

Conceptos generales en bioestadística y epidemiología clínica: estudios observacionales con diseño de cohorte

General concepts in biostatistics and clinical epidemiology: Observational studies observational studies with cohort design

Gabriel Lazcano^a, Cristian Papuzinski^{a,b}, Eva Madrid^{a,b}, Marcelo Arancibia^{a,b,*}

^a Cátedra de Metodología de la Investigación Científica, Escuela de Medicina, Universidad de Valparaíso, Viña del Mar, Chile

^b Centro Interdisciplinario de Estudios en Salud (CIESAL), Universidad de Valparaíso, Valparaíso, Chile

*Autor corresponsal marcelo.arancibiame@uv.cl

Citación Lazcano G, Papuzinski C, Madrid E, Arancibia M. General concepts in biostatistics and clinical epidemiology: observational studies with cohort design. *Medwave* 2019;19(11):e7748

Doi 10.5867/medwave.2019.11.7748

Fecha de envío 19/10/2019

Fecha de aceptación 29/11/2019

Fecha de publicación 16/12/2019

Origen Este artículo es parte de una colección de "Notas metodológicas" acordada entre Medwave y la Cátedra de Metodología de la Investigación Científica de la Escuela de Medicina de la Universidad de Valparaíso.

Tipo de revisión Con revisión por pares externa con tres revisores, a doble ciego

Palabras clave observational study, cohort studies, relative risk, incidence, bias, epidemiology, biostatistics

Resumen

Los estudios con diseño de cohorte evalúan la relación entre una exposición y la ocurrencia o no de un evento de interés, comenzando el análisis desde la exposición. Habiendo sido diseños muy utilizados en algunas áreas de la medicina, como la descripción de factores de riesgo cardiovascular o los efectos de la radiación ionizante en humanos, representan una herramienta con características atractivas debido a su adaptabilidad a numerosos contextos, sobre todo en el estudio de exposiciones de baja ocurrencia. Esta revisión es la cuarta entrega de una serie metodológica sobre conceptos generales en bioestadística y epidemiología clínica desarrollada por la Cátedra de Metodología de la Investigación Científica de la Escuela de Medicina de la Universidad de Valparaíso, Chile. En este artículo, se abordan conceptos teóricos generales sobre los estudios de cohorte, considerando aspectos históricos, generalidades sobre la construcción de un estudio utilizando este diseño, presentando distintas variantes y diseños derivados de interés, y potenciales sesgos a los que se puede ver enfrentado el investigador.

Ideas claves

- Los estudios de cohorte analizan la relación entre una exposición y un desenlace prospectivamente, por lo que podrían estimar directamente el riesgo y establecer causalidad.
- Se realizan especialmente durante exposiciones de baja ocurrencia.
- El desarrollo de un estudio de cohorte puede ser costoso y difícilmente reproducible. Son sensibles a los sesgos de confusión y selección, pero existen métodos para evitarlos.

Introducción

El término cohorte tiene su antecedente histórico en las *cohors* romanas, subunidades que componían una legión, cuya creación se suele atribuir a la reforma impulsada por el Cayo Mario en el siglo II antes de Cristo, misma que permitiría acelerar la transición desde una milicia a un ejército profesional¹. Epidemiológicamente, podemos encontrar un antecedente primigenio en las tablas de vida, instrumentos desarrollados en el siglo XVII que buscaban documentar causas de muerte y estimar mortalidad; la primera fue publicada en 1662 por John Graunt, haciendo uso de la amplia documentación disponible por las recurrentes epidemias, formato que sería refinado por Edmond Halley en 1693². Ya a finales del siglo XIX, el surgimiento de la industria de seguros moderna y la necesidad de proyectar el riesgo de sus asegurados, hizo surgir diseños de seguimiento equivalentes a las cohortes prospectivas actuales, a razón de describir tanto la historia natural como los efectos de diversas intervenciones en pacientes estadounidenses con tuberculosis a comienzos del siglo XX³⁻⁵. Posterior a la Segunda Guerra Mundial, estudios emblemáticos por lo vasto y rico de sus datos, moldearon tanto a la cohorte contemporánea en sí, como al entendimiento básico del concepto de “riesgo”⁶: el Estudio Framingham sobre riesgo cardiovascular⁶, aún en desarrollo; el *British Doctor’s Study*, que entregó evidencia importante respecto al riesgo de desarrollar cáncer pulmonar producto del consumo de tabaco⁷; el seguimiento de sobrevivientes japoneses a ambas bombas atómicas⁸ y la cohorte de mineros de uranio de la meseta de Colorado⁹, que en conjunto proveen gran parte del conocimiento actual sobre los efectos de la radiación en seres humanos¹⁰.

Sea en su origen militar –considerando experiencia y funciones dentro de la legión– o en su concepción epidemiológica actual –agrupando, por ejemplo, según la exposición a ciertas noxas o intervenciones médicas–, es posible ejemplificar el concepto prototípico de las cohortes: individuos agrupados según sus características hacia un destino común, concepto a desarrollar en esta revisión.

Este artículo corresponde a la cuarta entrega de una serie metodológica de seis revisiones narrativas acerca de tópicos generales en bioestadística y epidemiología clínica, las que explorarán artículos publicados disponibles en las principales bases de datos y textos de consulta especializados. La serie está orientada a la formación de estudiantes de pre y posgrado, y es realizada por la Cátedra de Metodología de la Investigación Científica de la Escuela de Medicina de la Universidad de Valparaíso, Chile. Por lo tanto, el objetivo de este manuscrito es abordar los principales conceptos teóricos y prácticos de los estudios de cohorte.

Conceptos preliminares

El estudio de cohorte corresponde a un diseño observacional, longitudinal y analítico: a partir de una exposición que ocurre “naturalmente” (sin intervención de los investigadores) se realizará un seguimiento en el tiempo, permitiéndonos evaluar las diferencias (o ausencia de éstas) en la ocurrencia de un evento o desenlace utilizando pruebas de hipótesis estadística^{11,12}.

En comparación a los estudios de casos y controles, metodología observacional abordada en la entrega anterior de esta serie, que poseen

un desenlace ya conocido para el cual se estudiarán sus probables exposiciones, en las cohortes son las exposiciones las que separan a los grupos hacia sus destinos potencialmente distintos. De lo anterior se extrae otra diferencia fundamental: la direccionalidad, ya que los estudios de casos y controles analizan sus datos desde el efecto hacia la causa, mientras que la cohorte opera en forma prospectiva: desde la causa hacia el efecto¹³.

La recolección de datos en una cohorte puede ocurrir tanto de manera prospectiva, como retrospectiva y ambidireccional/ambispectiva, lo que no afecta la direccionalidad analítica ya descrita. En la primera, se inicia el seguimiento con una exposición, esperando a la ocurrencia o no de un desenlace en el futuro; en la variedad retrospectiva, tanto exposición como desenlace ya han ocurrido desde el punto de vista del investigador. La variedad ambidireccional corresponde a un caso intermedio, en el cual se estudia una exposición ocurrida en el pasado y un desenlace aún por suceder; esta modalidad es útil en la evaluación de eventos que demoran en ocurrir, o exposiciones que pudieran desencadenar múltiples desenlaces de interés^{14,15} (Ejemplo 1).

Ejemplo 1. Un estudio con diseño de cohorte ambispectiva¹⁶, desarrollado en China, buscó estudiar los efectos adversos de la terapia utilizada en la tuberculosis resistente a múltiples fármacos. En la fracción retrospectiva, se enrolaron los registros de 751 pacientes que recibieron tratamiento entre mayo de 2009 y julio de 2013, realizándose en forma prospectiva una evaluación de seguimiento en diciembre de 2016.

Las cohortes contemplan un elemento de causalidad o relación temporal de causa y efecto ausente tanto en los estudios transversales como en los casos y controles^{14,17}. Así, resulta como requisito ineludible que tanto sujetos expuestos como no expuestos, al inicio del seguimiento, no posean el desenlace de interés. En el caso de una patología, todos serán “sanos” para dicho cuadro.

Resulta inevitable resaltar cómo dicha temporalidad va en directa concordancia con el criterio homónimo planteado por Sir Austin Bradford Hill. Estos nueve conceptos: fuerza de asociación, consistencia, especificidad, temporalidad, gradiente biológico, plausibilidad, coherencia, evidencia experimental y analogía, si bien no deben ser pretendidos como una lista de verificación rígida, ofrecen un marco de trabajo todavía vigente en la investigación epidemiológica, y son adaptables a un mundo altamente tecnificado en que la biología molecular y la genómica ofrecen explicaciones mecanicistas que complementan y fortalecen los hallazgos encontrados a nivel epidemiológico¹⁸. De dichos criterios, la temporalidad es el único considerado estrictamente esencial para hablar de causalidad entre una exposición y un desenlace¹⁹.

Cabe enfatizar un concepto central: del mismo modo como los estudios de casos y controles son el mejor diseño para evaluar desenlaces de muy baja ocurrencia, los estudios de cohorte constituyen el mejor diseño para evaluar exposiciones de baja ocurrencia, tales como desastres ambientales.

Construcción de un estudio de cohorte

Las preguntas esenciales que deben responderse a la hora de diseñar un estudio de cohorte son^{12,20-22}: 1. ¿Quién está en riesgo?; 2. ¿Quién debe considerarse expuesto?; 3. ¿Quién constituye un control (no expuesto) apropiado?; 4. ¿Cómo se objetivará la ocurrencia de un evento? Y 5. ¿Cómo se realizará el seguimiento de los participantes?

La primera pregunta hace referencia a la susceptibilidad de los participantes de presentar el evento. Si bien contaremos con individuos expuestos y no expuestos, y ninguno de ellos debe poseer el evento de interés al comienzo del seguimiento, todos deben ser teóricamente capaces de desarrollarlo. Esto va de la mano con la tercera interrogante, en cuanto los individuos no expuestos deben presentar características similares (edad, sexo, etnia, nivel socioeconómico, entre otras) a los expuestos, a fin de reducir potenciales sesgos. El origen del grupo no expuesto puede ser interno (mismo lugar y tiempo en relación a los expuestos) o externo (un hospital distinto o estadísticas nacionales disponibles); generalmente, es deseable un origen interno¹².

La segunda y cuarta interrogante hacen referencia a una característica requerida en cualquier estudio, independientemente de su diseño: un protocolo robusto ha de contener definiciones precisas sobre lo estudiado. Esto implica, en el sentido práctico, el uso de criterios clínicos actualizados y claramente establecidos, así como explicitar todo instrumento, escala, puntaje y valores discriminatorios que fueran a usarse. Además, deben considerarse las particularidades propias de la población en seguimiento, si es que pudieran afectar la medición del evento de interés^{11-13,23}.

Con respecto al seguimiento, el contar con métodos robustos, adecuados y aceptables para la población en estudio es de extrema importancia, debido a que las pérdidas de seguimiento son una potencial fuente de sesgo al desbalancear los grupos de la cohorte y censurar la ocurrencia de posibles eventos. Por tanto, algunas características para tomar en cuenta son^{20-22,24} (Ejemplo 2):

- Forma de realización del seguimiento: presencial o a distancia.
- Medio: llamada telefónica, encuestas electrónicas o en papel, entrevistas presenciales.
- Extensión y periodicidad de cada control: potencial desinterés en responder encuestas muy largas o de asistir muy seguido a un centro de salud.
- Necesidad de toma de muestras o realización de mediciones: considerar disponibilidad logística de los centros adscritos al estudio.
- Mecanismos de redundancia: formas alternativas de comunicarse con el paciente en caso de que los medios usuales de contacto cambien o no estén disponibles.
- Costo: asegurar sustentabilidad del estudio para el tiempo de seguimiento planeado.

Ejemplo 2. El estudio UK COSMOS²², desarrollándose actualmente en Reino Unido, busca evaluar los efectos de la telefonía móvil sobre la salud de sus usuarios. Los autores, enfrentados a una baja tasa de respuesta en las primeras fases de reclutamiento, deciden dar énfasis al uso de medios digitales, como cuestionarios vía web y recordatorios mediante mensajes de texto. Si bien a través de estas medidas logran mejorar la captación de personas, refieren verse enfrentados a nuevos desafíos propios de las tecnologías escogidas, como el que los sitios web tuvieran un diseño atractivo y optimizado para su visualización en teléfonos móviles y tablets.

Tipos de diseños de estudio de cohorte: ventajas, desventajas y potenciales usos

Considerados los elementos propios a la formulación de un estudio de cohorte, vale la pena plantear una nueva pregunta: ¿Es un estudio de cohorte el diseño correcto para evaluar la hipótesis de investigación? o, en forma más amplia, ¿cuál es el diseño más apropiado para responder a mi pregunta de investigación? Desde ese punto de vista, los estudios de cohorte gozan de una interesante capacidad de adaptación en base a sus modalidades clásicas –según la temporalidad de la recogida de datos–, contando además con diseños derivados capaces de responder a escenarios específicos.

Recapitulando algunos elementos ya señalados, son ventajas propias del diseño de cohorte su direccionalidad de análisis prospectiva (desde exposición hacia desenlace), temporalidad que permite inferir, dado un estudio correctamente diseñado, una relación de causa-efecto entre exposición y desenlace, de encontrarse una asociación estadística significativa. Con ello, será posible la estimación de la incidencia y de los riesgos absolutos y relativos^{12,23,25}.

Estudio de cohorte prospectiva

Este diseño tiene innumerables ventajas, ya que, una vez construida la muestra de sujetos expuestos y no expuestos, se realizan mediciones prospectivas de un evento que está por suceder: el desenlace es un evento futuro, cuya ocurrencia es desconocida al inicio del estudio. Una fortaleza de esta modalidad es la medición directa de la exposición, permitiendo establecer una temporalidad clara respecto a la ocurrencia del desenlace y de la evolución de la enfermedad o condición, y asegurando la medición y la calidad de la recolección de datos, que no resulta registro-dependiente. Además, se impide que el conocimiento del desenlace afecte el registro de las exposiciones. Constituye el mejor diseño para evaluar exposiciones de rara ocurrencia, especialmente cuando se refieren a daño (exposiciones de tipo ambiental, desastres o efectos por uso), considerando que el único otro diseño que puede aportar datos de causalidad son los ensayos clínicos, que en estos casos no sería ético aplicar. De la misma manera, las cohortes prospectivas permiten describir y estudiar la historia natural de una enfermedad.

Una de las principales desventajas de este diseño se relaciona con su relativo alto costo económico, lo que lo vuelve ineficiente para evaluar eventos de rara ocurrencia y alejados en el tiempo. Esto implica además que su reproducibilidad en estudios futuros sea muy dificultosa. Adicionalmente, no se libra de la debilidad común a todos los

diseños observacionales, en comparación a sus pares experimentales, en cuanto al efecto que pueden tener las variables de confusión en la distorsión de los resultados^{11,17,23}.

Estudio de cohorte retrospectiva

En esta variedad, tanto exposición como desenlace ya han ocurrido, por lo que debe recurrirse a registros ya existentes para reconstruir la secuencia de eventos: ambos corresponden a eventos pasados, conocidos desde el inicio del estudio. Además de las fortalezas propias del diseño de cohorte, resultan más baratos y ágiles de realizar que las versiones prospectivas. Sin embargo, dependen de que los datos consultados estén registrados en forma prolija (o siquiera existan), habiendo limitado control sobre la construcción de la muestra y la calidad del seguimiento. Va a ser la modalidad de elección para evaluar eventos raros, y son una alternativa atractiva en ambientes hospitalarios o institucionales, en que es factible encontrar registros de buena calidad²³.

Estudio de cohorte ambispectiva o ambidireccional

Esta tercera variante implica recolectar datos preexistentes, documentando una exposición (e incluso eventos) ya ocurrida, para continuar el seguimiento en forma prospectiva. Como modelo híbrido, posee ventajas y desventajas de los dos diseños ya explicados¹⁵. Resulta particularmente útil al analizar exposiciones que pueden tener desenlaces a corto y largo plazo, o más de un desenlace¹². Otro escenario que considerar es la ocurrencia de crisis sanitarias, como son brotes o epidemias, eventos que por su naturaleza implican una exposición inesperada, que inicialmente debe ser recogida retrospectivamente. También, en estricto rigor, los estudios de hábitos de estilo de vida pueden catalogarse como diseños ambispectivos, en cuanto la exposición –probablemente aún en desarrollo– ya se encuentra presente al inicio del seguimiento²⁶.

Estudio de cohorte de un grupo

Corresponde a una conceptualización que considera que los elementos esenciales de una cohorte son: muestreo por exposición, seguimiento en el tiempo, y cálculo de riesgo absoluto, no siendo condición *sine qua non* la existencia de un grupo de control no expuesto²⁵. Bajo definiciones clásicas, esto correspondería a una serie de casos²⁷, no obstante, aquella metodología posee un muestreo basado en el desenlace, pudiendo considerar o no una exposición específica en conjunto, sin mediar seguimiento ni siendo posible el cálculo de riesgo absoluto²⁵; además, las series de caso son atemporales, es decir, pueden reunir distintos casos ocurridos en temporalidades distintas. La cohorte de un grupo puede resultar útil para evaluar intervenciones nóveles en muestras reducidas. La inexistencia de un grupo no expuesto impide realizar tanto pruebas de hipótesis como calcular riesgos relativos, pero si es posible registrar la incidencia de eventos de interés, como podrían ser efectos adversos (Ejemplo 3).

Ejemplo 3. Un estudio de diseño de cohorte de un grupo²⁸, realizado en Italia, documentó las tasas de éxito y complicaciones ocurridas al utilizar una técnica sin colgajo para la instalación de implantes de carga inmediata en pacientes con maxilar edéntulo. En el estudio fueron reclutados 33 pacientes, los que fueron controlados al momento de la instalación y tras 12 meses de seguimiento, evaluándose la estabilidad de los implantes.

Estudio de cohorte múltiple

La idea anterior puede extenderse a estudios en que distintas muestras están sometidas a exposiciones diferentes, o a distintos niveles de una misma exposición (pudiendo considerar un grupo no expuesto)^{23,29}. Al haber más de un grupo en seguimiento, se posibilita el cálculo de riesgo relativo y la realización de pruebas de asociación estadística (Ejemplo 4).

Ejemplo 4. Un estudio de doble cohorte³⁰ evaluó la técnica ideal para el manejo quirúrgico mínimamente invasivo de pacientes con prolapso rectal. Se plantean dos cohortes, cada una con una técnica distinta: rectopexia ventral laparoscópica, llevada a cabo en un centro de Países Bajos, y rectopexia mediante resección laparoscópica, realizada en Estados Unidos. Esto se realizó en función de la técnica preferida en los respectivos continentes de los centros participantes del estudio.

Uso de cohortes externas

Puede considerarse un caso especial de la cohorte múltiple, en que, ante la imposibilidad de obtener un grupo de control, se recurre a fuentes de información externas preexistentes, tales como censos, registros poblacionales u otros estudios similares. Otra opción posible es obtener una muestra de individuos en otro centro o institución de características similares. Una desventaja esperable son las diferencias en las características de base que pudieran presentar los individuos. El uso de cohortes externas puede considerarse frente a exposiciones raras, o ante la imposibilidad de identificar un grupo de control durante el desarrollo del estudio^{12,23}.

Estudio de casos y controles anidado en cohorte

Corresponde a la realización de un estudio de casos y controles a partir de los individuos de la cohorte que presenten el desenlace. Esta metodología fue abordada en una entrega previa de la serie metodológica.

Estudio de casos-cohorte

Metodología derivada de los estudios de casos y controles, pero realizada a partir de una cohorte, desde la cual se elaboran dos grupos: un grupo de casos, que incluye a todos quienes desarrollen el evento de interés hasta un determinado momento del seguimiento, y una subcohorte muestreada aleatoriamente desde la cohorte general, independientemente de si corresponden a casos o no³¹. Resulta un diseño de particular utilidad en grandes cohortes con múltiples eventos de interés, en que recolectar datos para cada evento por separado resulta ineficiente, presentando una ventaja comparativa a los casos y controles anidados en cohorte al poder reutilizar la subcohorte en cada uno de los desenlaces evaluados (un estudio de casos y controles anidado requeriría construir un nuevo grupo de control para cada evento distinto). Una ventaja a destacar es que, debido a que los

controles representan a la cohorte en general, se cuenta con una base para la estimación de la incidencia y prevalencia en la población desde la cual fueron obtenidos³¹⁻³³ (Ejemplo 5).

Ejemplo 5. El proyecto MORGAM³⁴ corresponde a una colaboración internacional que busca estudiar la relación entre ciertos fenotipos y genotipos con la ocurrencia de eventos cardiovasculares. Los casos identificados fueron comparados a una subcohorte construida aleatoriamente. La eficiencia del diseño escogido salta a la luz: puesto que no resultaba costo efectivo realizar una genotipificación de toda la cohorte, el construir una subcohorte que puede ser reutilizada para varios desenlaces resultó de particular interés para los investigadores.

Medidas de frecuencia y asociación

Las medidas de asociación pueden ser expresadas en forma de probabilidades (riesgos) o tasas (incidencia). El cálculo de riesgo computará la proporción entre individuos que presenten el evento de interés y el total de individuos durante el tiempo total de seguimiento. El cálculo de incidencia incluirá una unidad de tiempo en forma explícita en las operaciones, expresando la “velocidad” a la que ocurre un evento¹⁵.

El riesgo calculado en forma individual para cada uno de los grupos que compongan una cohorte será llamado riesgo absoluto, pudiendo expresar las asociaciones entre ellos mediante el cálculo de razones (razón de riesgos o riesgo relativo) o diferencias (reducción de riesgo absoluto, diferencia de riesgo o riesgo atribuible)³⁵:

$$\text{Riesgo absoluto (R)} = \frac{\text{Individuos que presentan el evento}}{\text{Individuos en seguimiento en dicho grupo}}$$

$$\text{Riesgo relativo (RR)} = \frac{R(\text{expuestos})}{R(\text{no expuestos})}$$

$$\text{Riesgo atribuible (RA)} = R(\text{expuestos}) - R(\text{no expuestos})$$

El cálculo de la incidencia es matemáticamente similar, pero incluyendo en el denominador el tiempo en riesgo de presentar el evento:

$$\text{Incidencia} = \frac{\text{Individuos que presentan el evento}}{\text{Individuos en seguimiento} \times \text{tiempo en riesgo}} \text{ o } \frac{\text{Individuos que presentan el evento}}{\sum \text{tiempo en riesgo}}$$

Si el tiempo en riesgo es igual para todos los miembros de la cohorte, el valor obtenido corresponde a la tasa de incidencia; de presentar tiempos diferentes, el cálculo tiene la misma forma general, pero sumando los tiempos en riesgo individualizados de cada persona, llamándose en ese caso tasa de densidad de incidencia. Algunos textos mencionan además la incidencia acumulada o proporción de incidencia, que es matemáticamente igual al cálculo de riesgo absoluto y no corresponde a una tasa en sentido estricto^{35,36}.

Al ser un valor único y estático, la tasa de incidencia resulta en una potencial pérdida de información para reportar fenómenos cuya ocurrencia es variable en el tiempo. En dichos casos, la utilización de técnicas de análisis de supervivencia, permiten estudiar el tiempo de seguimiento libre del evento estudiado en una o más poblaciones. Particularmente, la construcción de tablas de vida o de curvas de Kaplan-Meier (análisis actuarial), contribuirá a la riqueza de la información presentada³⁷. Es posible encontrar en la literatura diferencias (o a veces confusiones) en las definiciones operacionales de estas medidas, por lo que se recomienda describirlas apropiadamente cuando sean usadas^{38,39}. Los estudios de supervivencia permiten también comparar las curvas de supervivencia, es decir, realizar un análisis inferencial para estimar diferencias entre ellas, mediante métodos paramétricos y no paramétricos. Los estudios de cohorte serán el diseño metodológico adecuado para los análisis de supervivencia³⁹.

Sesgos y estrategias de control

Las fuentes de sesgo en los estudios con diseño de cohorte son variadas, dadas por su naturaleza de estudio observacional en conjunto al seguimiento propio de un estudio longitudinal. Desde el punto de vista de la constitución de la muestra, destaca el fenómeno de sesgo de selección por pérdida de seguimiento. Un primer paso es abordar el problema desde el diseño, considerando los elementos mencionados en la sección “Construcción de un estudio de cohorte”. Cuando las pérdidas resultan inevitables, es posible utilizar la técnica de ponderación de la probabilidad inversa del riesgo de censura, en la cual, mediante la estimación de la probabilidad de pérdida de seguimiento, se construye un coeficiente de corrección que busca modelar el comportamiento de la muestra sin pérdidas, permitiendo corregir la estimación de riesgo⁴⁰. Mediante simulaciones se ha descrito que pérdidas desde el 20% resultan en un riesgo de sesgo importante, por lo que debe haber especial énfasis en evitarlas⁴¹. También en el campo del sesgo de selección, se reconoce el llamado “sesgo del trabajador sano”, en el cual se compara el riesgo de presentar una condición entre una cohorte de trabajadores (“expuestos”) y una cohorte obtenida desde la población general (“no expuestos”). Puede suceder que la medida de asociación resultante de la cohorte de trabajadores aparezca como un factor protector (riesgo relativo menor a 1) e incluso de menor magnitud que el riesgo estimado para la población general, por lo que podría interpretarse que ese trabajo en específico protegería a sus empleados de presentar el desenlace estudiado. No

obstante, esto podría explicarse debido a la óptima condición de salud (trabajador joven, sin patologías crónicas, “sano”) que tienen aquellos trabajadores y que ha sido corroborada tras un proceso de selección laboral. Este sesgo podría evitarse mediante comparaciones entre estratos de la cohorte de trabajadores, seleccionadas según cantidad de horas laborales semanales, tipo de trabajo realizado, entre otros⁴².

Los sesgos de información a los que deberíamos prestar atención varían según la modalidad de estudio de cohorte escogida. En una modalidad retrospectiva o en la fracción retrospectiva de un estudio ambispectivo, son potenciales fuentes de sesgo la calidad y legibilidad de los registros consultados, así como los mismos participantes, si es que se utilizan cuestionarios o entrevistas para recabar datos complementarios a la exposición. En el primer caso, deben reportarse rigurosamente los datos perdidos; desde el punto de vista del análisis, se describen tres grandes estrategias²⁴: 1. Ignorar el dato puntual faltante: el individuo se incluye solo en los análisis para los datos que sí posee; 2. Censurar el registro completo: independiente del dato faltante, se omite al individuo de todos los análisis y 3. Asignar un valor promedio o basal. Todas estas estrategias sesgan la muestra de formas impredecibles, por lo que deben estar claramente descritas y haber sido protocolizada de antemano. Respecto a la aplicación de cuestionarios o entrevistas, los instrumentos utilizados deben ser claros, en lenguaje entendible por los pacientes; además, quien aplica el instrumento o realiza la entrevista debe desconocer el estado de exposición del participante^{22,43}.

En estudios prospectivos o en la fracción prospectiva de un diseño ambispectivo, errores en la medición pueden llevar a clasificaciones erróneas tanto del estado de exposición del participante como de la ocurrencia del evento de interés; esto debe ser manejado desde el punto de vista del diseño, con criterios claramente definidos desde lo clínico y lo operacional (valores umbral, escalas a utilizar, métodos de laboratorio, etc.)^{14,26,44}.

Tampoco es ajeno a los estudios de cohorte el fenómeno de confusión, ya abordado en entregas anteriores de esta serie^{45,46}. Pudiera resultar de particular interés en el estudio con diseño de cohorte la utilización de técnicas de análisis estratificado y regresión estadística multivariada para, mediante modelamiento matemático, examinar y controlar el efecto de cada una de las variables registradas, identificando potenciales variables de confusión⁴³.

Consideraciones finales

Los estudios de cohorte tienen un fuerte antecedente histórico dentro del desarrollo de la medicina moderna y contemporánea¹⁰; no obstante, distan de ser una metodología anticuada, contando con

numerosas herramientas que permiten refinar los resultados obtenidos, beneficiándose enormemente de nuevos desarrollos en el campo de la bioestadística y del enorme poder de cómputo disponible actualmente, permitiendo una ágil implementación y aplicación de estos. Así, algunos autores destacan que, si bien el uso de estas herramientas debe ser analizado caso a caso y no ser aplicadas por mera novedad, su no utilización se debe comúnmente a la costumbre o desconocimiento, y no a una falta de factibilidad o soporte tecnológico acorde^{32,40,47}.

Este diseño metodológico ha permitido el estudio de asociaciones de gran relevancia para la salud pública y la medicina en general, especialmente en lo relativo a exposiciones a daño. Son el diseño de elección en el caso de estudiar la incidencia, la sobrevida a un fenómeno, los factores protectores y de riesgo y, a diferencia de otros diseños observacionales, permite el establecimiento de una relación causal, ya que el sentido del análisis de datos da cuenta del criterio de temporalidad, único que no puede estar ausente al hablar de hipótesis causales, esto es: desde la causa hacia el efecto (Figura 1).

Notas

Roles y contribuciones de los autores

CP, EM y MA son académicos de la Cátedra de Metodología de la Investigación Científica, en la que se circunscribe el desarrollo de la presente serie metodológica como una actividad investigativa de los ayudantes alumnos del curso. Todos los autores contribuyeron en la planificación y escritura del manuscrito original. GL, CP y MA aportaron el desarrollo de la Introducción, Conceptos preliminares, Construcción de un estudio de cohorte y Tipos de diseños de estudios de cohorte. GL, CP, EM y MA ayudaron en el desarrollo de Medidas de frecuencia y asociación y Sesgos y estrategias de control. GL aportó los Ejemplos del artículo. GL, CP y MA realizaron la Figura 1.

Financiamiento

Los autores declaran que no hubo fuentes externas de financiamiento.

Conflictos de intereses

Los autores completaron la declaración de conflictos de interés de ICMJE y declararon que no recibieron fondos por la realización de este artículo; no tienen relaciones financieras con organizaciones que puedan tener interés en el artículo publicado en los últimos tres años y no tienen otras relaciones o actividades que puedan influenciar en la publicación del artículo. Los formularios se pueden solicitar contactando al autor responsable o al Comité Editorial de la Revista.

Aspectos éticos

Este estudio no requirió la evaluación por parte de un comité de ética-científica, debido a que trabajó sobre fuentes secundarias.

Figura 1. Infografía sobre los estudios de cohorte. Fuente: diseñada por los propios autores.



Referencias

1. Goldsworthy AK. The complete Roman army. Thames & Hudson 2003. | Link |
2. Bacaër N. Halley's life table. En: A Short History of Mathematical Population Dynamics. London: Springer London; 2011: 5-10. | CrossRef |
3. Brown L, Pope EG. The postdischarge mortality among the patients of the Adirondack Cottage Sanitarium. Ethical Publishing Company; 1904.
4. Sartwell PE. The time factor in studies of the outcome of chronic disease. *Am Rev Tuberc*. 1951 May;63(5):608-12. | PubMed |
5. Comstock GW. Early studies of tuberculosis. *Natl Cancer Inst Monogr*. 1985 May;67:23-7. | PubMed |
6. Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J 3rd. Factors of risk in the development of coronary heart disease--six year follow-up experience. The Framingham Study. *Ann Intern Med*. 1961 Jul;55:33-50. | PubMed |
7. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. *Br Med J*. 1950 Sep 30;2(4682):739-48. | PubMed |
8. Ozasa K, Grant EJ, Kodama K. Japanese Legacy Cohorts: The Life Span Study Atomic Bomb Survivor Cohort and Survivors' Offspring. *J Epidemiol*. 2018 Apr 5;28(4):162-169. | CrossRef | PubMed |
9. Waxweiler RJ, Roscoe RJ, Archer VE, et al. Mortality follow-up through 1977 of the white underground uranium miners cohort examined by the United States Public Health Service [on line]. | Link |
10. Samet JM, Muñoz A. Evolution of the cohort study. *Epidemiol Rev*. 1998;20(1):1-14. | PubMed |
11. Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. *Lancet*. 2002 Jan 5;359(9300):57-61. | PubMed |
12. Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. *Lancet*. 2002 Jan 26;359(9303):341-5. | PubMed |
13. Carlson MD, Morrison RS. Study design, precision, and validity in observational studies. *J Palliat Med*. 2009 Jan;12(1):77-82. | CrossRef | PubMed |
14. Lu CY. Observational studies: a review of study designs, challenges and strategies to reduce confounding. *Int J Clin Pract*. 2009 May;63(5):691-7. | CrossRef | PubMed |
15. Madrid Aris E, Martínez Lomakin F. Moving towards a destination: considerations about cohort studies in less than 1000 words. *Medwave*. 2014 Jan 2;14(1):e5877. | CrossRef | PubMed |
16. Zhang Y, Wu S, Xia Y, Wang N, Zhou L, Wang J, et al. Adverse Events Associated with Treatment of Multidrug-Resistant Tuberculosis in China: An Ambispective Cohort Study. *Med Sci Monit*. 2017 May 18;23:2348-2356. | PubMed |
17. Silva LC. [Temporal sequence in observational studies to establish causality]. *Medwave*. 2014 May 2;14(4):e5944. | CrossRef | PubMed |
18. Fedak KM, Bernal A, Capshaw ZA, Gross S. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol*. 2015 Sep 30;12:14. | CrossRef | PubMed |
19. Höfler M. The Bradford Hill considerations on causality: a counterfactual perspective. *Emerg Themes Epidemiol*. 2005 Nov 3;2:11. | PubMed |
20. Seubsman SA, Kelly M, Sleigh A, Peungson J, Chokkanapitak J, Vilainerun D. Methods used for successful follow-up in a large scale national cohort study in Thailand. *BMC Res Notes*. 2011 May 27;4:166. | CrossRef | PubMed |
21. Fawcett J, Garrett N, Bates MN. Follow-up methods for retrospective cohort studies in New Zealand. *Aust N Z J Public Health*. 2002;26(3):256-61. | PubMed |
22. Toledano MB, Smith RB, Brook JP, Douglass M, Elliott P. How to Establish and Follow up a Large Prospective Cohort Study in the 21st Century--Lessons from UK COSMOS. *PLoS One*. 2015 Jul 6;10(7):e0131521. | CrossRef | PubMed |
23. Hulley SB, Cummings SR, Newman TB. Designing Cross-Sectional and Cohort Studies. In: *Designing Clinical Research*. Philadelphia: 2013:85-96.
24. Kang H. The prevention and handling of the missing data. *Korean J Anesthesiol*. 2013 May;64(5):402-6. | CrossRef | PubMed |
25. Dekkers OM, Egger M, Altman DG, Vandembroucke JP. Distinguishing case series from cohort studies. *Ann Intern Med*. 2012 Jan 3;156(1 Pt 1):37-40. | CrossRef | PubMed |
26. Sessler DI, Imrey PB. *Clinical Research Methodology 2: Observational Clinical Research*. Anesth Analg. 2015 Oct;121(4):1043-51. | CrossRef | PubMed |
27. Song JW, Chung KC. Observational studies: cohort and case-control studies. *Plast Reconstr Surg*. 2010 Dec;126(6):2234-42. | CrossRef | PubMed |
28. Cannizzaro G, Leone M, Esposito M. Immediate functional loading of implants placed with flapless surgery in the edentulous maxilla: 1-year follow-up of a single cohort study. *Int J Oral Maxillofac Implants*. 2007 Jan-Feb;22(1):87-95. | PubMed |
29. Mathes T, Pieper D. Clarifying the distinction between case series and cohort studies in systematic reviews of comparative studies: potential impact on body of evidence and workload. | CrossRef | PubMed |
30. Formijne Jonkers HA, Maya A, Draaisma WA, Bemelman WA, Broeders IA, Consten EC, et al. Laparoscopic resection rectopexy versus laparoscopic ventral rectopexy for complete rectal prolapse. *Tech Coloproctol*. 2014 Jul;18(7):641-6. | CrossRef | PubMed |
31. Prentice RL. A Case-Cohort Design for Epidemiologic Cohort Studies and Disease Prevention Trials. *Biometrika* 1986;73:1. | CrossRef |
32. Vojvodic M, Shafarenko M, McCabe SJ. Case-Cohort Studies: Design and Applicability to Hand Surgery. *J Hand Surg Am*. 2018 Aug;43(8):762-767. | CrossRef | PubMed |
33. Cologne J, Preston DL, Imai K, Misumi M, Yoshida K, Hayashi T, et al. Conventional case-cohort design and analysis for studies of interaction. *Int J Epidemiol*. 2012 Aug;41(4):1174-86. | CrossRef | PubMed |
34. Kulathinal S, Karvanen J, Saarela O, Kuulasmaa K. Case-cohort design in practice - experiences from the MORGAM Project. *Epidemiol Perspect Innov*. 2007 Dec 4;4:15. | PubMed |
35. Porta M. *A Dictionary of Epidemiology*. Oxford University Press. | CrossRef |
36. Tenny S, Bektor SW. Incidence. 2019 Jan 2. *StatPearls*. Treasure Island (FL): StatPearls Publishing; 2019 Jan-. | PubMed | Link |
37. Kartsonaki C. Survival analysis. *Diagnostic Histopathol* 2016;22:263-70. | CrossRef |
38. Spronk I, Korevaar JC, Poos R, Davids R, Hilderink H, Schellevis FG, et al. Calculating incidence rates and prevalence proportions: not as simple as it seems. *BMC Public Health*. 2019 May 6;19(1):512. | CrossRef | PubMed |
39. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: basic concepts and first analyses. *Br J Cancer*. 2003 Jul 21;89(2):232-8. | PubMed |
40. Howe CJ, Cole SR, Lau B, Napravnik S, Eron JJ Jr. Selection Bias Due to Loss to Follow Up in Cohort Studies. *Epidemiology*. 2016 Jan;27(1):91-7. | CrossRef | PubMed |
41. Kristman V, Manno M, Côté P. Loss to follow-up in cohort studies: how much is too much? *Eur J Epidemiol*. 2004;19(8):751-60. | PubMed |

42. Chowdhury R, Shah D, Payal AR. Healthy Worker Effect Phenomenon: Revisited with Emphasis on Statistical Methods - A Review. *Indian J Occup Environ Med.* 2017 Jan-Apr;21(1):2-8. | CrossRef | PubMed |
43. Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet.* 2002 Jan 19;359(9302):248-52. | CrossRef | PubMed |
44. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Selection bias and information bias in clinical research. *Nephron Clin Pract.* 2010;115(2):c94-9. | CrossRef | PubMed |
45. Barraza F, Arancibia M, Madrid E, Papuzinski C. General concepts in biostatistics and clinical epidemiology: Random error and systematic error. *Medwave.* 2019 Aug 27;19(7):e7687. | CrossRef | PubMed |
46. Cataldo R, Arancibia M, Stojanova J, Papuzinski C. General concepts in biostatistics and clinical epidemiology: Observational studies with cross-sectional and ecological designs. *Medwave.* 2019 Sep 25;19(8):e7698. | CrossRef | PubMed |
47. Xue X, Agalliu I, Kim MY, Wang T, Lin J, Ghavamian R, et al. New methods for estimating follow-up rates in cohort studies. *BMC Med Res Methodol.* 2017 Dec 1;17(1):155. | CrossRef | PubMed |

Correspondencia a

Angamos 655, Edificio R2
Oficina 1124
Reñaca, Viña del Mar
Chile



Esta obra de Medwave está bajo una licencia Creative Commons Atribución-No Comercial 3.0 Unported. Esta licencia permite el uso, distribución y reproducción del artículo en cualquier medio, siempre y cuando se otorgue el crédito correspondiente al autor del artículo y al medio en que se publica, en este caso, Medwave.