

# Diseño de los estudios de investigación. Debilidades y fortalezas

Types of research study designs. Strengths and weakness.

**Bottaro FJ**

*Servicio de Clínica Médica - Hospital Británico  
Coordinador del Comité de Revisión Institucional - Hospital Británico*

*febottaro@hotmail.com*

*Fecha de recepción: 10/04/2014  
Fecha de aprobación: 14/04/2014*



COMO SE LEE  
UN ARTÍCULO  
CIENTÍFICO

HEMATOLOGÍA, Vol. 18 N° 1: 74-83  
Enero - Abril 2014

*“Todo se resume en una sentencia sencilla: existen buenas y malas maneras de hacer las cosas. Usted puede practicar el tiro 8 horas diarias, pero si la técnica es errónea, sólo se convertirá en un individuo que es bueno para tirar mal”.*

**Michael Jordan**

**Palabras clave:** diseño, cohortes, caso-control, ensayo clínico controlado, corte transversal.

**Keywords:** design, cohorts, case-control, randomized controlled trial, cross-sectional.

En esta entrega analizaremos los diseños de estudios de investigación más comunes ventajas y debilidades de cada uno y sus usos más frecuentes.

No todos los diseños podrán mostrarnos o evidenciar las mismas cosas. La capacidad de un estudio de hallar una asociación entre dos variables, como también de ser susceptible a determinados sesgos, dependerá en gran medida del diseño del mismo.

## Identificación del diseño de un estudio

Identificar el diseño del estudio que analizaremos, y si es el adecuado para responder a la pregunta de investigación planteada es uno de los primeros pasos

en la evaluación del mismo.

Hay dos categorías básicas de estudios que son fácilmente identificables: *descriptivos y longitudinales*. Los primeros son estudios observacionales que comprenden la descripción de uno o varios casos clínicos agrupados por alguna condición; y los denominados “estudios de corte transversal” que describen las características de una población en un momento determinado (algo similar a lo que sería obtener una “foto” de una población). Los estudios *longitudinales* pueden ser tanto estudios observacionales o de intervención que son realizados en un período de tiempo determinado (ver **Tabla 1**).

**Tabla 1:** Características de diferentes tipos de diseños de estudios de investigación

DISEÑO	CARACTERÍSTICA DEL ESTUDIO		
	Randomizado	Prospectivo	Grupo control o comparador
ECA	<b>Si</b>	<b>Si</b>	<b>Si</b>
Ensayo clínico no aleatorizado con grupo control histórico o actual	<b>No</b>	<b>Si</b>	<b>Si</b>
Estudio de cohortes	<b>No</b>	<b>Si/No</b>	<b>Si</b>
Casos y controles	<b>No</b>	<b>No/Si</b>	<b>Si</b>
Corte transversal	<b>No</b>	<b>No</b>	<b>Si</b>
Serie de casos	<b>No</b>	<b>No/Si</b>	<b>No</b>
Reporte de un caso	<b>No</b>	<b>No</b>	<b>No</b>

ECA: ensayo clínico aleatorizado

En negrita la opción más importante

### Estudios descriptivos

Su objetivo es el de describir un paciente, grupo de pacientes o población determinada en un momento específico. Su finalidad es la de comunicar a la comunidad médica las características de una nueva enfermedad, signo clínico, síntoma, tratamiento o test diagnóstico. Los *reportes de casos* describen características de un paciente individual, mientras que las *series de casos* describen características de un grupo de pacientes que comparten una condición. En los estudios de *corte transversal*, aspectos interesantes de una población determinada, incluyendo causas y efectos potenciales, son todos observados en un mismo momento determinado.

#### Reporte de caso y Series de casos

Si bien representan el escalón más bajo de la Medicina Basada en la Evidencia, este tipo de publicaciones ha sido la principal manifestación en la literatura médica por mucho tiempo.

Los nuevos tratamientos o procedimientos diagnósticos descritos en un estudio sin un grupo control también caen dentro de esta categoría de *series de casos*.

Algunos reportes de casos puede alertar a la comunidad médica de una nueva enfermedad que esté

adquiriendo relevancia. Por ejemplo, el síndrome de inmunodeficiencia adquirida (SIDA) fue inicialmente identificado cuando los primeros pacientes fueron reportados en dos series de casos en 1981. Una de las series correspondía a un grupo de pacientes homosexuales previamente sanos que habían desarrollado neumonía por *Pneumocystis jirovecii* (*Pneumocystis carinii* en ese entonces)<sup>(1)</sup> y la segunda, un grupo de hombres con sarcoma de Kaposi.<sup>(2)</sup>

Este fue el comienzo de la epidemia del SIDA, un hecho que no era evidente al momento de la publicación de estas observaciones; pero que sin embargo alertaron a médicos en otras regiones al observar casos similares.

Como la mayoría de los reportes de casos son descripciones de enfermedades muy raras o presentaciones muy poco frecuentes de enfermedades más comunes, es muy poco probable que se repitan en períodos breves de tiempo. A pesar de cierta coherencia fisiopatológica, la asociación observada entre dos variables en estas publicaciones puede ser solamente por azar. Por lo que si bien son útiles poseen limitado valor. En la mejor de las situaciones, estos estudios descriptivos pueden sugerir hipótesis de

estudio a futuro. Sin embargo, al igual que los estudios de corte transversal, poseen ciertas ventajas que son interesantes. Son baratos, relativamente fáciles de realizar con los registros médicos existentes, y cualquiera que observe nuevas presentaciones de una enfermedad o casos interesantes puede escribir y publicar un reporte de caso. No obstante, sus debilidades equilibran sus ventajas. Estos estudios no proveen explicación alguna de lo que reportan, ni pueden establecer asociaciones de causalidad (causa-efecto). Por esta razón nunca debemos cambiar nuestra práctica médica basados únicamente en estudios de este grupo.

### Estudios de corte transversal

Típicamente son estudios descriptivos que observan una muestra proveniente de una población para determinar por ejemplo cuantos individuos en esta población están afectados por una determinada condición. Estos estudios registran eventos y observaciones en un momento determinado, en una población determinada. Son muy útiles como estudios exploratorios iniciales para rastrear o clasificar aspectos de una determinada enfermedad. Es posible que describan asociaciones entre variables que fueron observadas, pero sin poder establecer una vinculación de "causa-efecto" entre estas. Si están bien hechos, estos estudios pueden establecer prevalencia, pero como la observación se realiza en un momento dado determinado no es posible calcular incidencia con los datos que derivan de estos. Estos estudios son lo más parecidos a obtener una "fotografía" de una población. En ella veremos múltiples asociaciones sin poder saber cuál de las variables "asociadas" es causa o consecuencia de la otra.

Los estudios de corte transversal son especialmente susceptibles a ciertos sesgos. Por ejemplo, podemos estar ante una situación en la que una variable que impresiona ser la causa de un evento es en realidad una consecuencia de este, o solo están asociadas. Esto puede ocurrir cuando un factor de riesgo está fuertemente asociado con una enfermedad y se cree que su aparición es previa a la aparición de esta; pero sin embargo, en realidad solo afecta la duración o el pronóstico de la misma. Un ejemplo de esta situación es la asociación existente entre el antígeno HLA-a2 y la leucemia linfoblástica aguda (LLA) en niños que fue descrita en un estudio de corte transversal. Tras esta publicación se asumió a este

antígeno como factor de riesgo para ocurrencia de la LLA. Varios estudios posteriores hallaron que sobrevivientes de la enfermedad a largo plazo poseían el antígeno, mientras que la ausencia del mismo era indicativa de muerte temprana. El HLA-A2 no era un factor de riesgo para el desarrollo de la enfermedad, sino un indicador de buen pronóstico.<sup>(3)</sup>

### Estudios longitudinales

Hay tres tipos de estudios longitudinales: estudios de casos y controles, estudios de cohortes y ensayos clínicos. Estos estudios son analíticos o estudios inferenciales, en la medida en que buscan hallar una asociación estadística entre dos variables (factor de riesgo y evento).

Para poder establecer esta asociación entre dos variables es necesaria la existencia de un grupo control para poder comparar la frecuencia de aparición de un evento en uno de los grupos en relación al otro. La selección de un grupo control adecuado es vital para poder hallar asociaciones y dependerá de cada diseño la elección de los controles (discutiremos más adelante como seleccionar adecuadamente los controles en un estudio de investigación).

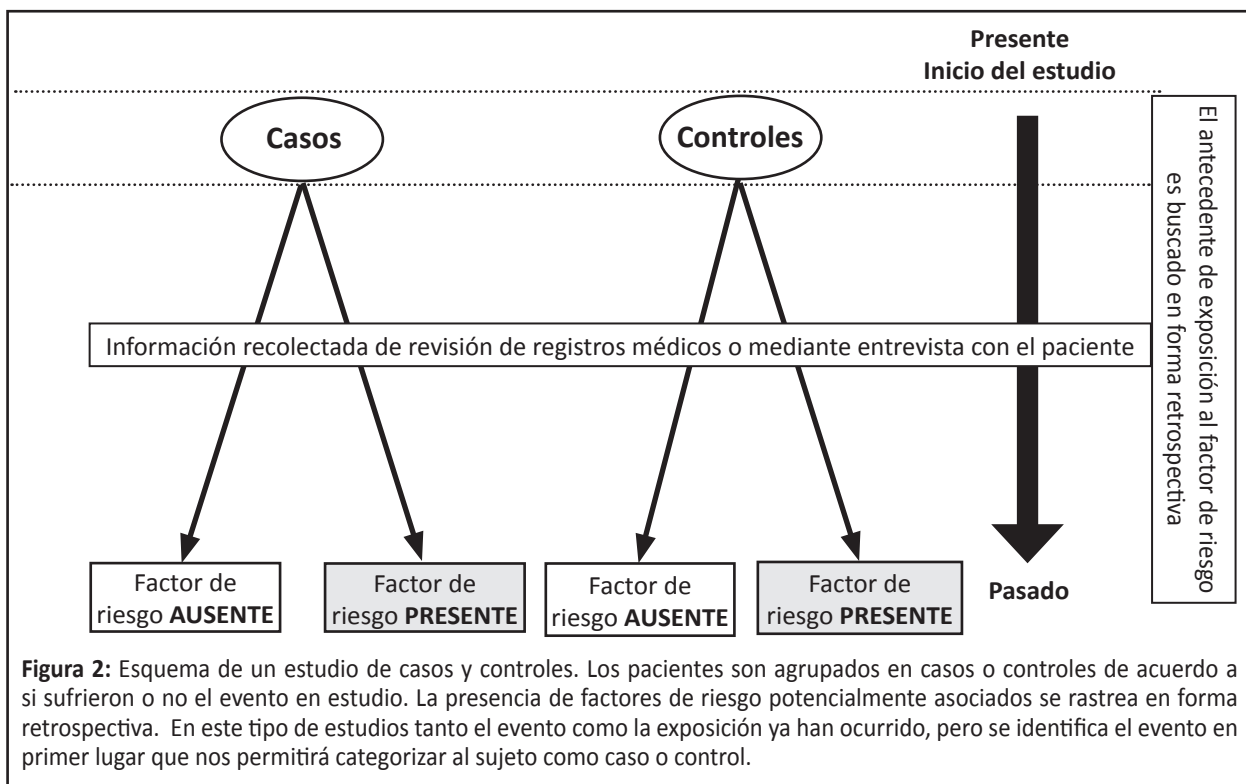
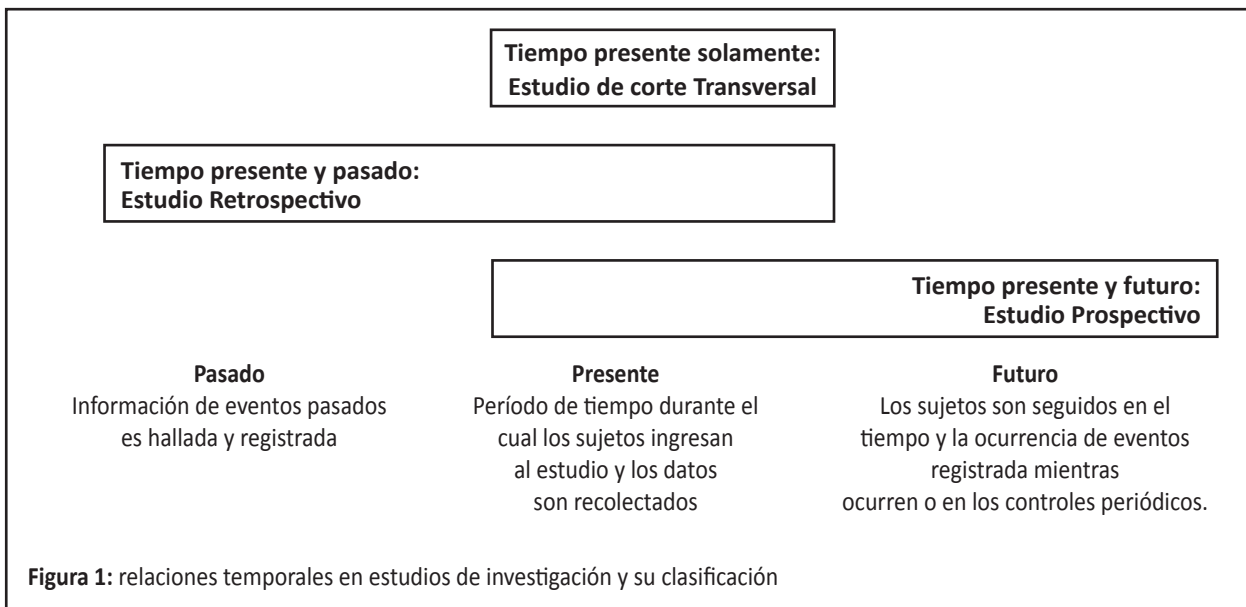
#### Estudios de casos y controles

Estos estudios suelen ser habitualmente retrospectivos (ver Figura 1). La pregunta que intentan responder este tipo de estudios es: ¿cuáles son los factores de riesgo asociados al desarrollo de este evento.

Se caracterizan en que los sujetos son inicialmente seleccionados de acuerdo a la ausencia (controles) o presencia (casos) de un evento determinado de interés. De esta manera, los sujetos son agrupados al comienzo de la investigación como casos o controles dependiendo de haber padecido o no el evento en estudio, y luego se intenta determinar la existencia de factores de riesgo que pudieran estar asociados con el evento mencionado (ver Figura 2).

Este tipo de estudios son especialmente útiles como forma primaria de estudiar un a enfermedad rastreando potenciales factores de riesgo revisando los antecedentes del sujeto de investigación.

Debido a que el número de casos y controles es arbitrariamente fijado por el investigador no es posible obtener datos de prevalencia ni incidencia de estos estudios. Por esta razón, la medida de asociación entre variables que se observan en este tipo de estudios es el odds ratio, no pudiéndose utilizar en ningún caso el riesgo relativo (ver artículo anterior de la serie



para una descripción de los mismos<sup>(4)</sup>). La fortaleza de los estudios de casos y controles reside en que son estudios relativamente fáciles de conducir, baratos y rápidos de realizar. Son estudios ideales para investigar enfermedades o eventos de muy baja incidencia ya que el estudio comienza cuando el evento o la enfermedad ya han ocurrido o es identificable. Una vez

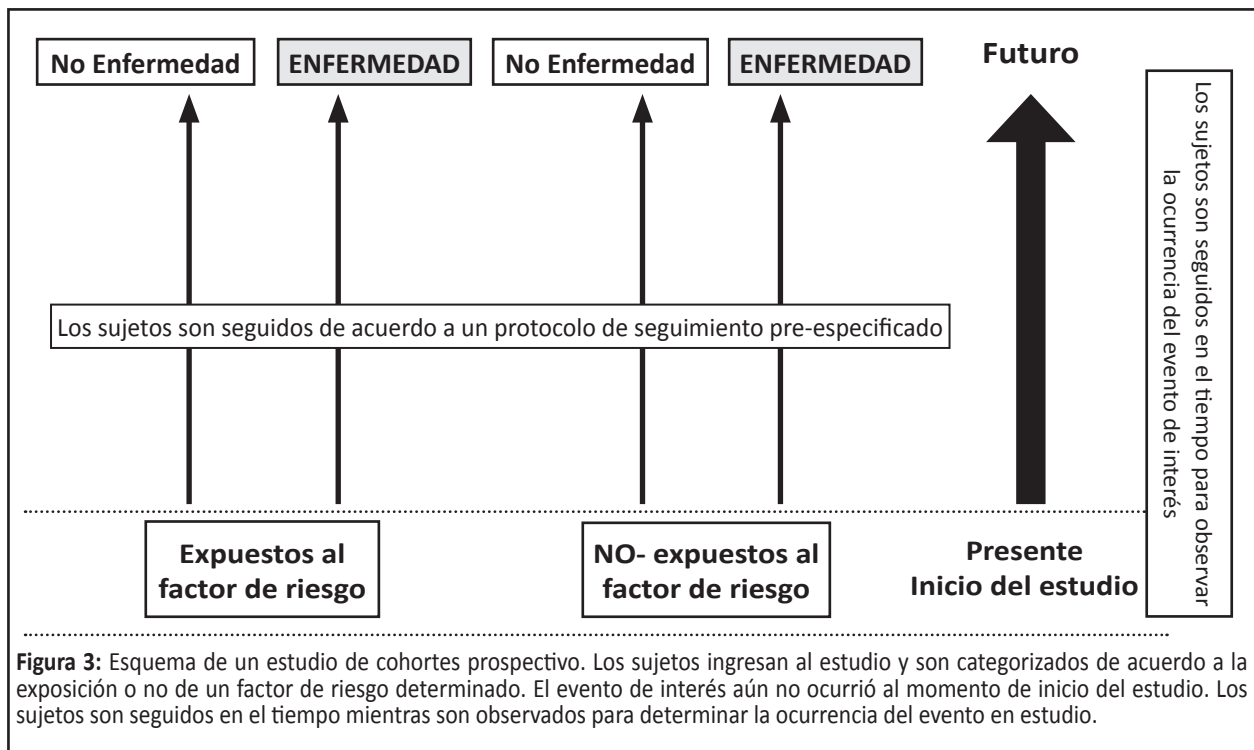
seleccionado los casos, se selecciona al grupo de sujetos que servirán como población control. Podrán ser 1 o varios controles por cada caso. Idealmente los controles debieran ser similares a los casos excepto por la presencia del evento índice o enfermedad en estudio. Son buenos estudios exploratorios y pueden evaluar varios factores de riesgo en un solo estudio.

Los resultados pueden servir para luego generar hipótesis de posteriores estudios con diseños más sólidos. Este tipo de estudios es parte de la evolución natural de una serie de casos, a la que luego de su recolección se la comparó con un grupo de controles seleccionado. Sin embargo, desafortunadamente existen varias desventajas de este tipo de diseño: el riesgo de error y/o sesgos en la información que se obtiene vinculada con la exposición, la potencial pobre calidad de la información que se recolecta dada la naturaleza retrospectiva de estos estudios, y las dificultades en seleccionar un grupo control adecuado. Un buen ejemplo de este tipo de estudios al evaluar una nueva enfermedad fue una investigación que se llevó a cabo hace 4 décadas ante el inesperado aumento de casos de cáncer de vagina en mujeres jóvenes. Esta condición, habitualmente descripta como rara y más frecuente en pacientes añosas, se presentó en 7 mujeres de entre 15 y 22 años en un lapso de 3 años en un hospital de Boston. Estos 7 casos (sumados a un octavo adicional) fueron estudiados, para lo cual se seleccionó cuatro controles por cada uno de estos casos que consistían en mujeres nacidas dentro de un período de 5 días de diferencia de los casos en los mismos hospitales que estas. Fue planeada una entrevista detallada para ser llevada a cabo de la misma manera en las 40 mujeres

que se evaluaron (8 casos y 32 controles). De estas entrevistas surgió una diferencia en el uso por parte de las madres de las pacientes durante el embarazo de dietilestilbestrol (DES) a favor en el grupo de los casos. Esta asociación entre uso de DES en el embarazo y desarrollo de cáncer de vagina en las hijas fue confirmada en estudios posteriores.<sup>(5)</sup>

**Estudios de cohortes**

Habitualmente se los denomina estudios prospectivos, porque usualmente son realizados desde el presente siguiendo los sujetos en el tiempo. Sin embargo, también pueden realizarse en forma retrospectiva. Los estudios de cohortes consisten en un grupo de sujetos que son seleccionados y clasificados de acuerdo a la presencia o ausencia de determinados factores de riesgo o exposición (Figura 3) y son seguidos en el tiempo para determinar cuales de ellos desarrolla un evento (punto final del estudio) de interés o enfermedad. La probabilidad de desarrollar este evento es la incidencia o riesgo, y puede ser calculado para ambos grupos, los expuestos y los no expuestos. Los estudios de cohortes pueden ser estudios muy poderosos ya que pueden establecer incidencia de una enfermedad, así como relaciones de causa-efecto, pudiendo evidenciar si la causa antecede al efecto.



**Figura 3:** Esquema de un estudio de cohortes prospectivo. Los sujetos ingresan al estudio y son categorizados de acuerdo a la exposición o no de un factor de riesgo determinado. El evento de interés aún no ocurrió al momento de inicio del estudio. Los sujetos son seguidos en el tiempo mientras son observados para determinar la ocurrencia del evento en estudio.

Sin embargo, no pueden “manipular” la “causa” (son estudios observacionales) y por lo tanto no pueden probar que alterando la “causa” se obtenga un efecto diferente.

Los estudios de cohortes son estudios ideales para contestar preguntas de investigación vinculadas a etiología, daño o pronóstico ya que recolectan información de una forma metódica, objetiva y uniforme. Los investigadores pueden establecer criterios de ingreso al estudio, que mediciones se realizarán y cual es la mejor de manera de llevarlas a cabo. De esta forma se eliminan ciertos sesgos (sesgo de memoria) que pueden ser un problema en estudios observacionales. Este diseño es especialmente útil para cuando se estudian eventos de incidencia relativamente frecuente.

La principal limitación de los estudios de cohortes es que suelen muy caros, demandan un tiempo considerable y una logística que debe ser especialmente considerada (la evaluación de la ocurrencia de los eventos debe ser metódica, uniforme y sistemática). Este es un problema importante cuando se estudian enfermedades raras o poco comunes, ya que será muy difícil asegurar un seguimiento de buena calidad (que no haya pérdidas de seguimiento, que las mediciones o controles que se planificaron se cumplan adecuadamente) y un número aceptable de pacientes que permita hallar diferencias clínicas o estadísticamente significativas entre el grupo de pacientes expuestos y no-expuestos al factor de riesgo en estudio. Además, debemos tener en cuenta que como las cohortes son categorizados desde el inicio del estudio en base a la presencia o no de determinado factor de riesgo o exposición, no son buenos estudios para descubrir nuevos factores de riesgo.

En la valoración de estos estudios adquiere radical importancia la interpretación de potenciales *confundidores*. Las variables confundidoras son factores que afectan tanto al factor de riesgo como a su efecto (evento o punto final). Pueden afectar al grupo de sujetos expuestos y no-expuestos de forma diferente, llevando a un sesgo o error en las conclusiones. Por ejemplo, en un estudio se identificó una asociación positiva entre consumo de helados y aumento de la mortalidad por ahogamiento. Un investigador apresurado podría asumir que el consumo de helados es la explicación del aumento de la tasa de muertes por ahogamiento y podría elucubrar una dudosa hipótesis del porque ocurre esta observación. Sin embargo,

es evidente que el consumo de helados posiblemente está asociado a la época del año y la temperatura diaria, siendo este factor el principal determinante del aumento en la cantidad de personas que deciden bañarse en fuentes de agua y por lo tanto, inexorablemente aumenten la probabilidad de ahogamiento. Si bien este ejemplo impresiona grosero, la presencia de los confundidores y su efecto es mucho más sutil en la vida cotidiana y se debe ser muy cuidadoso en la interpretación de los resultados.

Otro aspecto de vital importancia en el desarrollo de estudios de cohortes es la *pérdida de seguimiento*, que puede derivar en una pérdida de información o datos vinculados con la ocurrencia del evento o punto final de interés. Inclusive se debe valorar si la pérdida al seguimiento puede haber sido por condiciones directamente relacionadas con el estudio. En la práctica se tolera, aproximadamente, hasta un 20% de pérdida de seguimiento de sujetos participantes, sin embargo este valor debe tomarse más como una guía que como un valor absoluto. Un porcentaje de pérdida menor de 20% puede introducir un sesgo en la interpretación de los resultados de un estudio si la razón de la pérdida está relacionada con el factor de riesgo o exposición que se está estudiando.

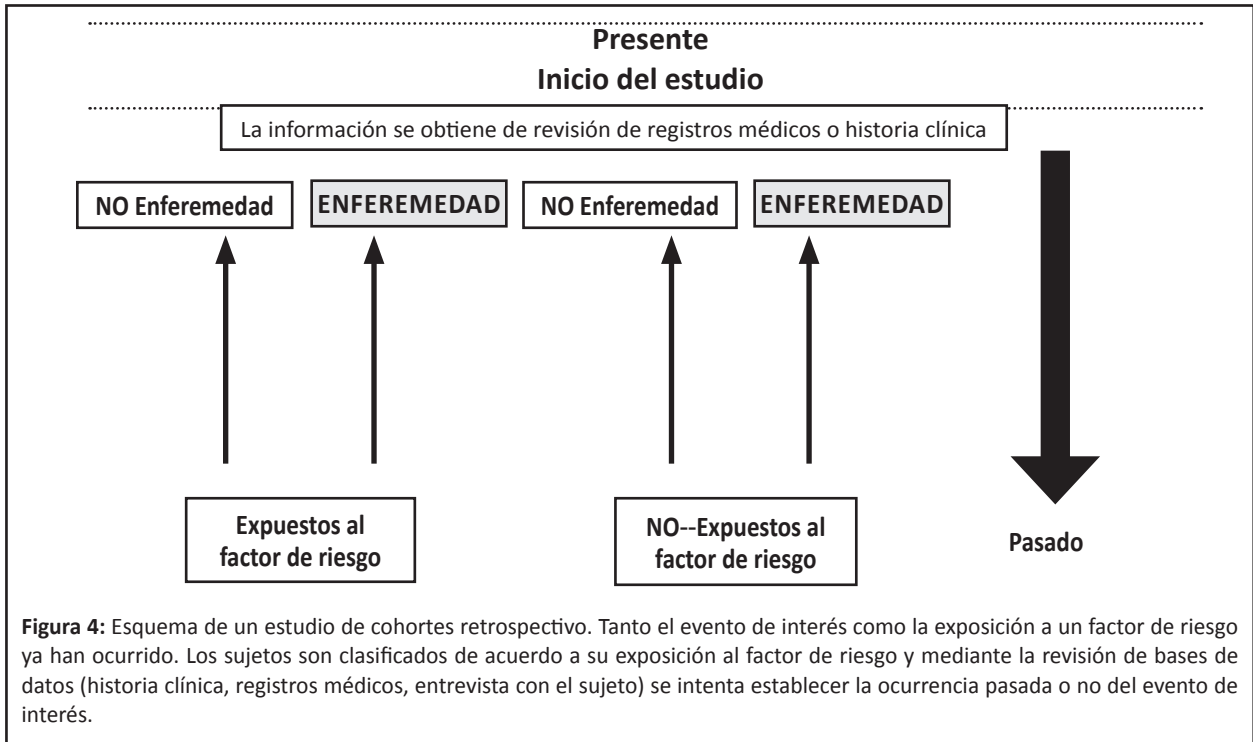
Existe un tipo especial de estudios de cohortes que se denominan en muchas ocasiones como estudios de cohortes retrospectivos (ver **Figura 1** y **Figura 4**). Es esencialmente un estudio que comienza en el presente y utiliza datos de eventos que ocurrieron en el pasado. Los sujetos siguen siendo separados en cohortes de acuerdo a la presencia o ausencia del factor de riesgo o exposición en estudio, a pesar que este no sea la principal razón por la que le paciente ingresó al estudio. Poseen las mismas fortalezas y debilidades que un estudio de cohortes convencional, solo que la calidad de registro de los datos que determinarán si el sujeto de investigación sufrió o no el evento o punto final adquiere un valor mucho más relevante.

Para recordar o tener en cuenta: los sujetos de investigación en los estudios de casos y controles son inicialmente agrupados de acuerdo a la presencia o ausencia del evento o punto final de interés. Estos estudios nunca pueden determinar prevalencia, ni incidencia (es decir no podemos estimar riesgos de ocurrencia de eventos) ya que la relación entre el número de casos y controles está fijada de modo arbitrario por el investigador y no es reflejo de los



que potencialmente podría ocurrir en la realidad. La medida de asociación entre variables cuando se comparan los casos y los controles en estos estudios es el odds ratio. Los sujetos de investigación en los estudios de cohortes son inicialmente agrupados de acuerdo a la presencia o ausencia de determinado

factor de riesgo, independientemente que el grupo haya sido constituido de forma retrospectiva o prospectiva. Estos estudios pueden establecer incidencia y la medida de asociación entre variables cuando se comparan los expuestos versus no-expuesto, por lo tanto, es el riesgo relativo.



### Ensayos clínicos

Un ensayo clínico es un estudio de cohortes en el que un investigador interviene manipulando la presencia o ausencia de un factor de riesgo, usualmente una intervención terapéutica o diagnóstica. Estos estudios son experimentos humanos y usualmente se denominan **estudios de intervención** o **intervencionales**, en contraposición a los estudios de cohortes y de casos y controles en los que no hay intervención intencional. Un ejemplo de un ensayo clínico es un estudio que evalúa el efecto en el riesgo de infarto de miocardio de la dieta mediterránea. Por lo que los sujetos que deseen participar en el estudio serán divididos en dos grupos: un grupo a los que se les administrará dieta mediterránea y un segundo grupo que será alimentado con una dieta tradicional. Ambos grupos serán seguidos en el tiempo para observar la ocurrencia de infarto (ver Figura 5).

Los ensayos clínicos se caracterizan por la presencia de un grupo control que es idéntico al grupo "experimental" o "tratamiento" excepto por la exposición

a la intervención en estudio. Idealmente, un sujeto de investigación participante en estos estudios debería poseer igual posibilidad que sea asignado al grupo control como al grupo intervención para limitar la presencia de posibles sesgos y/o efecto de confundidores. La única forma de asegurarse esta condición es que la asignación a los grupos de los sujetos participantes sea determinada por el azar. Esto se consigue por un proceso que se denomina "aleatorización" o "randomización". Existen diversos métodos de aleatorización que deben ser evaluados al valorar críticamente un ensayo clínico.

A su vez, también es ideal en este tipo de diseño que la asignación a cada uno de los grupos (control y experimental) no sea conocida por los investigadores del estudio, y de ser posible, tampoco por los sujetos participantes. Cuando esta condición ocurre, el estudio se denomina "ciego", y puede ser doble ciego (si tanto investigadores como pacientes desconocen el grupo al que fueron asignados estos últimos) o

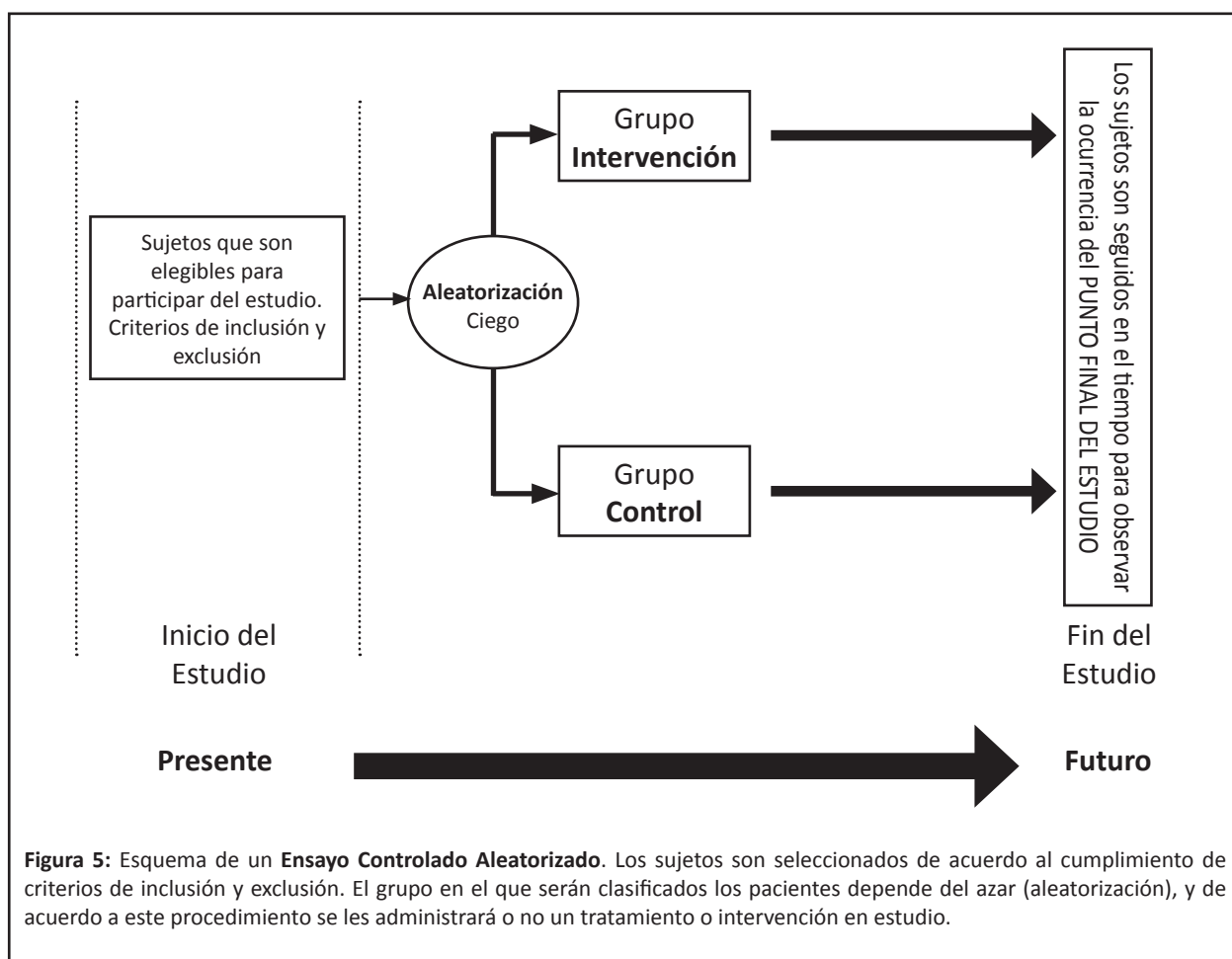
simple ciego (si al menos uno de los dos conoce el grupo de asignación).

Este tipo de diseños es el único diseño que puede contestar en forma inequívoca preguntas sobre causalidad. Pueden demostrar que la causa y el efecto están asociados y que se debe a una razón más allá de la chance o azar, que la causa precede al efecto, y que la modificación de la causa puede resultar en modificaciones del efecto. A su vez es el único diseño que cuando la asignación a los grupos de intervención o control es aleatorizado puede controlar la totalidad de las variables confundidoras que pueden estar presentes. Sin embargo, estos estudios distan mucho de ser perfectos. Suele ser muy caro desarrollarlos y la logística de su implementación puede ser extremadamente dificultosa. Por esta razón, es la industria farmacéutica la que muchas veces es la asume la res-

ponsabilidad de su conducción y financiamiento; lo que puede ser en algunos casos una potencial fuente de sesgo, ya que es parte por demás interesada en los resultados de la investigación.

También en estos estudios puede jugar un importante papel las pérdidas al seguimiento y la potencial baja adherencia a cumplir con los procedimientos del estudio, incluyendo la toma de la medicación. Debemos tener en cuenta que el hecho mismo que se trate de un ensayo clínico no garantizará la validez interna y/o externa de los resultados; como así tampoco, si bien se toman todos los recaudos posibles, que sus resultados sean solo por efecto del azar y no realmente por el efecto de la intervención sobre el evento o punto final.

Algunas características de los diseños que repasamos se resumen en la **Tabla2**.



**Figura 5:** Esquema de un **Ensayo Controlado Aleatorizado**. Los sujetos son seleccionados de acuerdo al cumplimiento de criterios de inclusión y exclusión. El grupo en el que serán clasificados los pacientes depende del azar (aleatorización), y de acuerdo a este procedimiento se les administrará o no un tratamiento o intervención en estudio.



**Tabla 2:** Resumen de características de los diseños de investigación.

Diseño	Ensayo clínico controlado	Estudio de cohortes	Estudio de casos y controles	Estudio de corte transversal
<b>Pregunta</b>	¿Cuál es el efecto de esta intervención?	¿Cuál es el efecto de esta exposición?	¿Cuáles son las causas de este evento?	¿Cuán común es esta condición? ¿existe alguna asociación entre condición y exposición?
<b>Aplicabilidad</b>	Intervención controlada de potencial beneficio	Cualquier exposición para la cual se pueden hallar número de sujetos expuesto puede ser estudiada y se puede medir el evento	Cualquier evento para el cual se puede hallar un grupo de casos y controles apropiados, y los factores de exposición pueden ser valorados retrospectivamente	Cualquier exposición, condición o asociación que es razonablemente común y para lo cual la evaluación en un momento único determinado puede ser suficiente
	Método primario para estudiar nuevas terapias	Método primario para estudiar exposiciones nuevas o raras	Método primario para estudiar eventos o puntos de final nuevos o raros	Método primario para evaluar prevalencia
<b>Fortalezas</b>	Permite aleatorización (mejor forma de controlar confundidores). Permite evaluación doble ciega (principal método para controlar potenciales sesgos).	Permite la evaluación de múltiples puntos finales. Tiempo entre causa y efecto se puede establecer. Todas las mediciones de riesgo se pueden realizar. La exposición es evaluada previo al evento (se reduce sesgo). Puedo calcular incidencia y prevalencia.	Usualmente puede ser realizado con una cantidad moderada de sujetos. Es realizable aún con números de sujetos pequeños. Pueden evaluarse múltiples factores de riesgo o exposición. El método retrospectivo es rápido.	Se pueden obtener muestras representativas de una población. Los métodos pueden ser estandarizados y confiables. Puede ser repetida fácilmente.
<b>Debilidades</b>	Limitaciones éticas, problemas de organización e implementación, muy caros.	Usualmente requiere gran cantidad de pacientes. Se requiere de largos períodos de seguimiento en algunos casos	La naturaleza retrospectiva limita la información sobre la exposición y es susceptible de sesgo. Un grupo control adecuado puede ser difícil de hallar o identificar	Pérdida de la dimensión temporal limita la interpretación de causalidad. Ineficiente para condiciones o exposiciones infrecuentes o raras

**Declaración de conflictos de interés:**

El autor declara no poseer conflictos de interés.

**Bibliografía**

1. Centers for Disease Control (CDC). Pneumocystis pneumonia-Los Angeles. MMWR Morb Mortal Wkly Rep. 1981;30(21):250-2.
2. Centers for Disease Control (CDC). Update on Kaposi's sarcoma and opportunistic infections in previously healthy persons-United States. MMWR Morb Mortal Wkly Rep. 1982;31(22):294, 300-1.
3. von FlieDner VE, Sultan Khan Z, Jeannet M. HLA-A and HLA-B antigens in acute leukemia: A2--B12 phenotypes correlate with longer survival in acute myelogenous leukemia. Acta Haematol. 1981;65(2):73-8.
4. Bottaro FJ. Claves para la interpretación de conceptos estadísticos en estudios de investigación. Hematología. 2013;17(3):299-305.
5. Herbst AL, Ulfelder H, Poskanzer DC. Adenocarcinoma of the vagina. Association of maternal stilbestrol therapy with tumor appearance in young women. N Engl J Med. 1971;284(15):878-81.