



# Los ensayos clínicos

## Los orígenes

Margarita García  
Unidad de Investigación Clínica  
Institut Català d'Oncologia



# Los ensayos clínicos

## Los orígenes

- ¿por qué hacemos ensayos clínicos?
- ¿por qué randomizamos?
- ¿para qué enmascaramos?
- ¿por qué controlamos?
- ¿por qué utilizamos placebo?



# Los ensayos clínicos

## Los orígenes

- ¿por qué no nos fiamos de los controles históricos?
- ¿por qué decimos que no podemos extrapolar más los resultados?
- ¿por qué no analizamos sólo aquellos pacientes que han sido tratados?



# ¿por qué hacemos ensayos clínicos?



En el siglo XVIII uno de los mayores problemas de la Royal Navy era el escorbuto

En 1747, a bordo de un buque llamado Salisbury, el cirujano **James Lind** decidió tratar a 12 marineros afectados:

- 25 gotas de elixir vitriol a dos de ellos
- un cuarto de sidra la día a otros dos
- 3 cucharadas de vinagre diarias a otros dos
- media pinta de agua de mar a dos más
- nuez moscada 3 veces al día a dos
- dos naranjas y un limón al día a 2

Este experimento cambió la política británica de alimentación en los barcos

*Lind, James (1753). A Treatise on the Scurvy. London: A. Millar.*

# Fuerza de la evidencia científica según los distintos tipos de estudios clínicos (Byar, 1978)



## ✓ Publicación de un caso

Única demostración de que es posible un acontecimiento clínico de interés

## ✓ Series de casos

Una demostración de ciertos acontecimientos clínicos posiblemente relacionados pero sujetos a grandes sesgos de selección

## ✓ Análisis de bases de datos

El tratamiento no está determinado por un diseño experimental sino por factores tales como preferencia del médico o el paciente. Es improbable que los datos hayan sido recogidos específicamente para evaluar la eficacia

## ✓ Estudios observacionales

El investigador aprovecha exposición "natural" o tto seleccionado y elige un grupo comparativo

## ✓ Ensayo clínico controlado

Se diseña para asignar el tratamiento. Se evalúan los objetivos de forma cuidadosa y los análisis son planeados a priori.

## ✓ Ensayos clínicos replicados

Verificación independiente

Byar DP. On combining information : historical controls, overviews, and comprehensive cohort studies. *Recent Results in Cancer Research* (1978 )111: 95-98 .



## Medicina basada en la evidencia

La Medicina Basada en la Evidencia (MBE; en inglés, *Evidence Based Medicine* (EBM)), es la utilización consciente, explícita y juiciosa de la mejor evidencia científica disponible para tomar decisiones sobre el cuidado y manejo de los pacientes individuales.

*Sacket DL et al. Evidence based medicine: what it is and what it isn't, BMJ 1996; 312: 71-72*





# ¿por qué randomizamos?

La técnica de la randomización había sido usada en agricultura en experimentos descritos por Fisher en 1926.

“... A real experimenter, in fact, so far from being willing to introduce an element of chance into the formation of his scientific conclusions, has been steadily exerting himself, in the planning of his experiments, and in their execution, to decrease or to eliminate [by randomization]\* all the causes of fortuitous variation which might obscure the evidence.”



*Statistical Methods and Scientific Inference, Edinburgh: Oliver and Boyd, 1956, p.97; \*3rd edition, 1973.*



# ¿por qué randomizamos?



El primer científico en introducir la asignación aleatoria (randomización) para evaluar un tratamiento fue Sir Austin Bradford Hill.

La técnica de la randomización había sido usada en agricultura en experimentos descritos por Fisher en 1926.

Aunque la consideraba, Bradford Hill no la recomendó inicialmente, pero sí en 1947, invocando no sólo su capacidad de evitar el error del azar, sino el sesgo de selección.



El primer estudio en que se utilizó fue un ensayo para evaluar la eficacia de la inmunización contra la tos ferina.

Sin embargo, este artículo fue publicado 3 años después del mucho más conocido de la estreptomycin en el tratamiento de la tuberculosis pulmonar

*Medical Research Council. Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment of pulmonary tuberculosis. [British Medical Journal](#) 1948;ii: 769-82*

*Medical Research Council whooping-Cough Immunization Committee. The prevention of whooping-cough by vaccination. [British Medical Journal](#) 1951;i: 1463-71.*



# Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials

Gordon C S Smith, Jill P Pell

## Abstract

**Objectives** To determine whether parachutes are effective in preventing major trauma related to gravitational challenge.

**Design** Systematic review of randomised controlled trials.

**Data sources:** Medline, Web of Science, Embase, and the Cochrane Library databases; appropriate internet sites and citation lists.

**Study selection:** Studies showing the effects of using a parachute during free fall.

**Main outcome measure** Death or major trauma, defined as an injury severity score > 15.

**Results** We were unable to identify any randomised controlled trials of parachute intervention.

**Conclusions** As with many interventions intended to prevent ill health, the effectiveness of parachutes has not been subjected to rigorous evaluation by using randomised controlled trials. Advocates of evidence based medicine have criticised the adoption of interventions evaluated by using only observational data. We think that everyone might benefit if the most radical protagonists of evidence based medicine organised and participated in a double blind, randomised, placebo controlled, crossover trial of the parachute.

## Introduction

The parachute is used in recreational, voluntary sector, and military settings to reduce the risk of orthopaedic, head, and soft tissue injury after gravitational challenge, typically in the context of jumping from an aircraft. The perception that parachutes are a successful intervention is based largely on anecdotal evidence. Observational data have shown that their use is associated with morbidity and mortality, due to both failure of the intervention<sup>1,2</sup> and iatrogenic complications.<sup>3</sup> In addition, "natural history" studies of free fall indicate that failure to take or deploy a parachute does not inevitably result in an adverse outcome.<sup>4</sup> We therefore undertook a systematic review of randomised controlled trials of parachutes.

## Methods

### Literature search

We conducted the review in accordance with the QUOROM (quality of reporting of meta-analyses) guidelines.<sup>2</sup> We searched for randomised controlled trials of parachute use on Medline, Web of Science, Embase, the Cochrane Library, appropriate internet sites, and citation lists. Search words employed were "parachute" and "trial." We imposed no language restriction and included any studies that entailed jumping from a height greater than 100 metres. The

accepted intervention was a fabric device, secured by strings to a harness worn by the participant and released (either automatically or manually) during free fall with the purpose of limiting the rate of descent. We excluded studies that had no control group.

### Definition of outcomes

The major outcomes studied were death or major trauma, defined as an injury severity score greater than 15.<sup>5</sup>

### Meta-analysis

Our statistical approach was to assess outcomes in parachute and control groups by odds ratios and quantified the precision of estimates by 95% confidence intervals. We chose the Mantel-Haenszel test to assess heterogeneity, and sensitivity and subgroup analyses and fixed effects weighted regression techniques to explore causes of heterogeneity. We selected a funnel plot to assess publication bias visually and Egger's and Begg's tests to test it quantitatively. Stata software, version 7.0, was the tool for all statistical analyses.

## Results

Our search strategy did not find any randomised controlled trials of the parachute.

## Discussion

### Evidence based pride and observational prejudice

It is a truth universally acknowledged that a medical intervention justified by observational data must be in want of verification through a randomised controlled

Department of Obstetrics and Gynaecology, Cambridge University, Cambridge CB2 2QQ

Gordon C S Smith  
*Professor*

Department of Public Health, Greater Glasgow NHS Board, Glasgow G5 8YU

Jill P Pell  
*Assistant*

Correspondence to:  
G C S Smith  
gcs2@cam.ac.uk

BMJ 2005;327:1459-61



Parachutes reduce the risk of injury after gravitational challenge, but their effectiveness has not been proved with randomised controlled trials

Smith GCS. BMJ 2005; 527:1459-61 .



2 motivos para randomizar:

- Eliminar *el sesgo de selección*.
- Controlar *los factores que influyen en el resultado*, relacionados con:
  - \* las características personales,
  - \* la enfermedad,
  - \* los cuidados aplicados.

*Hay consenso en la aplicación de esta metodología*



## EL SESGO DE SELECCIÓN

- Efecto bien conocido de la selección en la entrada de pacientes en un estudio
- Típico de fase II: tasas que no se confirman en fase III
- Se elimina con la randomización
- Es un riesgo potencial

### “Interferon Gamma-1b compared with Placebo in Metastatic Renal-Cell Carcinoma”

1. Los datos de fase II demostraron un 22% de TR y un 7% de RC
2. Los datos de este estudio fueron: 4.4% de TR (95%CI:1.4-11.5) y 3.3 % de RC

*Gleave et al. N Engl J Med 1998;338:1265-71*



## ¿Por qué y para qué randomizar?

- Es necesario que los dos grupos de tratamiento sean comparables, para factores pronósticos así como otros que puedan influir en el resultado.
- La variabilidad biológica puede hacer aparecer diferencias entre los dos grupos de tratamiento debidas sólo al azar.

**LA RANDOMIZACIÓN PUEDE ASEGURAR QUE LAS DIFERENCIAS QUE SE PUEBAN ENCONTRAR SE DEBAN SÓLO AL EFECTO DEL TRATAMIENTO**

- Es bien conocido que los estudios randomizados muestran menores efectos del tratamiento que los que usan controles históricos.



## ¿para qué enmascaramos?



A. MESMER



París, 1780 el famoso doctor Franz Mesmer utilizaba un fluido único universal, que en los animales se denominaba el magnetismo animal, para curar todo tipo de enfermedades, habiendo conseguido cientos de adeptos en poco tiempo.



## ¿para qué enmascaramos?

- Tal era su éxito y el peligro que representaba, que el propio rey Luis XVI convocó una comisión de expertos para comprobar qué había de cierto en todo aquello.
- La comisión real era liderada por Benjamin Franklin, y algunos de sus miembros eran Antoine de Lavoisier, Jean-Sylvain Bailly y Joseph-Ignace Guillotin.
- Ellos debían averiguar si, el más tarde llamado mesmerismo, tenía efectividad real o era sólo una ilusión de la mente.
- Entre los muchos experimentos que realizaron: tapar a los pacientes para que no viesen si estaban o no recibiendo magnetismo.
- Resultado: sus efectos sólo se manifestaron en aquellos pacientes a los que se dijo que estaban recibiendo el magnetismo, fuera cierto o no, por lo que quedó en evidencia que el efecto era un producto de la sugestión.

Este fue el primer experimento de la historia que utilizó una técnica de enmascaramiento (real en esta ocasión) y de él procede la denominación de “ciego” para definir al paciente, al médico o al estadístico, que no sabe qué tratamiento está recibiendo.

<http://en.wikipedia.org/wiki/Mesmerism>



Figure 1: The authors: double blinded versus single blinded



Figure 2: The authors blinded and masked

*Schulz and Grimes. Lancet, 2002*



# ¿por qué controlamos?



**biology**  
**letters**

*Biol. Lett.* (2006) 2, 36–38  
doi:10.1098/rsbl.2005.0411  
Published online 6 December 2005

## Video-induced yawning in stumptail macaques (*Macaca arctoides*)

Annika Paukner\* and James R. Anderson

Department of Psychology, University of Stirling, Stirling,  
FK9 4LA, UK

\*Author for correspondence (ap14@stir.ac.uk).



Figure 1. Frame sequences from sample videos. (a) Yawn sequence, (b) control sequence.

*Biol Lett* (2006)2, 36-38



**¿por qué utilizamos placebo?**

## EL EFECTO PLACEBO

El efecto "total" de una droga es la suma del efecto "activo" más el efecto placebo

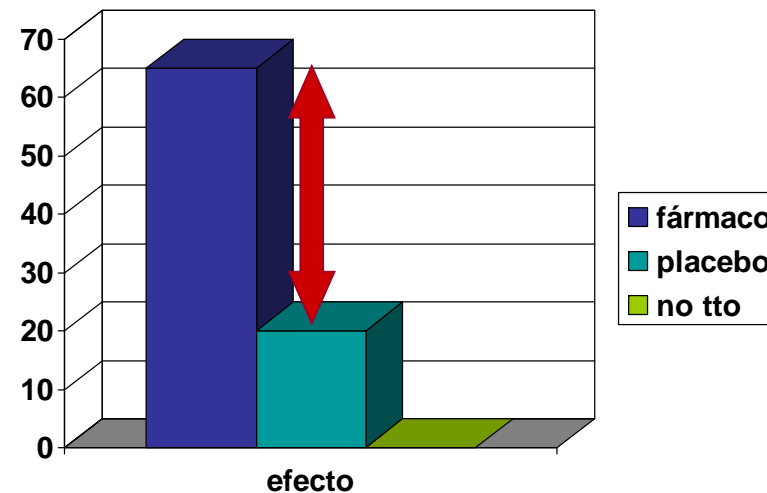
Una sustancia o procedimiento inactivo

Un efecto no relacionado con lo que hace una sustancia

- Uso generalizado del placebo en la Medicina pre-guerra
- *Beecher, 1955, JAMA: "The Powerful Placebo"*: de 1082 pacientes tratados con placebo, 35.2% experimentaron beneficios terapéuticos
- Detrás de esta etiqueta podrían estar: historia natural de una enfermedad, regresiones espontáneas, regresión a la media, cuidados de rutina de médicos o enfermeras, descanso, dieta, ejercicio, relajación, relación médico-paciente, condicionamientos clásicos, comportamientos aprendidos, expectativas, imaginación, creencias del médico y/o paciente...
- Hay diferencias culturales y sociales: en el tratamiento de la HTA el efecto en alemanes es menor que en cualquier otra parte del mundo
- Hay revisiones sistemáticas acerca de la existencia de un efecto placebo real, con resultados contradictorios, pero parece que no hay evidencia de efectos clínicamente importantes, excepto en el tratamiento del dolor y en evaluaciones subjetivas realizadas por los propios pacientes



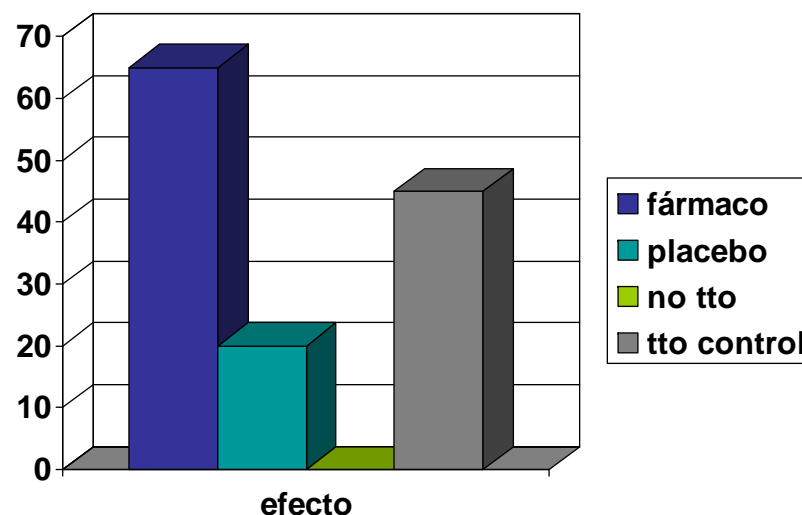
- En investigació clínic, se utilitza el placebo com a tractament “control”.
- Suele ser en forma de pastillas, pero puede ser una simulación con un aparato, una cirugía simulada...
- Tienen un efecto positivo en aquellos pacientes que piensan que están siendo tratados con drogas activas, por encima de aquellos que saben que no reciben tratamiento





### Ventajas del placebo:

- Facilita la aplicación del enmascaramiento y de la ocultación de la secuencia de aleatorización
- Disminuye la probabilidad de presentación del sesgo de información (por parte del participante o del observador) y del sesgo de selección
- Aumenta la validez interna de un ensayo



- Los estudios que ocultan la asignación obtienen estimaciones de efecto inferiores
- Suponen menores magnitudes de la diferencia, un tamaño de muestra más grande
- El uso de un grupo placebo requiere más muestra que un grupo sin tratamiento y menos que un grupo con control activo



## Problemas del placebo

- Problemas éticos
  - problemas de aceptabilidad por parte del médico y del paciente (más abandonos)
  - problema de validez externa
  - dificultades en el análisis estadístico
- 
- No se puede justificar cuando existe un tratamiento efectivo contra la enfermedad
  - El problema ético que comporta no se puede soslayar con una buena información al paciente de que puede que no reciba un tratamiento efectivo y la firma del consentimiento, ni sugiriendo que un periodo con placebo no será dañino.
  - Los controles con placebo deben ser ética y científicamente justificados



**¿por qué no nos fiamos de los  
controles históricos?**

# FENOMENO DE WILL ROGERS

- Descrito por Feinstein, 1985.
- Will ROGERS: “Cuando los Okies dejaron Oklahoma y se fueron a California, elevaron el nivel medio de inteligencia de los 2 estados”
- Advierte del peligro de los controles históricos.



*AR Feinstein The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. NEJM 1985;312:1604-1608.*





# Understanding variations in survival for colorectal cancer in Europe: a EURO CARE high resolution study

G Gatta, R Capocaccia, M Sant, C M J Bell, J W W Coebergh, R A M Damhuis, J Faivre, C Martinez-Garcia, J Pawlega, M Ponz de Leon, D Pottier, N Raverdy, E M I Williams, F Berrino

*Gut* 2000;47:533–538

## Colorectal cancer survival in Europe: the Will Rogers phenomenon revisited

---

M SHAHRIER

D J AHNEN

*Gastroenterology Division, Department of Medicine,  
University of Colorado School of Medicine, Colorado, USA  
and  
Department of Veterans Affairs Medical Center,  
Denver, Colorado, USA  
dennis.ahnen@uchsc.edu*



**¿por qué decimos que no podemos  
extrapolar los resultados?**

## “Why do old men have big ears?”

*Heathcote BMJ 1995;311:1668*

- Miden las orejas de 206 pacientes consecutivos
- Obtienen u
- Concluyen: año)

Length of ear (mm)



0.22 mm al

edad)

3 mm

2 mm

100

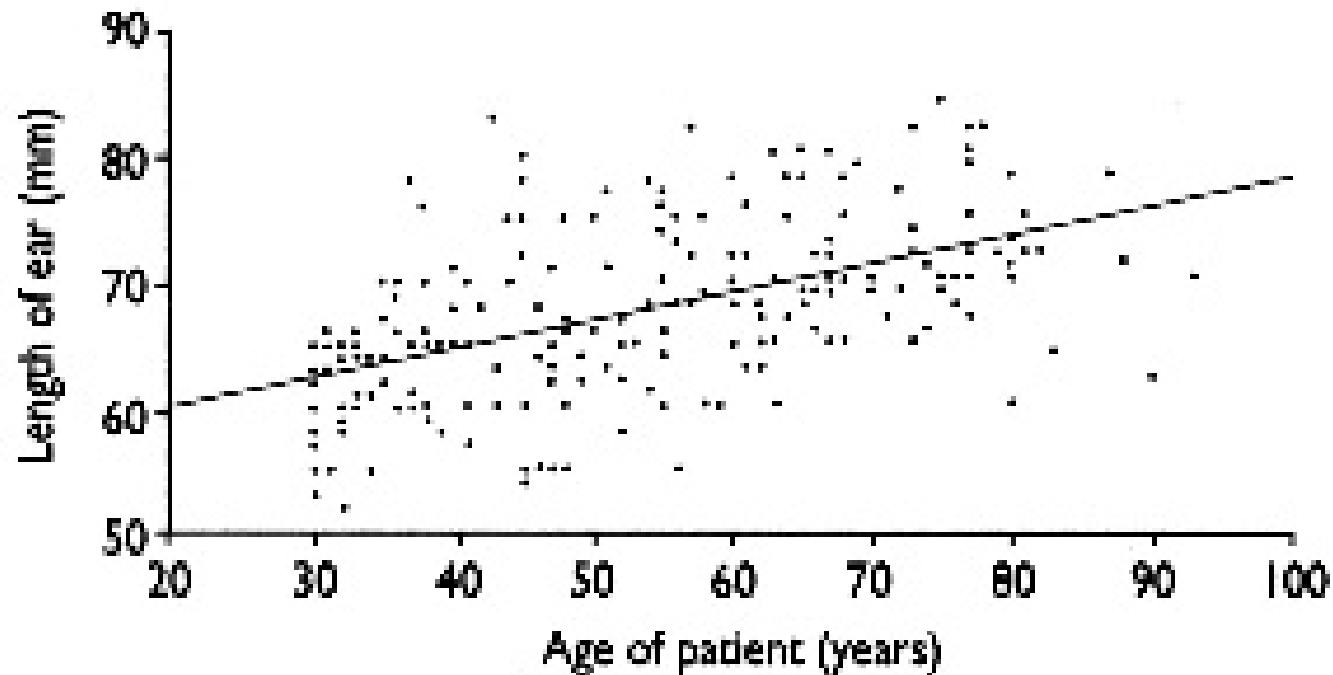
## “Generalisation and extrapolation”

*Altman and Bland, BMJ 1998,317:409-410*

## “Why do old men have big ears?”

*Heathcote BMJ 1995;311:1668*

- Miden las orejas de 206 pacientes consecutivos
- Obtienen una ecuación de regresión lineal:
- Concluyen: a mayor edad, mayores orejas (crecimiento medio de 0.22 mm al año)



### “Generalisation and extrapolation”

*Altman and Bland, BMJ 1998,317:409-410*



**¿por qué no analizamos sólo aquellos  
pacientes que han sido tratados?**

## Historia del “Coronary Drug Project”



- Estudio del impacto del clofibrato sobre la supervivencia a 5 años después de un ataque cardíaco.
- Clofibrato versus placebo
- Se concluyó que no había efecto del tratamiento.

Se decidió realizar un reanálisis basado en el tratamiento real

• La tasa de mortalidad en los que tomaron todo su tratamiento fue un 40% menor de los que no lo tomaron.

• Los que tomaron adecuadamente su placebo también tenían ventaja en supervivencia

Es la adherencia con el tratamiento lo que conlleva un mejor pronóstico

Ilustra la “Falacia del análisis según tratamiento real”

*Rachel Nowak. Clinical Trial Monitoring: Hit or Miss? Science 1994;264:1534-1241.*



## **Influence of Adherence to Treatment and Response of Cholesterol on Mortality in the Coronary Drug Project**

The Coronary Drug Project Research Group

The Coronary Drug Project was carried out to evaluate the efficacy and safety of several lipid-influencing drugs in the long-term treatment of coronary heart disease. The five-year mortality in 1103 men treated with clofibrate was 20.0 per cent, as compared with 20.9 per cent in 2789 men given placebo ( $P = 0.55$ ). Good adherers to clofibrate, i.e., patients who took 80 per cent or more of the protocol prescription during the five-year follow-up period, had a substantially lower five-year mortality than did poor adherers to clofibrate (15.0 vs. 24.6 per cent;  $P = 0.00011$ ). However, similar findings were noted in the placebo group, i.e., 15.1 per cent mortality for good adherers and 28.3 per cent for poor adherers ( $P = 4.7 \times 10^{-16}$ ). These findings and various other analyses of mortality in the clofibrate and placebo groups of the project show the serious difficulty, if not impossibility, of evaluating treatment efficacy in subgroups determined by patient responses (e.g., adherence or cholesterol change) to the treatment protocol after randomization.

*N Engl J Med. 1980; 303:1038–41*

## El efecto Hawthorne:

1924-1932, Hawthorne Works, fábrica de componentes eléctricos

- Varios experimentos para mejorar la productividad
- En todos, un observador se sentaba y miraba los cambios: hasta 23 cambios en la iluminación, el ambiente, la comida...
- En todos los casos se incrementó la producción

El efecto consiste en que cuando alguien se sabe observado, modifica su conducta, a favor del efecto que se espera





Research article

Open Access

## The Hawthorne Effect: a randomised, controlled trial

Rob McCarney<sup>1</sup>, James Warner\*<sup>1</sup>, Steve Iliffe<sup>2</sup>, Robbert van Haselen<sup>3</sup>,  
Mark Griffin<sup>2</sup> and Peter Fisher<sup>4</sup>

### Abstract

**Background:** The 'Hawthorne Effect' may be an important factor affecting the generalisability of clinical research to routine practice, but has been little studied. Hawthorne Effects have been reported in previous clinical trials in dementia but to our knowledge, no attempt has been made to quantify them. Our aim was to compare minimal follow-up to intensive follow-up in participants in a placebo controlled trial of Ginkgo biloba for treating mild-moderate dementia.

**Methods:** Participants in a dementia trial were randomised to intensive follow-up (with comprehensive assessment visits at baseline and two, four and six months post randomisation) or minimal follow-up (with an abbreviated assessment at baseline and a full assessment at six months). Our primary outcomes were cognitive functioning (ADAS-Cog) and participant and carer-rated quality of life (QOL-AD).

**Results:** We recruited 176 participants, mainly through general practices. The main analysis was based on Intention to treat (ITT), with available data. In the ANCOVA model with baseline score as a co-variate, follow-up group had a significant effect on outcome at six months on the ADAS-Cog score ( $n = 140$ ; mean difference =  $-2.018$ ; 95%CI  $-3.914, -0.121$ ;  $p = 0.037$  favouring the intensive follow-up group), and on participant-rated quality of life score ( $n = 142$ ; mean difference =  $-1.382$ ; 95%CI  $-2.642, -0.122$ ;  $p = 0.032$  favouring minimal follow-up group). There was no significant difference on carer quality of life.

**Conclusion:** We found that more intensive follow-up of individuals in a placebo-controlled clinical trial of Ginkgo biloba for treating mild-moderate dementia resulted in a better outcome than minimal follow-up, as measured by their cognitive functioning.

**Trial registration:** Current controlled trials: ISRCTN45577048

BMC Medical Research Methodology 2007, 7:30

## El efecto *Observer-expectancy*:

El investigador produce una influencia inconsciente en los participantes de un experimento

## El efecto Jonh Henry:

En un experimento controlado, si el control sabe de su estatus como miembro del grupo control y puede compararse con los del grupo de tratamiento, mejora su actitud para evitar esa “desventaja”

## El efecto Pigmalion:

A mayor expectación que muestra el observador, mejores resultados

Son una amenaza para la validez interna de un estudio

Se evitan con un diseño experimental doble ciego



La necesidad del uso de la mayor parte de las herramientas mencionadas viene determinada por la poca capacidad humana para evaluar los resultados de las intervenciones de forma objetiva.