguale numero di trattati e di non trattati.

La variabilità intrinseca del fenomeno e delle scale di misurazione causerà due errori  $e_1$  ed  $e_2$ :

$e_1$  = errore tra i soggetti che non hanno avuto esiti positivi (Ae B)  
 $e_2$  = errore tra i soggetti che hanno avuto esiti positivi (C).

L'errore  $e_1$  è causa di falsi positivi mentre l'errore  $e_2$  è causa di falsi negativi dove per falsi positivi si intende soggetti che appaiono avere risposto positivamente al trattamento quando ciò è falso; e per falsi negativi si intende soggetti che appaiono non avere risposto positivamente al trattamento quando ciò è falso.

Considerato  $e_1$  ed  $e_2$  come proporzioni di errori la tabella iniziale risulterà così modificata:

Esito positivo	Trattamento	
	Si	No
No	$A - e_1 * A + e_2 * C$	$B - e_1 * B$
Si	$C + e_1 * A - e_2 * C$	$e_1 * B$

L'efficacia misurata in queste condizioni sarà pertanto:

$$\text{efficacia} = C + e_1 * A - e_2 * C - e_1 * B$$

dove  $e_2 * C$  e  $e_1 * B$  tenderanno a far diminuire l'efficacia misurata rispetto a quella reale e dove  $e_1 * A$  tenderà invece ad aumentarla. Ma  $e_1 * A$  è minore di  $e_1 * B$  perché A è minore di B e quindi C (efficacia reale) risulta essere maggiore dell'efficacia misurata in condizioni di variabilità degli esiti non attribuibile all'intervento in esame.