

# Metodología de investigación y lectura crítica de estudios

## Experimentos clínicos aleatorizados, o cómo evaluar las intervenciones

Álvaro J. Ruiz Morales<sup>1</sup>  
Lina María Vargas Abello<sup>2</sup>

### Resumen

*Introducción:* Para la evaluación de una intervención (medicamento, terapia, técnica quirúrgica) se utilizan los experimentos clínicos aleatorizados, estudios que controlan de manera automática muchas fuentes de sesgos. *Desarrollo:* Deben tener siempre un grupo control; la asignación a los grupos de tratamiento debe hacerse mediante el azar, e idealmente, para la evaluación, debe cegarse al sujeto de estudio, al evaluador y a quien hace el análisis. Tanto la intervención experimental como la de control deben haber sido seleccionados con bases sólidas, y debe haber una incertidumbre genuina en relación con su efectividad y seguridad. Debe procurarse que las variables de medición del desenlace sean sólidas y objetivas (duras) o variables blandas que hayan sido obtenidas o transformadas mediante métodos ya definidos. Si no es posible cegar al sujeto o cegar la intervención, puede recurrirse a utilizar el método de doble simulación o a variables de desenlace duras.

**Palabras clave:** Experimento clínico aleatorizado, doble ciego, variables duras, variables blandas, asignación, aleatorización.

**Title:** Randomized Clinical Trials in the Evaluation of Interventions

### Abstract

*Introduction:* Randomized Clinical Trials are used to evaluate interventions (drugs, therapies or surgical techniques), since they control for many possible sources of bias. *Development:* Clinical Trials must have a control group, assignment to treatment groups must be at random and it is ideal to blind the study subject, the researcher and the analyst for the evaluation. Both the experimental and the control interventions must be selected with solid bases, and there should be a genuine uncertainty about their effectiveness and safety. Outcome variables should be hard and objective, and if soft variables are used, they must be transformed using predefined methods to make them as objective as possible. If it is not possible to mask the subject or the intervention, double dummy methods should be employed or hard outcomes used.

**Key words:** Randomized clinical trial, double-blind, hard outcomes, soft outcomes, assignment, randomization.

- .....
- <sup>1</sup> Médico. MSc, FACP. Profesor titular, Departamento de Epidemiología Clínica y Bioestadística y Departamento de Medicina Interna. Facultad de Medicina, Pontificia Universidad Javeriana.
  - <sup>2</sup> Médica. Asistente de Investigación, Departamento de Epidemiología Clínica y Bioestadística, Facultad de Medicina, Pontificia Universidad Javeriana, Bogotá.

## **Introducción**

La aparición de un nuevo anti-depresivo exige la evaluación tanto de su utilidad clínica como de su seguridad. El diseño de investigación que se use para evaluarlo debe ser estricto para que ofrezca información confiable, que encuentre de manera válida los aspectos positivos y negativos, y que además sea práctico para que permita su uso por parte del clínico.

El uso de un medicamento en una población, sin un grupo control, puede ofrecer información que no necesariamente sea aplicable a otras poblaciones, porque puede tratarse de un grupo muy seleccionado de pacientes, que no representen al común de aquellos con depresión, o bien el investigador puede, en su interés por el estudio, influir en los resultados.

El experimento clínico aleatorizado es un diseño específico para intervenciones, que tiene como ventaja ofrecer controles automáticos para muchos de los posibles errores que pueden ocurrir en la investigación con pacientes.

## **Definición**

### *Estudios experimentales*

Un estudio es experimental cuando el investigador decide cuál es la intervención que va a ser evaluada. Si se quiere saber los efectos de consumir alcohol, ha sido el paciente quien ha tomado la decisión de consumirlo; si se busca

el impacto de fumar cigarrillos, ha sido el sujeto quien ha decidido fumar. Por el contrario, en un experimento, el investigador es quien decide cuáles son las intervenciones que evaluará y que serán asignadas posteriormente a los participantes.

El experimento –parecido a los realizados en física y química (de allí su nombre)– tiene implícito el concepto de investigar experimentando.

Las primeras etapas de investigación de un medicamento involucran su uso en animales de experimentación, bajo una rígida reglamentación ética. Pero para que los hallazgos tengan utilidad clínica, las intervenciones deben haberse probado en seres humanos, también con normas éticas que buscan proteger al individuo, asegurarse de que los beneficios son claramente superiores a los riesgos y que éstos están claramente justificados.

A pesar de la reglamentación, de la preparación y de la conducción cuidadosas, es imposible eliminar los riesgos, por lo que debe haber claridad en que la experimentación en seres humanos no está exenta de riesgos, que tienen su justificación cuando se piensa que puede haber beneficio potencial para los participantes en el estudio y para la producción de conocimiento generalizable.

### *Necesidad de grupo control*

La medición y la evaluación implican, siempre, una comparación

con un patrón. En las intervenciones, es útil saber cuánto mejor (o peor) es el medicamento nuevo cuando se lo compara con la intervención comúnmente utilizada o idealmente con la mejor disponible. Hay varias razones por las que el uso de un grupo control es indispensable en los experimentos, como se ve en la Tabla No. 1.

Tabla No.1. Razones para la necesidad de grupo control en los experimentos

- |   |
|---|
| <ul style="list-style-type: none"><li>- Efecto placebo</li><li>- Regresión a la media</li><li>- Efecto Hawthorne (o halo)</li></ul> |
|---|

### *Efecto placebo*

Se describió el *efecto placebo* como el efecto producido por la administración, por cualquier vía, de una sustancia inerte o sin propiedades farmacológicas, y cuyo efecto aparece por sugestión secundaria a la administración. Dado que por este efecto los sujetos pueden tener cambios en las variables evaluadas, la presencia de un grupo control, que también tendrá el mismo efecto, permite que se anulen mutuamente, puesto que estará presente, por igual, en los dos grupos.

### *Regresión a la media*

Algunas características de los seres vivos (frecuencia cardíaca, por ejemplo) cambian (hacia los extremos) cuando se las evalúa por primera vez (aparece taquicardia o bradicardia), y las siguientes mediciones se van devolviendo progresi-

vamente al valor verdadero. Por este efecto, llamado *regresión a la media*, puede haber cambios artificiales, que dependen de regresión, al valor real, de la desviación inicial.

### *Efecto Hawthorne*

Se describe como los cambios en conducta y comportamiento ocurridos por efecto de que el sujeto sea consciente de que está siendo observado.

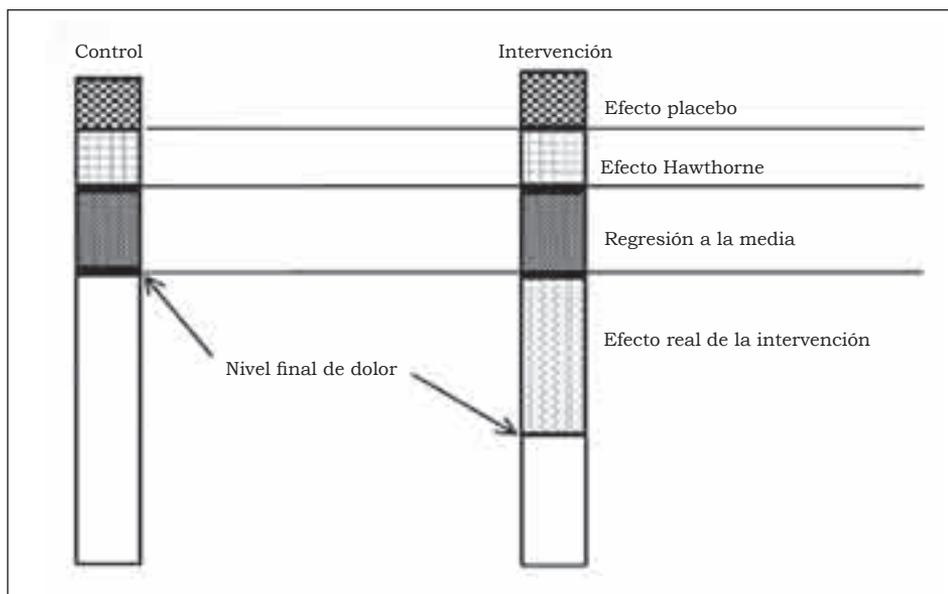
La suma de los tres efectos anteriores puede producir resultados que induzcan, equivocadamente, a concluir efectividad de un tratamiento. Para evitar este error, el experimento exige la presencia de un grupo control. Al comparar dos grupos, uno con la intervención que se experimenta y otro con la intervención base (o control), se pueden anular mutuamente estos tres efectos, puesto que cada uno de ellos ocurrirá, por igual, en los dos grupos comparados, como puede verse en la Figura No.1.

### **Selección de la población**

La *población de estudio*, aquella en la que se realiza el experimento, proviene de una población mayor (llamada *población blanco*), a la que representa y a la que luego se querrá extrapolar los resultados.

Para la selección de la población se utilizan los criterios de selección, que definen quiénes podrán participar en el estudio. Estos criterios son:

Figura No. 1. Impacto en una escala visual análoga, para medir dolor, de los posibles efectos



### Criterios de inclusión

Permiten la caracterización del paciente prototipo que se quiere incluir. Como ejemplo, en el estudio de un nuevo antihipertensivo, se incluirá a: 1. Pacientes hipertensos; 2. mayores de 40 años; 3. con hipertensión arterial primaria.

### Criterios de exclusión

No siempre el paciente que cumple los criterios de inclusión quiere o puede participar. Puede ser que no esté interesado, o que no acepte firmar consentimiento, o que tenga alguna contraindicación para uno de los medicamentos, o que tenga alguna enfermedad que dificulte la evaluación posterior. En estos casos, aunque el paciente cumpla los

criterios de inclusión, se le excluye del proceso de selección.

Cabe anotar que los criterios de exclusión no son los opuestos a los de inclusión.

### Utilidad de los criterios de selección

Estos criterios son útiles para definir, por una parte, la población que ingresa al estudio. Por otra parte, son el fundamento para decidir a quiénes se pueden extrapolar los resultados del experimento. Si un paciente hubiera podido ingresar al estudio (porque cumple los criterios de inclusión y no tiene ninguno de los de exclusión), se le pueden extrapolar los resultados. (ver Tabla No. 2).

Tabla No. 2. Usos de los criterios de selección

1. Definir y caracterizar la *población de estudio*
2. Definir y caracterizar la *población a la cual se puede extrapolar*

Tabla No. 3. Ventajas de la aleatorización

1. Se evitan sesgos en la asignación
2. Se logran grupos que son comparables en los aspectos conocidos
3. Se logran grupos comparables también en aspectos no conocidos

## Aleatorización

En la asignación a los grupos de tratamiento pueden introducirse sesgos si no se hace mediante un método independiente del investigador, ya que este podría tender a asignar a los pacientes más graves a uno de los tratamientos, o a asignar a los potencialmente mejores respondedores al tratamiento que más apropiado le parezca.

Por esta razón, la asignación debe ser hecha por el azar, a lo que se denomina aleatorización. Este proceso tiene varias ventajas, que pueden verse en la Tabla No.3.

### Sesgos en la asignación

Un sesgo (desviación sistemática) puede ocurrir cuando los pacientes con más severidad en su enfermedad son asignados a un tratamiento, o cuando un grupo tiene los pacientes mayores, o cuando los mejores respondedores están en mayoría en alguno de los grupos de comparación. La aleatorización hace que no haya ninguna posible influencia de parte de los investigadores en la asignación.

## Comparabilidad de los grupos

Si bien la aleatorización no puede garantizar el perfecto equilibrio entre los grupos, casi siempre se logra que los grupos sean comparables tanto en las variables conocidas (peso, edad, sexo, frecuencia de diabetes, etc.) como en las variables que no se midan (proteína C reactiva, homocisteína, cromosoma Philadelphia).

### Evaluación de resultados

En la evaluación de los desenlaces estudiados, el paciente, el investigador y quien haga el análisis pueden influir en los resultados por sus expectativas, sus prejuicios o por deseo de mejoría.

Para evitar esta influencia, se recomienda cegar (enmascarar) al sujeto de investigación, de manera que conozca que se están probando dos intervenciones, sepa cuáles son, pero no pueda saber cuál específicamente está recibiendo. A esto se le conoce como *estudio ciego*. También puede cegarse al investigador o evaluador, que solamente recolecta la información sin conocer cuál medicación está recibiendo el sujeto. En este caso, se habla de estudio *doble-cego*.

## **Selección de las intervenciones**

Esta es la característica que define al experimento clínico, puesto que es decisión del investigador cuáles intervenciones se utilizarán en el estudio, y es muy frecuente que el paciente no hubiera estado expuesto al menos a la intervención experimental de no ser por su participación en el estudio.

### *Intervención experimental*

Debe tener alguna base fisiopatológica, lógica biológica o antecedentes en estudios en animales, y debe existir, por parte de los investigadores, una incertidumbre genuina, es decir seguridad de que no hay evidencia suficiente para establecer, antes del estudio, si es benéfica o nociva.

Esta incertidumbre es la justificación ética para la realización del estudio, dado que sin saber sobre su efectividad o seguridad, el uso de un medicamento tiene serias limitaciones.

### *Intervención de control*

Debe cumplir requisitos básicos, como ser la mejor intervención que exista, que a la vez esté disponible y que sea aplicable. El medicamento no sólo debe ser la intervención óptima en el mundo científico (mejor intervención), sino que debe existir como opción terapéutica en el ámbito de realización del estudio (disponible). No sería

ético hacer un experimento clínico con un medicamento, aunque sea el mejor existente, que no estará disponible en el país luego del estudio. Y finalmente el medicamento debe poder ser aplicable, en términos prácticos, logísticos y económicos.

Si no se cumplen estas condiciones, podría ocurrir que un estudio concluya que el medicamento nuevo, A, es mejor que el medicamento B, que, sin embargo, no es el que se utiliza generalmente, o no es el mejor disponible. La interpretación de estos resultados y su aplicabilidad se dificultan enormemente.

### *Utilización de placebo*

En los casos en los que *no haya una intervención con efectividad demostrada* debe utilizarse, como comparador, una sustancia inerte, que permita mantener el ciego. Por consideraciones éticas, el uso de placebo cuando existe una intervención de utilidad conocida está proscrito, puesto que se priva al grupo control del beneficio de un medicamento reconocido.

Hay un uso apropiado del placebo aún si hay un tratamiento reconocido. Si se va a evaluar un medicamento que se use como adición a una terapia de base, todos los sujetos del estudio deben recibir la terapia basal (por ejemplo un antiepiléptico), uno de los grupos recibe la adición del nuevo y el otro grupo recibe placebo.

### *Estudios doble-simulados*

Si se quiere comparar dos intervenciones que tengan formas de administración diferentes, por ejemplo un medicamento inyectado y uno oral, se puede recurrir a una doble simulación: un grupo recibe el medicamento activo oral y adicionalmente una inyección de placebo; y al otro grupo se le administra por vía oral un placebo y el medicamento activo inyectado. De esta manera, todos los pacientes del estudio reciben una intervención oral y una inyectada, y se puede mantener el placebo.

### **Evaluación**

Las variables con las que se evalúan la efectividad y seguridad de la intervención pueden clasificarse, de acuerdo con la objetividad e instrumentos para su recolección, en tres grupos:

#### *Variables de medición primariamente objetiva (variables duras)*

Estas variables pueden ser medidas con instrumentos precisos, sus resultados no dependen de circunstancias externas y son repetibles. También se las conoce en la literatura como *variables duras*. Como ejemplos, el colesterol sérico, la creatinina en orina, la presión arterial de oxígeno.

#### *Variables de medición primariamente subjetiva (variables blandas)*

Se trata de aspectos cuya medición es más compleja, puesto que no hay instrumentos definidos o criterios inequívocos para su evaluación o interpretación, como en el caso de la disnea, edema de miembros inferiores, hallazgos radiológicos o dolor torácico. También son variables en las que la influencia externa o de las circunstancias puede ser significativa (como ejemplo, la percepción de dolor postcesárea puede ser diferente según las condiciones del producto del embarazo).

Habitualmente estas variables, que son más complejas de evaluar, se miden a través de cuestionarios o encuestas, que deben cumplir condiciones de validación para la población en la que se aplica.

#### *Variables de medición aparentemente objetivas (variables semiduras)*

En algunos casos, la variable parece ser dura, por tratarse de mediciones numéricas, como en el caso de días de estancia hospitalaria, o días de intubación o de ventilación mecánica. En estos casos debe tenerse precaución especial, puesto que la presentación de datos de un estudio con promedios de días de estancia en cuidado intensivo puede parecer muy objetiva, y sin embargo en la vida real la decisión de trasladar a un paciente fuera de cuidado intensivo depende de mu-

Figura No. 2. Escala visual análoga para evaluación de dolor



chos factores independientes de la condición médica, como aspectos personales del paciente o su familia, de la aseguradora o pagador, del hospital o de las necesidades de cuidado intensivo.

### **Importancia del tipo de variables según la medición**

Si bien es deseable que las variables utilizadas en una investigación sean duras, esto en la vida real no es posible, ya que en la medicina clínica se utilizan principalmente variables blandas, como dolor, calidad de vida, disnea, fuerza muscular, calidad de los reflejos osteotendinosos, tos, somnolencia diurna, soplos cardíacos o reflujo.

En muchos casos se pueden hacer transformaciones para que la medición de una variable blanda sea más confiable, como en el uso de una escala visual análoga para evaluación de dolor. No sólo puede hacerse un poco más objetiva la medición, sino que sirve para hacer análisis de la evolución del dolor en el mismo sujeto (Figura No. 2).

Con frecuencia en medicina, particularmente en las áreas

quirúrgicas, no se puede cegar la intervención o es imposible enmascarar al evaluador. En estos casos es aún más importante que se busquen variables duras como desenlace, ya que la combinación de un estudio abierto (no cegado) con variables blandas reúne dos fuentes potenciales enormes de sesgos en la evaluación.

En una futura entrega se discutirán aspectos de tamaño de muestra, análisis estadístico y evaluación crítica de un experimento, así como de aspectos éticos.

### **Referencias**

1. Shein-Chung Chow, Jen-Pei Liu. Design and Analysis of Clinical Trials: Concepts and Methodologies (Wiley Series in Probability and Statistics). 2a Edición. Wiley-Interscience, New York, 2003.
2. Steven Piantadosi. Clinical Trials: A Methodologic Perspective Second Edition (Wiley Series in Probability and Statistics). Wiley-Interscience, New York, 2005.
3. Robert H Fletcher, Suzanne W Fletcher. Clinical Epidemiology: The Essentials. Lippincott Williams & Wilkins, Philadelphia, 2005.
4. R Brian Haynes, David L Sackett, Gordon H Guyatt, Peter Tugwell. Clinical Epidemiology: How to To Clinical Practice Research. Lippincott Williams & Wilkins, Philadelphia, 2005.

5. Alvaro Ruiz Morales, Luis Morillo Zárate. *Epidemiología Clínica: Investigación Clínica Aplicada*. Editorial Médica Panamericana. Bogotá, 2004.
6. Lawrence M Friedman, Curt D Furberg, David L De Mets. *Fundamentals of Clinical Trials*. 3a Edición. Springer, New York 1998.

*Recibido para evaluación:* 15 de enero de 2007  
*Aprobado para publicación:* 23 de febrero de 2007

Correspondencia  
Álvaro Ruiz Morales  
*Departamento de Epidemiología Clínica y Bioestadística*  
*Facultad de Medicina, Pontificia Universidad Javeriana*  
*Bogotá, Colombia*  
*Correo electrónico:* [aruiz@javeriana.edu.co](mailto:aruiz@javeriana.edu.co)