

Il reporting di trial randomizzati controllati. Parte 1

Francesco Burrai

Professore a contratto, Università di Bologna, Bologna

THE REPORTING OF RANDOMIZED CONTROLLED TRIALS. PART 1

Abstract. Introduction. This paper is the first of two articles about the fundamental problem regarding the proper use of randomized controlled trials (RCT) in Italy in the research fields of nephrology and dialysis. The gold standard for measuring the effectiveness of an health care intervention is the RCT. Numerous RCTs show several methodological gaps, particularly evident in terms of the quality of trial reporting.

Objective. To indicate a guideline for trial reporting that can be useful both to authors, for a proper and rigorous structuring and conduct of the trial, and readers, providing the methodological tools for a critical reading of any RCT.

Materials and Methods. The reporting tool is the methodological standard ensuring the rigor of an RCT. The methodology used is that of analysing, in a concise but complete manner, the basic steps of a reporting that a trial must respect.

Results. Compliance with a guideline of RCT reporting shows results that are unambiguous, clear, not manipulated or distorted, and produced by a sound and rigorous methodology in each stage of the RCT, thus allowing a proper understanding of the results and their possible generalization.

Conclusions. The authors should systematically use the standard guidelines for the production of RCTs, respecting each indication, in order to ensure a useful and effective progress of the scientific knowledge, as well as the ethics and professionalism of the research.

Key words: Random control trial, Research, Reporting

Conflict of interest: None.

Financial support: None.

Accettato: 30 Maggio 2014



Francesco Burrai

Introduzione

Per valutare l'efficacia di un intervento assistenziale e sanitario, il *gold standard* è rappresentato dal *trial* controllato randomizzato, meglio conosciuto con il termine in inglese *Random Controlled Trial* (RCT). La metodologia RCT è considerata l'unica capace di ridurre notevolmente o di eliminare vari *bias* che invece sono presenti in altre metodologie di studio. Ovviamente, RCT che non vengono condotti in maniera rigorosa presenteranno diversi *bias* nei vari *step* del *trial*, inficiando negativamente il valore dei risultati. Appare fondamentale, da parte degli Autori, la strutturazione di un RCT che abbia un *reporting* veritiero, chiaro e trasparente, permettendo ai lettori la comprensione del *trial* in ogni sua fase, in definitiva sull'utilità e sulla generalizzabilità o meno dei risultati di un *trial*. *Trial* che non hanno un rigore metodologico, che riportano informazioni incomplete, selettive, non chiare e non trasparenti, violano

non solo qualsiasi valore scientifico del *trial*, ma non sono né eticamente corretti né moralmente accettabili. Due classici esempi sono, per esempio, nella fase di analisi, la scelta di metodologie statistiche sbilanciate verso una maggiore probabilità di dare significati importanti a favore dell'intervento oppure, nella fase dei risultati, non presentare valori non favorevoli all'intervento. Obiettivo di questo lavoro è quello di analizzare in maniera sintetica, ma completa, i passi e le fasi fondamentali che un *trial* deve rispettare, indicando le informazioni utili per un suo corretto *reporting*. In questo primo articolo analizzeremo la sezione del titolo, l'*abstract*, l'introduzione, gli obiettivi o le ipotesi, i metodi, il disegno, i pazienti, il *setting*, gli interventi, gli *outcome*, la dimensione del campione, la randomizzazione e i metodi statistici.

Il titolo

L'indicizzazione degli articoli sui *database* elettronici, come per esempio *PubMed*, è legata spesso alla presenza nel titolo

del lavoro del termine “*randomized*”. Tale termine può avere la funzione strategica di una *key word* inserita nel titolo e il mancato inserimento potrebbe creare problemi nell’indicizzazione e non facilitare l’identificazione (1).

L’abstract

La presenza di un *abstract* presenta due funzioni fondamentali:

- 1) i lettori possono non avere la possibilità di accedere alla versione *full text* dell’articolo, dunque reperiscono informazioni esclusivamente dal testo dell’*abstract* (2);
- 2) i lettori possono avere la possibilità di accedere al *full text* e decidono di farlo solo se le informazioni dell’*abstract* sono interessanti per la loro ricerca e, dunque, procedono con la lettura integrale del *full text*, il quale, a volte, presenta dei costi di *download*.

Un *abstract* deve avere la capacità di sintetizzare adeguatamente, in modo chiaro, rappresentativo, sufficiente e trasparente le informazioni riguardanti la metodologia e i risultati di un *trial*. Per ottenere questo risultato le informazioni devono essere congruenti con il *full text*, accurate e non presentare informazioni che non sono presenti nel *full text*; inoltre, devono evitare omissioni di risultati negativi. La metodologia più efficace per offrire al lettore uno strumento di valutazione rapida e corretta di un lavoro e per evitare interpretazioni non corrette (3, 4) consiste nel presentare un *trial* sotto forma di *abstract* strutturato. Un *abstract* strutturato è formato da sezioni riguardanti il disegno dello studio, metodi, risultati e conclusioni, cioè dalle informazioni fondamentali sul disegno di ricerca, sulla sua conduzione, sull’analisi e sull’interpretazione dei risultati (5).

L’introduzione

In questa sezione, gli Autori procedono con una revisione sistematica della letteratura per mostrare la presenza o meno di *trial* simili ed evidenziarne i risultati significativi (6). La revisione della letteratura è generalmente eseguita a livello infermieristico su questi principali *database*: *PubMed*, *CINAHL*, *Cochrane*. L’accuratezza della revisione scientifica, che deve essere formata da una bibliografia recente, indica il grado di conoscenza da parte degli Autori del miglior *background* scientifico a livello internazionale. Una profonda revisione della letteratura è legata alla giustificazione della necessità di un nuovo *trial* e ha implicazioni non solo di risorse economiche, umane e di tempo, ma anche etiche, legate alla Dichiarazione di Helsinki, in cui si evidenzia la non eticità nell’espone i pazienti a potenziali rischi in studi che non hanno nessuna utilità scientifica (7, 8).

Accanto alla revisione, deve essere presente il razionale del *trial* e il suo schema. Un razionale può avere una valenza pragmatica, in cui per esempio si portano informazioni che guidano l’assistenza infermieristica confrontando due tipologie di interventi infermieristici con i relativi benefici o effetti negativi, oppure una valenza esplicativa, in cui per esempio si evidenziano informazioni relative alla valutazione di una

possibile influenza di un intervento infermieristico su un *outcome* assistenziale.

Gli obiettivi o le ipotesi

Uno o più obiettivi rappresentano le domande e i quesiti a cui un *trial* deve rispondere. Una o più ipotesi rappresentano invece aspetti più specifici rispetto agli obiettivi, che sono sottoposti a un’analisi statistica. Generalmente, gli obiettivi sono legati allo studio degli effetti di un determinato intervento infermieristico.

I metodi

Nel capitolo metodi, gli Autori devono presentare le informazioni del *trial* nelle sezioni disegno del *trial*, pazienti, interventi, *outcome*, dimensione del campione, randomizzazione, cecità e metodi statistici.

Il disegno

Un disegno classico di un *trial* presenta la metodologia di pianificazione rispetto, soprattutto, alla randomizzazione e all’allocazione dei pazienti a uno dei due gruppi paralleli. I *trial* generalmente presentano una randomizzazione bilanciata ovvero un rapporto di allocazione dei pazienti 1:1 tra i due gruppi.

I pazienti

Per i pazienti è fondamentale stabilire quali sono i criteri su cui si basano i ricercatori per selezionare i pazienti del *trial*, i quali sono fondamentali per l’interpretazione del *trial*, e quali sono il *setting* e le aree geografiche in cui sono stati raccolti i dati. I criteri di arruolamento indicano l’eleggibilità del tipo di paziente idoneo a entrare nel *trial*. La correttezza e la chiarezza dei criteri di eleggibilità dei pazienti sono strategici per individuare la dimensione di generalizzabilità e applicabilità dei risultati e la rilevanza per l’assistenza infermieristica (9). Tali criteri includono o eludono i pazienti all’arruolamento nel *trial*, dunque precedono il processo di randomizzazione e si dividono in criteri di inclusione, che includono nell’arruolamento soggetti che possiedono caratteristiche specifiche legate alla tipologia della patologia o a un tipo di assistenza, e in quelli di esclusione, che escludono dall’arruolamento i soggetti che possono presentare rischi legati all’intervento o che non incontrano norme etiche o legali, come per esempio il non accettare il consenso informato. In questa sezione è importante anche evidenziare il metodo di reclutamento, che può essere rappresentato per esempio da una partecipazione volontaria oppure da una selezione su variabili specifiche.

Il setting

Anche le informazioni riguardanti il *setting* e le aree geografiche in cui sono stati raccolti i dati sono importanti nella definizione di applicabilità e generalizzabilità del *trial* e, dunque, della sua validità esterna. Queste informazioni sono fondamentali per il

lettore, il quale ha gli strumenti critici essenziali per capire se i risultati del *trial* possono essere utili per il proprio contesto assistenziale (10). Infatti, la tipologia del *setting* è influenzata da svariate variabili, di tipo sanitario, organizzativo, sociale, culturale ed economico.

Importante specificare chi eroga l'intervento, in quale contesto e *setting* assistenziale, in quale nazione, città, ospedale, unità operativa o ambulatorio, se in assistenza domiciliare, in quante sedi e se è mono o multicentrico.

Gli interventi

In un *trial* con metodologia a gruppi paralleli, la descrizione degli interventi deve essere eseguita accuratamente non solo per il gruppo sperimentale, ma anche per il gruppo di controllo (11). Gli Autori devono descrivere dettagliatamente gli interventi, la tipologia, le modalità tecniche assistenziali, i tempi, la durata e se ci sono condizioni in cui non è possibile applicare l'intervento. In molti RCT il gruppo di controllo riceve solo l'assistenza *standard* senza l'applicazione sperimentale; in tal caso, è necessario descrivere in cosa consiste tale assistenza infermieristica.

Gli outcome

La potenza informativa di un *trial* è legata alla comparazione dei risultati tra i due gruppi paralleli, quello sperimentale e quello di controllo. Tale comparazione tra i due gruppi avviene attraverso la valutazione dei risultati mostrati dagli *outcome* o *endpoint*. L'*outcome* individua una specifica dimensione assistenziale misurabile. La misurazione di un *outcome* deve essere eseguita attraverso strumenti validati, generalmente scale e Linee Guida, le quali devono essere accuratamente descritte nella loro provenienza e nelle loro caratteristiche, anche statistiche (12, 13). In base alla rilevanza che gli Autori attribuiscono agli *outcome*, questi vengono classificati in primari e secondari. Si raccomanda l'individuazione di un solo *outcome* primario. L'*outcome* primario rappresenta un esito pre-definito che per i ricercatori ha una valenza di estremo interesse, dunque di maggiore rilevanza dal punto di vista assistenziale rispetto ad altri esiti secondari, e come tale deve essere ben definito nel *trial*. Esiti che non presentano lo stesso livello di rilevanza sono definiti *outcome* secondari, i quali, nei *trial*, sono rappresentati in diverso numero, in base alla complessità del *trial*. Particolare attenzione deve essere posta alla giusta individuazione dell'*outcome* primario, al quale è legato il calcolo della dimensione del campione. Fondamentale è la corretta descrizione degli *outcome* primari e secondari, perché permette il loro eventuale utilizzo in altri studi da altri Autori. Negli studi RCT longitudinali in cui si studiano gli effetti degli interventi nel tempo, gli Autori devono specificare per tutti gli *outcome* in quale momento temporale avvengano le misurazioni ovvero definire quali sono i *time point*.

La dimensione del campione

La definizione del numero dei pazienti che costituirà il campione è legata al concetto di potenza statistica dello studio, ovvero

alla presenza di una determinata numerosità dei pazienti per avere un'elevata probabilità di rilevare differenze cliniche statisticamente significative. Sembra chiaro che, se l'Autore è interessato a stimare piccole differenze cliniche, maggiore deve essere la numerosità del campione, dunque esiste un rapporto inversamente proporzionale tra la dimensione del campione e la dimensione delle differenze cliniche che si vogliono stimare. Dunque, il *trial* deve presentare una potenza statistica adeguata, il che significa che non dovrebbe presentare una numerosità campionaria sottodimensionata, perché questo inficia notevolmente la significatività dei risultati. Infatti dimensioni piccole del campione presentano il rischio di affermare che non esistono differenze statisticamente significative tra i gruppi paralleli, conclusione errata, perché la *sample size* è troppo piccola per poter giustificare tale affermazione (14), dunque hanno una potenza statistica insufficiente per stimare effetti clinicamente significativi. Per dare al lettore gli strumenti critici per valutare se il *trial* presenta la dimensione campionaria prevista, gli Autori devono indicare la metodologia di stima della dimensione del campione, che può basarsi generalmente o sull'*outcome* primario del *trial* o su studi precedenti, considerando la percentuale di incidenza degli abbandoni, indicando il livello di significatività e la potenza in percentuale. *Software* utilizzati per il calcolo della dimensione del campione sono facilmente reperibili su *Internet*.

La randomizzazione

Il processo di randomizzazione si basa sul concetto che i pazienti devono essere allocati ai due gruppi paralleli in maniera casuale e imprevedibile come mostrato in Figura 1. La casualità (*random*) e l'imprevedibilità sono fenomeni che garantiscono al *trial* un notevole riparo da eventuali *bias* nell'assegnazione dei pazienti ai gruppi.

Infatti ogni paziente eleggibile sottoposto al processo di randomizzazione, dunque a una metodologia che utilizza la casualità e l'imprevedibilità, presenta una probabilità X di essere assegnato a uno dei gruppi del *trial*.

La sequenza di allocazione casuale dei pazienti ha il principale obiettivo di eliminare qualsiasi fenomeno deterministico e non casuale e, per ottenere tale risultato, i ricercatori spesso utilizzano *software* che producono una lista di numeri casuali, come *SATA*[®] o *SPSS*[®] oppure una tabella di numeri *random*. Accanto al termine randomizzazione spesso si trova il termine semplice, che indica che l'allocazione casuale dei pazienti nei due gruppi presenta un rapporto 1:1. Prodotta la lista di allocazione ai gruppi paralleli, è fondamentale specificare nel *trial* come viene nascosta e occultata all'infermiere che procede all'arruolamento dei pazienti l'informazione riguardante il gruppo al quale il paziente successivo verrà assegnato. Questo processo di occultamento della lista di randomizzazione garantisce che qualsiasi decisione inerente l'inclusione o la non inclusione del paziente sia eseguita sempre, senza conoscere l'assegnazione successiva (15). Come si può intuire, in questa fase si cerca di mantenere attive e preservate la casualità e l'imprevedibilità dell'allocazione da eventuali *bias* prodotti da un intervento umano di tipo manipolativo, concentrandosi sulla fase di garanzia

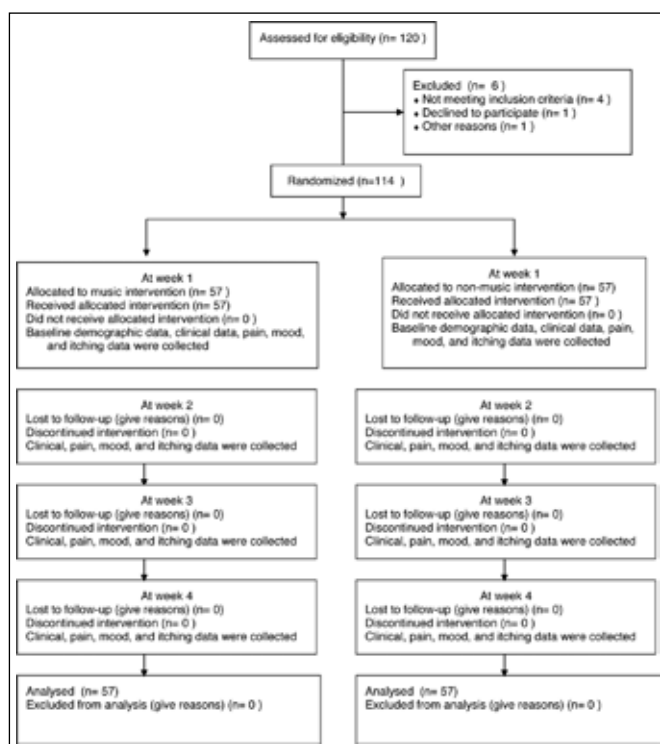


Fig. 1 - Flow chart.

della selezione e della sequenza di allocazione sino all'assegnazione. Per garantire tale processo, diversi *trial* utilizzano parti terze nella fase di occultamento della lista di randomizzazione, come centri esterni.

Un secondo processo di occultamento delle informazioni viene attuato nella fase degli interventi. In questa fase, denominata *blinding*, blindata, i soggetti coinvolti nel *trial* non hanno le informazioni riguardanti gli interventi assegnati. I soggetti che possono essere mantenuti *blinded* sono i pazienti, gli infermieri, i sanitari, i professionisti che hanno raccolto dati, i professionisti che valutano gli esiti e gli statistici.

Questa non conoscenza permette di preservare eventuali *bias* dovuti alla consapevolezza di tali informazioni. Per esempio, pazienti che sanno che il tipo di intervento a cui sono sottoposti è innovativo possono rispondere in maniera più favorevole, infermieri che conoscono il tipo di intervento possono alterarlo o manipolarlo, professionisti che raccolgono gli esiti, conoscendo gli interventi, possono raccoglierti in maniera diversa, ripetere le misure che hanno dato valori anomali o influenzare i pazienti, professionisti che valutano gli esiti possono introdurre valutazioni distorte oppure statistici possono eseguire calcoli solo su casi favorevoli o con metodologie statistiche non idonee, scegliendo statistiche mirate su esiti favorevoli (16-18).

Metodi statistici

Nel *trial* è indispensabile descrivere le metodologie adottate per l'analisi dei dati. Generalmente, le statistiche utilizzate nei *trial* stimano tra i gruppi l'efficacia di un intervento infermieristico.

Esistono due metodi statistici per presentare la stima dell'efficacia: indicazione dei limiti di confidenza e significatività statistica. I limiti di confidenza (IC), convenzionalmente posti al 95%, sono definiti come il *range* dei valori entro cui è contenuto, con una probabilità del 95%, il valore reale.

La significatività statistica è indicata con il *P-value*, il quale rappresenta la probabilità che il valore dell'efficacia dell'intervento sia dovuto alla casualità, dunque indica la significatività o non significatività statistica di un risultato.

Per convenzione, il limite soglia, nella stragrande maggioranza dei *trial*, è $p < 0.05$. Valori di p superiori a 0.05 indicano che il risultato non è statisticamente significativo. I *trial* dovrebbero indicare i valori di p in maniera precisa, per esempio $p = 0.0001$, e non imprecisa, $p < 0.005$.

Conclusioni

In questo primo articolo abbiamo analizzato le sezioni di un RCT che riguardano il titolo, l'*abstract*, l'introduzione, gli obiettivi o le ipotesi, i metodi e gli *outcome*. Nel secondo articolo analizzeremo le successive sezioni, ovvero, nella sezione dei risultati, la *flow chart* del flusso dei pazienti, le date e il periodo di reclutamento, le caratteristiche demografiche e cliniche, i risultati per gli *outcome* e i risultati per gli effetti indesiderati e avversi. Nella sezione discussione analizzeremo i limiti di un *trial*, la generalizzabilità dei risultati, il confronto dei risultati, la registrazione del *trial* e, infine, il finanziamento di un *trial*.

Procederemo, poi, a una fase di conclusione, riportando le indicazioni fondamentali per la strutturazione di un *trial* e gli aspetti metodologici ed etici.

Riassunto

Introduzione. Questo lavoro è il primo di due articoli che affronteranno la problematica fondamentale dell'utilizzo corretto della metodologia RCT nella ricerca in Italia nel campo nefrologico e dialitico. Il *gold standard* per misurare l'efficacia di un intervento assistenziale e sanitario è rappresentato dal *trial* controllato randomizzato. Numerosi RCT presentano ancora oggi però diversi *gap* metodologici, evidenziati a livello della qualità del *reporting* del *trial*.

Obiettivo. Mostrare una Linea Guida per il *reporting* di un *trial*, utile per gli Autori per una strutturazione e conduzione corrette e rigorose del *trial* e per i lettori per avere gli strumenti metodologici per una lettura critica di qualsiasi RCT.

Strumenti e Metodi. Il *reporting* è lo strumento metodologico *standard* per garantire la rigorosità di un RCT. La metodologia utilizzata è quella di analizzare in maniera sintetica, ma completa, le fasi fondamentali che compongono un *reporting* che un *trial* deve rispettare.

Risultati. Il rispetto di una Linea Guida del *reporting* di un RCT mostra risultati non ambigui, chiari, non manipolati o distorti, prodotti da una metodologia corretta e rigorosa in

ogni fase del RCT, permettendo una comprensione corretta dei risultati e la loro possibile generalizzazione.

Conclusioni. Gli Autori dovrebbero usare sistematicamente le Linee Guida *standard* per la produzione di RCT, rispettandone ogni indicazione, per garantire un'utile ed efficace progressione della conoscenza scientifica nonché l'eticità e la serietà della ricerca.

Parole chiave: Random Controlled Trial, Ricerca, Reporting

Dichiarazione di conflitto di interessi: L'Autore dichiara di non avere conflitto di interessi.

Contributi economici agli Autori: L'Autore dichiara di non aver ricevuto sponsorizzazioni economiche per la preparazione dell'articolo.

Indirizzo degli Autori:

Prof. Francesco Burrai
Professore a contratto
Università di Bologna
Viale Ercolani 6
40138 Bologna
francesco.burrai2@unibo.it

Bibliografia

1. Dickersin K, Manheimer E, Wieland S, Robinson KA, Lefebvre C, McDonald S. Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. *Eval Health Prof* 2002; 25 (1): 38-64.
2. Barbour V, Chinnock P, Cohen B, Yamey G. The impact of open access upon public health. *Bull World Health Organ* 2006; 84 (5): 339.
3. Ioannidis JP, Evans SJ, Gøtzsche PC, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004; 141 (10): 781-8.
4. Ioannidis JP, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. *JAMA* 2001; 285 (4): 437-43.
5. Haynes RB, Mulrow CD, Huth EJ, Altman DG, Gardner MJ. More informative abstracts revisited. *Ann Intern Med* 1990; 113 (1): 69-76.
6. Savulescu J, Chalmers I, Blunt J. Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. *BMJ* 1996; 313 (7069): 1390-3.
7. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992; 327 (4): 248-54.
8. Fergusson D, Glass KC, Hutton B, Shapiro S. Randomized controlled trials of aprotinin in cardiac surgery: could clinical equipoise have stopped the bleeding? *Clin Trials* 2005; 2 (3): 218-29.
9. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet* 2005; 365 (9453): 82-93.
10. Weiss NS, Koepsell TD, Psaty BM. Generalizability of the results of randomized trials. *Arch Intern Med* 2008; 168 (2): 133-5.
11. Mc Dowell I. *Measuring Health: A Guide to Rating Scales and Questionnaires*. 3rd ed. New York: Oxford University Press 2006.
12. Streiner DL, Norman GR. *Health measurement scales: A practical guide to their development and use*. 3rd ed. Oxford: Oxford University Press 1995.
13. Altman DG, Bland JM. Absence of evidence is not evidence of absence. *BMJ* 1995; 311 (7003): 485.
14. Chalmers TC, Levin H, Sacks HS, Reitman D, Berrier J, Nagalingam R. Meta-analysis of clinical trials as a scientific discipline. I: Control of bias and comparison with large cooperative trials. *Stat Med* 1987; 6 (3): 315-28.
15. Gøtzsche PC. Believability of relative risks and odds ratios in abstracts: cross sectional study. *BMJ* 2006; 333 (7561): 231-4.
16. Wood L, Egger M, Gluud LL, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008; 336 (7644): 601-5.
17. Guyatt GH, Pugsley SO, Sullivan MJ, et al. Effect of encouragement on walking test performance. *Thorax* 1984; 39 (11): 818-22.
18. Noseworthy JH, Ebers GC, Vandervoort MK, Farquhar RE, Yetisir E, Roberts R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994; 44 (1): 16-20.