

# Anexo 11. Valoración de la calidad de los distintos tipos de estudios

## Ensayos Clínicos

### Definición

Un ensayo clínico aleatorizado (ECA) es un estudio experimental en el que se aplica una intervención a un grupo de pacientes (grupo experimental) y el resultado se compara con el obtenido en un grupo idéntico de pacientes (grupo control) que recibe otra intervención. Los pacientes son asignados de forma aleatoria al grupo experimental y al grupo control con la misma probabilidad, de forma que ambos grupos sean similares en todo (características de los pacientes, factores pronósticos conocidos o no, otros tratamientos recibidos, etc.), excepto en la intervención que reciben. De esta forma, la diferencia observada podrá atribuirse al tratamiento recibido, con un margen de error establecido, y siempre que el ensayo esté libre de sesgos.

El ECA se considera el *patrón oro* para valorar la eficacia de las intervenciones. Los ECA que se evalúan con más frecuencia en GPC son ensayos en fase III o en fase IV.

### Aspectos metodológicos

- **Variables de resultado.** Se debe tener en cuenta si los resultados se miden mediante variables intermedias (subrogadas) o variables finales. El ensayo debe definir claramente cuál es la variable principal y cuáles las secundarias; éstas últimas deben estar predefinidas en el diseño del estudio. El ensayo puede no tener potencia estadística suficiente para detectar diferencias estadísticas significativas en variables secundarias; por otro lado, es necesario ser cautos ante hallazgos inesperados procedentes de variables secundarias. Si las variables de resultado son compuestas, su validez dependerá de que se cumplan tres premisas: que los distintos componentes tengan la misma importancia clínica, que la frecuencia con la que ocurren sea similar y que la intervención afecte de igual forma a todos los componentes.
- **Aleatorización.** La fortaleza principal de un ensayo clínico es la adecuada aleatorización. Si no hay ninguna indicación de que se ha efectuado una aleatorización, el estudio debería ser rechazado. Es importante que además exista “Ocultación de la Secuencia de Aleatorización” (OSA), es decir, que quien realiza la investigación no sepa qué rama le va a corresponder a cada sujeto ni pueda influir de manera consciente o inconsciente en el orden de aleatorización. El método de ocultación debe estar claramente descrito. Es importante también comprobar si existen diferencias en las características basales de los grupos.
- **Enmascaramiento o “doble ciego”.** Idealmente, en un ECA ni el paciente ni el investigador deberían saber qué tratamiento está recibiendo el grupo experimental o el grupo control. Suele ser frecuente en ECA con fármacos. Es especialmente importante cuando las variables que se miden son subjetivas (como calidad de vida, dolor...) o cuando es posible que, al margen de la intervención que se estudia, puedan añadirse otros tratamientos concomitantes. Si por razones éticas o

técnicas el ensayo no puede ser *doble ciego* (por ejemplo, intervenciones quirúrgicas, educativas, un programa de ejercicio físico), al menos el evaluador de los resultados debería desconocer la intervención asignada (“evaluación ciega por terceros”).

- El seguimiento de los pacientes. Siempre se deben cuantificar y describir las causas de la pérdida de pacientes. El porcentaje de pérdidas puede ser muy variable, en función de la patología estudiada, del tipo de intervención o de la duración del estudio. Es importante que los motivos de las pérdidas afecten de la misma forma al grupo experimental y al control, y que el porcentaje de abandonos no sea muy elevado. Si en una de las ramas hay más pérdidas debido a efectos adversos o a falta de eficacia, los resultados estarán sesgados; si las causas de las pérdidas se distribuyen de la misma forma en las dos ramas su influencia será menor. Es importante que, al analizar los resultados, se siga el principio del “análisis por intención de tratar” (AIT), de forma que cada paciente se contabilice en la rama a la que fue aleatorizado/a, aunque no completen el estudio. La excepción a esta regla son los estudios de equivalencia y de no inferioridad, donde se deben presentar tanto el AIT como el análisis por protocolo.
- El análisis de subgrupos: Los resultados provenientes de análisis de subgrupos son más fiables si se cumplen los siguientes requisitos:
  - o La hipótesis específica se había planteado en el diseño del estudio y es una de las pocas hipótesis que se quiere contrastar.
  - o La magnitud del efecto es grande.
  - o El efecto del tratamiento es altamente significativo (cuanto más bajo es el valor de la  $p$  más creíble será la diferencia).
  - o Los resultados son consistentes con los de otros estudios.
  - o Existe evidencia científica indirecta a favor de los resultados encontrados (plausibilidad biológica).
  - o Más creíbles si se observa un beneficio en la población global del estudio.

## Revisiones sistemáticas

### Definición

Es una revisión acerca de una cuestión formulada claramente, y que utiliza métodos sistemáticos y explícitos para identificar, seleccionar y evaluar críticamente la investigación relevante, así como para obtener y analizar los datos de los estudios que son incluidos en la revisión. Aunque lo más habitual son las revisiones sistemáticas de ensayos clínicos, existen revisiones sistemáticas de otros diseños de estudios (como casos-control, cohortes, estudios de pruebas diagnósticas, o RS que incluyen distintos diseños sobre el mismo tema).

Se pueden utilizar o no métodos estadísticos (metaanálisis) para combinar los resultados de los estudios incluidos, dando una estimación global del efecto. La unidad de análisis de un metaanálisis son los estudios individuales.

## Aspectos metodológicos

- **Búsqueda y selección de la literatura.** La principal fortaleza de una revisión sistemática es que localiza y sintetiza toda la evidencia científica relevante sobre una pregunta concreta, por lo que es fundamental que la búsqueda sea lo más exhaustiva posible utilizando las bases de datos más específicas sobre el tema tratado. Es deseable que ésta se complemente con una búsqueda de la literatura gris, búsqueda manual en revistas clave, rastreo de las referencias incluidas, contacto con expertos en el tema, con la industria o con las agencias reguladoras. Los criterios de inclusión y exclusión de los estudios deben ser claros, de acuerdo al protocolo de la revisión. Se debe evaluar la calidad de los estudios individuales incluidos.
- **Sesgo de publicación.** Para saber si los estudios encontrados son representativos de toda la investigación realizada, pueden utilizarse técnicas gráficas (como el *funnel plot*) o estadísticas (Rosenthal). Si el número de estudios incluidos es pequeño, estas técnicas tienen un valor limitado.
- **Heterogeneidad.** Es la variabilidad o las diferencias entre los estudios incluidos en la revisión y es un aspecto clave para decidir si es o no razonable combinar los resultados. La heterogeneidad estadística se refiere a diferencias en los efectos reportados, la heterogeneidad metodológica a diferencias en el diseño y calidad de los estudios y la heterogeneidad clínica a diferencias entre los estudios referidas a características clave de los participantes, a intervenciones o a medidas de resultado. Es importante que se analicen las posibles causas de heterogeneidad en función de los factores mencionados. Si éstos difieren sustancialmente, lo más recomendable puede ser no agregar los resultados. Los test estadísticos de heterogeneidad se utilizan para valorar si la variabilidad en los resultados de los estudios (la magnitud de los efectos) es mayor que la esperada por el azar. Estos test tienen una baja potencia estadística cuando el número de estudios es pequeño. Si se sospecha o se demuestra heterogeneidad se debe intentar explicarla realizando un análisis de sensibilidad, en función de aspectos como la calidad de los estudios, las características de los pacientes o de las intervenciones, etc. En ocasiones la heterogeneidad no explicada puede determinar que la síntesis mediante metaanálisis no sea apropiada, realizando entonces una RS que sintetiza los resultados de los estudios individuales, sin agregarlos. En presencia de heterogeneidad, si se opta por combinar los resultados, se debe emplear un modelo de análisis de efectos aleatorios (más conservador) en lugar de un modelo de efectos fijos.

## Pruebas diagnósticas

### Definición

Los estudios sobre pruebas diagnósticas pueden cumplir dos objetivos. El primero es valorar el impacto de una o de varias estrategias diagnósticas en las decisiones clínicas o en los resultados en pacientes. Esta evaluación se realiza mediante ensayos clínicos o estudios comparativos no experimentales. Este tipo de planteamiento, aunque es el ideal, está disponible en muy pocas ocasiones.

El segundo objetivo, tradicionalmente más frecuente, es conocer la capacidad diagnóstica de una prueba (capacidad para clasificar a una persona como sana o enferma). En este apartado nos referimos a este segundo objetivo. Su diseño se basa en una comparación entre la prueba que se estudia y el patrón oro (*gold standard*), que se aplican a un conjunto de pacientes, evaluándose los resultados en términos de sensibilidad, especificidad, valores predictivos o cocientes de probabilidad.

### Aspectos metodológicos

- Selección de los pacientes. El estudio debe incluir el espectro de pacientes más parecido posible al que se pretende aplicar la prueba (por ejemplo, incluyendo pacientes con distintos grados de severidad). Un sesgo muy frecuente consiste en comparar la prueba en pacientes que se sabe tienen la enfermedad y en individuos que no la tienen. Este diseño “casos-control” sobreestima el rendimiento de la prueba, que ha de incluir pacientes con sospecha de la enfermedad.
- Independencia entre la prueba y el patrón oro. Se debe aplicar el patrón oro a todos los sujetos del estudio, no sólo a los que han presentado resultados positivos con la prueba a estudio. A veces, cuando este patrón es una prueba invasiva, se utiliza un patrón de referencia distinto en los que presentan un resultado positivo y negativo, lo que conduce a una sobreestimación del rendimiento de la prueba a estudio.
- Enmascaramiento en la interpretación de las pruebas. La prueba a estudio debe valorarse sin conocer los resultados de la prueba de referencia y viceversa.

## Estudios de Cohortes

### Definición

Los estudios de cohortes son un tipo de estudio observacional en el que se hace un seguimiento a un grupo de personas (una cohorte) durante un tiempo determinado con el objetivo de evaluar la aparición de un determinado evento.

Estos estudios se pueden utilizar con fines diferentes: para determinar si la frecuencia de una enfermedad o acontecimiento es diferente dependiendo de la exposición a un factor, por ejemplo, estudiar la frecuencia de enfermedad coronaria en personas que están expuestas al humo del tabaco (fumadores pasivos); para valorar la evolución de la enfermedad mediante el seguimiento de una cohorte (estudios de pronóstico).

Los estudios de cohortes pueden ser útiles para medir la efectividad de una intervención, cuando no se dispone o no es aceptable la realización de ensayos clínicos, pero están más sujetos a sesgos que éstos últimos.

El estudio de cohortes puede ser prospectivo o retrospectivo. Los prospectivos están menos sujetos a sesgos, los retrospectivos dependen mucho de la calidad de los registros utilizados.

### Aspectos metodológicos

- Un estudio de cohortes debe incluir una buena descripción de la población que se estudia (incluyendo claramente el grado de evolución de la enfermedad), y de los

criterios de selección (por ejemplo, si los pacientes provienen del ámbito hospitalario es probable que tengan un peor pronóstico que si proceden de atención primaria). Asimismo, debe describir de forma detallada la exposición, los efectos esperados y las medidas que se utilizan para evaluarlos.

- En los estudios de pronóstico es importante que los pacientes sean homogéneos en cuanto a las características de presentación de la enfermedad y el momento de la evolución de ésta.
- Cuando se comparan dos cohortes, sus características deben ser lo más similares posible, salvo en la exposición al factor que se pretende estudiar.
- El periodo de seguimiento debe ser suficientemente largo para que dé tiempo a la aparición de los efectos y a tener en cuenta las pérdidas y sus causas.
- Es necesario analizar siempre la presencia de sesgos y considerar y controlar los factores de confusión y modificadores del efecto (interacciones), tanto en el diseño como en el análisis de los resultados (características de los pacientes, tratamientos recibidos durante el seguimiento que pueden modificar la historia natural de la enfermedad, presencia de otros factores de riesgo, etc.).

## Estudios de Casos y controles

### Definición

Los estudios de casos y controles son un tipo de estudio observacional en el que se identifica un grupo de personas que tienen el efecto o la enfermedad de interés (casos) y se compara con otro grupo de sujetos que carecen de ella (controles). Se realiza una investigación en ambos grupos hacia atrás en el tiempo buscando la exposición o el factor que se sospecha que pueda tener relación con el efecto de interés. Habitualmente se utilizan registros o historias clínicas como fuentes de información.

Generalmente los estudios de casos o controles se utilizan para analizar las causas de un problema (etiología) y son también una alternativa a otros diseños (como ensayos clínicos para valorar efectos adversos de las intervenciones o estudios de cohortes para valorar factores de riesgo de enfermedades) cuando se quieren estudiar efectos adversos o enfermedades poco frecuentes o con largos períodos de latencia.

### Aspectos metodológicos

- Un aspecto metodológico básico que condiciona la validez de un estudio de casos y controles es que los controles deben representar a la población de la que provienen los casos. Por ello el proceso de definición y selección de los casos y de los controles es un aspecto clave del cual depende en gran medida la calidad de la evidencia científica aportada por el estudio.
- Los criterios de selección y exclusión deben ser aplicados por igual a los casos y a los controles.
- Se deben estudiar y ajustar los factores de confusión y los modificadores de efecto.
- Se deben describir adecuadamente los registros y fuentes de datos utilizados.

## Evaluación económica

### Definición

La evaluación económica es un conjunto de técnicas que utilizan métodos y teorías de las ciencias de la salud y la economía, desarrolladas para evaluar los aspectos económicos de la asistencia sanitaria, es decir, los costes y los beneficios o consecuencias de diferentes intervenciones, y que proporciona datos útiles para la toma de decisiones sanitarias.

La evaluación económica puede realizarse desde la perspectiva de la sociedad, del sistema sanitario y del paciente. El punto de vista que se asuma en el análisis es importante porque influye en las medidas de costes y efectos que se van a considerar.

La metodología empleada es muy variable, se puede realizar evaluación económica en el marco de estudios epidemiológicos observacionales (con información procedente de historias clínicas, registros y bases de datos administrativas), mediante modelización (como análisis de decisión o los modelos de Markov), o a través de ensayos clínicos que incorporen una evaluación económica. Estos últimos tienen la ventaja de que tanto los costes como los efectos se miden de forma prospectiva.

El análisis de la incertidumbre puede realizarse mediante análisis de sensibilidad o mediante análisis estadísticos como el muestreo repetitivo o *bootstrap*, el método de Fieffer o el análisis probabilístico con técnicas de simulación de Monte Carlo.

Algunas organizaciones como NICE han incorporado la evaluación económica en la valoración de las tecnologías sanitarias y han establecido el método que se empleará (habitualmente un análisis de tipo coste-utilidad, y una perspectiva social).

Un aspecto sin resolver en las evaluaciones económicas es la transferibilidad de los resultados a contextos sanitarios y países diferentes.

### Aspectos metodológicos

- Se debe especificar claramente cuál es la perspectiva del análisis (el sistema sanitario, la sociedad, etc.).
- Se debe justificar adecuadamente la elección del diseño del estudio.
- El estudio debe incluir todos los costes y efectos que son relevantes desde el punto de vista elegido y tanto los costes como los efectos deben medirse y valorarse de forma adecuada. Las asunciones realizadas deben estar suficientemente explicadas y justificadas. Se debe establecer de forma explícita la tasa de descuento utilizada.
- Es importante que se realice un análisis de la incertidumbre mediante el análisis de sensibilidad u otras técnicas (como los modelos probabilísticos) para comprobar qué repercusión tienen en los resultados las variaciones en los parámetros sujetos a incertidumbre.

## Series de casos

### Definición

Las series de casos clínicos son un tipo de diseño observacional basado en la identificación y descripción de un conjunto de casos clínicos que aparecen en un intervalo de tiempo. Puede tener carácter prospectivo o retrospectivo.

La fuente de información es la observación clínica de un grupo de pacientes que tienen un diagnóstico similar, en los cuales se puede valorar la evolución de una enfermedad o la respuesta a un tratamiento.

A pesar de ser uno de los estudios realizados con mayor frecuencia en la investigación clínica, su utilidad es bastante limitada puesto que los resultados se refieren a un grupo limitado de personas, carecen de grupo de comparación o control y no sirven para probar hipótesis de investigación

Desde el punto de vista de calidad de la evidencia científica se considera que las series de casos aportan baja calidad. Sin embargo, en ocasiones, según las características de la investigación, puede ser el único diseño disponible, por lo que es importante tener en cuenta los diferentes criterios que ayudan a verificar la calidad de la evidencia científica aportada por este tipo de estudios.

### Calidad de la evidencia científica

- Debe incluir una descripción detallada de la enfermedad que se estudia, de los criterios diagnósticos y del método de selección de pacientes. Es importante conocer las características de las personas incluidas y excluidas de la serie de casos, la gravedad de la enfermedad y la comorbilidad.
- Se debe describir la fuente de obtención de los datos.
- La relación temporal ha de estar adecuadamente documentada.
- Los resultados deben estar bien descritos y es necesario comprobar que las conclusiones provengan de los resultados obtenidos del estudio y no de otras fuentes de información.
- Las series de casos son más sólidas si la recogida de datos es prospectiva y a partir de un protocolo estandarizado de recolección de la información.