

CARACTERÍSTICAS DE LOS TIPOS DE ESTUDIOS MÁS FRECUENTES EN INVESTIGACIÓN CLÍNICA

Dr. Santiago Perez Lloret^{1,2}, Dr. Daniel Vigo³, Dr. Claudio González⁴

¹SECCIÓN DE MOVIMIENTOS ANORMALES, DEPARTAMENTO DE NEUROLOGÍA, FLENI.

²CENTRO DE FARMACOLOGÍA CLÍNICA (FLENI-MRC), FLENI.

³LABORATORIO DE NEUROFISIOLOGÍA, FACULTAD DE MEDICINA, UBA.

⁴DEPARTAMENTO DE FARMACOLOGÍA. INSTITUTO UNIVERSITARIO CEMIC

DIRECCIÓN: FLENI. MONTAÑESES 2325. C1428AQK. BUENOS AIRES // E-MAIL DE CONTACTO: splloret@fleni.org.ar

RESUMEN

Los estudios de investigación son la mínima expresión del método científico. La medicina se vale de la evidencia aportada por los estudios de investigación clínica para posibilitar el proceso de toma de decisiones. Estos estudios pueden clasificarse, inicialmente, en experimentales y observacionales. En los primeros, el investigador interviene deliberadamente sobre el sistema observado y la asignación de los sujetos a cada tratamiento o intervención puede realizarse por métodos azarosos o no. Los estudios clínicos aleatorios son los diseños más parecidos a los experimentos de la investigación básica. En cambio, en los estudios observacionales, el investigador observa lo que ocurre en una muestra de sujetos, los cuales recibieron uno, otro o ningún tratamiento. Dichos estudios pueden clasificarse en analíticos o descriptivos de acuerdo a si tienen o no un grupo control, respectivamente. Finalmente, dentro de los estudios analíticos se incluyen a los estudios de cohortes, de casos y controles y a los diseños transversales. Cada uno de estos diseños tiene ventajas y desventajas, referidos principalmente a los costos, capacidad de responder preguntas y presencia de sesgos, de ahí que sea indispensable para el investigador conocerlos en detalle.

Palabras Clave: Investigación biomédica- Investigación empírica- Estudios epidemiológicos- Estudios de casos y controles- Estudios de cohorte- Estudios transversales

ABSTRACT

Clinical research studies are the minimal expression of the scientific method. Medicine uses evidence generated by clinical research in order to facilitate decision-making processes. These studies can be classified initially in experimental and observational. In the former, the researcher intervenes deliberately on the system under observation, and assignation of subjects to each treatment or intervention may be carried out using a random selection process or not. Randomized clinical trials are most similar in design to basic research experiments. In observational studies however, the investigator observes what happens to a sample of subjects, who may or may not receive a given treatment, or receive a different one altogether. These studies may be classified as analytical or descriptive, depending on whether or not they have a control group. Finally, included among analytical studies are: cohort studies, case-control studies, and those with a cross-sectional design. Each of these designs has advantages and disadvantages, mainly referred to costs, ability to answer questions and proneness to methodological biases, reason for which it is very important that investigators be very familiar with all of them.

Key Words: Biomedical Research- Empiric Research- Epidemiologic Studies- Case-Control Studies- Cohort Studies- Cross-sectional Studies

CALIDAD DE LA EVIDENCIA MÉDICA

Hemos visto que la investigación provee evidencia que el médico debe combinar su experiencia con los deseos del paciente para tomar una decisión.¹ Como todo sujeto que toma decisiones, el médico adopta conductas en un ámbito de riesgo: todos los tratamientos tienen tanto efectos adversos como beneficiosos, pero no administrarlos a tiempo puede ser grave y las herramientas diagnósticas imperfectas pueden llevar a tomar innecesariamente conductas de riesgo, por ejemplo, en sujetos libres de la patología sospechada. En estas situaciones, las decisiones deben sustentarse firmemente en el peso de la evidencia (*weight of evidence*), es decir en la información disponible al momento.²

Desde el punto de vista científico, el concepto engloba a los diferentes resúmenes ordenados de la evidencia, como pueden ser las revisiones sistemáticas o los meta-análisis y a los métodos disponibles para producirlos.³ Sin embargo, el progreso diario de la ciencia produce un amplio caudal de evidencia cuya calidad suele ser heterogénea.⁴ ¿A qué nos referimos cuando hablamos de la calidad de la evidencia? Es la serie de atributos que nos permiten confiar en menor o mayor grado en los resultados obtenidos en los diferentes estudios.

Por esta razón, diseñar un estudio implica planificar la manera en que los diferentes sesgos serán evitados y cómo podrán interpretarse los resultados a pesar de los sesgos que no pudieron ser eliminados. Así, la calidad de un estudio está determinada por la forma en que dicho estudio fue diseñado y conducido. Más aún, así como la anatomía determina la fisiología, el diseño básico del estudio deter-

mina la máxima calidad que se podrá alcanzar con ese diseño. Esto ha permitido elaborar jerarquías de calidad de la evidencia en función del tipo de estudio considerado, como se muestra en la tabla 1.²

TIPOS DE DISEÑOS EN INVESTIGACIÓN CLÍNICA

Previamente hemos visto que, con fines operativos, podemos considerar que un factor es causa de un evento cuando la frecuencia del evento es mayor en el grupo de sujetos expuestos al factor en comparación con el grupo no expuesto —procuraremos aquí evitar discusiones filosóficas más profundas acerca del concepto de causalidad.⁵⁻⁷ Por esta razón, todos los estudios que investiguen causalidad deberán registrar y comparar la frecuencia del evento de interés en los grupos de sujetos expuestos y no expuestos al factor de riesgo.^{4,6-9}

Los diseños que analizaremos diferirán principalmente en la manera en que se seleccionan los grupos de sujetos.⁴ Podemos comenzar por dividir los estudios en aquellos que son experimentales u observacionales como se muestra en la figura 1. En los primeros, el investigador interviene deliberadamente sobre el sistema observado y la asignación de los sujetos a cada tratamiento o intervención puede realizarse por métodos azarosos o no.⁴ Los estudios clínicos aleatorios son los diseños más parecidos a los experimentos de la investigación básica. En cambio, en los estudios observacionales, el investigador observa lo que ocurre en una muestra de sujetos, los cuales recibieron uno, otro o ningún tratamiento —en este caso, el observador no interviene deliberadamente sobre el sistema observado.⁴ Este sería el caso si, por ejemplo, se deseara estudiar el riesgo de cáncer de mama en mujeres que reciben anticonceptivos: se podría comparar la incidencia de cáncer en mujeres que recibían o no anticonceptivos al momento de realizarse una mamografía. En términos generales, podemos afirmar que los estudios observacionales tienden a mostrar más sesgos que los experimentales.¹⁰

Los estudios observacionales pueden clasificarse en analíticos o descriptivos de acuerdo a si tienen o no un grupo control, respectivamente^{4,6} (figura 1).

Los estudios descriptivos pueden ir desde el registro de un caso hasta una serie de casos y

Tabla 1. Sistema de calificación de la evidencia en función de su calidad

Grado	Evidencia obtenida de
I	Al menos 1 estudio clínico aleatorio controlado bien diseñado
II-1	Estudios clínicos no aleatorios bien diseñados
II-2	Estudios de cohortes o caso-control bien diseñados
II-3	Series temporales múltiples con o sin intervención o de experiencias clínicas no controladas importantes
III	Estudios descriptivos, recomendaciones de comités de expertos, opiniones personales

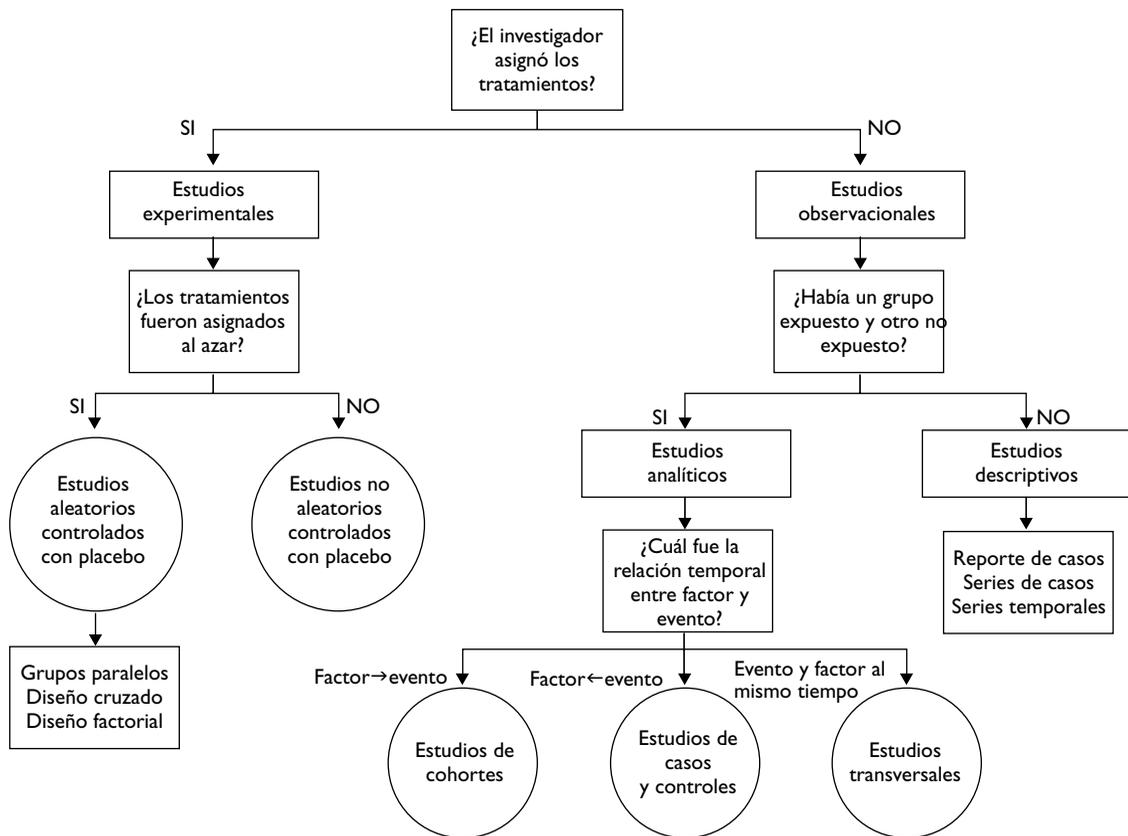


Figura 1. Clasificación de los principales tipos de estudios en investigación clínica.

no representan en realidad un verdadero estudio científico, principalmente porque en general, la observación no cumple con los criterios que hemos elaborado con anterioridad. Sin embargo, su valor para la medicina es muy elevado¹¹. Baste con citar el reconocimiento del SIDA a partir del reporte de casos de pacientes jóvenes con sarcoma de Kaposi,^{12,13} o el establecimiento de uno de los modelos animales de enfermedad de Parkinson más influyentes a partir del informe de casos de pacientes con la enfermedad, que habían consumido heroína contaminada con MPTP.^{14,15}

Finalmente, los estudios analíticos son aquellos en el que un grupo de sujetos que presentan el factor de riesgo que se está estudiando se compara con otro sujeto que no lo presenta. Este tipo de estudios son frecuentes en medicina, por su relativa simplicidad y bajo costo y como hemos visto, son también los que más sesgos presentan.¹⁰

Ellos se clasifican en relación al punto de partida que se emplea^{4,6} (figura 1). En los estudios de cohortes, se seleccionan pacientes con o sin el fac-

tor de riesgo y se investiga cuántos sujetos, en cada grupo, desarrollaron el evento. Es decir el tiempo fluye hacia delante (figura 2). Por el contrario, para los estudios de caso-control, se seleccionan sujetos

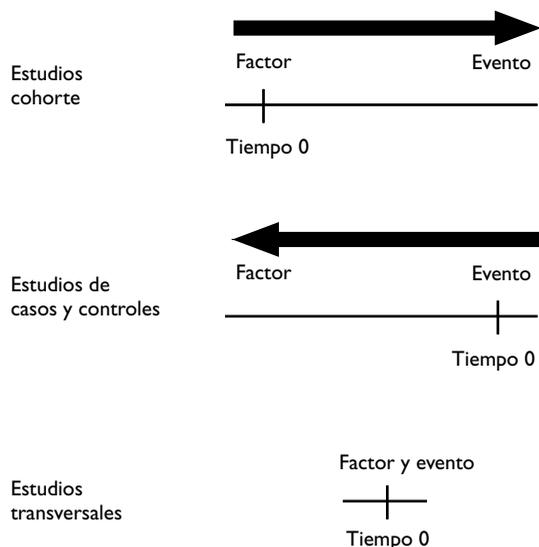


Figura 2. Flujo del tiempo en los estudios observacionales analíticos.

con o sin el evento de interés y se busca determinar cuántos de ellos estuvieron expuestos o no al factor de riesgo de interés. Observemos, que aquí hemos comenzado por el final, por lo cual podemos decir que el tiempo fluye en reversa. ¿Cuál es la ventaja de comenzar nuestro análisis desde el evento? Los estudios de cohortes pueden volverse muy engorrosos cuando el evento de interés es poco frecuente. En estos casos, se prefiere seleccionar directamente los casos y compararlos contra los controles. Más adelante analizaremos las ventajas y desventajas de cada uno de ellos.

Por último, los estudios transversales (llamados a veces de sección cruzada) son aquellos en los cuales la presencia del factor de riesgo y el evento se determinan al mismo tiempo. Por ejemplo, podemos determinar que los sujetos obesos son aquellos que realizan menor nivel diario de actividad. Pero, ¿qué ocurrió primero? Podría bien resultar que la falta de actividad física haya conducido a la obesidad, pero también podría ser posible que la obesidad haya tenido otra causa, y a su vez haya sido la causa de la inactividad. Es muy difícil establecer la relación temporal entre la causa y el efecto, por lo cual esta clase de estudios no nos es útil para evaluar relaciones causales.

ESTUDIOS CLÍNICOS ALEATORIOS CONTROLADOS (ECAC)

En este tipo de diseños, un investigador comienza por seleccionar una muestra de pacientes que cumpla con una serie de criterios de inclusión y exclusión. (figura 3) Una vez seleccionados, los sujetos son asignados por un método aleatorio a uno u otro grupo de tratamiento.¹⁶ Un grupo será tratado con el método terapéutico en estudio, ya sea un medicamento o un procedimiento quirúrgico, por ejemplo, mientras que el otro recibirá placebo, es decir una versión inactiva del medicamento en estudio o la mejor alternativa terapéutica disponible. Estos sujetos serán seguidos en el tiempo y se contará la cantidad de sujetos que desarrollaron el evento, en cada grupo (figura 3).

¿Por qué asignamos aleatoriamente los sujetos a los grupos? Desde el punto de vista médico quizás resultaría mucho más lógico asignar a los sujetos en peores condiciones al grupo control, si el medicamento estudiado no es del todo conocido, mientras que aquellos en mejor estado recibirían el me-

dicamento experimental. De la misma manera, si la enfermedad fuera mortal, los pacientes con riesgo inminente de muerte podrían ser los candidatos ideales para recibir el medicamento en estudio.

Sin embargo, el criterio médico, en este caso, se opone al criterio científico.^{4,6,7,17} Veamos por qué. Tomemos el primer caso, bien podría suceder que tras el análisis de los datos se observara que la tasa de curación fue mayor en el grupo tratado que en el de control. Una conclusión posible sería que el medicamento fue eficaz, aunque también podríamos sospechar que esta diferencia se relaciona con el hecho que la enfermedad era más leve en los sujetos asignados al grupo control.

En el segundo caso comentado, podría suceder que la tasa de mortalidad fuera similar en los grupos tratado y no tratado. Bien podríamos cuestionar esta supuesta igualdad, aduciendo que en realidad, la mortalidad del grupo más severamente afectado hubiera sido mayor que la del otro grupo, si no hubiera estado recibiendo el medicamento en estudio.

Toda diferencia entre los grupos estudiados, por ejemplo en la edad, gravedad de la enfermedad, género, etc, genera una diferencia entre los grupos, que afecta la capacidad del estudio de concluir que el tratamiento vigente es la causa de la curación/supervivencia de los sujetos tratados.^{6,17} Hemos visto que estos sesgos de selección generan los sesgos

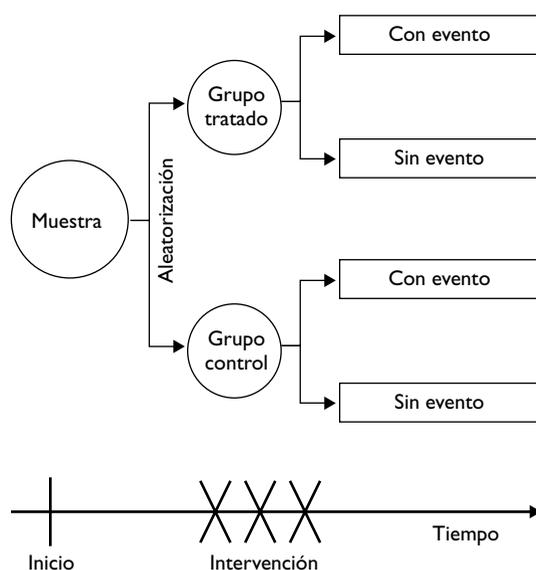


Figura 3. Esquema básico de un estudio clínico aleatorio controlado.

por factor confundente. La aleatorización, al eliminar cualquier razón por la que un sujeto debería recibir un tratamiento con más probabilidad que el otro, es la mejor forma de reducir o eliminar el sesgo de selección. Por extensión, sólo los ECAC están potencialmente exentos de sesgos de selección, razón por la cual son el estándar de oro de la “medicina basada en la evidencia”, pues proveen la evidencia de mayor calidad.

Adicionalmente, la asignación aleatoria posibilita cegar al sujeto experimental y/o al investigador respecto del tratamiento.^{18,19} Es decir, puede haber estudios aleatorios abiertos, pero, es imposible diseñar un estudio doble ciego que no sea aleatorio.¹⁹ El desconocimiento del tratamiento por parte del sujeto y del investigador, en primer lugar, permite que las expectativas de ambos no afecten las evaluaciones y que dichas evaluaciones y las conductas médicas que se pudiera tomar, no se vean afectadas por el conocimiento del tratamiento recibido.

La mayoría de los ECAC se conducen en el ámbito de la medicina asistencial, pero, por sus características, generan una suerte de microentorno que difiere radicalmente del asistencialismo.²⁰ En primer lugar, el empleo de criterios de selección estrictos, con el fin de homogeneizar la muestra estudiada, impide ver el impacto de las características particulares de los pacientes sobre el efecto del medicamento. En segundo lugar, la contención médica constante, estimula la adhesión al tratamiento, lo cual no siempre ocurre en el ámbito usual. De todo lo cual surge que la credibilidad de los resultados de los ECAC, puede aplicarse a un grupo selecto de pacientes.²⁰ En otras palabras, la validez externa ha sufrido en pos de maximizar la validez interna del estudio. En condiciones ideales deseáramos un estudio en donde ambos tipos de validez fueran máximos, pero por lo general, siempre debemos comprometer uno de ellos (figura 4). ¿Qué es preferible, la validez interna o la externa?

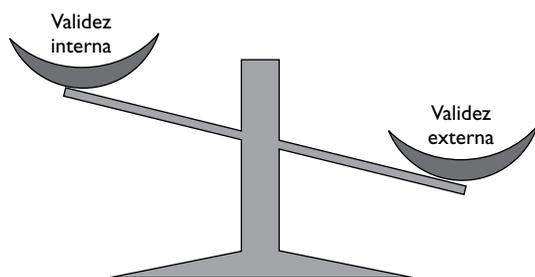


Figura 4. Compromiso entre la validez interna y externa de un estudio.

Si bien ésta no es una pregunta fácil, la mayoría de las veces debemos preferir estudios con gran validez interna.⁶ De otra forma, podríamos cometer el riesgo de poder aplicar ampliamente resultados de poca credibilidad. No obstante, debemos tener en mente que ésta es la principal limitación actual de la “medicina basada en la evidencia”: la imposibilidad de tener evidencia aplicable a la mayoría de los pacientes atendidos.

ESTUDIOS DE COHORTE

Los estudios de cohortes son los estudios observacionales por excelencia.⁹ En ellos una cohorte o grupo de sujetos es seguido prospectivamente durante un periodo determinado. Dentro de la cohorte algunos sujetos eventualmente se expondrán al factor de riesgo mientras que otros no. Luego, al final del periodo de seguimiento, contaremos cuántos en cada grupo tuvieron el evento de interés y cuántos no (figura 5).

Tomemos un ejemplo. En la década del '40 el Consejo de Investigaciones Médicas del Reino Unido encaró la investigación de los efectos nocivos, pues había cierta sospecha acerca de su asociación con ciertos tipos de cáncer, pero no había evidencia sólida al respecto. Para el estudio se enroló una cohorte de 34 mil profesionales de la medicina, de ambos sexos, de los cuales el 83% afirmó ser tabaquista. En un periodo de seguimiento de 5 años, se observó que los sujetos fumadores presentaban mayor frecuencia de cáncer pulmonar e infarto de miocardio en comparación al porcentaje de sujetos

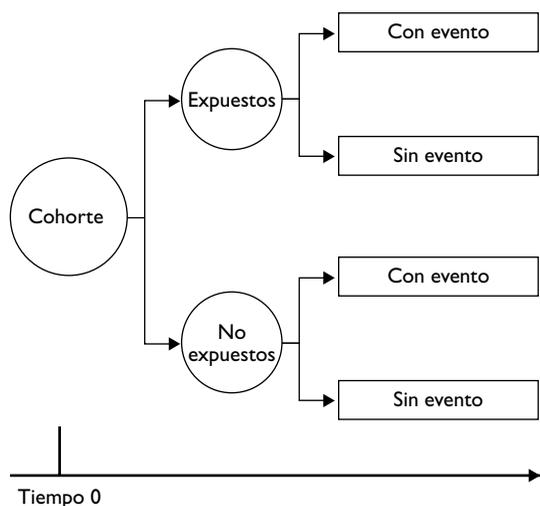


Figura 5. Esquema básico de un estudio de cohortes.

no fumadores.²¹ Los autores, Doll y Hill, concluyeron que el tabaquismo era un factor causal de infarto de miocardio y cáncer pulmonar.

Éste es quizás el mejor ejemplo del valor del estudio de cohortes para la medicina. Desde ya que hubiera sido imposible estudiar el problema mediante un estudio aleatorio controlado, pues habría sido ética y fácticamente imposible obligar a un grupo aleatorio de médicos a consumir tabaco. Como éste, podemos citar otros ejemplos de estudios que han tenido gran impacto en el conocimiento de las causas de las principales enfermedades humanas, entre los que se encuentran el estudio de “anticonceptivos orales”, el estudio de Framingham o el estudio de las “enfermeras en Estados Unidos”.²²⁻²⁴

Ahora bien, veamos las principales limitaciones. En el “estudio del efecto nocivo de los anticonceptivos orales” (ref) se estudió una cohorte de 46 mil mujeres durante 25 años. Las mujeres que recibían anticonceptivos tuvieron mayor frecuencia de accidente cerebrovascular (ACV). Sin embargo, tiempo después se observó que en realidad la frecuencia de ACV en mujeres no fumadoras que recibían anticonceptivos era similar a las mujeres no tratadas.²⁵ Podría postularse que las mujeres que ingerían anticonceptivos podrían tener ideas más liberales y verse más inclinadas a adquirir el hábito del tabaquismo. Más allá de las explicaciones, vemos que entre los grupos de mujeres comparadas había otras diferencias, más allá de la exposición a los anticonceptivos. Estas diferencias surgen como un sesgo de selección, constituyendo en este caso el tabaquismo un factor confundente en la verdadera relación entre los anticonceptivos y el ACV, que parecería ser nula.

ESTUDIOS DE CASOS Y CONTROLES

En este tipo de estudios, se selecciona a los sujetos por el evento y se determina la frecuencia de la exposición.⁸ (Figura 6). Consideremos el siguiente ejemplo. Supongamos que queremos estudiar la relación entre cáncer de pulmón y tabaquismo pero nos resultaría imposible encarar un estudio con 30 mil participantes. Si decidiéramos estudiar menos sujetos, probablemente nos resultaría muy difícil reunir suficiente cantidad de eventos como para poder analizarlos estadísticamente. En cambio, podríamos seleccionar un grupo de sujetos con cáncer de pulmón, a los que llamaremos

“casos” y un grupo de sujetos sin la enfermedad, a los cuales llamaremos “controles”. En cada grupo investigaremos cuántos sujetos eran tabaquistas previamente al diagnóstico de la neoplasia. Probablemente observaríamos que los sujetos con cáncer de pulmón reportarían tabaquismo con mayor frecuencia. Sin embargo, ¿qué tan creíbles serían estas afirmaciones? Analicemos el siguiente ejemplo. En un estudio sueco acerca de la relación entre la inducción de abortos y el desarrollo de cáncer de mama, se investigó la frecuencia de abortos inducidos en mujeres con o sin cáncer mamario mediante cuestionarios o buscando dicha información en los registros del sistema de salud. Se observó que muchas de las mujeres sin cáncer de mama (controles), no reportaban abortos previos, que por otro lado sí aparecían en los registros. Por el contrario, todas las mujeres con cáncer de mama habían reportado los abortos (ref28-case-control). En muchas ocasiones, los casos recuerdan más circunstancias de su vida o son menos proclives a ocultarlas que sus controles. Este sesgo se conoce como “por olvido” y es muy habitual en los estudios que se basan en cuestionarios.⁸ Aquellos estudios basados en datos concretos, por ejemplo, extraídos de la historia clínica, pueden tener problemas, igualmente. Muchas veces se omite la información negativa, por lo que no se puede distinguir cuándo realmente el factor de riesgo no estaba presente o cuándo el médico sencillamente olvidó registrar un dato, que podría haber positivo.

Otro gran problema de este tipo de estudios es la selección de los controles adecuados.^{8,17} En tér-

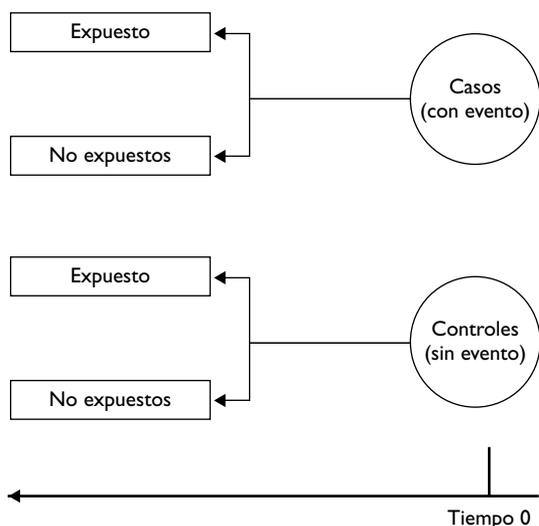


Figura 6. Esquema básico de un estudio de casos y controles.

minos generales, podemos afirmar que ellos deben ser sujetos que, de haber sufrido la enfermedad, hubieran integrado el grupo de los casos. Si se toman casos de una unidad coronaria de un hospital que recibe derivaciones de toda la provincia, los controles son seleccionados de la guardia del mismo hospital. La guardia sólo recibe casos provenientes de la ciudad, por lo que, si el factor de riesgo fuera, por ejemplo, un factor ambiental, muchos casos, al habitar en otras regiones, probablemente no hubieran estado expuestos a él, subestimando una relación causal que podría perfectamente ser real.

ESTUDIOS TRANSVERSALES

En este tipo de estudios, se registra en un grupo de sujetos la frecuencia de un evento y su factor de riesgo al mismo tiempo.⁴ Consideremos los siguiente ejemplos. Un investigador demuestra que los sujetos con trastornos de sueño presentan niveles elevados de presión arterial. También se observa en sujetos con esclerosis múltiple mayor frecuencia de sedentarismo. Sin embargo, ¿cuál fue primero? En el primer caso, bien podría ser que la hipertensión dañe zonas cerebrales encargadas de la regulación del sueño, o también que el trastorno de sueño se asocie a hipertensión por hiperactivación simpática nocturna. Igualmente, en

el segundo ejemplo, la esclerosis múltiple podría generar una imposibilidad de realizar actividad o la inactividad podría ser un factor de riesgo. Como vemos, en este tipo de estudios es muy difícil establecer la línea del tiempo, por lo cual no pueden ser empleados para evaluar causalidad.

CONCLUSIÓN

El peso de la evidencia es el conjunto de evidencias que rodea un tema determinado y al formarla debemos ser cuidadosos con la calidad de la evidencia seleccionada, pues de eso depende el éxito de la actividad médica. El diseño de los estudios de investigación clínica es uno de los principales determinantes de la calidad del estudio. Mientras que los estudios clínicos aleatorios controlados son los que menos sesgos presentan, la posibilidad de aplicarlos a grandes grupos de pacientes es limitada. Los estudios observacionales son aquellos en los que el investigador no determina qué sujeto recibe cuál tratamiento. Si bien están abiertos a muchos tipos de sesgos, son muy útiles en medicina porque permiten analizar situaciones fuera del alcance de los estudios clínicos aleatorios controlados. En las tablas 2 y 3 se pueden observar cuadros resumen con las principales diferencias entre ellos ●

Tabla 2. Resumen comparativo de los tipos principales de diseños en investigación clínica

Investigación clínica	Experimentales	Observacionales	
Pregunta	De intervención Cuál es el resultado de la intervención	Caso-control Cohorte ¿Podría ser la exposición al factor la causa del evento?	Transversales Cuál es la situación actual
Objetivos principales	Comparar el efecto de una intervención en un grupo expuesto vs control	Identificar relaciones entre la exposición a un factor y la ocurrencia de un evento	Establecer un diagnóstico de situación
Hipótesis de causalidad		SI	NO
Comparación por grupos		SI	SI o NO
Selección de los grupos	Asignación grupal aleatoria o no	Los grupos vienen preformados	
Intervenciones	Hay intervención	No se hacen intervenciones	
Posibilidad de ciego	SI	Sólo del observador	
Problemas éticos importantes	SI	Menos (confidencialidad)	
Evaluación de temporalidad		SI	NO
Evaluación de Asociación		SI	
Evaluación de causalidad	La asociación causal se establece con la mayor claridad posible	La asociación causal no puede establecerse tan claramente como en otros diseños	No tiene alcance para comprobar causalidad ni permite hacer predicciones

Tabla 3. Resumen comparativo de las principales características de los estudios de cohortes y caso-control

	Caso-control	Cohorte
Pregunta	Cuáles son los factores relacionados a un evento determinado	Cuáles son eventos relacionados a un factor determinado
Selección de los sujetos	Por la presencia del evento	Por la presencia del factor de riesgo
Medida directa de prevalencia del evento	NO	SI
Medida de riesgo para el evento	Odds ratio	Relative Risk
Ventajas más relevantes	Útil para estudio eventos poco frecuentes	Información más completa
	Útil cuando tiempo al evento es largo	Control de factores confundentes
	Requiere pocos recursos	Permite el cálculo de prevalencia del evento
Mayores desventajas	Importante efecto del sesgo de recuerdo	Poco útiles para eventos raros
	Mayores posibilidades de otro tipo de sesgos	No útil cuando tiempo al evento es largo
	No permite cálculo de prevalencia del evento	Alto sesgo por abandono del estudio (drop-out)
	Control incompleto de factores confundentes	Requiere muchos recursos
Situaciones en los que conviene emplearlo	Evento poco frecuente. Al inicio de investigaciones causales	Con eventos frecuentes o exposición poco frecuente. Para evaluación causal más precisa

REFERENCIAS

- Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Evidence-Based Medicine: How to Practice and Teach EBM. 2000;
- US Preventive Services Task Force. Guide to clinical preventive services. 1996; Second:
- Weed DL. Weight of evidence: a review of concept and methods. Risk Anal 2005; 25: 1545-57.
- Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. Lancet 2002; 359: 57-61.
- Coggon DI, Martyn CN. Time and chance: the stochastic nature of disease causation. Lancet 2005; 365: 1434-7.
- Elwood JM. Critical Appraisal of Epidemiological Studies and Clinical Trials. 1998;
- Grimes DA, Schulz KF. Bias and causal associations in observational research. Lancet 2002; 359: 248-52.
- Schulz KF, Grimes DA. Case-control studies: research in reverse. Lancet 2002; 359: 431-4.
- Grimes DA, Schulz KF. Cohort studies: marching towards outcomes. Lancet 2002; 359: 341-5.
- Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. BMJ 1998; 317: 1185-90.
- Grimes DA, Schulz KF. Descriptive studies: what they can and cannot do. Lancet 2002; 359: 145-9.
- Gottlieb GJ, Ragaz A, Vogel JV, Friedman-Kien A, Rywlin AM, Weiner EA, et al. A preliminary communication on extensively disseminated Kaposi's sarcoma in young homosexual men. Am J Dermatopathol 1981; 3: 111-4.
- Hymes KB, Cheung T, Greene JB, Prose NS, Marcus A, Ballard H, et al. Kaposi's sarcoma in homosexual men-a report of eight cases. Lancet 1981; 2: 598-600.
- Wright JM, Wall RA, Perry TL, Paty DW. Chronic parkinsonism secondary to intranasal administration of a product of meperidine-analogue synthesis. N Engl J Med 1984; 310: 325.
- Langston JW, Ballard P, Tetrud JW, Irwin I. Chronic Parkinsonism in humans due to a product of meperidine-analog synthesis. Science 1983; 219: 979-80.
- Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. Lancet 2002; 359: 515-9.

17. Strom BL. *Pharmacoepidemiology*. 2000;
18. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet* 2002; 359: 614-8.
19. Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet* 2002; 359: 696-700.
20. Sanson-Fisher RW, Bonevski B, Green LW, D'Este C. Limitations of the randomized controlled trial in evaluating population-based health interventions. *Am J Prev Med* 2007; 33: 155-61.
21. DOLL R, HILL AB. Lung cancer and other causes of death in relation to smoking; a second report on the mortality of British doctors. *Br Med J* 1956; 2: 1071-81.
22. Beral V, Hermon C, Kay C, Hannaford P, Darby S, Reeves G. Mortality associated with oral contraceptive use: 25 year follow up of cohort of 46 000 women from Royal College of General Practitioners' oral contraception study. *BMJ* 1999; 318: 96-100.
23. Seman LJ, DeLuca C, Jenner JL, Cupples LA, McNamara JR, Wilson PW, et al. Lipoprotein(a)-cholesterol and coronary heart disease in the Framingham Heart Study. *Clin Chem* 1999; 45: 1039-46.
24. Belanger CF, Hennekens CH, Rosner B, Speizer FE. The nurses' health study. *Am J Nurs* 1978; 78: 1039-40.
25. Burkman RT. Cardiovascular issues with oral contraceptives: evidenced-based medicine. *Int J Fertil Womens Med* 2000; 45: 166-74.