



UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH



Tema 10:

# Ensayo clínico

## Azar, riesgos de sesgo, ética

Erik Cobo y Joan Carles Oliva

Jordi Cortés, José Antonio González y Marta Vilaró

Enero 2014

Departament d'Estadística  
i Investigació Operativa

UNIVERSITAT POLITÈCNICA DE CATALUNYA



MEDICINA  
CLINICA



## Ensayo clínico

Presentación .....	3
1. Características fundamentales .....	4
1.1. El ensayo clínico es experimental .....	4
1.2. El ensayo clínico es ético .....	5
1.3. El ensayo es clínico .....	5
1.4. El ensayo clínico es prospectivo .....	6
1.5. El ensayo clínico es comparativo .....	6
1.6. El ensayo clínico estima efectos .....	8
1.7. El ensayo clínico puede apoyar decisiones .....	9
1.8. El ensayo clínico es controlado .....	10
2. El azar .....	10
2.1. Obtención al azar: representatividad .....	11
2.2. Asignación al azar (aleatorización): comparabilidad .....	12
3. Riesgos de sesgo .....	14
3.1. Sesgo de selección .....	14
3.2. Sesgo de ejecución .....	15
3.3. Sesgo por atrición .....	15
3.4. Sesgo del informe selectivo .....	15
3.5. Sesgo de evaluación .....	16
3.6. Sesgo impredecible .....	16
4. Ajuste: el control metodológico .....	17
4.1. Un buen control metodológico evita los riesgos de sesgo .....	19
4.2. Un buen control aumenta la precisión de la estimación .....	20
4.3. Métodos de control o ajuste .....	22

4.3.1. Restricción .....	23
4.3.2. Estudio de subgrupos .....	24
4.3.3. Modelado .....	26
4.3.4. Control global (minimizació).....	28
4.4. Enmascaramiento .....	29
5. Ética, azar y papel del científico .....	32
6. Regresión a la media .....	35
7. Tipos de ensayos clínicos.....	39
7.1. Según el objetivo.....	39
7.2. Según lo avanzado de la propuesta de intervención.....	40
7.3. Según el diseño .....	40
7.4. Según el método de asignación.....	41
Soluciones a los ejercicios .....	42

## Presentación

Estamos ya en condiciones de estudiar la metodología del ensayo clínico para poder estimar el efecto de una intervención, cómo sortear sus riesgos de sesgo y ejecutarlo respetando los derechos humanos.

**Contribuciones:** Basados en la versión previa de EC, JAG y PM, lo re-elaboraron EC y JCO; y lo revisaron JC, JAG, y MV.

# 1. Características fundamentales

## 1.1. El ensayo clínico es experimental

Un ensayo clínico es un experimento que asigna seres humanos a intervenciones en comparación.

La característica fundamental del ensayo clínico es que el investigador determina el grado de exposición a la causa en estudio: el investigador decide el valor de la causa en estudio.



### Recuerde

Un **EC** asigna intervenciones a los participantes.

**Nota técnica:** Es equivalente asignar “intervenciones a participantes” que asignar “participantes a intervenciones”.

**Ejemplo 1.1:** Asignamos una de las intervenciones en estudio a cada paciente que acepta participar.

**Ejemplo 1.2:** Repartimos (asignamos) los trabajadores de cierta entidad entre los protocolos alternativos de seguimiento y control sanitario que deseamos comparar.



### Recuerde

Distinga entre intervenciones (**X**) asignables por el investigador y condiciones (**Z**) con las que se presentan las unidades.

**Historieta:** El pasado, Z, me esclaviza; el futuro, X, me libera.



### Ejercicio 1.1

Un estudio pretende evaluar el efecto que, sobre la variable respuesta Tensión Arterial (TA), tiene el número de visitas de seguimiento. Suponga:

- Un estudio experimental que asigna a cada paciente un determinado número de visitas, y
- Un estudio observacional que simplemente observa el número de visitas espontáneamente solicitadas por el paciente.

1) ¿El número de visitas es **X** o **Z** en cada estudio?

2) Ambos estudios demuestran una relación significativa entre número de visitas y TA: 5 mmHg menos por cada visita adicional (IC<sub>95%</sub>: 3 a 7 mmHg). Interprete este mismo resultado en ambos estudios.



### Recuerde

Una variable susceptible de ser intervenida, debe ser “asignable”.

## 1.2. El ensayo clínico es ético

Nos hemos dotado de derechos y principios éticos que nos protegen. Los Comités de Estudios de Investigación Clínica (CEICs) vigilan que se respeten.

**Ejemplo 1.3:** Sólo causas con efectos potencialmente positivos, los “tratamientos”, son éticamente asignables a seres humanos (principio de no maleficencia).

**Nota:** El investigador que recluta voluntarios puede decir, por ejemplo: “Mire, Vd. va a recibir todo el conjunto de intervenciones efectivas que gracias a la generosidad de pacientes con su enfermedad, la Ciencia ha seleccionado previamente. Ahora estamos estudiando, en condiciones muy controladas y reguladas, el balance entre beneficios y riesgos de una nueva y prometedora posibilidad. Si Vd. se ofrece voluntario, tendrá una posibilidad del 50% de acceder a ella. Para alcanzar el máximo rigor científico, ni Vd. ni yo sabremos durante el estudio si la recibe o no, lo que nos obliga a utilizar un simulador del tratamiento. Por supuesto, tanto si recibe Vd. esta nueva opción, o el simulador al que llamamos placebo, en ambos casos estará tratado dentro del máximo nivel técnico y científico”.

Por ello, previamente al inicio del estudio, un comité de ética debe velar por la veracidad de esta afirmación.

**Nota:** El apartado de este capítulo profundiza en la ética del Ensayo Clínico.

## 1.3. El ensayo es clínico

Lo más frecuente es evaluar la eficacia y la seguridad de un tratamiento con un fármaco. Pero “intervención” puede indicar cualquier maniobra o producto, químico o biológico, administrados con intención preventiva o curativa.



### Ejercicio 1.2

Poner un par de ejemplos de EC: 1) en el que no intervengan fármacos; y 2) cuya intención sea preventiva.

#### 1.4. El ensayo clínico es prospectivo

Su objetivo es responder la pregunta de un clínico, o un paciente, sobre cómo una intervención cambiará su futuro.

**Ejemplo 1.4:** Si me tomo una aspirina, ¿se me irá el dolor de cabeza?

#### 1.5. El ensayo clínico es comparativo

**Historieta:** Le preguntan a una estadística, “¿es guapa tu pareja?”; y responde, “¿comparada con la de quién?”

Fould (1985) [ [Pocock](#), p 53] revisó 52 ensayos no comparativos en psiquiatría y encontró que 44 (85%) concluyeron que existía éxito terapéutico; mientras que de 20 ensayos con un grupo de referencia, sólo 5 (25%) encontraron dicho éxito.



### Ejercicio 1.3

¿Qué puede explicar estas discrepancias?

**Ejemplo 1.5:** La artritis reumatoide es una enfermedad crónica que tiende a empeorar con el paso del tiempo. Si administramos una nueva terapia y hacemos una comparación antes-después (pre-post), el efecto medido podría ser incluso negativo: al final del ensayo, los pacientes pueden estar incluso peor que al inicio. Sin embargo, el tratamiento en investigación podría ser positivo, pues tiene el mérito de enlentecer la progresión de la enfermedad. Este beneficio sólo se puede poner de manifiesto con un comparador.



**Figura 1.1** Hay más casualidades que causalidades

**Nota:** Para poder afirmar que la clase le cambió, es preciso asumir que si no hubiera ido a clase no lo sabría. Como este hecho no ha sido observado, no hay evidencia de que la relación observada (“antes=no sabe, luego=sí”) sea causal.



#### Ejercicio 1.4

Una comparación antes-después muestra que la presión sistólica bajó 5mmHg en los pacientes asignados a la intervención A. Comente algunas explicaciones alternativas al descenso de la presión sistólica.

#### Ejercicio 1.5

Proponga intervenciones de control o alternativas a tomar una aspirina en un vaso de agua.



#### Recuerde

La pregunta sobre el futuro lleva implícita una intervención de referencia.

**Historieta:** Tip y Coll están en el safari fotográfico de su vida. Bajan a recepción vestidos con botas y cargados con cámaras y objetivos. Pero les avisan de que un león anda suelto y que ellos deben asumir los riesgos, por lo que les piden que firmen un acuerdo informado. Tip sale corriendo y vuelve con bambas, un ligero atuendo deportivo y una mini-cámara. Coll le dice riendo, “serás tonto, ¿crees que así correrás tú más que el león?”; y Tip le responde, “no amigo mío, por favor no te confundas, yo sólo quiero correr más que tú”.



### Recuerde

Debe precisar tanto la intervención en estudio como la alternativa a la que pretende sustituir.

**Contra-ejemplo 1.6:** En los niveles iniciales de investigación, un estudio de factibilidad podría centrarse en un solo grupo, que recibe la [intervención](#) en estudio.

## 1.6. El ensayo clínico estima efectos



### Definición

**Efecto** es la diferencia entre la respuesta Y que se observa cuando se asigna A y la que se observa cuando se asigna B.

Estimar efectos precisa inferencia causal. Los estudios observacionales sin comparador requieren el razonamiento [contrafáctico](#) (de no haberse observado la causa A, el efecto Y no se habría producido); pero en los estudios experimentales, las [respuestas](#) a las causas en comparación son potencialmente observables y comparables.

**Ejemplo 1.7:** “Among persons at high cardiovascular risk, a [Mediterranean diet](#) supplemented with extra-virgin olive oil or nuts reduced the incidence of major cardiovascular events”. Como esta aparición de eventos cardiovasculares es observable en todos los pacientes reclutados, no hay necesidad de argumento contrafáctico.

**Nota técnica:** Veremos más adelante que cada paciente será asignado a una sola intervención y, por tanto su respuesta bajo la otra intervención no se observará, será un dato ausente (“missing”). Pero como la asignación al grupo de tratamiento es al azar, se trata de “missings at random”, cuya ausencia no sesga los resultados.

Observe que hemos definido “efecto de A relativo a B en la respuesta Y”. Además, convendrá especificar la población de pacientes P a la que aplican los resultados.

**Ejemplo 1.7 (cont.):** Efecto de la dieta mediterránea (A) relativo al consejo para reducir grasas (B) en los eventos cardiovasculares mayores (respuesta Y) en pacientes con diabetes tipo II o, al menos, 3 riesgos cardiovasculares mayores (población P).



### Recuerde

Comparar las respuestas entre 2 grupos aleatorizados permite estimar el efecto.



### Ejercicio 1.6

Durante un año, en 2580 pacientes, eliminaron el cabello de la zona de la incisión quirúrgica mediante un proceso nuevo. Luego, . compararon la tasa de infección de la cicatriz con 17424 pacientes vistos en los 7 años anteriores. Como no hubo cambio significativo en la tasa de infección, que se mantuvo en torno al 1% ([Surg Gynecol Obstet 1986; 162:181](#)), los autores dijeron “dado que no hemos visto diferencias, el uso de controles históricos no compromete las conclusiones”. Busque diferencias entre los dos grupos en comparación que puedan comprometer esta interpretación.

### Ejercicio 1.7

¿Qué problemas pueden haber con el uso de *controles de la bibliografía*? ¿Y con los *controles del mismo hospital*?



### Recuerde

La máxima calidad científica requiere un comparador interno y simultáneo.

## 1.7. El ensayo clínico puede apoyar decisiones

El futuro de una intervención puede pivotar sobre ECs: sólo si el resultado es positivo, la intervención pasa a formar parte del arsenal terapéutico. En este caso, los riesgos de decisiones erróneas deben estar controlados. El entorno de decisión de Neyman-Pearson, con riesgos alfa y beta acotados, es el preferido por las agencias de regulación de medicamentos.

Ello requiere que el criterio para tomar decisiones esté perfectamente definido. En el caso más general, descansa en **1** análisis principal de **1** variable principal de evolución.

**Nota:** Por supuesto, el EC puede desear aportar conocimiento sobre otros temas, usualmente recogidos en objetivos y variables secundarios. Además, otros objetivos secundarios pueden comprobar la robustez de los resultados (¿se habría llegado a la misma decisión con otro análisis o variable?) o la validez de las premisas (¿permite el análisis de subgrupos sostener que el mismo efecto aplica a todos los casos?).

### 1.8. El ensayo clínico es controlado

**Nota:** En un ensayo clínico el término ‘control’ puede tener dos connotaciones: (1) el grupo sometido a la intervención de referencia, comparativa, como en ‘control histórico’, ‘control interno’, etc.; y (2) el hecho de controlar las restantes variables. Para distinguir, usamos comparador o referencia para el primer uso.

El ensayo clínico debe garantizar que la única diferencia entre los grupos en comparación son los tratamientos. En caso contrario, la comparación de los grupos estaría afectada, sesgada, por estas otras diferencias.

**Historieta:** Un niño juguetón y algo sádico quita una pata a una araña, y la suelta mientras le dice “anda”. Observa que la araña se escapa corriendo. Lo repite hasta que la araña se queda sin patas y no se marcha, por lo que concluye “cuando le quitas todas las patas se vuelve sorda”.



#### Recuerde

Sólo podemos “concluir” que el tratamiento es eficaz, si éste es la única diferencia entre los grupos. Si hay más diferencias, requiere “interpretar” y “discutir”.

El ensayo clínico “controla” que no aparezcan riesgos de sesgo durante la selección, seguimiento y evaluación de los participantes. Antes de estudiar cómo lo hace y cuáles son los riesgos de sesgo, veamos las contribuciones del azar.

## 2. El azar

La inferencia estadística puede recurrir al azar de dos formas: para seleccionar a las unidades o para asignar los tratamientos. Como se muestra a continuación, el primer caso persigue representatividad; y el segundo, comparabilidad.

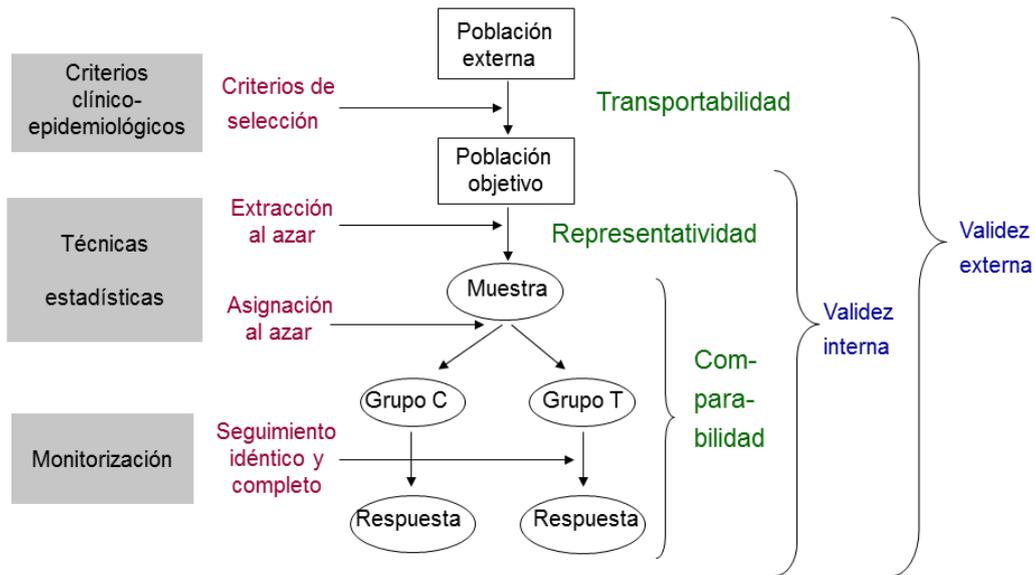


Figura 2.1 Dos aportaciones del azar a la validez: representatividad y comparabilidad.



### Ejercicio 2.1

De acuerdo con la Figura 2.1, diga si las siguientes afirmaciones son ciertas.

- a) Puede haber validez interna sin que exista validez externa.
- b) Puede haber validez externa sin que exista validez interna.
- c) Para que los grupos sean comparables debe haber extracción al azar.
- d) Un seguimiento idéntico y completo garantiza comparabilidad.
- e) El azar dota de transportabilidad.

### 2.1. Obtención al azar: representatividad

Ya se dijo que, si los elementos de la muestra eran seleccionados al azar, la teoría de probabilidad permite cuantificar (error estándar, intervalos de confianza) la incertidumbre al inferir al conjunto de toda la población los resultados obtenidos en la muestra.

**Ejemplo 2.1:** Al realizar la planificación sanitaria de una comunidad se recurre a una muestra aleatoria para conocer el estado de salud de sus ciudadanos.



**Recuerde**

La teoría de probabilidad permite cuantificar las oscilaciones de los resultados debidas al azar del muestro.



**Ejercicio 2.2**

Para realizar el estudio del Ejemplo 2.1, a partir de un listado de todos los elementos de la población, se seleccionan al azar una serie de individuos en los que se evalúa su nivel de salud. Dado que es un derecho no participar, ¿qué consecuencias puede tener que algunos casos rehúsen?



**Recuerde**

La obtención al azar otorga representatividad a los resultados.

La selección al azar de los participantes requiere disponer de (un listado de) todos los miembros elegibles de la población, lo que puede ser factible en sociología, política o salud pública, pero no en clínica. Por eso, los ensayos clínicos NO se basan en la selección o extracción al azar.

**2.2. Asignación al azar (aleatorización): comparabilidad**

Los ensayos clínicos SÍ se basan en la asignación al azar.

Una vez se han seleccionado aquellos individuos que formarán parte del estudio, el proceso de asignación aleatoria del tratamiento genera dos muestras de esta misma población. Dado que ambas muestras provienen, de manera aleatoria, de la misma población, la distribución de todas las variables es la misma a nivel poblacional.

Como todas las terceras variables están igualmente distribuidas en los grupos, no existe ninguna que pueda explicar diferencias entre los grupos, por lo que la asignación aleatoria del tratamiento ofrece una estimación insesgada del efecto (causal).



**Recuerde**

Asignar al azar otorga comparabilidad de los grupos y permite una estimación insesgada del efecto.



### Ejercicio 2.3

Sobre la asignación al azar, diga cuáles de las siguientes afirmaciones son correctas:

- a) Las dos muestras provienen de la misma población.
- b) A nivel poblacional, las dos muestras son idénticas.
- c) Los valores observados en las dos muestras son idénticos.



**Figura 2.2** Papel del azar. Un profesional que desee ‘controlar’ un elevado número de variables con pequeños efectos en la respuesta puede recurrir al azar para conseguir un reparto similar del conjunto de todas ellas, de la misma forma que al freír unos champiñones introducimos entropía, azar, ruido, para que se hagan homogéneamente. ¡Y funciona!

**Nota técnica:** Al ser muestras aleatorias, aplica todo lo estudiado en inferencia. Si, por ejemplo, la variable respuesta es numérica y estamos interesados en comparar medias, la diferencia de las medias muestrales ( $\bar{y}_1 - \bar{y}_2$ ): (1) tiene como valor esperado la diferencia poblacional entre medias [  $E(\bar{y}_1 - \bar{y}_2) = \mu_1 - \mu_2$  ], es decir, no hay sesgo porque el conjunto de todos los estudios posibles apunta en la dirección correcta; y (2) su imprecisión es la suma de las oscilaciones de las dos medias: [  $V(\bar{y}_1 - \bar{y}_2) = \sigma_1^2/n_1 + \sigma_2^2/n_2$  ], lo que proporciona su error estándar.

Por tanto, el resultado de un estudio aleatorizado concreto (1) apunta en la dirección correcta; y (2) permite cuantificar su desvío esperado del auténtico valor.



### Recuerde

Un estudio aleatorizado (bien diseñado y ejecutado) carece de sesgo.

### Recuerde

El azar proporciona la base para calcular la incertidumbre originada por estudiar un subconjunto de casos.



### Ejercicio 2.4

Sobre la estimación del efecto, si hay asignación al azar, son correctas:

- a) Diferentes asignaciones (o reparto de los casos) darán lugar a diferentes estimaciones.
- b) La oscilación de las posibles estimaciones es desconocida.
- c) La distancia entre el verdadero efecto y la estimación en una asignación concreta es estimable.
- d) La distancia entre el verdadero efecto y el promedio de las estimaciones de todas las posibles asignaciones es conocida.

## 3. Riesgos de sesgo

Para tener una estimación insesgada del efecto, además de asignar al azar, hay que hacer bien todo el estudio. Repasemos, cronológicamente, algunas [oportunidades](#) que tenemos para hacerlo mal.

### 3.1. Sesgo de selección

Hay sesgo de [selección](#) cuando las intervenciones se asignan a grupos que proceden de poblaciones diferentes.

**Ejemplo 3.1:** Asignar los pacientes más graves a la intervención nueva.

**Nota:** Un clínico que desea actuar de esta manera no cree en el estudio y no debería participar en el mismo.

Se previene con la asignación al azar y oculta: la aleatorización garantiza que los grupos en comparación proceden de la misma población; y mantenerla oculta hasta el momento de la intervención minimiza las posibilidades de excluir pacientes de forma direccionada.

**Contra-ejemplo 3.2:** Un investigador que excluye los pacientes más graves del estudio para tratarlos, si estuviera permitido, fuera del mismo, reduce la extensión de la población estudiada y por tanto atenta a la validez externa (transportabilidad), pero no a la interna (comparabilidad).

### 3.2. Sesgo de ejecución

Hay sesgo de [ejecución](#) (“performance”: actuación, rendimiento, acción, o desempeño) cuando los grupos en comparación reciben diferentes cuidados (aparte de los que se comparan).

**Ejemplo 3.3:** En un estudio que sólo enmascara la evaluación de la respuesta final, dar más medicación de rescate a los pacientes del grupo de referencia.

Se previene con el enmascaramiento.

### 3.3. Sesgo por atrición

Hay sesgo por [atrición](#) o desgaste cuando se pierden casos. La aleatorización garantiza grupos comparables al inicio del estudio. Para mantener esta propiedad es necesario mantener hasta el final del estudio a todos los casos.

**Contra-ejemplo 3.4:** Un paciente o un clínico pueden decidir interrumpir la intervención en estudio. Es su derecho. Pero este hecho es, en sí mismo, un resultado muy importante que debe ser reportado, por lo que debe mantenerse al paciente en el estudio.



#### Recuerde

No confunda abandonar el tratamiento en estudio [con](#) abandonar el estudio.

Se evita con un seguimiento completo.

### 3.4. Sesgo del informe selectivo

**Ejemplo 3.5:** Escoger aquella variable de la evolución o aquel análisis estadístico que más favorecen las [expectativas](#) o los resultados deseados.

La obligación de publicar los protocolos pretende evitar el cambio de variable y análisis principal. Aun así, conviene publicar un plan de análisis estadístico [especificado](#) sin ambigüedades.

**Nota:** Una práctica pasada consistía en desvelar el grupo (A o B) de cada paciente pero no el tratamiento de cada grupo (experimental o referencia) y escribir 2 informes (uno asumiendo que el experimental era A; y otro, B) y, una vez firmados ambos, desvelar cuál era el grupo experimental. Pero los eventos adversos permitían desvelar el tratamiento de cada grupo, Por ello, se abandonó este método.

### 3.5. Sesgo de evaluación

#### Ejemplo 3.6: Ser más “generoso” al valorar la respuesta de los tratados.

El enmascaramiento previene el riesgo de sesgo durante la evaluación.

Los diseños PROBE, enmascaran la evaluación de la respuesta para evitar este riesgo de sesgo.

### 3.6. Sesgo impredecible

Los sesgos comentados hasta el momento explican estimaciones optimistas de los efectos de las intervenciones. Como la estadística acompaña las medidas del efecto con estimaciones de su incertidumbre, también debemos estudiar qué pasa con las estimaciones de los errores típicos.

El sesgo impredecible o imprevisible aparece cuando la imprecisión del estudio es mayor que la cuantificada por el error estándar.



#### Recuerde

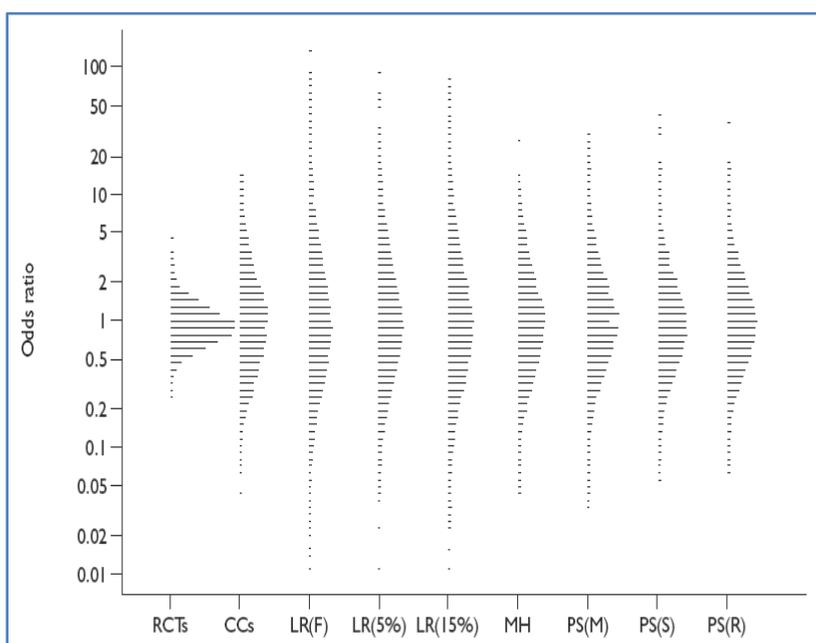
Si hay sesgo impredecible, la estimación obtenida del error estándar es optimista.

Ello que implica que el auténtico valor estará más allá de los límites del IC más veces de las  $\alpha$  aceptadas: es decir, la cobertura prometida, por ejemplo, del 95%, es mayor que la real.



#### Recuerde

Si hay sesgo impredecible, la cobertura del IC es menor de la especificada.



**Figura 3.1** Comparación de métodos de ajuste de casos mixtos aplicado a los resultados de estudios controlados de 14 regiones en el IST. RC: resultados de los ECAS correspondientes; CCs: comparadores concurrentes sin ajuste, LR(F): ajuste con regresión logística completa, LR(5%): ajuste con regresión logística por pasos  $p_r = 0,05$ ; LR(15%): ajuste de la regresión logística por pasos  $p_r = 0,15$ ; MH: ajuste por estratificación de Mantel-Haenszel; PS(M): ajuste por emparejamiento del “propensity score”; PS(S): ajuste por estratificación del “propensity score”, PS(R): ajuste por regresión del “propensity score”.

La Figura 3.1 tomada de [Deeks et al](#), muestra que el uso de técnicas de ajuste no soluciona este problema.

**Recuerde**

La asignación al azar proporciona estimaciones no sesgadas de las medidas del efecto y de su incertidumbre.

Todo desvío, incorrección o error en un EC puede contribuir a [aumentar](#) la incertidumbre, resultando en IC excesivamente optimistas ya que sólo recogen la oscilación aleatoria pura.

**Ejercicio 3.1**

¿Cuáles de los 6 sesgos comentados pueden aparecer en un estudio no enmascarado? Ponga ejemplos de qué podría pasar. ¿Y en un estudio PROBE?

**Ejercicio 3.2**

¿Por qué conviene NO excluir del ensayo un paciente que debe abandonar el tratamiento (p.e., por falta o de eficacia o de tolerabilidad)?

**Ejercicio 3.3**

¿Cómo podemos comprobar que unos autores no presentan, de todos los posibles análisis, aquellos que les convienen?

**Ejercicio 3.4**

Intente explicar con sus propias palabras qué es el sesgo impredecible.

**Ejercicio 3.5**

Ponga un ejemplo hipotético de cómo se puede producir cada tipo de sesgo.

## 4. Ajuste: el control metodológico

**Nota:** El término control tiene muchas acepciones. En un EC puede referirse: (1) al grupo o (2) a la intervención de referencia; (3) al riguroso seguimiento de los pacientes por los investigadores; (4) a la regulación legal; (5) a la garantía de calidad que proporcionan la monitorización y la gestión de datos; y (6) a la metodología que permite mejorar las propiedades estadísticas.

En un ensayo con 2 casos y 2 intervenciones, cualquier diferencia entre los casos puede ser una explicación alternativa a las diferencias observadas. Y por tanto, ese estudio tiene un valor muy limitado.

**Ejemplo 4.1:** Supongamos una población con sólo 2 casos (Mar y Pep), que asignamos al azar a las 2 intervenciones en comparación (A y B). Hay 2 asignaciones posibles: (1) Mar con A y Pep con B; y (2) Mar con B y Pep con A. El promedio de los efectos estimados en los 2 posibles estudios con estos 2 casos coincide con el efecto poblacional, pero cada uno de los 2 estudios diferirá del auténtico valor por 2 motivos, uno aleatorio y otro sistemático. El primero es el resultado de la variabilidad entre casos y podemos cuantificarlo con el error típico. El segundo, en cambio, incluye todas las diferencias “reproducibles”: por ejemplo, que las mujeres vivan unos 5 años más que los hombres. Un observador crítico encontrará muchas otras diferencias entre Mar y Pep a las que podría atribuir las diferencias observadas.

Pero el conjunto de todos los estudios posibles apunta en la dirección correcta: es insesgado.



**Recuerde**

El azar garantiza que el conjunto de la ciencia avanza en la dirección correcta.

Pero un buen investigador no tiene suficiente con el hecho de que la ciencia avance. Él quiere que su estudio concreto sea convincente. Quiere que su estudio acierte en sus estimaciones.



**Recuerde**

Aunque la asignación aleatoria permite estimaciones insesgadas a lo largo de todos los experimentos, no garantiza que cada posible experimento esté equilibrado para toda tercera variable.



**Figura 4.1** Límites del azar. Si una variable importante debe quedar equilibrada entre los grupos, el azar no lo garantiza, especialmente si la muestra es pequeña. El azar que le sirve al cocinero para ‘equilibrar’ los champiñones, no funciona con los bistecs. Puede argumentarse que, si se dispusiera de muchos bistecs, la mitad de ellos se harían por su cara proximal y la otra mitad por la distal, llegando a un equilibrio ‘a largo plazo’. Ciertamente, habría ‘equilibrio’ para el conjunto de los comensales. Pero cada uno de ellos se comería un bistec que tendría una cara cruda y la otra hecha. De la misma manera, si un investigador quiere garantizar que, en su estudio concreto una variable importante quede equilibrada, no debe descansar en el azar.

#### 4.1. Un buen control metodológico evita los riesgos de sesgo

Delante los posibles riesgos de sesgos, ¿qué puede hacerse? Muy sencillo: no dejar en manos del azar aquellas terceras variables que, si resultaran desequilibradas, podrían comprometer las conclusiones del estudio. Nótese que las variables no observables (pongamos  $V$ ) nunca comprometerán la credibilidad del estudio (porque no puede observarse su desequilibrio). Pero entre las variables observables ( $Z$ ), aquellas relacionadas con la respuesta, pueden comprometer la credibilidad de los resultados y, por tanto, no conviene dejarlas en manos del azar.

**Ejemplo 4.2:** Si el nivel inicial de la enfermedad es una variable crucial, que determina claramente la evolución de los pacientes, conviene controlarla.

**Contra-Ejemplo 4.3:** La situación personal, familiar y profesional es un conjunto de muchas variables que podrían influir, como mucho, un poco cada una de ellas, en la presión arterial. Ajustar por cada una de ellas haría inviable el estudio.



**Figura 4.2.** Límites del control. Pretender controlar cada pequeña variable haría inviable el estudio. Si se definen bloques de edad, género, color de los ojos y preferencias sexuales dentro de cada centro, siempre faltarán casos en alguno de los bloques y nunca se podrá cerrar el estudio.



### Recuerde

Las terceras variables observables se controlan con las técnicas de diseño y de ajuste; y las no observables, con la asignación al azar.



### Ejercicio 4.1

De las siguientes expresiones, remarque aquellas que son ciertas.

- Si realizo al azar la asignación del tratamiento, las dos muestras, que vienen de la misma población, tienen una idéntica distribución poblacional.
- Si realizo al azar la asignación del tratamiento, las dos muestras, que vienen de la misma población, toman idénticos valores en ambas muestras.
- Si realizo al azar la asignación del tratamiento, la estimación es insesgada.
- Si realizo al azar la asignación del tratamiento, no debe preocuparme ningún posible desequilibrio entre las muestras, ya que la estimación es insesgada.
- Si mediante las técnicas de ajuste controlo a todas las terceras variables observables, la asignación al azar ya no aporta nada al diseño.
- Si se acepta que la ciencia empieza por descubrir las variables importantes, las técnicas de control (que controlan a las variables conocidas  $Z$ ) tienen un papel más relevante que la asignación al azar (que controlan a las desconocidas  $V$ ).

## 4.2. Un buen control aumenta la precisión de la estimación

Un control o un ajuste bien planificado y ejecutado suele conllevar un beneficio en la precisión de la estimación.

**Ejemplo 4.4:** Cabe esperar menos diferencias entre dos pacientes de la misma edad, género y centro, que entre dos pacientes de diferentes edades, géneros y centros.

La comparación directa de dos pacientes de las mismas características, al estar menos influida por fluctuaciones aleatorias, aportará más información.

**Ejemplo 4.5:** Vimos 2 pruebas para comparar 2 medias, para datos independientes y apareados. En ambas, el numerador proporciona la misma señal. Pero la imprecisión calculada en el denominador era diferente: cuanto mayor era la correlación entre las dos observaciones, mayor era el beneficio por aparear los casos.

**Nota técnica:** Vimos que una gran correlación entre las dos determinaciones corresponde a una situación con mucha variabilidad entre-casos y poca variabilidad intra-casos.



### Recuerde

Existe mayor eficiencia en la prueba de comparación de medias con datos apareados que independientes.

### Recuerde

El ajuste permite controlar la variabilidad de la variable  $Z$  que se deja fija. Eso implica, en un buen diseño, mayor eficiencia estadística, por lo que un mismo número de casos permite un menor error de estimación y un intervalo de confianza más estrecho.

Además, controlar mediante el diseño permite equilibrar el número de casos en cada grupo.

**Ejemplo 4.6:** Un estudio de factibilidad desea asignar 12 pacientes, 6 hombres y 6 mujeres, a dos intervenciones. Si se desea comparar medias, se sabe que los hombres y las mujeres difieren en la variable respuesta, y se puede asumir que las varianzas son iguales, la máxima eficiencia se obtiene asignando 3 de cada género a cada intervención.

Un ajuste no planificado en el protocolo puede tener consecuencias imprevisibles sobre la eficiencia. Puede conllevar un beneficio o, por el contrario, aumentar considerablemente el error típico de estimación de los parámetros.

**Ejemplo 4.7:** El mismo estudio de factibilidad anterior decide controlar por género en el análisis en lugar de en el diseño. Podría haber sucedido que la intervención A la formaran 5 hombres y 1 mujer. Y al revés para la B. Las comparaciones A frente a B, serían 5 a 1 para los hombres y 1 a 5 para las mujeres, con mayor imprecisión que si hubieran sido 3 a 3.



### Ejercicio 4.2

Siguiendo con el ejemplo anterior, suponga que: desea conocer el efecto sólo en los hombres; dispone de 8 casos; y la varianza de la variable cuantitativa respuesta es 1. Calcule el error estándar de la estimación del efecto de A frente a B en los hombres si los 8 observados han quedado repartidos entre las intervenciones de las 2 formas siguientes: (i) 4 a 4; y (ii) 7 a 1. ¿Cuánto mayor es el error en el 2º caso?

**Nota técnica:** Estos beneficios son muy claros si se analiza una respuesta numérica (p.e., comparación de medias). Veremos que la situación es más sofisticada en otros casos (p.e., comparación de proporciones, análisis de supervivencia, etc.).

### 4.3. Métodos de control o ajuste

Hay 4 grandes grupos de métodos para el control metodológico: restricción, análisis de subgrupos, modelado; y ajuste global. Todos ellos pueden decidirse en la fase de diseño o en la fase de análisis, pero sólo un ajuste decidido independientemente de los resultados está protegido del riesgo de sesgo del informe selectivo. Por ello, es importante especificar el método de ajuste durante el diseño del estudio.



#### Recuerde

Especificar el control en el diseño garantiza que los resultados no guían el método (sesgo del informe selectivo).

La Tabla 4.1 resume las ventajas e inconvenientes de las diferentes posibilidades para el control metodológico y los nombres que reciben si se decide en el momento del diseño o durante el análisis.

**Nota técnica:** En el entorno del EC, el azar proporciona la propiedad de no sesgo y el control aumenta la precisión (eficiencia). En los estudios observacionales el ajuste tiene otra interpretación, ya que también persigue “controlar” el sesgo.

El curso de observacionales abordará más afondo estos métodos. Veámoslos brevemente.

Opción	Fase	Nombre	Ventajas	Inconvenientes
Restricción	Diseño	Criterios de elección	Control completo Barato	Reduce la población objetivo Número de variables limitado
	Análisis	Análisis de un subgrupo	Simple de diseñar Simple de analizar	Posible confusión residual (si las restricciones son amplias)
Subgrupos	Diseño	Bloques (apareamiento)	Potencia Eficiencia	Logística más sofisticada Coste
	Análisis	Estratificación (apareamiento)	Sin premisas Directa Cálculo simple	Dispersión de casos en estratos Diferentes estratificaciones Difícil ‘síntesis’
Modelado Estadístico	Diseño	Modelado	Factible con pocos casos Redondea efectos menores	Muchas premisas Elección del modelo
	Análisis	Covarianza, regresión, otros	Permite predicciones Permite variables continuas Permite varias variables	Elección de variables Interpretación Parametrización del <i>software</i>
Ajuste global	Diseño	Minimización	Permite considerar varias Z	Logística sofisticada
	Análisis	Pareja óptima	No reduce población objetivo	

**Tabla 4.1** Opciones para el control [Adaptada de [Kleinbaum et al](#)]



**Figura 4.3** Control por variables conocidas. Un profesional que desee tener ‘controlada’ una variable conocida y medida sin error, utilizará un sistema que le garantice que queda igualmente distribuida en los grupos en comparación, de la misma forma que al freír un bistec nos aseguramos que cada lado se fría justo en la cantidad que deseamos.

### 4.3.1. Restricción

La forma más sencilla de control es la **restricción** de la población en estudio, bien sea mediante la definición de criterios de selección en el protocolo, bien sea mediante el estudio de un subgrupo en el momento del análisis.

**Ejemplo 4.8:** Se limita el estudio a casos de una cierta edad y de un hospital concreto. Ahora, estas dos variables ya no pueden ser una explicación alternativa a la relación observada entre el tratamiento X y la respuesta Y.

Su gran ventaja es la sencillez de aplicación. Su inconveniente es la reducción de la población objetivo, tanto para obtener casos para el estudio como para aplicar después sus resultados.

**Ejemplo 4.9:** En el ejemplo anterior, si un profesional de otro centro desea utilizar los resultados de ese estudio deberá analizar con detalle si existen características diferenciales entre ambos centros que puedan comprometer aplicar los efectos observados en el otro centro.



#### Ejercicio 4.3

En un estudio experimental muy controlado, se ha logrado establecer el efecto de un tratamiento que disminuye el área necrosada tras un infarto. Los criterios de inclusión de dicho estudio exigían que el paciente fuera tratado en un plazo de tiempo muy limitado. Por esta razón, el protocolo del centro en el que Ud. trabaja no incluye este tratamiento para aquellos casos que superan este límite de tiempo. Ud. desea tratar este paciente. ¿Qué argumentos deben utilizarse?



### Recuerde

Los criterios de selección son la aplicación de la restricción decidida en la fase de diseño del estudio.

**Lectura:** La declaración [CONSORT](#) recomienda no distinguir entre criterios de inclusión y exclusión y llamarlos criterios de selección o elegibilidad. Pero otras, como [STARD](#), siguen distinguiendo.

Si una variable  $Z$  dispone de un amplio rango de variación y los criterios de elegibilidad son generosos, puede persistir una cierta confusión residual.

**Ejemplo 4.10:** Si la edad tiene un efecto claro en la respuesta, que hace muy diferente un caso con 18 años de uno con 70, limitar el estudio a pacientes de 18 a 70 puede ser correcto, pero no suficiente.

### 4.3.2. Estudio de subgrupos

La siguiente herramienta de control es la definición de subgrupos, que reciben el nombre de **bloques** si se especifican en el diseño y de **estratos** si se definen en el análisis.

**Ejemplo 4.11:** En un ensayo clínico multicéntrico se especifica en el protocolo que se desea tener controlada la variable centro y la asignación aleatoria se planifica de forma que en cada centro exista el mismo número de casos en cada grupo de tratamiento. Se dice que cada centro es un bloque y que el estudio está balanceado por centro.

**Ejemplo 4.12:** En un estudio sobre el efecto de los programas de atención a domicilio en la supervivencia de las personas mayores que se quedan solas, se decide estratificar en el análisis por el nivel de actividades iniciales que realiza cada persona.

Una única estimación del efecto que represente al efecto común de cada subgrupo será siempre más fácil de aplicar, pero requiere asumir que el efecto es el mismo en todos los subgrupos.



### Recuerde

Sea parsimonioso al exponer sus resultados. Si es razonable, proporcione una única medida del efecto que sirva para todos los subgrupos: bloques o estratos.

**Lectura:** No hay nada más tedioso que un informe que va presentando resultados por subgrupos cuando el efecto es muy similar en todos ellos.

En el curso de observacionales veremos métodos para poner a prueba si existe un único efecto común a todos los subgrupos. Observe que si tiene sospechas de que el grupo puede modificar el efecto, quizás sea más práctico restringir el estudio a un subgrupo concreto. Y si no las tiene, el estudio común de varios subgrupos partirá de esta premisa de que el efecto es común a todos los subgrupos. Y las premisas no se someten a prueba, que se reserva para la hipótesis. Considerar a la homogeneidad del efecto como una premisa, o a su heterogeneidad como una hipótesis, podría explicar los matices diferenciales entre los consejos de [Trials](#), [Lancet](#) y [NEJM](#) sobre cómo interpretar el análisis de subgrupos.

**Ejemplo 4.13:** Si el estado pandrial puede modificar el efecto del tratamiento, no tiene sentido combinar los resultados obtenidos en pacientes en ayunas con los de pacientes en plena digestión. Es más simple y directo definir al estado pandrial como criterio de selección o como condición del estudio. Si conviene, más adelante ya se estudiará la posible interacción o se complementará el estudio realizado con otro de la condición no considerada.



#### Ejercicio 4.4

Imagine dos estudios multicéntricos, ambos de  $n=200$ , para comparar un tratamiento  $t$  con un comparador  $c$ . El estudio A ha sido diseñado para recoger información de cinco centros. Cada uno aporta 40 casos, 20 en cada brazo de tratamiento. El estudio B se inicia en veinte centros, de los que uno aporta 160 casos; y los otros 19 centros, los 40 restantes; destacando 10 centros que sólo aportan un caso cada uno. ¿Cuáles de las siguientes afirmaciones son ciertas?

- El estudio A seguro que evita la posible confusión por la variable centro, ya que está equilibrado o balanceado.
- La razón entre el número de casos tratados con  $t$  y con  $c$  es constante para todos los centros del estudio A. Esto implica que la variable centro y la variable tratamiento son independientes por lo que sus efectos no se pueden confundir.
- En el estudio B, hay 10 centros que no pueden aportar, por sí solos, información del efecto del tratamiento y por tanto no podrán constituir, cada uno de ellos, un bloque separado.

- d) El estudio A por su equilibrio en número de casos parece que será más eficiente, proporcionando intervalos de confianza más estrechos.
- e) El estudio B al abarcar más centros, es más extrapolable a otros centros nuevos.
- f) El estudio B no es un estudio multicéntrico sólido.

El apareamiento es un caso extremo en el que cada bloque o estrato está compuesto por dos observaciones que comparten cierta característica común que, por otro lado, les diferencia de las otras parejas. El grado de conexión entre estas parejas, naturales o artificiales, puede variar.

**Ejemplo 4.14:** Los diseños con intercambio del tratamiento (“*cross-over*”) consisten en aplicar un tratamiento en un periodo y otro tratamiento en otro periodo, cambiando el orden de aplicación en diferentes pacientes. Cada paciente define una pareja de observaciones y constituye un (mini) bloque.

**Ejemplo 4.15:** Olvide ahora los ensayos clínicos que estiman efectos. En un estudio etiológico de búsqueda de causas X, a cada paciente que ha desarrollado cierta enfermedad (Y+) se le busca una pareja sana (Y-) que tenga los mismos valores en las terceras variables Z. Cada pareja constituye un (mini) grupo.

### 4.3.3. Modelado

La siguiente opción de control es el **modelado** estadístico, del que las diferentes técnicas de regresión son el procedimiento más habitual. El protocolo puede especificar en mayor o menor grado el modelo que se utilizará para realizar el ajuste.

**Ejemplo 4.16:** Al comparar 2 antihipertensivos (X), puede especificar que el análisis de la respuesta (por ejemplo, Y=PAD a las 12 semanas) se ajustará mediante un modelo de regresión lineal que incluye el valor inicial o basal (Z) de la PAD. Nótese que la relación entre PA basal (Z) y final (Y) no es objetivo del estudio (un investigador experto en el tema conoce la magnitud de esta relación con anterioridad). El auténtico objetivo es conocer el efecto del tratamiento X en la respuesta Y ajustando por la condición Z. [Los modelos de [regresión](#) se estudian en el curso ‘observacional’.]

Cuanto más concreto sea el protocolo sobre el modelo que se utilizará para realizar el ajuste o control metodológico, mayor será el carácter confirmatorio del estudio.

Un **ejemplo** (4.17) muy sencillo de modelo que se especifica completamente en el protocolo es el llamado estudio del cambio o diferencia entre valores finales y basales —o iniciales.

**Nota técnica:** Algunos paquetes estadísticos dan una falsa sensación de facilidad en el uso de estos modelos. La selección de variables a ser incluidas no es un tema trivial, especialmente en presencia de colinealidades o correlaciones entre las variables a ser introducidas. Dejar el control para el final es muy desaconsejable, ya que el margen de maniobra es siempre inferior.



### Recuerde

Considere al inicio del estudio el método para controlar las terceras variables.



### Ejercicio 4.5

En el estudio del efecto de un tratamiento X en la PAD de las 12 semanas, se ha realizado un modelo de regresión que incluía la PAD en la semana inicial. Los estudios que se comentan a continuación han llegado al modelo por diferentes caminos. Ordénelos de mayor a menor protección ante el riesgo de sesgo del informe selectivo (“*data driven analysis*”). Y remarque cuáles representan un salto importante en credibilidad.

- Tal y como especificaba el protocolo, la respuesta estudiada ha sido el cambio o diferencia entre la PAD final y la inicial.
- Con la  $PAD_{12}$  como respuesta, se ha introducido la  $PAD_0$  en el modelo con un coeficiente que ya venía especificado en el protocolo.
- Con la  $PAD_{12}$  como respuesta y tal como especificaba el protocolo, se ha introducido la  $PAD_0$  en el modelo y se ha dejado al programa que escoja el coeficiente por el procedimiento habitual (que minimiza el error de predicción de la respuesta).
- Como a) pero sin que figurara en el protocolo.
- Como b) pero sin que figurara en el protocolo.
- Como c) pero sin que figurara en el protocolo.

g) Tal como decía el protocolo, entre todos los posibles modelos, se ha seleccionado aquél en el que más significativa era la variable tratamiento.

#### 4.3.4. Control global (minimización)

Finalmente, el cuarto método de control considera simultáneamente varias variables en el ajuste. En un diseño experimental con aleatorización, persigue mejorar el equilibrio proporcionado por la asignación al azar. Requiere una logística más sofisticada ya que para aleatorizar, el investigador debe proporcionar los valores de las variables por las que se quiere ajustar. En cambio, como garantiza que se registra al paciente antes de asignarlo, protege del sesgo de selección.

**Ejemplo 4.18:** En el estudio de la PAD se quiere tener controladas edad, centro, estrés y sedentarismo. En el momento de introducir un nuevo paciente en el estudio, el investigador proporciona los datos de estas variables al centro de aleatorización que, sin romper el enmascaramiento del investigador, asigna el nuevo caso al grupo que consigue un mayor equilibrio en estas variables.

La minimización permite tener equilibrados los grupos sin necesidad de recurrir a criterios de inclusión más restrictivos, ni a la definición de muchos subgrupos, lo que podría comprometer el reclutamiento de casos.

**Nota técnica:** Los algoritmos habituales de minimización suelen ponderar por separado a las variables consideradas en el proceso y persiguen equilibrio en cada una de ellas.

La minimización puede combinarse también con la asignación aleatoria y el enmascaramiento de investigadores y pacientes.

**Ejemplo 4.19:** Un investigador interesado en ajustar por edad, centro, estrés y sedentarismo (estas dos últimas medidas en escalas de 0 a 10) ha determinado en estudios previos que la relación de estas 4 variables con la PAD es la siguiente: cada decena de años supone 5 mmHg más, los hombres tienen 5 mmHg más; cada punto de sedentarismo son 2 mmHg más y cada punto de estrés son 3 mmHg más. Así, un índice global combina estos valores en cada paciente y el próximo se asigna al grupo que hace más similares en ese centro las medias de este índice en los dos grupos. Es óptimo, no para cada variable  $Z$  por separado sino para el conjunto de todas ellas. Es decir, pretende compensar los desequilibrios de unas variables con los de otras.

También puede usarse en los estudios observacionales.

**Ejemplo 4.20:** En la misma búsqueda anterior de causas X, y de la misma forma que antes, a cada paciente que ha desarrollado cierta enfermedad (Y+) se le busca una pareja sana (Y-). La diferencia ahora es que la pareja debe tener un valor próximo en los valores de las terceras variables Z, sea en cada una de ellas, como en el Ejemplo 4.18 o en su combinación en un índice como en el Ejemplo 4.19.



### Recuerde

Sólo el control bien planificado puede garantizar dos beneficios: ausencia de sesgo y mayor eficiencia de las estimaciones.

## 4.4. Enmascaramiento

**Historia (real):** Una comisión de la academia francesa de ciencias, encabezada por Benjamin Franklin, hizo el primer (documentado) estudio enmascarado que desmontó creencias sobre el mesmerismo y el magnetismo animal.

Toda imprecisión en la recogida de datos implica pérdida de información y, en consecuencia, limita la eficiencia estadística, aumentando el error de las estimaciones. Pero además, si el error que se comete está ligado a ambas variables en estudio (causa X y respuesta Y), la estimación de la relación entre ambas, estará sesgada.

**Ejemplo 4.21:** Conocer el tratamiento recibido puede originar puntuaciones sistemáticamente diferentes de la respuesta.

**Historieta:** Podría ser que la capacidad de recordar los hábitos sexuales pasados dependa de tener el sida. Un paciente con varias parejas previas puede no acordarse de ningún partenaire si no ha desarrollado el sida; pero, en caso de contagio, se *acordará* muy bien de cada uno de ellos.

En un ensayo clínico, los datos de todas las variables se recogen en el mismo momento de aparición y su riesgo de sesgo es menor. Pero ciertas decisiones en el seguimiento o ciertos criterios de evaluación podrían originar sesgos. Para minimizar estos riesgos de sesgos, los ensayos clínicos recurren al **enmascaramiento** del tratamiento, de forma que se pueda administrar, seguir y evaluar **a ciegas**, sin que sea posible identificar los tratamientos en comparación. Para permitirlo, los estudios de intervenciones farmacológicas (químicas o biológicas) administran un **placebo** que emula al tratamiento en estudio. Si se trata de una intervención quirúrgica u otro tipo de maniobra, se procede a fingirla con un simulador (*'sham procedure'*).

**Nota:** Un buen simulador (placebo) debe ser idéntico en todo: apariencia, sabor, olor,...

Si se están comparando dos productos que se administran por vías diferentes puede recurrirse al doble placebo o **doble simulador** (*double dummy*) que consiste en administrar a cada grupo su tratamiento más el simulador del otro.

Conviene mantener el enmascaramiento hasta que finalice la limpieza de la base de datos, que se “blinda” para no permitir modificaciones posteriores.

**Nota:** Diga “a ciegas”, “enmascarado” o “cegado” y evite “ciego”. [“blind” ≠ “blinded”]

**Historieta:** ¿Apoyan los estudios con investigadores ciegos, evaluadores ciegos y pacientes ciegos la tesis de Ernesto Sábato en su “informe sobre ciegos”?

**Historia (real):** Los que describen más a sus pacientes como enmascarados y menos como ciegos son los [oftalmólogos](#).



#### Ejercicio 4.6

A continuación tiene una lista de personas que participan en un ensayo. Diga: (1) quiénes conviene que desconozcan el tratamiento (futuro, presente o pasado) del paciente; y (2) a quiénes de éstos puede aplicar el término ‘doble ciego’. Listado: (a) voluntario que recibe intervención; persona que (b) recluta pacientes; (c) evalúa criterios de selección; (d) obtiene consentimiento informado; (e) administra tratamiento farmacológico; (f) realiza intervención o maniobra (quirúrgica, fisioterapéutica, etc); (g) realiza seguimiento de los pacientes; (h) obtiene la variable de evolución; (i) introduce datos; (j) monitoriza datos; (k) diseña análisis estadístico; (l) ejecuta análisis estadístico; y (m) interpreta resultados.

**Lectura:** El ítem 11c de la extensión de [Consort para intervenciones no farmacológicas](#) dice: “*Whether or not those administering co-interventions were blinded to group assignment.*”



#### Recuerde

Especifique quiénes están enmascarados.

**Lectura:** [Isabelle Boutron et al.](#) recopilan métodos de enmascaramiento usados para intervenciones no farmacológicas.

**Recuerde**

Ocultar la intervención evita riesgos de sesgo. Siempre.

Poner una máscara al investigador no asegura que no pueda ver el tratamiento administrado.

**Ejemplo 4.22:** Al comparar un inhibidor de la multikinasa con placebo, la tabla de eventos adversos muestra 9 variables con diferencias significativas, como alopecia con un 27% en los tratados y un 3% de los controles.

**Historieta:** Si este investigador no veía que los pacientes tratados tenían rubor y se quedaban calvos es que estaba realmente ciego y no sólo enmascarado

**Ejercicio 4.7**

¿Cómo cree que debería haber sido la discusión de las limitaciones en el artículo del Ejemplo 4.22? ¿Lo hacen los autores? ¿Reciben alguna carta posterior?

**Ejercicio 4.8**

En las intervenciones que implican cambios de hábitos por el paciente, es imposible que éste desconozca el grupo de tratamiento al que pertenece. Si debe hacer ejercicio o dieta, tiene que saberlo. Y también el profesional que le asesora durante el periodo del estudio. Discuta quiénes conviene que permanezcan enmascarados y porqué.

**Ejercicio 4.9**

Imagine un estudio en el que un grupo recibe consejo dietético y el otro no, con evaluación a ciegas. ¿Cuáles de las siguientes expresiones son ciertas?

- El efecto directo de la intervención y los posibles efectos indirectos originados por conocer el paciente y la intervención que se le aplica están confundidos, en el sentido que no pueden descomponerse: el estudio proporciona una estimación de la suma de ambos efectos.
- Desde el punto de vista de adquisición de conocimiento, sería interesante poder distinguir entre los efectos directos e indirectos de la intervención.

- c) El sesgo originado por conocer el paciente su tratamiento elimina la utilidad posterior de los resultados.
- d) Dado que un paciente siempre sabrá si está haciendo ejercicio o no, preocuparse por este posible sesgo es irrelevante a efectos prácticos.



**Recuerde**

Valore si se desveló el tratamiento y discuta sus posibles implicaciones.

## 5. Ética, azar y papel del científico

[La Declaración de Helsinki](#) (DH) recoge los principios éticos que rigen toda experimentación con seres humanos para proteger sus derechos, seguridad y bienestar. Son:

*Confidencialidad:* exige la preservación del anonimato en la recogida, gestión y comunicación de la información.

*Autonomía:* requiere el respeto a la individualidad y deseos de cada uno: los participantes deben ser capaces de tomar una decisión informada sobre su inclusión en la investigación, lo que requiere el consentimiento informado. Las personas con autonomía disminuida tienen derecho a una protección adicional. El paciente debe ceder su derecho a decidir al proceso de aleatorización

*Beneficencia:* se apoya en el principio hipocrático de no maleficencia, es decir, de no causar daño: no asignar una intervención por debajo de las recogidas en las guías de práctica clínica.

*Equidad:* el peso de la investigación descansa por igual en todos los ciudadanos (justicia distributiva).

Nótese que aleatorización incluye dos significados: “asignación” y “azar”. Ambos con implicaciones éticas. (1) Ambas pautas de tratamiento deben ser éticamente asignables a los pacientes; y (2) los voluntarios deben renunciar mediante el consentimiento informado a su derecho a decidir (principio de autonomía).

**Ejemplo 5.1:** El famoso ensayo clínico sobre la estreptomicina, realizado en los años cuarenta por [Daniels y Hill](#), asignaba los pacientes al azar al grupo con antibiótico más reposo o bien al grupo tratado únicamente con reposo. Dos consideraciones apoyan a la asignación y a su aleatoriedad. La primera, común a casi todos los estudios, es que los (posibles) efectos positivos se compensan con los (posibles) efectos negativos y, en ausencia

de conocimiento sólido, ambos tratamientos son éticamente asignables (“*equipoise*”). La segunda consideración es un aspecto poco conocido de este ensayo: el British Council sólo disponía de suficiente estreptomina para tratar a unos 150 enfermos, ante 300 candidatos. Y propusieron asignarlo al azar para garantizar el principio ético de equidad.

**Recuerde**

Hay “*equipoise*” si no hay evidencia científica a favor de una intervención.

**Recuerde**

Asignar al azar apoya el principio de equidad.

**Recuerde**

El mayor compromiso ético viene impuesto porque ambas intervenciones deben ser asignables.

**Ejercicio 5.1**

¿Es ético el uso de placebo?

**Ejercicio 5.2**

En el ejemplo previo de la Estreptomina, si un paciente renuncia a participar en el estudio ¿qué principio de ética podemos comprometer si lo tratamos fuera del estudio?

**Recuerde**

Tratar fuera del estudio puede atentar al principio de equidad.

Si hay duda científica sobre una intervención, la actitud más ética es empezar lo antes posible un estudio bien diseñado (“asignar al azar el primer paciente”) que permita que el máximo número de pacientes se beneficien de sus resultados.

En ocasiones, el pronóstico es malo incluso con el ‘mejor cuidado de salud’. Una nueva intervención de la que no se conocen bien sus efectos (positivos y negativos), puede cumplir con *equipoise* clínica, pero abre una esperanza a los pacientes que quizás prefieran algo a lo que agarrarse. La duda es si queremos una sociedad en la que los pacientes tienen acceso a todas las

intervenciones imaginables o, por el contrario, sólo a aquellas que han demostrado un balance de efectos beneficioso.

[Rawls](#) sugiere considerar la situación de partida o posición original: antes de saber el resultado del reparto de papeles que en la vida nos toca desempeñar, es decir, independientemente de si somos el paciente actual, el paciente futuro o el investigador, ¿en qué sociedad preferimos vivir: una que haga ensayos clínicos aleatorizados y aprenda rápidamente o una que no los haga y no pueda seleccionar las intervenciones con perfil favorable de beneficios y riesgos?

La [FDA](#) ha apostado por aprender rápidamente y dificulta el acceso fuera de los ensayos clínicos.

[S. Senn](#) propone:

- 1) Hacer EC cuando el investigador crea mejor el nuevo Tratamiento
- 2) El nuevo tratamiento sólo es accesible en el EC.
- 3) El investigador debe creer que ningún paciente sale perjudicado por entrar en el EC.
- 4) La experimentación continúa hasta que: (a) el investigador deja de creer en el estudio en marcha; (b) el investigador convence a la sociedad de su eficacia.

Finalmente, conviene recordar que los pacientes han aportado sus propios datos para la investigación. Y, por tanto, la sociedad a la que representan debe conocer sus resultados y conclusiones. Sir Ian Chalmers, fundador de la Cochrane, resaltó en los [90](#) y [aún](#) insiste en que una comunicación deficiente (carencia de informes y publicaciones, difusión incompleta o sesgada,...) de los resultados no es ética.



### Ejercicio 5.3

¿En qué párrafos de la DH se habla de la obligación de publicar?

### Ejercicio 5.4

¿Qué dice SPIRIT sobre la política de publicación? ¿Qué poder da a los CEIC?



### Recuerde

Son obligaciones éticas de investigadores y estadísticos garantizar que:

- 1) El diseño permitirá responder la pregunta
- 2) Ningún paciente saldrá perjudicado por participar

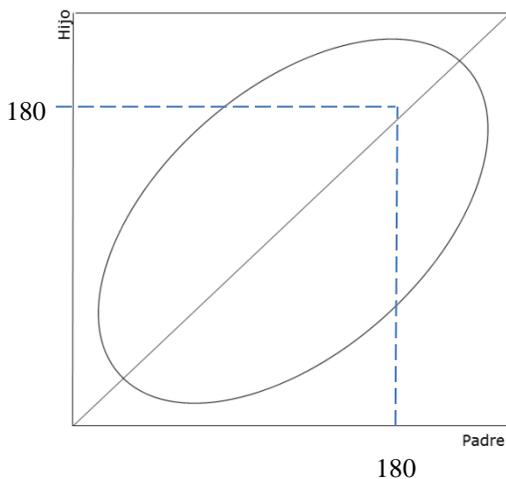
- 3) Se empieza lo antes posible
- 4) Los resultados se comunicarán de forma clara y rápida

## 6. Regresión a la media

Regresión hacia la media es el hecho de que los valores extremos observados una primera vez tienden a no ser tan extremos en una medición posterior.

**Historieta:** El curandero Asclepio conoce bien las dolencias crónicas de su conciudadano Alejandro. Sabe que son muy variables, con épocas buenas y malas que se alternan. También sabe que Alejandro acude a él cuando está en una de las épocas malas —a la que seguirá, algún día, una buena. Por eso, le receta algo inofensivo y le pide paciencia hasta que surja efecto. En la siguiente crisis, le dice: “bueno, esto ya hizo su efecto un tiempo, ahora deberemos cambiar a algo más fuerte”.

**Historieta:** La mitad de los internistas americanos [reconocen](#) una práctica similar

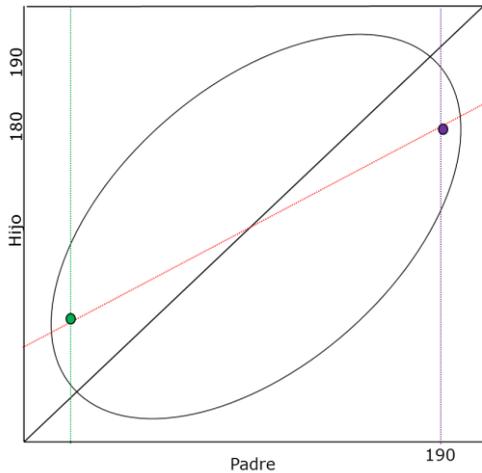


**Figura 6.1** Altura de padres e hijos: recta identidad. La nube representa todos los casos y la variabilidad: en muchas parejas de padres e hijos, su altura no coincide. En cambio, la diagonal representa la “recta identidad” en la que se sitúan aquellas parejas en las que si coincide la altura del hijo con la del padre.

[Galton](#), al estudiar la herencia, observó una relación no determinista entre la altura de progenitores y descendientes, como la que muestra la Figura 6.1, una relación positiva, pero muy alejada de la diagonal (recta identidad), que hubiera implicado que todos los hijos

tenían la misma altura que sus padres. Por debajo de la diagonal, se encuentran los casos con padres más altos; y, por encima, los de hijos más altos.

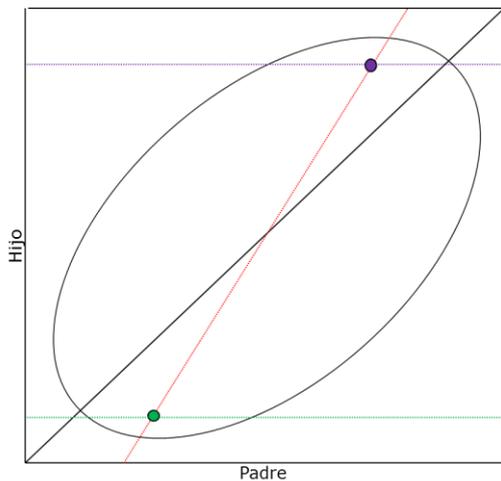
Si se concentraba en estudiar aquellas parejas en las que los padres eran extremadamente altos (como los seleccionados por la línea púrpura en la Figura 6.2), se encontró que los hijos, en promedio, eran más bajos: estaban más cerca de la media de la altura. Además, en el otro lado, también estaban más cerca de la media los hijos de padres extremadamente bajos (línea verde), que eran más altos que sus padres.



**Figura 6.2** Altura de padres e hijos: recta de regresión. El método de mínimos cuadrados (ver curso observacionales) traza la recta que pasa por el promedio de la altura de los hijos para cada valor de la altura de los padres. El punto lila marca el promedio de la altura de los hijos (180 cms) cuyos padres tienen una altura de 190 cm. Y el verde, que los hijos de padres bajos tienen mayor promedio. Como siempre las alturas promedio de los hijos están más cerca del centro que las de sus padres,

Galton habló de “regreso a la mediocridad”. A este fenómeno, absolutamente cierto, Galton lo denominó ‘regresión a la media’. Sin embargo, deducir que la dispersión disminuirá y,

en unas cuantas generaciones, todos nos pareceremos más, es falso: ya puede verse en las figuras anteriores, que la dispersión de la altura de los hijos es la misma que la de los padres, ya que hijos de padres no tan extremos ocupan ahora estos lugares.



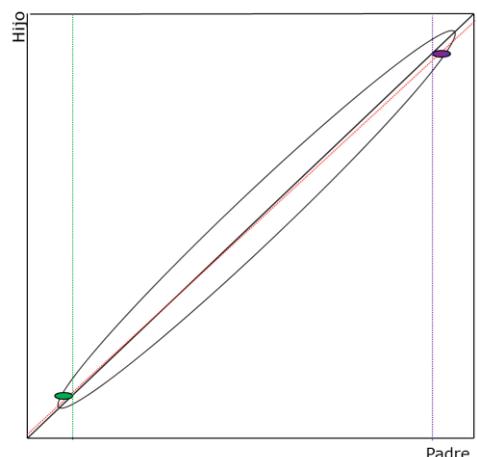
**Figura 6.3** Altura de padres e hijos: recta de regresión de padres según hijos. Para hijos de un altura dada, el promedio de las alturas de sus padres está más cerca de la media.

Más aún, la paradoja llega al extremo de que si invertimos los papeles de ambas alturas y estudiamos la altura promedio de los padres de hijos extremos observamos lo mismo fenómeno pero al revés: también sus padres tienen, en promedio, alturas más cerca de la media. La Figura 6.3 muestra que los padres de hijos muy altos (línea púrpura) tienen en promedio alturas más bajas; y los padres de hijos muy bajos tienen también alturas menos extremas.

de hijos muy altos (línea púrpura) tienen en promedio alturas más bajas; y los padres de hijos muy bajos tienen también alturas menos extremas.

**Figura 6.4** A menor variabilidad, menor efecto de regresión a la media

Este fenómeno es menos marcado cuando más se aproxima a 1 la correlación entre ambas variables (Figura 6.4). O, dicho al revés, para que aparezca la regresión a la media, es preciso un componente aleatorio, variabilidad, en la altura que no se repita en ambas generaciones (“independiente”).



**Nota técnica:** Nos hemos centrado en comentar la

variabilidad y cómo evolucionan los ‘casos’ extremos, pero no hemos mostrado ni comentado las medias de las alturas de ambas generaciones que pueden perfectamente ser idénticas (como los ejemplos de las figuras).

**Ejemplo 6.1:** Genotipo y fenotipo de la altura. Ciertas características se heredan, pero otras son el resultado de infinidad de factores como la alimentación, ejercicio y reposo de cada día durante el crecimiento.

**Ejemplo 6.2:** La repetición de un examen. Un buen examen y una buena corrección ordenan bien a los aspirantes y deja poco al azar. Cuanto más nos alejemos de este ideal, mayor será la influencia de la suerte en la puntuación final. Pero la suerte, el azar, no tienen memoria, no se repiten. Si un ejercicio de evaluación tiene mucho azar (poco repetible o fiable, en términos técnicos), menos se repetirá el resultado previo. En este caso, aquél alumno que tuvo la buena suerte de quedar el primero, cabe esperar que se sitúe por debajo en la siguiente evaluación; y el último, que mejore. Ambos casos, tan solo porque su suerte (buena y mala) no se repetirá. Sí que se repetirá aquello que es constante, su base: el bueno seguirá siendo bueno y el malo, también, pero ninguno tan extremo como la vez anterior.



### Ejercicio 6.1

Imaginemos en el ejemplo del examen repetido que la tendencia fuera a premiar al que queda primero, y castigar al último. ¿Qué veríamos al observar la evolución en el siguiente examen? ¿Cómo lo interpretaríamos?

### Ejercicio 6.2

La presión arterial no es 100% repetible, no hay correlación perfecta entre dos determinaciones en las mismas condiciones. Explique qué pasará al repetir la determinación al que tenga valores muy extremos.

Sea la repetición de una medida  $Y$ , por ejemplo, la PAD, en 2 ocasiones, realizadas de forma que podamos asumir independientes los errores de medida:  $Y_1$  y  $Y_2$ . Para seleccionar pacientes con PAD alta, definimos un punto  $A$  que está por encima de la media de la primera determinación [ $A > E(Y_1)$ ], entonces, asumiendo distribución normal bivariante con varianzas iguales, el valor esperado del cambio o diferencia entre la segunda y la primera determinación vale:

$$E[(Y_2 - Y_1) | Y_1 = A] = (A - E(Y_1))(\rho - 1)$$

**Ejemplo 6.3:** La PAD tiene una media de 80 y una SD de 10 mmHg; si la correlación entre 2 determinaciones vale 0.75 y seleccionamos a los pacientes con PAD igual a 100 mmHg en la primera determinación, ¿cuánto cabe esperar que baje en la segunda determinación?

$$E[(Y_2 - Y_1) | Y_1 = 100] = (100 - 80) (0.75 - 1) = 20 (-0.25) = -5$$

El valor esperado de la próxima repetición será 5 mmHg inferior.



### Ejercicio 6.3

¿Cuál sería la magnitud del fenómeno de regresión a la media si la correlación valiera 1? ¿Y si valiera 0? Interprete. ¿Cómo cambia si A se aleja o se acerca a la media?

### Ejercicio 6.4

Ahora seleccionamos por debajo de la media, a pacientes que están muy-muy sanos en sus parámetros: ¿qué cabe esperar en la siguiente determinación?

**Historieta:** La salud es un estado transitorio que no presagia nada bueno.



### Recuerde

La regresión a la media será mayor cuanto: (1) mayor sea la distancia del punto de corte A de la media; y (2) menor sea la repetitividad de la variable en estudio.

### Recuerde

Mayor fiabilidad, menor problema de regresión a la media.

**Ejemplo 6.4:** Una forma de aumentar la fiabilidad de un proceso de medida consiste en realizar medidas repetidas del mismo proceso y tomar el valor medio (quizás habiendo excluido los 2 extremos). Al aumentar la fiabilidad, disminuirá el fenómeno de regresión a la media.



### Ejercicio 6.5

Vd. quiere valorar cuánto hay de verdad en la “hipertensión de bata blanca” y cuánto hay de regresión a la media. ¿Cómo lo estudiaría?



### Ejercicio 6.6

Se incluye pacientes en ECAs porque tienen cifras alteradas, no normales. Cabe esperar por tanto, cierta regresión a la media. Si, como antes, selecciona a pacientes con PAD igual a 100 ( $SD=10$ ,  $E(Y)=80$  y  $\rho=0.75$ ), razone qué sucederá cuando estudie (1) el “cambio” final-basal en el grupo tratado; (2) su diferencia entre ambos grupos; o bien, (3) la diferencia entre los valores finales en ambos grupos. Formalmente, sean las causas T y C y las PAD inicial, Z, y final, Y. Comente los valores esperados de (1)  $(Y_T - Z_T)$ ; (2)  $(Y_T - Z_T) - (Y_C - Z_C)$ ; y (3)  $(Y_T - Y_C)$



### Recuerde

Tener un grupo comparativo evita los problemas de regresión a la media



### Ejercicio 6.7

Vd. quiere distinguir entre el llamado “efecto placebo” y la regresión a la media. ¿Cómo lo estudiaría?

## 7. Tipos de ensayos clínicos

### 7.1. Según el objetivo

Al aconsejar una nueva intervención, podemos distinguir entre (1) añadirla, sin eliminar nada, como un componente más de la mejor guía disponible; o (2) reemplazar a uno de los componentes de esa guía. En el primer caso, se debe demostrar que el tratamiento T en estudio (la mejor guía más el nuevo componente) es mejor que su comparador C (la mejor guía). En el segundo caso, aunque sería mejor *ganar*, basta con demostrar equivalencia: que la nueva opción (mejor guía con A pero sin B) *empata* con la previa (mejor guía con B pero sin A).



### Ejercicio 7.1

¿Estos objetivos se responden mejor desde la perspectiva de la inferencia (Fisher) o desde la de decisión (Neyman-Pearson)?

### 7.2. Según lo avanzado de la propuesta de intervención

Una primera gran distinción es si el objetivo del estudio es confirmar o explorar. Una hipótesis y un protocolo previos y completamente especificados permiten demostrar o confirmar una intervención y estimar la magnitud de su efecto. Dentro de los estudios exploratorios encontramos los de [factibilidad](#) que podrían incluir tan solo el nuevo tratamiento en estudio, y los pilotos, que pueden llegar a ser una prueba práctica final del estudio confirmatorio antes de emprenderlo.

Una clasificación muy utilizada tiene en cuenta la fase de desarrollo de un fármaco. Una vez se dispone de un producto químico o biológico ya “inventado”, su desarrollo para personas requiere estudiar, por este orden, los niveles tolerados por el organismo (fase I); la eficacia a nivel exploratorio (fase II) y confirmatorio (fase III); y la efectividad (fase IV).

Fase	Objetivo	Voluntarios	Tamaño	Duración
I	Tolerabilidad y cinética	Sanos	Unidades	Días
II	Exploración de Eficacia	Pacientes	Decenas	Semanas
III	Confirmación de Eficacia	Pacientes	Centenas	Meses
IV	Efectividad y seguridad	Pacientes	Millares	Años

**Tabla 7.1:** Ensayos clínicos según la fase de desarrollo del fármaco.



#### Ejercicio 7.2

Ojee la figura de la página 7 del documento [ICH E8](#) y diga si fase (I, II, III, IV) aplica mejor al tipo de ensayo o al momento del desarrollo del fármaco.

Un ensayo de demostración de eficacia (fase III) actúa como una guía o [pivote](#) que decide el futuro del producto (acceso al mercado o vuelta a I+D o...).

### 7.3. Según el diseño

Veamos brevemente algunos diseños que estudiaremos más adelante.

Los pacientes pueden asignarse de forma individual o en grupo ([cluster](#)).

**Ejemplo 7.1:** Un estudio para valorar el efecto de una intervención educativa para prevenir el embarazo adolescente podría asignar clases de estudiantes a la intervención de referencia (información clásica) o a la intervención en estudio (material específico).

En el estudio usual, los pacientes solo reciben un tratamiento en estudio, pero ciertas condiciones crónicas podrían permitir que reciban varias intervenciones en diferentes tiempos o lugares: ensayos con intercambio o [cross-over](#).

**Ejemplo 7.2:** Para valorar el efecto de una intervención sobre el asma, cada paciente se asigna a la secuencia AB o BA, siendo A y B las dos intervenciones en comparación.

Nótese que se trata de un caso particular de estudio apareado.

En el estudio usual, el número final de casos se decide de antemano, pero algunos diseños [secuenciales](#) permiten parar el estudio cuando se ha alcanzado suficiente información para concluir, o que la intervención es eficaz, o que el presente estudio no podrá demostrarlo.

Los ECAs [pragmáticos](#) pretenden valorar el efecto en condiciones lo más parecidas a la clínica real.

#### 7.4. Según el método de asignación

La Tabla 7.2 resume algunas estrategias de asignación al azar del tratamiento.

Asignación	Elección del criterio de asignación de tratamientos
Aleatorización simple	Cada paciente es asignado a un grupo con la misma probabilidad $\Pi$
Aleatorización balanceada por bloques	Cada 'n' casos, el número de pacientes en cada grupo se equilibra. Los bloques se definen por terceras variables (orden, centro, edad,...)
Minimización	Se definen una serie de variables iniciales que se desea tener equilibradas y, conocidos los datos del paciente, se le asigna con mayor probabilidad al grupo que hace mínimo el desequilibrio.
Aleatorización según severidad (cut-off)	Se procede en función de los valores de severidad. Por ejemplo, los casos severos se comparan con un activo y los leves con placebo
Aleatorización y libre elección ( <i>comprehensive cohort</i> )	Se dispone de datos de pacientes que han aceptado la asignación aleatoria y de datos de pacientes que han elegido libremente el tratamiento
Apostar por el mejor ( <i>Play the winner</i> )	Se da mayor probabilidad de asignar un paciente al grupo con mejores resultados hasta el momento

**Tabla 7.2.** Posibilidades de asignación del tratamiento

## Soluciones a los ejercicios

- 1.1.** 1) En un estudio experimental es X y en un estudio observacional es Z.  
2) En un estudio experimental cada visita adicional disminuye 5 mmHg la TA (IC<sub>95%</sub>: 3 a 7 mmHg). En un estudio observacional cada visita adicional se acompaña de 5 mmHg de una disminución de 5 mmHg de la TA. Antes de asegurar que esta relación es causal, necesitamos un diseño experimental, ya que queda por comprobar: (a) que esta relación no puede ser explicada por terceras variables; y (b) que los pacientes cumplirán con el régimen de visitas asignado.
- 1.2.** Cirugía, radioterapia, consejos médicos, dieta ejercicio, tratamientos domiciliarios,...
- 1.3.** Diversas causas alternativas pueden explicar este fenómeno: se puede haber asignado el nuevo tratamiento a los pacientes con mejor pronóstico (sesgo de selección); la evaluación puede haber sido optimista, por el propio entusiasmo del experimentador (sesgo de evaluación); o pueden haberse eliminado del juicio final aquellos pacientes que no han respondido porque “es evidente que en ellos el tratamiento no puede funcionar” (sesgo por atrición).
- 1.4.** Reposo, concienciación de los pacientes, otras intervenciones, calibrado de los aparatos de medida, sesgos del observador, etc.
- 1.5.** Podría ser o bien no tomar nada o bien tomar un vaso de agua sin la aspirina.
- 1.6.** Los pacientes bajo la intervención alternativa obtenidos de experiencias anteriores a la realización del estudio en marcha se denominan *controles históricos*. El problema del uso de este tipo de “controles” es que no se puede asegurar una *comparación justa*: como los dos grupos pueden diferir en alguna característica que no está bajo el control del experimentador cualquier diferencia observada entre grupos no puede ser atribuida al tratamiento. Otro motivo que invalida las comparaciones frente a los controles históricos son los cambios, muchas veces no documentados, en el entorno experimental:
- Uso de información retrospectiva de baja calidad.
  - Cambios en la evaluación de la respuesta: criterios, evaluadores, tecnología,...
  - Nivel de atención médica y de salud en general distinta.
  - Criterios de parada del tratamiento en el grupo histórico distintos a los aplicados al grupo bajo el nuevo tratamiento.
- 1.7.** Todas estas formas de control suelen englobarse en el término Control Externo y no son formas adecuadas para estimar, sin sesgo, el tratamiento en investigación. Como un “control” externo implica muy poco “control” de todas las condiciones del estudio mejor usen “comparador” externo.
- 2.1.** (a) Sí; (b) No; (c) No; (d) No, falta asignación; (e) No, es el estudio, teórico, de los criterios de selección.
- 2.2.** Evidentemente comprometen la representatividad de la muestra. Convendría estudiar si tienen características diferenciales con los que sí que han aceptado participar.
- 2.3.** (a) Sí; (b) Sí; (c) No.
- 2.4.** (a) Sí; (b) No, ya que es lo que proporciona el error estándar; (c) No, ya que lo que estima el error estándar es la distancia promedio o esperada (preguntas anterior y siguiente), pero la de un estudio concreto es desconocida; (d)

Sí. [Nota: si desea ser más preciso recuerde que el error estándar no estima esta distancia sino la raíz cuadrada del promedio de las distancias cuadradas.]

- 3.1.** Un estudio no enmascarado está afectado por todos los riesgos de sesgo: (1) el de selección por incluir a pacientes diferentes en los dos grupos); (2) el de ejecución (p.ej., mejor trato a los del grupo control por misericordia o mejor trato a los tratados por la ambición de los investigadores de demostrar sus hallazgos); (3) el de atrición por eliminar casos de forma diferencial en los grupos (p.e., los curados en el grupo de referencia y los no curados en el experimental); (4) el del informe selectivo, por escoger aquella variable o aquél análisis más conveniente; (5) el de evaluación (p.ej., con una respuesta basada en una escala de valoración subjetiva se podía, incluso de manera inconsciente puntuar distinto según el tratamiento recibido) y el impredecible, porque las diferencias podrían ir en cualquier sentido, resultando en mayor incertidumbre que la que recogen las medidas basadas en oscilaciones únicamente aleatorias (p.e., el error estándar). Los estudios PROBE sólo protegen contra el riesgo de sesgo de evaluación.
- 3.2.** Excluirlo implica perder la comparabilidad de grupos, generando riesgo de sesgo por atrición. Además, el clínico que quiera usar las recomendaciones del estudio, necesita conocer cuántos pacientes abandonaron y porqué.
- 3.3.** Revisando el protocolo y el plan de análisis estadístico puede verse si el análisis se ajusta a lo planeado.
- 3.4.** El error estándar sólo cuantifica la variabilidad en los resultados originada por la aleatoriedad del muestreo. Si hacemos las cosas mal, los resultados aún serán más erráticos, en cualquier sentido. Los valores verdaderos del efecto estarán fuera de los límites del IC<sub>95%</sub> más de un 5% de veces. A diferencia del resto de sesgos, que apuntan en una dirección, ahora aumentan los resultados fuera del IC en ambos lados.
- 3.5.** Dos ejemplo de cada, para comparar (A) dos tipos de dietas para el sobrepeso; y (B) un activo con un placebo.
- Sesgo de selección: (A) se asigna la dieta 1 a pacientes con más de 65 años y la dieta 2 a pacientes con menos de 65 años. Los grupos no son comparables. (B) Al ver que a un paciente grave le toca el grupo placebo, el médico le pide que vuelva más tarde, mientras busca e invita a un paciente muy leve.
- Sesgo de ejecución: (A) los pacientes que reciben la dieta 1 tienen más visitas de control que los de la dieta 2. (B) A los pacientes con placebo se les da enseguida la medicación de rescate.
- Sesgo de atrición: (A) se excluyen del estudio a los pacientes que engordan con una de las dietas. (B) Eliminamos del estudio a los pacientes que no toleran el tratamiento activo.
- Sesgo del informe selectivo: (A) reportar la variable peso a la semana que hace mayores las diferencias; y (B) escoger el análisis que ofrece los resultados deseados.
- Sesgo en la evaluación: (A) valorar más positivamente los resultados de la dieta 1. (B) Usar una variable que permita subjetividades: p.e., periodo libre de enfermedad y explorar más a fondo a los tratados con placebo.
- Sesgo impredecible: (A) si no controlamos el nivel de ejercicio, los resultados se pueden desequilibrar hacia cualquier lado; y (B) si perdemos algunos casos, el desequilibrio puede favorecer a cualquier tratamiento.
- 4.1.** Son falsas la b), la d) y la e). Para saber si es cierta la f) haría falta confirmar la premisa de que el conocimiento de las variables predictoras avanza de las más importantes a las menos importantes.

- 4.2. (i)  $EE=1 \cdot \sqrt{(1/4+1/4)} \approx 0.71$ ; (ii)  $EE=1 \cdot \sqrt{(1/1+1/7)} \approx 1.07$ . El error estándar del caso no balanceado es más de un 50% superior.
- 4.3. ¡Atención! Esta pregunta tiene dos trampas (mis sinceras disculpas). La primera es que la estadística aborda la evidencia empírica, método inductivo de inferencia. Así, la respuesta formal de un estadístico debe ser “es que he repetido el estudio con un mayor lapso de tiempo hasta el tratamiento y me da lo mismo”. Si Vd. desea utilizar el razonamiento lógico-deductivo clásico, debe basarse en consideraciones que exceden nuestro campo. La segunda trampa es que este capítulo discute un tema científico, de adquisición de conocimiento: ¿qué es lo que sabemos sobre un cierto tema? En cambio, la pregunta hace referencia a una decisión concreta: tratar o no tratar a ese paciente. Y la decisión requiere otra metodología. Por supuesto que el conocimiento (bien empírico o bien deductivo) sobre la eficacia influye en esta decisión, pero no es el único factor. Hay que añadir todas las consideraciones sobre seguridad (en general, con una transportabilidad más argumentable) y, sobre coste. Nótese que dejar al paciente ejercer su derecho a decidir conlleva que él, o ella, pueda considerar además sus propias preferencias personales –menos argumentable que puedan ser compartidas por todo el grupo de pacientes con la misma patología.
- 4.4. Sólo es falsa la e), ya que los resultados de este estudio se basan, en el fondo, en un único centro. Nótese que las respuestas a) y b) son dos formas correctas de expresar lo mismo.
- 4.5. Ya están bien ordenadas. a) y b) ofrecen la misma credibilidad. Igual que d) y e). Los dos saltos más importantes son de c) a d) porque deja de estar especificado en el protocolo y el último, ya que ese criterio es una auténtica barbaridad, algo así como si dijéramos: “no pares de analizar los datos hasta que den el resultado que yo quiero”, buen ejemplo de lo que significa “sesgo del informe selectivo”.
- 4.6. (1) Conviene que todos, pero puede ser imposible para f y no tener sentido para l y m, cuyos desvíos pueden ser fácilmente detectados. Y (2) ‘doble ciego’ tuvo su utilidad cuando sólo había 2 protagonistas en un ensayo clínico. Demos las gracias por los servicios prestados: pero ahora CONSORT pide especificar quién está enmascarado.
- 4.7. Deberían por lo menos mencionar que el investigador podía haber adivinado en muchos casos el tratamiento recibido y defender, si pudieran sostenerlo, que este hecho no podía haber originado ninguna diferencia en tratamientos adicionales, ni al determinar el estado vital ni tampoco, más delicado, el nivel de progreso ‘libre de enfermedad’. Nótese que no encuentran diferencias en la variable más objetiva, estado vital, y en cambio sí las ven en ‘libre de progresión’ algo más delicada, ya que un buen clínico, que crea en el producto, explorará más a fondo al paciente tratado con placebo. No, no mencionan esta gran limitación, al revés, “double-blind” está en la primera línea del resumen. Hasta donde hemos visto, no reciben cartas luego, aunque sí hacen ellos una pequeña [corrección](#).
- 4.8. Todo el resto, en especial quiénes tomen decisiones sobre intervenciones adicionales y los evaluadores.
- 4.9. La única falsa es la c). El argumento correcto es el d). Nótese que, por el contrario, en el caso de un fármaco, es imprescindible poder distinguir ambos efectos
- 5.1. El placebo es tan sólo un simulador para enmascarar el tratamiento. La pregunta correcta es si el tratamiento que recibe el grupo control está de acuerdo con los protocolos profesionales. Recuérdese que al grupo “placebo” no se le suprimen los tratamientos de conocida eficacia. Por eso, muchos autores prefieren llamarlos estudios “add-on” para recordar que lo que se hace es añadir el tratamiento en estudio.

- 5.2. El de equidad, ya que corremos el riesgo de que algunos grupos sociales tengan mayor acceso a la intervención en estudio.
- 5.3. En el párrafo 30.
- 5.4. El punto 31 dice que el protocolo debe especificar la política de publicaciones. Si el protocolo no deja claro que los resultados se publicarán independientemente de los resultados y que el patrocinador y el financiador no podrán influir en el redactado, el CEIC puede negar la autorización del estudio.
- 5.5. El punto 31 dice que el protocolo debe especificar la política de publicaciones. Si el protocolo no deja claro que los resultados se publicarán independientemente de los resultados y que el patrocinador y el financiador no podrán influir en el redactado, el CEIC puede negar la autorización del estudio.
- 6.1. Observaríamos, quizás con frustración, ¡que el premiado empeora y el castigado mejora! Un observador que no tenga en cuenta la variabilidad, interpretaría que “la letra con sangre entra”.
- 6.2. El que tenga valores muy extremos, tendrá tendencia a repetir valores extremos, sí, pero menos extremos: aquél con valores extremadamente bajos, subirá algo su PA; y aquél con los valores más altos, bajará algo. Sus posiciones extremas serán ocupadas por otros casos y la variabilidad de todos los casos no cambiará (en ausencia de un efecto estabilizador como el reposo, relajación, etc.).
- 6.3.  $E[(Y_2 - Y_1) | Y_1 = 100] = (100 - 80) (1 - 1) = 20 (0) = 0$ ;  $E[(Y_2 - Y_1) | Y_1 = 100] = (100 - 80) (0 - 1) = 20 (-1) = -20$ ; A medida que se acerca a la media, la magnitud del efecto es menor.
- 6.4. Se espera que estén otra vez con valores bajos, pero no tanto, algo más cerca de la media.
- 6.5. Muy difícil —o imposible. Una primera aproximación sería repetir la determinación en todos los casos y mirar si la distribución ha cambiado. El problema es que el reposo (o algo parecido) puede haber originado tanto un descenso de la media como de la variabilidad. Si fuera posible, el diseño ideal debería repetir lo anterior en 2 grupos, uno “expuesto a la bata blanca” (o sus componentes hipertensores) y el otro no. Pero si uno de los componentes es la propia ansiedad del paciente.... Si se le ocurre un diseño, puede, o discutirlo en el foro, o ¡hacerlo!
- 6.6. (1) la variable “cambio” final-basal en el grupo tratado ( $Y_T - Z_T$ ) estará afectada por el fenómeno de regresión a la media  $E[(Y_T - Z_T) | Y_T = 100] = (100 - 80) (0.75 - 1) = 20 (-0.25) = -5$ .
- Ahora bien tanto (2) su diferencia entre ambos grupos,  $(Y_T - Z_T) - (Y_C - Z_C)$ , como (3) la diferencia entre los valores finales en ambos grupos,  $(Y_T - Y_C)$  están afectadas por igual por el fenómeno de regresión a la media, por lo que, en ausencia de efecto del tratamiento y de sesgos, tendrán el mismo valor esperado en ambos grupos.
- 6.7. Una opción (¿ética?) sería realizar un ensayo clínico con tres grupos: el grupo experimental, un grupo comparador que recibe placebo, y un tercer grupo que saben que no son tratados. Un [estudio](#) que compara los 2 últimos grupos encuentra mejor evolución en el tratado con placebo.
- 7.1. Nótese que deseamos responder cómo proceder delante de un paciente (¿qué intervención aconsejarle?), por lo tanto, el paradigma de decisión de NP será más adecuado.
- 7.2. La figura muestra cómo ambos están relacionados, pero que aplica al desarrollo de cierta indicación del producto. Así, una nueva indicación puede requerir un nuevo estudio de confirmación de eficacia aunque ya existan de seguridad previos.