

# Estudios de cohorte

ALEJANDRO MARTÍNEZ JUÁREZ<sup>a</sup>

## RESUMEN

Entre los estudios observacionales, aquéllos con mayor fortaleza son los estudios de cohorte. Éstos comprenden diferentes estrategias, en las cuales un grupo de sujetos son evaluados en repetidas ocasiones, a lo largo de un periodo de tiempo elegido por el investigador y determinado, tanto por la prevalencia, como por la historia natural de la enfermedad. A pesar de que sus resultados pueden ensombrecerse por la influencia de variables confusoras y de diferentes tipos de sesgo, representan una estrategia de gran relevancia en la investigación de factores involucrados en la cadena causal de problemas médicos. El presente artículo pretende ser una revisión introductoria para conocer las características principales de los estudios de cohorte y proveer una referencia rápida y breve de las principales fuentes de sesgo en la aplicación médica de este tipo de estudios.

**PALABRAS GUÍA:** *Estudios de cohorte, estudios observacionales, sesgo.*

## INTRODUCCIÓN

El nombre de este grupo de estudios deriva del latín *cohors, cohortis* (séquito o agrupación), que era el término romano para nombrar a un grupo de soldados, comúnmente 500, que marchaban juntos. En estudios médicos, para integrar una cohorte deben reclutarse (no necesariamente en la misma ocasión) a individuos con características similares. Los participantes no deben manifestar el evento de interés (enfermedad, curación, cambio en algún estudio paraclínico, etc.) al inicio del seguimiento; pero deben encontrarse en riesgo de presentarlo a corto, mediano o largo

plazo, de acuerdo con las expectativas del estudio.

Para su inclusión, los individuos son clasificados de acuerdo con las características que podrían guardar relación con el resultado (potenciales factores de riesgo o presunto factor causal), por lo que se va construyendo un grupo expuestos y otro de no expuestos. Posteriormente, cada sujeto se sigue en el tiempo, a través de evaluaciones repetidas, en búsqueda de la ocurrencia del evento de interés. El seguimiento de cada sujeto de estudio se continúa hasta que ocurre una de las siguientes condiciones:

- Se manifiesta el evento de estudio. En general, cuando ocurre esta condición el individuo deja de contribuir a la cohorte; pero si el evento de desenlace puede ocurrir nuevamente en el mismo sujeto, independientemente de haberse presentado previamente (evento recurrente), el sujeto puede reingresar a la cohorte si mantiene las características que lo ponen en riesgo de desarrollar el desenlace en cuestión.

<sup>a</sup> Investigador del Departamento de Genética, Instituto Nacional de Perinatología Isidro Espinosa de los Reyes.

Correspondencia:  
Dr. Alejandro Martínez Juárez

Departamento de Genética, Instituto Nacional de Perinatología Isidro Espinosa de los Reyes. 2do Piso Torre de Investigación.  
Montes Urales No. 800, Col. Lomas de Virreyes, C.P. 11000, México, D.F.

Recibido: 25 de agosto de 2008.  
Aceptado: 7 de octubre de 2008.

- Muerte del sujeto de estudio.
- Se pierde al sujeto de estudio durante el seguimiento.
- El estudio termina.

De acuerdo con su temporalidad, los estudios de cohorte se consideran que son de tipo longitudinal, porque los sujetos son seguidos en el tiempo y se les realizan al menos dos mediciones durante el lapso que dura el estudio.

En relación con el tipo de población, las cohortes pueden ser fijas (también llamadas cerradas) o dinámicas. Las cohortes fijas son aquéllas que no consideran la inclusión de población en el estudio, más allá del periodo de reclutamiento fijado por los investigadores. Las dinámicas son aquellas cohortes que consideran la entrada y salida de nuevos sujetos en el estudio, durante la fase de seguimiento.

De acuerdo con la intervención, los estudios de cohorte se clasifican como observacionales. Si bien son semejantes a los ensayos clínicos aleatorizados (ECA), en el sentido de que los sujetos de estudio se siguen en el curso de la exposición hasta la aparición del evento de estudio, la diferencia principal radica en que en los ECA, el investigador asigna la exposición; mientras que en los estudios de cohorte el investigador sólo observa a los sujetos después de ocurrida la exposición, la cual se da de manera natural.

Con base en su propósito principal, pueden ser de dos tipos: descriptivos, típicamente para describir la ocurrencia y características de ciertos desenlaces a través del tiempo, y analíticos para analizar asociaciones entre variables predictoras y desenlaces de interés.

Dependiendo de la relación temporal del inicio del estudio respecto a la ocurrencia del evento, los estudios de cohorte pueden ser prospectivos o retrospectivos (históricos). Sus características principales se detallan a continuación.<sup>1-4</sup>

## **ESTUDIOS DE COHORTE PROSPECTIVO**

En un estudio de cohorte prospectivo, el investigador inicia ensamblando una muestra de sujetos. El investigador mide características en cada sujeto que podrían predecir los

desenlaces subsecuentes y sigue a estos sujetos con mediciones periódicas de los desenlaces de interés.

Este diseño es útil para evaluar incidencias y para investigar las causas potenciales de alguna condición. La medición de los niveles del factor predictor antes de que ocurra el resultado establece la secuencia temporal de las variables y previene que las mediciones del factor predictor sean influidas por el conocimiento del resultado. El diseño prospectivo también permite al investigador medir las variables de manera más completa y precisa de lo que es posible retrospectivamente. Esto es importante para aquellos factores que son difíciles de recordar de manera precisa por el sujeto de estudio, como los hábitos dietéticos.

Una debilidad particular del diseño prospectivo es su costo e ineficiencia para estudiar desenlaces raros, ya que en estos casos deben seguirse grandes números de personas durante largos períodos de tiempo para observar el número suficiente de desenlaces que produzcan resultados significativos. Por ello, este diseño es más eficiente para el estudio de fenómenos en donde los desenlaces sean más comunes e inmediatos.<sup>1,4-6</sup>

## **ESTUDIOS DE COHORTE RETROSPECTIVOS**

En este diseño el ensamblaje de la cohorte, las mediciones basales y el seguimiento han ocurrido en el pasado. Estos estudios sólo son posibles si existen datos adecuados disponibles que permitan establecer las características de la exposición y los desenlaces de interés en una cohorte de sujetos que ha sido construida para otros propósitos. Se puede decir que los estudios de cohorte retrospectivos reconstruyen la experiencia de la cohorte en el tiempo. Frecuentemente se utilizan en las investigaciones de exposición ocupacional, donde se reconstruye la exposición y frecuencia de eventos en un grupo ocupacional o empresa durante un periodo determinado.

Estos estudios tienen muchas de las fortalezas de los estudios de cohorte prospectivos y tienen la ventaja de ser mucho menos costosos y de consumir menos tiempo, debido a que los sujetos ya están captados, las



mediciones ya se han hecho y el periodo de seguimiento ya ha ocurrido.

Las desventajas principales son el control limitado que el investigador tiene sobre el método de muestreo de la población y sobre la naturaleza y la calidad de las variables predictoras. La validez del estudio dependerá en gran medida de la calidad de los registros utilizados, debido a que los datos existentes pueden ser incompletos, imprecisos o medidos de manera que no son ideales para responder la pregunta de investigación. Además, cuando se estudian de manera retrospectiva enfermedades fatales, las variables predictoras del sujeto fallecido sólo pueden ser reconstruidas a partir de fuentes indirectas, tales como los expedientes médicos o el interrogatorio a amigos y familiares.<sup>1-3,5</sup>

### **ESTUDIO DE CASOS Y CONTROLES ANIDADO EN UNA COHORTE Y ESTUDIO DE CASOS-COHORTE**

Un diseño de casos y controles “anidado” en una cohorte permite resolver preguntas de investigación con la metodología de un estudio de casos y controles, pero a partir de una cohorte formada con otro fin, siempre y cuando las variables de interés puedan evaluarse al final del seguimiento de la cohorte. El investigador inicia con una cohorte disponible con suficientes casos al final del seguimiento. Posteriormente, se identifican todos aquellos sujetos que han desarrollado el desenlace (los casos). Luego, se selecciona una muestra aleatoria de los sujetos que no desarrollaron el desenlace (los controles); se puede incrementar el poder del estudio seleccionando dos o tres controles por cada caso y pareando por características que pudieran influir en el desenlace, tales como edad y sexo. Posteriormente se recuperan las muestras, imágenes o registros (expedientes) que fueron colectados antes de que ocurrieran los desenlaces, se miden las variables predictoras y se comparan sus niveles en casos y controles.

El estudio de casos-cohorte tiene el mismo diseño excepto que los controles son una muestra aleatoria de todos los miembros de la cohorte sin importar los desenlaces. Esto significa que

habrá algunos casos entre la muestra seleccionada como grupo de comparación, quienes también aparecerán entre los casos y serán analizados como tales. Este diseño tiene la ventaja de que los controles representan la cohorte en general y, por lo tanto, proporcionan una base para la estimación de incidencias y prevalencias en la población de la cual fueron obtenidos. Más importante, significa que esta muestra de la cohorte puede ser utilizada como el grupo de comparación para más de un tipo de desenlace.

Ambos diseños son especialmente útiles para estudiar variables predictoras cuya medición es costosa, como la cuantificación de algunos metabolitos. En estos diseños se almacenan materiales de algunos o todos los individuos de la cohorte (sangre, suero, leucocitos, biopsia, ADN, etc.) desde el inicio del estudio y se preservan para su análisis posterior, sólo en caso de presentar el evento de interés o ser seleccionado como sujeto control.

Además de los ahorros en costos al no realizar las mediciones en la cohorte completa, estos diseños preservan todas las ventajas de los estudios de cohorte que resultan de colectar las variables predictoras antes que hayan ocurrido los desenlaces y evitan los sesgos potenciales de los estudios de casos y controles convencionales, que extraen a los casos y controles de diferentes poblaciones y no pueden hacer mediciones en los sujetos que han muerto.

Su principal desventaja es que muchas preguntas de investigación no son compatibles con la estrategia de almacenar materiales de una muestra de sujetos de estudio para su análisis posterior. También, cuando los datos están disponibles para la cohorte entera sin costo adicional, no se obtiene ningún beneficio si se estudia una muestra de controles, sino que se debe usar la cohorte entera.

Éstos son diseños tan buenos que un investigador que planea un estudio prospectivo debería considerar siempre la posibilidad de preservar muestras biológicas y almacenar imágenes o registros que implican mediciones caras para su análisis subsiguiente con diseños de casos y controles anidados o de casos-cohorte. Se deben asegurar que las condiciones de

almacenamiento preservarán las sustancias de interés durante muchos años y considerar separar especímenes para mediciones periódicas para confirmar que los componentes permanecen estables. También resulta útil colectar nuevas muestras de información durante el periodo de seguimiento que puedan usarse en las comparaciones por casos-controles.<sup>1,6,7</sup>

## **SELECCIÓN DE LA COHORTE**

La población usada para integrar una cohorte varía según la exposición que será objeto de estudio. En caso de exposiciones comunes (como la dieta o los estilos de vida), la cohorte puede conformarse con muestras de la población general, mientras que si se trata de exposiciones raras o muy particulares es posible que se necesite formarla a partir de grupos ocupacionales específicos, como pueden ser los trabajadores de una determinada fábrica o sector industrial o grupos profesionales determinados como médicos, enfermeras o de profesionales de la educación.

Antes de que pueda identificarse a las personas expuestas es necesario definir claramente las características de la exposición, tales como el nivel y duración mínimas para considerar a un sujeto como expuesto. Adicionalmente, pueden fijarse criterios de selección, pero lo más importante es que los sujetos estén libres del evento de estudio al inicio de su seguimiento. La decisión de incluir o excluir a determinados sujetos de la población de estudio dependerá de la exposición y del resultado que interesa, así como de la factibilidad de realizar las mediciones y de la representatividad deseada. En algunos casos es necesario contar con más de un grupo de exposición.

Los sujetos no-expuestos deben ser similares a los expuestos en todos sus aspectos, excepto en que no han estado sometidos a la exposición que se estudia. El proceso de captación de los sujetos no-expuestos debe ser el mismo que se aplicó a los sujetos expuestos. Deben tener, además, el mismo riesgo potencial de presentar el evento de estudio, y tener las mismas oportunidades que los expuestos de ser diagnosticados con el desenlace en estudio.<sup>3,4</sup>

## **MEDICIÓN DEL DESENLACE**

Los eventos de estudio pueden ser:

- Evento simple, fijo en el tiempo (muerte o incidencia de enfermedad). En este caso, al observar el evento en cada unidad de análisis el seguimiento termina.
- Eventos múltiples (enfermedades recurrentes, sintomatología o eventos fisiológicos). Al presentar el evento el individuo deja de estar en riesgo por lo que ya no cumple con el criterio de permanencia en la cohorte. Se puede reiniciar el seguimiento cuando se restablece el riesgo, es decir, cuando hay curación y el individuo vuelve a estar en riesgo de presentar el evento.
- Modificación de medida eje (modificación de la función en el tiempo hacia un aumento o disminución; crecimiento), que son evaluados mediante tasa de cambio.
- Marcadores intermedios del evento (cuenta de metabolitos como marcadores de predisposición a enfermedad).<sup>2,3</sup>

## **SEGUIMIENTO**

El periodo de seguimiento puede abarcar años, meses, semanas o días, dependiendo de la frecuencia del evento estudiado. Dos momentos definen el periodo de seguimiento: el examen inicial (medición basal) y el final del seguimiento. El inicio del seguimiento depende del tipo de cohorte: si es cerrada o dinámica, ya que en el caso de esta última, el inicio del seguimiento se define para cada participante a través de un largo periodo de tiempo. El seguimiento, dependiendo del evento de interés, puede ser activo o pasivo. Activo es aquél en el que se utilizan contactos repetidos por diversos medios: nueva entrevista y obtención de muestras, cuestionarios o llamadas telefónicas. El seguimiento pasivo es el que se realiza mediante búsqueda sistemática en registros prestablecidos (registros de cáncer, hospitalarios, registro civil, entre otros).

La exposición y las covariables en estudio, así como los factores confusores y los modificadores de efecto pueden variar en el tiempo, por lo que es necesario considerar los factores por los cuales ocurren estas variaciones y su impacto en el

desenlace de interés para definir un mejor periodo de seguimiento. En las exposiciones fijas los factores no cambian a través del tiempo, tal es el caso del sexo y el lugar de nacimiento, entre otros.<sup>2-4,6</sup>

## PÉRDIDAS EN EL SEGUIMIENTO

Las pérdidas en el seguimiento pueden originarse principalmente por tres razones:

- Abandono del estudio.
- Muerte por otra causa al evento de interés.
- Pérdidas, “administrativas”, originadas por la terminación temprana del estudio por razones ajenas a las que se plantearon originalmente (ejemplo: agotamiento de la fuente de financiamiento).

Cuantificar las causas que producen pérdidas en el seguimiento es importante para evaluar la validez del estudio.<sup>2,3,6</sup>

## SESGO Y VALIDEZ EN ESTUDIOS DE COHORTE

A pesar de tener menor riesgo de presentar sesgos, los estudios de cohorte también pueden presentarlos de manera importante. Además, los estudios de cohorte retrospectiva son más vulnerables a los sesgos de selección, esto se debe a que al inicio de este tipo de estudios el evento ya ocurrió en un buen número de los participantes y a que esto puede influir en la probabilidad de participación en el estudio. Este tipo de sesgo en estudios retrospectivos es particularmente serio cuando los participantes conocen también su condición de exposición y cuando la presencia conjunta de estos eventos (exposición y enfermedad) motiva una participación diferencial en el estudio. A continuación se mencionan los sesgos más importantes a considerar en los estudios de cohorte.<sup>3,5,8-10</sup>

### Sesgos de selección

- **Sesgo de muestreo.** Ocurre cuando la muestra que integra la cohorte no representa de manera adecuada el espectro de características de la población blanco. Por ejemplo, cuando

la población en estudio se constituye con voluntarios. Para constituir una cohorte representativa de una determinada población se sugiere realizar el seguimiento de muestras representativas de la población general.

- **Sesgo por pérdidas durante el seguimiento.** Ocurre cuando los sujetos que abandonan el estudio durante el seguimiento difieren de los que permanecen en él. Ésta es la principal causa de sesgo de selección en los estudios de cohorte. Las pérdidas en el seguimiento no invalidan *per se* el estudio, pero los investigadores deben utilizar procedimientos para minimizar su ocurrencia y, en caso de que se presenten, considerar si afectan o no los resultados observados en el estudio. Para ello, se debe registrar información clave de los participantes que abandonen el estudio, en especial para investigar si el abandono tiene alguna relación con las exposiciones o con las enfermedades o eventos estudiados. Uno de los motivos más frecuentes de pérdidas en el seguimiento obedece a la movilidad de los participantes de la cohorte, que dificulta los contactos repetidos que se utilizan, tanto para realizar nuevas determinaciones de las exposiciones, como para conocer el desarrollo de las condiciones estudiadas (incidencia de enfermedades o muerte). Para mantener y maximizar la participación y minimizar las pérdidas en el seguimiento, es necesario implementar diferentes estrategias durante el desarrollo del estudio.
- **Efecto del trabajador sano.** Se observa frecuentemente en cohortes laborales cuando el grupo no expuesto queda constituido por la población general. La cohorte de trabajadores representa un grupo especial de personas de la población general: son personas en edad laboral y con capacidad física óptima, existe una autoselección de las personas que van a desarrollar tareas concretas en una empresa o área de trabajo en particular (fábrica, oficina, laboratorio, etc.), de manera que personas con determinados trastornos crónicos se autoexcluyen (o son excluidos por la dirección de la empresa) o pasan a trabajos menos pesados. Una manera de evitar este tipo de sesgo de selección es realizar comparaciones

internas en el seno de la cohorte, por ejemplo, entre los trabajadores de una misma empresa, pero con diferentes niveles de la exposición por laborar en diferentes áreas u horarios.

### Sesgos de información

- **Sesgo de seguimiento.** En estudios clínicos de seguimiento es frecuente que algunos participantes presenten condiciones comórbidas que generen, incluso de manera no apreciable para los investigadores, una mayor vigilancia o control de esos pacientes con relación a otros sujetos del estudio, aumentando artificialmente las posibilidades de diagnóstico de la condición de interés.
- **Sesgo del observador.** Ocurre cuando el propio investigador evalúa de forma diferencial la condición de interés, puesto que conoce las hipótesis bajo investigación o la historia de exposiciones de los participantes. Este sesgo se puede evitar manteniendo ciegos, tanto a los participantes, como a los observadores sobre la condición de exposición y la hipótesis de estudio.
- **Sesgo de clasificación no diferencial.** Se presenta cuando ocurre clasificación errónea de los participantes respecto a la existencia o a la cuantificación de la exposición estudiada o a la ocurrencia de la enfermedad o de la condición de interés.
- **Sesgo de medición.** Se produce por deficiencias en los instrumentos utilizados y en su modo de aplicación (cuestionarios, técnicas analíticas, biomarcadores, etc.) que pueden generar sub o sobreestimación de las mediciones de interés. Se clasifica en dos tipos: a) diferencial, cuando el error en la clasificación depende del valor de otras vari-

ables, por ejemplo del nivel socioeconómico, y b) no-diferencial, cuando el error no depende de las otras variables y el error afecta por igual a todos los participantes, independientemente de su exposición verdadera, por ejemplo un cuestionario que subestima una medición pero aplicado de la misma manera a todos los encuestados. Las posibles soluciones para minimizar los errores de medición aleatorios consisten en la validación de los instrumentos de medida utilizados, ya sean cuestionarios estructurados, pruebas psicométricas, instrumentos médicos (esfigmomanómetros, balanzas, y otros), técnicas de laboratorio, etcétera, junto con la implantación de protocolos de aplicación estrictos, previo entrenamiento y estandarización, sobre todo cuando los observadores son múltiples e, incluso, de diferentes centros y la realización de medidas repetidas en los mismos sujetos. Finalmente, la realización de una prueba piloto en condiciones reales y la implementación de controles de calidad continuados de la información recolectada pueden alertar sobre la aparición de este tipo de sesgos que difícilmente son controlables en las fases de análisis del estudio.

### CONCLUSIÓN

En general, los estudios de cohorte tienen amplias ventajas respecto a otro tipo de estudios epidemiológicos observacionales; sin embargo, sus resultados pueden ensombrecerse por las influencias de variables confusoras y de diferentes tipos de sesgo, por lo que su principal desventaja en comparación con los estudios de intervención es una menor potencia para demostrar una inferencia causal.

## ABSTRACT

Among observational studies, those with the best strength are cohort studies. These include different strategies which evaluate a group of subjects in several occasions through a established period of time chosen by the researcher and specified by the prevalence as well as the natural history of the disease. Even when their results can be shadowed by the influence of both confusing parameters and different kind of bias, they represent a strategy of great relevance in finding about the factors involved in the causal chain of medical problems. This article is meant to be an introductory brush up to know the main characteristics of cohort studies and to provide a quick and brief reference of the main sources of bias in the medical application of these kinds of studies.

**KEY WORDS:** *Cohort studies, observational studies, bias.*

## REFERENCIAS

1. Cummings SR, Newman TB, Hulley SB. Designing an Observational Study: Cohort Studies. In: Hulley SB, Cummings SR (eds.). Designing Clinical Research: An Epidemiologic Approach. 30th Ed. Lippincott Williams & Wilkins; 2006.
2. Hennekens CH. Cohort Studies. In: Hennekens CH, Buring JE (eds.). Epidemiology in Medicine. 1st Ed. USA: Little Brown and Company; 1987, p. 153-177.
3. Lazcano-Ponce E, Fernández E, Salazar-Martínez E, Hernández-Ávila M. Estudios de cohorte. Metodología, sesgos y aplicación. Salud Pública Mex 2000; 42(3): 230-41.
4. Thadhani R, Tonelli M. Cohort studies: marching forward. Clin J Am Soc Nephrol 2006; 1(5): 1117-23.
5. Mann CJ. Observational research methods. Research design 11: cohort, cross sectional and case-control studies. Emerg Med J 2003; 20(1): 54-60.
6. Gladman DO, Farewell VT. Longitudinal cohort studies. J Rheumatol Suppl 2005; 72: 30-2.
7. Sato T. Risk ratio estimation in case-cohort studies. Environ Health Perspect 1994; 102(Suppl. 8): 53-6.
8. Rochon PA, Gurwitz JH, Sykora K, Mamdani M, Streiner DL, Garfinkel S, et al. Reader's guide to critical appraisal of cohort studies: 1. Role and design. BMJ 2005; 330(7496): 895-7.
9. Pennefather PM, Tin W, Clarke MP, Dutton J, Fritz S, Hey EN. Bias due to incomplete follow up in a cohort study. Br J Ophthalmol 1999; 83(6): 643-5.
10. Sica GT. Bias in research studies. Radiology 2006; 238(3): 780-9.