



Revista alergia México

ISSN: 0002-5151

ISSN: 2448-9190

Colegio Mexicano de Inmunología Clínica y Alergia A.C.

Zurita-Cruz, Jessie Nallely; Villasís-Keever, Miguel Ángel
Principales sesgos en la investigación clínica
Revista alergia México, vol. 68, núm. 4, 2021, Octubre-Diciembre, pp. 291-299
Colegio Mexicano de Inmunología Clínica y Alergia A.C.

DOI: <https://doi.org/10.29262/ram.v68i4.1003>

Disponible en: <https://www.redalyc.org/articulo.oa?id=486771934008>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org



Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

Main biases in clinical research

Principales sesgos en la investigación clínica

Jessie Nallely Zurita-Cruz,¹ Miguel Ángel Villasís-Keever²

Abstract

In developing a research protocol, authors must consider the possible errors that may occur throughout the study. In clinical research, two types of biases are recognized: random errors and systematic errors; the latter are called biases. To date, dozens of biases have been described, which is why the purpose of this article is to describe the main biases that can occur in clinical research studies, as well as strategies to avoid them or to minimize their effects. Since there are several classifications, in order to provide a more practical overview in this review, the biases are grouped into three types: selection biases, information (or performance) biases, and confounding biases. In addition, to make it even more specific, we describe the biases considering the purpose of the research: prognosis, therapeutics, causality, and diagnostic test studies.

Key words: Random errors; Systematic errors; Biases; Clinical research; Review

Resumen

En la elaboración de un protocolo de investigación, los autores deben tomar en cuenta los posibles errores que puedan ocurrir a lo largo del estudio. En la investigación clínica se reconocen dos tipos: los errores aleatorios y los errores sistemáticos, estos últimos se denominan sesgos. A la fecha se han descrito decenas de sesgos, por lo que este artículo tiene como objetivo describir los principales sesgos que pueden ocurrir en los estudios de investigación clínica, así como la forma para evitarlos o minimizar sus efectos. En virtud de que existen varias clasificaciones, en la presente revisión y, a fin de disponer un panorama más práctico, los sesgos se agrupan en tres: sesgos de selección, sesgos de información (o ejecución) y sesgos de confusión. Además, para una descripción más específica, se toma en cuenta el propósito del estudio: pronóstico, terapéutica, causalidad y la evaluación de una prueba diagnóstica.

Palabras clave: Errores aleatorios; Errores sistemáticos; Sesgos; Investigación clínica; Revisión

¹Universidad Nacional Autónoma de México, Facultad de Medicina, Ciudad de México, México

²Instituto Mexicano del Seguro Social, Centro Médico Nacional Siglo XXI, Hospital de Pediatría, Ciudad de México, México

Correspondencia: Miguel Ángel Villasís-Keever.
miguel.villasis@gmail.com

Recibido: 29-10-2021

Aceptado: 29-10-2021

DOI: 10.29262/ram.v68i4.1003

Introducción

Desde hace cinco años, en los diferentes artículos sobre metodología de la investigación que se han publicado en *Revista de Alergia de México*, se han definido diversos aspectos para realizar estudios de investigación clínica de buena calidad. Así, se ha señalado la importancia de lo siguiente:

- Elegir el mejor diseño de investigación para responder a la pregunta de planteada.
- Identificar las características necesarias que ayudarán a distinguir adecuadamente la población de estudio.
- Calcular el número de participantes (tamaño de muestra) indispensable a fin de cumplir los objetivos de estudio.
- Evaluar con técnicas válidas las variables en investigación.
- Ejecutar cada etapa de la investigación con personal capacitado e instrumentos de medición óptimos, con el propósito de obtener datos confiables.
- Seleccionar adecuadamente la(s) prueba(s) estadísticas para presentar los resultados apropiadamente.^{1,2,3,4,5,6}

Una investigación de calidad es la que está libre de errores, para lograrlo cada uno de los seis puntos señalados deben ser considerados en la planeación de un estudio. A fin de conseguir un buen estudio, al elaborar el protocolo los autores deben tomar en cuenta los posibles errores que puedan cometerse durante la investigación. En general, se debe reconocer que en cualquier estudio pueden ocurrir errores aleatorios y errores sistemáticos; estos últimos se denominan sesgos. Lo anterior es trascendente cuando se revisa la publicación de los resultados de una investigación, ya que, si se detectan sesgos, es posible que no sea apropiado utilizar esa información para su aplicación en un paciente en particular en vista que el efecto a observar no será similar a lo señalado por los autores.^{7,8}

Este artículo tiene como objetivo describir los principales sesgos que pueden ocurrir en los estudios de investigación clínica, así como la forma de evitarlos o minimizar sus efectos.

Errores aleatorios

Los errores aleatorios se presentan en cualquier investigación y, en general, ocurren en el momento de realizar las mediciones, tanto por los instrumentos utilizados, como por las *variaciones normales* en los participantes de los estudios. Se considera que los datos obtenidos son reales cuando se toman en cuenta tres puntos principales:

- Se utiliza el instrumento óptimo para hacer la medición que se está evaluado.
- El instrumento se encuentra calibrado en el momento de realizar las mediciones.
- El personal que efectúa la medición tiene un método estandarizado.⁶

Por ejemplo, un lector podría asumir que el peso corporal obtenido será correcto, si en un protocolo se señala que para la medición se utilizará una báscula que se calibra cada seis meses, que el personal que lo realizará ha sido capacitado y que el participante en el estudio se medirá en la mañana, antes de consumir alimentos, con mínima ropa y sin zapatos.

A pesar de realizar los tres puntos señalados, es posible que después de hacer dos o más mediciones al mismo participante, repitiendo exactamente la misma técnica, se obtengan variaciones (mínimas) en los valores obtenidos. Esta variación es el error aleatorio. Regresando al ejemplo del peso; si en una primera medición se obtuvo un peso de 72.300 kg y en la segunda 72.250 kg, la variación será de 50 g, lo cual deberá considerarse como un error aleatorio.

En contraste con los sesgos, los errores aleatorios comprometen la confiabilidad, más que la validez. Esto significa que, los errores aleatorios se relacionan con la variabilidad de los resultados, lo cual se puede traducir como menor precisión. Esta variabilidad se puede reducir de diferentes formas. La primera es al disminuir la varianza (o variabilidad intrasujeto), por ejemplo, retomando la medición del peso, se recomienda llevar a cabo tres mediciones diferentes (usando la misma técnica) al mismo sujeto y, posteriormente promediarlas. De esta forma, al calcular el promedio se reduce la varianza. Otra manera de reducir la variabilidad es aumentando el tamaño de la muestra, ya que entre mayor sea el número de participantes la varianza disminuirá (menor variabilidad entre sujetos).^{6,9}

Errores sistemáticos o sesgos

Los sesgos son los errores que ocurren, generalmente, por problemas metodológicos y, a diferencia de los errores aleatorios, se pueden evitar.¹⁰ Para que un estudio esté libre de sesgo, es necesario identificarlos y eliminarlos durante la planeación y elaboración de un protocolo; sin embargo, es posible que no se pueda descartarlos, pero dado que ya se sabe de su existencia, los autores tienen la oportunidad de minimizar sus efectos en el momento de llevar a efecto el análisis estadístico, mediante el uso de modelos multivariantes, como el análisis de regresión logística, regresión lineal, o el de riesgos proporcionales de Cox.⁸

Cuando un estudio se realizó con uno o más sesgos, la validez de los hallazgos está en duda, ya que el efecto que tiene su presencia radica en que pueden subestimar o sobreestimar los resultados obtenidos. Retomando el ejemplo del registro del peso, en caso de que no se tome en cuenta que el participante consumió alimentos antes de su medición, la consecuencia será que el peso que se obtenga probablemente será más alto. Esto puede ser trascendente si el estudio tiene como objetivo demostrar la eficacia de una dieta para disminuir peso en sujetos con obesidad; en los resultados se puede

encontrar equivocadamente que con la dieta no hubo cambio en el peso de los participantes. En los estudios de causalidad, la subestimación o la sobrestimación se observa con la disminución o aumento de los valores de razón de momios (RM) o riesgo relativo (RR), respectivamente. Pero también puede suceder que se impida demostrar que existe asociación de un factor de riesgo con el desarrollo de una enfermedad o, por el contrario, encontrar una asociación espuria.

Aunque un estudio con sesgos conlleva a no confiar en los resultados del estudio porque puede ser que no sean verdaderos, existen situaciones particulares que esto no ocurre. A esta última situación algunos autores han optado por denominar el efecto de los sesgos en dos categorías: sesgos *diferenciales* y *no diferenciales*. Los primeros son los que, con la presencia de sesgo(s) origina que los resultados sean equivocados (*diferentes* a lo obtenido). Mientras que, en los sesgos no diferenciales, los resultados se podrán asumir como verdaderos; es decir, su presencia no modifica los resultados.¹¹

Clasificación de los sesgos

El estudio de los sesgos tiene varias décadas, por lo que se han identificado múltiples; para el año 2010, mediante una revisión de la literatura, se describieron 235 tipos de sesgos diferentes.¹² Por el número de sesgos, en el pasado ya se han descrito diferentes formas de clasificarlos;⁸ para los fines de la presente publicación, pero incluyendo exclusivamente los sesgos que ponen en riesgo la validez interna, los clasificaremos desde el punto de vista clínico-epidemiológico en tres grupos principales:^{9,10,11}

- a) *Sesgos de selección*. Este tipo de errores ocurren por diversas circunstancias, ya sea por la forma que fueron elegidos los participantes en el marco de muestreo (casos consecutivos, en lugar de haber sido mediante aleatorización). También cuando los criterios para ingresar al estudio no están suficientemente definidos o porque los investigadores no siguen apropiadamente los procedimientos diagnósticos; lo cual puede derivar clasificar equivocadamente a los participantes (mala clasificación), por ejemplo, identificar al participante como sanos, cuando en realidad es un enfermo, o viceversa.
- b) *Sesgos de información o ejecución*. Este grupo corresponde a los errores que suceden inmediatamente después de la selección de los participantes, es decir, en la fase de la recolección de los datos en un proyecto de investigación. En general, se refiere a cualquier condición que influya sobre la calidad de la obtención de la información o de los datos que constituyen las variables de interés. En este grupo se incluyen problemas relacionados con la falta de estandarización para la recolección de los datos, por parte de los investigadores, la utilización de instrumentos que no son los óptimos (o válidos) para las mediciones a realizar, o bien, a pesar de que los instrumentos

sean adecuados, no se calibren o tengan el mantenimiento requerido; o bien, generados por los participantes en el estudio, cuando no se consideran situaciones que afectan los resultados de estudios, por ejemplo, tiempo de ayuno insuficiente antes de tomar muestras sanguíneas, o la hipertensión de bata blanca.

Considerando esto último, los sesgos se pueden evitar cuando se dan instrucciones precisas a los participantes y al seguir estrictamente las pautas estipuladas en el protocolo de investigación.⁶ Además, es conveniente señalar que, en los sesgos de información también puede ocurrir la *mala clasificación*. Esto es trascendente en los estudios donde la variable dependiente principal es la condición clínica de los participantes, al término del estudio. Por ejemplo, cuando los investigadores desean identificar si hubo mayor tasa de curación por el empleo de un nuevo tratamiento en comparación con el tratamiento habitual, pero no tienen un proceso estandarizado o utilizan instrumentos inapropiados para determinar el efecto terapéutico, entonces la clasificación de un paciente como curado o no curado, puede ser errónea.

- c) *Sesgos de confusión*: Estos sesgos se presentan durante el análisis e interpretación de los datos, ya que se puede llegar a inferencias erróneas o espurias. Se pueden presentar en cualquier diseño de investigación. Por ejemplo, los autores pueden asumir que existe una asociación causal entre una determinada exposición (o factor de riesgo) y el evento de interés, pero sin considerar las variables de confusión, por lo cual, dicha asociación no es correcta.

Esta primera agrupación permite disponer de un panorama general para ubicar los múltiples sesgos. Sin embargo, en cualquiera de los tres grupos se integran aún múltiples sesgos, por lo cual, parece conveniente tratar de diferenciarlos aún más para disponer de una visión más práctica cuando se está elaborando un protocolo de investigación, o bien, en la lectura de los resultados de una investigación. De esta forma, a continuación nos concentraremos a describir los sesgos que consideramos más importantes, tomando en cuenta el propósito de la investigación, pero haciendo la aclaración de que existen mucho más, por lo que los lectores interesados deberán consultar otras fuentes.

Sesgos de acuerdo con el propósito de la investigación

Los diseños de acuerdo con el propósito del estudio son los siguientes:

- Estudios para determinar la eficacia/efectividad de una intervención (ya sea terapéutica, preventiva o educativa).
- Estudios para conocer los pronósticos de las enfermedades.

Cuadro 1. Principales sesgos de acuerdo con el propósito del estudio

Propósito de la investigación: mejor diseño	Sesgos de selección	Sesgos de información o ejecución
Pronóstico		
Estudio de una cohorte	• Susceptibilidad pronóstica • Tiempo de espera (<i>lead-time*</i>)	• De detección • De desgaste (<i>attrition*</i>)
Terapéutica		
Ensayo clínico aleatorizado	• Susceptibilidad pronóstica • Asignación	• De desgaste • De detección • Reporte
Causalidad		
Estudio de cohortes	• Mala clasificación	• De detección • Migración • Desgaste
Casos y controles	• Mala clasificación • Por tipo de admisión o de Berkson (<i>admission rate*</i>) • De prevalencia-incidencia o de Neyman	• De detección • Del recuerdo o memoria
Evaluación de una prueba diagnóstica		
Transversal analítico	Espectro de la enfermedad	• Verificación (parcial y deferencial) • Incorporación • Sospecha diagnóstica • Resultados no concluyentes

*Término usado en idioma inglés.

- Estudios para establecer las causas o factores de riesgo de las enfermedades.
- Estudios que evalúan la precisión o exactitud (sensibilidad y especificidad) de las pruebas diagnósticas.
- Estudios sobre frecuencia (prevalencia, incidencia) de las enfermedades.^{1,2}

Los sesgos particulares se describen tomando en cuenta los mejores diseños para resolver los propósitos. Por ejemplo, el ensayo clínico aleatorizado (ECA) es el ideal para determinar el beneficio de una intervención. En el Cuadro 1 se señalan los sesgos por el propósito del estudio.

Sesgos en estudios pronóstico

Los estudios de pronóstico tienen, en general, tres objetivos:

- Describir la evolución de un grupo de pacientes, a partir de que iniciaron a estar enfermos.

- Identificar los factores asociados con el pronóstico.
- Generar escalas de pronóstico. El estudio de una cohorte constituye el mejor diseño para identificar el pronóstico de las enfermedades; el estudio inicia con la identificación de pacientes que ya tienen alguna enfermedad, tratando de incluir sujetos en un mismo estadio de la enfermedad. Posteriormente los pacientes son vigilados durante el tiempo suficiente para que desarrollen la variable de resultado (por ejemplo, muerte/supervivencia, complicaciones, progresión de la enfermedad, curación, etcétera).¹³

Los cuatro sesgos (Cuadro 1) que consideramos son los más trascendentes en los estudios de pronóstico, son dos sesgos de selección y dos de ejecución.¹³ El sesgo de *susceptibilidad pronóstica* tiene que ver con el estadio o la gravedad de los pacientes en estudio. Cuando se incluyen pacientes con diferentes estadios o grados, entonces el pronóstico será diferente: cuanto más graves, mayor número de complicaciones

se observarán, o bien, en un tiempo más corto. Lo contrario puede ocurrir en etapas tempranas de la enfermedad. Sin embargo, los estudios con pacientes en una misma etapa de la enfermedad son más difíciles de realizar porque se requiere más tiempo para el reclutamiento. Ante esta situación, los resultados podrán considerarse como válidos, siempre y cuando, los autores describan independientemente el pronóstico de los pacientes de acuerdo con el estadio o gravedad que presentaban al momento del inicio del estudio, por ejemplo, mediante análisis estratificado.

Sesgo de tiempo de espera (lead-time bias). Los pacientes que se diagnostican mediante estudios de escrutinio es posible que tengan un mayor tiempo de supervivencia, en comparación con quienes se diagnostican habitualmente, es decir, cuando se presentan los primeros signos o síntomas de la enfermedad. Esta diferencia en el periodo de tiempo puede conducir a estimar erróneamente la probabilidad de cuánto se espera que un individuo pueda vivir o morir, una vez que se diagnostica la enfermedad.

El *sesgo de detección* se refiere a dos condiciones que ocurren durante la fase de selección de los pacientes a incluir en el proyecto de investigación. La primera condición para asegurar un estudio de buena calidad y el diagnóstico de la enfermedad en dicho estudio es que debe realizarse con los criterios más apropiados y actualizados. Mientras que la segunda condición puede ocurrir al inicio del seguimiento o vigilancia; en un estudio sin este sesgo, ningún paciente debe presentar la variable de desenlace en evaluación (por ejemplo, daño renal en pacientes con diabetes).

En cuanto a los dos sesgos de ejecución o información, es importante destacar que también puede estar presente el *sesgo de detección*. A diferencia de los descrito en el párrafo anterior, el error sistemático se presenta cuando se identifica, intencionadamente, la variable de desenlace en los pacientes con mayor susceptibilidad para presentarla. Por ejemplo, si en el momento de realizar la evaluación de una complicación se conoce que el paciente se encuentra en un estadio más avanzado de la enfermedad, será más probable que asuma que ya tiene la complicación. Esta situación sucede con mayor frecuencia cuando la forma de evaluar la variable de desenlace *no es dura* (como los resultados de un examen de laboratorio), y requiere la participación del investigador (por ejemplo, exploración física en la búsqueda de signos anormales, o la interpretación de estudios de rayos X). Para evitar este sesgo, todos los encargados de evaluar la variable de desenlace deberán estar cegados, es decir, desconocerán las condiciones de los pacientes, a fin de obtener información (más) objetiva.

Sesgo de desgaste (attrition): en los estudios donde se evalúa la evolución de los pacientes, es necesario que todos los participantes incluidos continúen su vigilancia durante el tiempo mínimo necesario para que aparezca la variable de resultado. En los estudios que requieren mayor tiempo de vi-

gilancia, hay más probabilidad de que los pacientes no regresen y, entonces, se desconocerá si desarrollaron o no la variable de resultado. Esto último se denomina pérdidas durante el seguimiento, lo cual corresponde al sesgo de desgaste; por supuesto, existe un gradiente ya que entre mayor número de pérdidas será menor la validez del estudio.

Sesgos en estudios sobre terapéutica

Como se comentó, el ECA es el mejor estudio para determinar la eficacia o efectividad de una maniobra o intervención; ya sea que la intervención tenga fines terapéuticos, preventivos o educativos. En los ECA, la forma de evaluar si dicha intervención sirve es al comparar la proporción tanto de efectos benéficos como de efectos colaterales entre un grupo que recibe (grupo experimental) y otro que no recibe (grupo control) la intervención en estudio. Cuando, en comparación al grupo control, en el grupo experimental se observa mayor número de beneficios y menor proporción de efectos adversos, entonces podremos confirmar que la intervención es efectiva o eficaz. Se debe recordar que la eficacia es cuando el grupo control recibe placebo; mientras que si recibe un tratamiento similar, entonces se denomina efectividad.¹⁴

En la Cuadro 1 se señalan los cinco sesgos específicos en los estudios sobre terapéutica: uno de selección y cuatro de ejecución.^{15,16} El de selección se puede denominar como *sesgo de susceptibilidad pronóstica*, lo cual significa la posibilidad de que los elegidos para estar en cualquiera de los dos grupos tengan ciertas características clínicas que los haga más o menos propensos para que en los participantes en el estudio se observen los beneficios de la intervención en estudio. Este sesgo se evita cuando se lleva a cabo una apropiada la aleatorización, es decir, otorgar la misma probabilidad de asignar a los participantes al grupo experimental o al grupo placebo. Se espera que la aleatorización haga homogéneos ambos grupos al momento del inicio de la investigación, para que las características clínicas sean similares. Sin embargo, se debe tomar en cuenta que a pesar de que los autores señalan que se realizó aleatorización es posible que no se cumpla; esto sucede cuando los investigadores conocen la secuencia de dicha aleatorización y, entonces pueden cambiar del grupo asignado a uno o más participantes, al suponer que puede ser mejor dar el tratamiento experimental o el tratamiento en comparación. Esto se denomina *sesgo de asignación (allocation, en inglés)*; para evitarlo, la mejor estrategia es tener una tercera persona que otorgue el tratamiento, tal y como fue planteado por la aleatorización.

Sesgo de desempeño. En un ECA, el efecto que de una intervención tiene que ver primordialmente con el cumplimiento terapéutico, de tal forma que los participantes sigan las indicaciones. Por ejemplo, si se trata de la evaluación de un fármaco, los participantes deberán tomar las dosis en el horario y número que fueron prescritas, durante todo el tiempo de duración del estudio. En caso de desviaciones, tanto en

el grupo experimental como en el control, los resultados no serán confiables. El sesgo se evita cuando los participantes están cegados (desconocen a qué grupo pertenecen) impide que se deje de tomar uno u otro tratamiento. Esto se puede lograr si ambos tratamientos o intervenciones son iguales (por ejemplo, en caso de fármacos, las tabletas serán del mismo color y sabor). También este cegamiento ayuda evitar que el personal médico (diferente al grupo de investigación) ofrezca otros tratamientos que puedan potenciar o disminuir el efecto(s) terapéutico(s) o adverso(s) en evaluación.

Sesgo de detección. Esto es similar a lo que se comentó en los estudios de pronóstico. Como se señaló, para evitar este sesgo es esencial el cegamiento de quienes evaluarán la(s) variable(s) de resultado. En los ECA se refiere específicamente a que los investigadores desconocerán si los participantes corresponden al grupo con la maniobra experimental o al grupo control.

Sesgo de desgaste. Es igual a los estudios de pronóstico; sin embargo, en los ECA se debe tomar en cuenta que, en los pacientes considerados como pérdidas, es indispensable conocer las razones de la discontinuación del tratamiento (tanto del grupo experimental como del control). Por ejemplo, los pacientes decidieron suspender el tratamiento por los efectos adversos de la intervención en evaluación, o porque los pacientes no mejoraron o se agravaron. Esto podría ser muy trascendente para la evaluación global de la intervención, por lo que, para evitar este sesgo, el análisis estadístico a realizar debe ser por *intención a tratar*. Esto significa que todos los participantes aleatorizados deben ser considerados que terminaron el estudio, tal como fue planeado. Cuando el análisis se lleva a cabo *por protocolo* (se incluye solo los participantes que terminaron apropiadamente), es posible que los resultados indiquen que la intervención sea efectiva o eficaz, lo cual podría ser incorrecto.¹⁷

Sesgo de reporte. Esto ocurre en el momento de terminar el estudio, cuando los autores deciden seleccionar algunas de las variables de resultado, en lugar de informar los resultados de todas las variables que fueron consideradas inicialmente en el protocolo. Generalmente, esta situación se presenta porque algunos datos obtenidos no corresponden con las hipótesis formuladas durante la planeación del estudio. Si bien, este fenómeno parece que ha ido disminuyendo, se debe recordar que, en el mundo, existe una tendencia a publicar únicamente lo estadísticamente significativo. Para verificar la existencia de discrepancias de los reportes finales con el protocolo, se pueden consultar en plataformas, como Clinical Trials (<https://clinicaltrials.gov>).

Sesgos en estudios sobre causalidad

Estos estudios tienen como propósito establecer la(s) posible(s) causa(s) de una enfermedad. Conviene señalar que los factores de riesgo también se pueden considerar como causa(s). Para este tipo de investigación, se reconocen dos

diseños como los mejores: estudios de cohortes, así como los de casos y controles. Para ejemplificar, usaremos el tabaquismo de los padres como factor de riesgo para el desarrollo de asma en niños. En los *estudios de cohortes*, los autores tratan de identificar si un factor puede producir una enfermedad, mediante el seguimiento de sujetos sanos. Al inicio, se establecen dos grupos de acuerdo con la presencia o no del factor en estudio: grupo expuesto (o cohorte) serán niños en edad preescolar que tienen el factor de riesgo (en el ejemplo, tabaquismo de los padres), y el grupo no expuesto (sin tabaquismo). Posteriormente, ambos grupos se vigilarán durante un tiempo —meses o años— para determinar si aparece la enfermedad (en nuestro ejemplo, asma).¹¹

En contraste, en los diseños de *casos y controles* la investigación se inicia en el desenlace, es decir, la enfermedad en estudio y no hay seguimiento de los participantes (estudio transversal). Los casos serán quienes ya tienen la condición (pacientes asmáticos) y los controles serán niños sanos. En general, la búsqueda del factor de riesgo es mediante la indagatoria de lo ocurrido en el pasado; en nuestro ejemplo, a los padres se le preguntará si han fumado y su intensidad.

En vista que estos dos diseños son muy distintos, en el Cuadro 1 se señalan los sesgos independientemente. Como se puede observar, en general, los sesgos en los estudios de cohorte son muy parecidos a los estudios de pronóstico, pero tienen particularidades que a continuación señalamos.

En los *estudios de cohortes* para determinar si un factor de riesgo incrementa la probabilidad del desarrollo de una enfermedad, en el inicio del estudio, los participantes generalmente son sujetos sanos. Uno de los sesgos de selección ocurre cuando uno o más participantes ya tienen la enfermedad en estudio (en nuestro ejemplo, asma), es decir, en estos sujetos ya ocurrió el desenlace antes de iniciar su vigilancia. Si este se presenta habrá una mala clasificación como señalar que un sujeto es sano cuando no lo es. Particularmente, este sesgo tendrá mayor trascendencia cuando hay mayor número de sujetos mal clasificados en el grupo expuesto. Retomando nuestro ejemplo, el efecto del tabaquismo sobre el asma estará sobrestimado.

El otro sesgo de selección es sobre la integración de los grupos, basado en la exposición. Si el grupo expuesto al factor de riesgo no se distingue apropiadamente del grupo no expuesto, también puede ocurrir sesgo de mala clasificación. De esta forma, se podrá aseverar erróneamente que habrá participantes que estén expuestos, y viceversa. En esta situación, este sesgo condiciona que la exposición se subestime y, entonces, no se observe el efecto real del factor. En nuestro ejemplo, el tabaquismo no será descrito como un factor de riesgo para el asma, o bien, el estimado de riesgo obtenido (RM o RR) será menor.

Sesgo de migración. En los estudios de cohortes, la asociación del factor de riesgo con el desarrollo de la enfermedad se establece cuando, en comparación con el grupo no

expuesto, existe una mayor proporción de pacientes, o bien, mayor valor de RM o RR en el grupo expuesto. Esto debería ser sí el factor de exposición se mantiene estable durante toda la vigilancia de los participantes. Sin embargo, es posible que, a pesar de que la hipótesis sea correcta, en los resultados no se encuentre dicha asociación, dado que hubo sujetos en la cohorte expuesta que migraron o pasaron al grupo no expuesto, y viceversa. En nuestro ejemplo, padres que fumaron dejan de hacerlo en algún momento después del inicio de la vigilancia. Como puede deducirse, la asociación se subestimará como consecuencia de la presencia de este sesgo.

Sesgo de desgaste. El efecto de este error es igual a lo comentado en los estudios sobre terapéutica y pronóstico, entre mayor número de participantes no llegue a la conclusión del estudio, menor será la validez. Se presentará más en estudios que requieren vigilancia a largo plazo para observar el desenlace; por ejemplo, se requieren años de vigilancia de pacientes fumadores, para observar casos de cáncer de pulmón asociados a al consumo de tabaco.

En cuanto a los estudios de *casos y controles*, también en el Cuadro 1 se señalan los principales sesgos. Es interesante conocer que, históricamente, este diseño es uno dónde se han descrito más sesgos; en particular, los relacionados con la manera de identificar la presencia o no de los factores de riesgo, dicho de otra manera, durante la ejecución del estudio.

Los sesgos de selección en casos y controles se relacionan con la posibilidad de una mala clasificación. Como en lo señalado en otros diseños, se puede clasificar equivocadamente a un participante, ya sea como caso (paciente *con* la enfermedad en estudio) o como control (paciente sano o *sin* la enfermedad en estudio). Igualmente, para impedir una mala clasificación, los investigadores deben conformar ambos grupos mediante métodos y criterios válidos, así como estandarizar al personal encargado de evaluar a los potenciales participantes en el estudio.

En cuanto a los sesgos de ejecución, debemos señalar que los incluidos en el Cuadro 1 no son todos los descritos en la literatura, sino que solamente los considerados más representativos. **Sesgo por tipo de admisión o Berkson** (*admission rate bias*). El sesgo ocurre cuando se toma como controles a pacientes hospitalizados, en lugar de ser controles de la comunidad de origen de los casos. En general se considera que, independientemente de la enfermedad de base, los pacientes hospitalizados (más cuando el estudio se realiza en la misma unidad médica) comparten muchas características, por lo que será difícil encontrar alguna asociación entre el factor de riesgo y la enfermedad, dado que los casos y los controles son muy parecidos entre sí. De esta forma, este sesgo subestima la asociación que está en investigación.

Sesgo de prevalencia-incidencia o sesgo de Neyman. El sesgo se presenta cuando se elige únicamente a los pacientes disponibles en el momento de realizar el estudio (casos prevalentes), dado que no se dispondrá de pacientes con todo el

espectro de la enfermedad. Este sesgo puede tener dos implicaciones: la primera, al no incluir pacientes fallecidos, por ser más graves o en etapas tardías, la enfermedad parecerá menos grave; la segunda, al no seleccionar pacientes que se mejoraron, entonces la enfermedad se considerará más grave. Este sesgo se puede evitar cuando se incluyen tanto casos prevalentes como casos nuevos (incidentes).

De los sesgos de ejecución, en los estudios de casos y controles se destaca el *sesgo de memoria o de recuerdo*. Cuando se interroga sobre la posible exposición al o los factores de riesgo, es más probable que quienes tienen la enfermedad (casos) lo(s) evoque(n) con más facilidad que quienes no la tienen (controles). En presencia de este sesgo, se sobrestima la asociación. Para tratar de minimizar el efecto de la memoria, la entrevista debe estar estandarizada y quienes interrogan a los participantes deberán estar cegados (desconocerán que se trata de caso o control). Otra forma de evitar el sesgo es identificar la presencia del factor de riesgo utilizando expedientes clínicos de los participantes, o bien, haciendo evaluaciones más objetivas. Un ejemplo de esto último es medir los metabolitos de la nicotina (cotinina) en sangre u orina, además de haber interrogado sobre la frecuencia y grado del tabaquismo, cuando se trata de asociar el consumo de tabaco con el asma.

Por último, el sesgo de detección también puede presentarse en ambos diseños, cuando los investigadores no están cegados en el momento de recolectar la información sobre el desenlace (estudios de cohortes) o de los factores de riesgo (casos y controles).

Sesgos en estudios para evaluación de pruebas diagnósticas

Ante un paciente, en la clínica es habitual solicitar estudios para confirmar (sensibilidad) o descartar (especificidad) alguna enfermedad. Entre otros, estos estudios pueden ser exámenes de laboratorio, radiografías, ultrasonido, o endoscopías. El diseño transversal y comparativo es el ideal cuando el propósito de una investigación es establecer el desempeño, precisión o la exactitud (*accuracy*) que tiene una prueba diagnóstica en evaluación, en términos de sensibilidad y especificidad. Estos estimados se establecerán al comparar los resultados de la prueba en evaluación, con lo que se obtendrán al aplicar un estándar de referencia, también llamado estándar de oro, el cual corresponde al mejor estudio para comprobar la presencia o no de una enfermedad.¹⁸

Al igual que otros diseños, los efectos que tienen los sesgos se observan en los resultados; para este tipo de investigación significa que los valores de sensibilidad o especificidad no serán correctos, ya sea porque se obtendrán datos que los sobre o subestimen. Con relación a este último respecto, los invitamos a revisar una publicación reciente de esta misma serie, en el cual se detallan y ejemplifican las consecuencias de los diferentes sesgos.¹⁹

Sesgo de espectro de enfermedad. En este tipo de investigación se requiere conformar dos grupos: uno con enfermos y otro con sujetos sin la enfermedad en estudio. El mejor estudio es el que se realiza en quienes tienen alguna sospecha (baja o alta) de una enfermedad, pero entre más diferentes sean los dos grupos (por ejemplo, enfermos graves vs. sujetos sanos) puede ocurrir sesgo, lo cual puede sobreestimar la sensibilidad y especificidad.¹⁷ También los valores pueden ser diferentes de acuerdo con el lugar dónde se realiza la investigación: la sensibilidad aumenta cuando se lleva a cabo exclusivamente en población de hospitales de referencia; en contraste, la especificidad se incrementa con poblaciones de unidades médicas de primer nivel.

Sesgo de verificación. En el estudio ideal, todos los enfermos y lo no enfermos se les efectuará tanto el estándar de oro como la prueba en evaluación (también llamada prueba índice). Este sesgo se presenta cuando no todos se les aplica el estándar de oro, de acuerdo con el protocolo; ante esta situación, no habrá certeza si los participantes tienen o no la enfermedad. Existen dos variantes de este sesgo, la primera es el *sesgo de verificación parcial* en dónde solamente se realiza el estándar de oro, si la prueba índice es positiva. La segunda es cuando a los participantes se les aplican diferentes pruebas consideradas como estándar de oro, lo que se denomina *sesgo verificación diferencial*.

Sesgo de incorporación. En un buen diseño, tanto el estándar de oro como la prueba índice deben ser independientes. Este sesgo se presenta en situaciones dónde la prueba índice forma parte del estándar de oro. Como consecuencia, los valores de sensibilidad y especificidad por obtener no serán confiables.

Sesgo por estándar de referencia inapropiado. Existen enfermedades para las cuales no se dispone de una prueba diagnóstica única que indique si un sujeto realmente tiene o no la condición en estudio, entonces, en el contexto clínico se hacen algunas aproximaciones (por ejemplo, lo que se conoce como prueba terapéutica; la enfermedad se confirma si el paciente mejora con el tratamiento, y viceversa). En este contexto, los valores de sensibilidad y especificidad de la prueba índice pueden tener variaciones importantes. Efectos similares se obtendrán en el *sesgo de resultados no concluyentes*, lo cual se refiere a que los valores de la prueba índice no discriminan apropiadamente los enfermos de los no enfermos.

Sesgo de sospecha diagnóstica. En los mejores estudios de prueba diagnóstica, los investigadores que evalúan la prueba índice deben ser diferentes de quienes evalúan el estándar de referencia. Ambos deben desconocer (cegamiento) el resultado de una y otra para que la interpretación de una u otra prueba sea lo más objetiva posible, ya que conocer con anticipación el resultado de la otra prueba puede influir para la decisión de confirmar o descartar la enfermedad.

Conclusiones

Los sesgos que ocurren en los estudios de investigación clínica hacen perder su validez interna, por lo que resultados obtenidos no podrán ser considerados como reales. Por lo anterior, cuando se elabora un protocolo siempre es necesario tenerlos en cuenta para tratar de evitarlos y mejorar la calidad de los estudios. Asimismo, los lectores de estudios de investigación deberán estar familiarizados con estos sesgos a fin de estar en posibilidad de juzgar la pertinencia de utilizar la información en la práctica cotidiana.

Referencias

1. Miranda-Navales MG, Villasís-Keever MÁ. El protocolo de investigación. Parte I. Rev Alerg Mex. 2015;62(4):312-7. Disponible en: <https://revistaalergia.mx/ojs/index.php/ram/article/view/129/258>
2. Villasís-Keever MÁ, Miranda-Navales MG. El protocolo de investigación II: los diseños de estudio para investigación clínica. Rev Alerg Mex. 2016;63(1):80-90. DOI: 10.29262/ram.v63i1.163
3. Arias-Gómez J, Villasís-Keever MÁ, Miranda-Navales MG. El protocolo de investigación III: la población de estudio. Rev Alerg Mex. 2016;63(2):201-206. Disponible en: <https://revistaalergia.mx/ojs/index.php/ram/article/view/181/309>
4. Villasís-Keever MÁ, Miranda-Navales MG. El protocolo de investigación IV: las variables de estudio. Rev Alerg Mex. 2016;63(3):303-310. DOI: <https://doi.org/10.29262/ram.v63i3.199>
5. Flores-Ruiz E, Miranda-Navales MG, Villasís-Keever MÁ. El protocolo de investigación VI: cómo elegir la prueba estadística adecuada. Estadística inferencial. Rev Alerg Mex. 2017;64(3):364-370. DOI: 10.29262/ram.v64i3.304
6. Villasís-Keever MÁ, Márquez-González H, Zurita-Cruz JN, Miranda-Navales G, Escamilla-Núñez A. El protocolo de investigación VII. Validez y confiabilidad de las mediciones. Rev Alerg Mex. 2018; 65(4): 414-421.
7. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. Lancet. 2017;390(10092):415-423. DOI: 10.29262/ram.v65i4.560
8. Lijndijk HJ, Page MJ, Burger H, Koolman X. Assessing risk of bias: a proposal for a unified framework for observational studies and randomized trials. BMC Med Res Methodol. 2020;20(1):237. Disponible en: <https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-020-01115-7>
9. Sessler D, Imrey P. Clinical research methodology 1: study designs and methodologic source of error. Anesth Analg. 2015;121(4):1034-1042. DOI: 10.1213/ANE.0000000000000815
10. Delgado-Rodríguez M, Llorca J. Bias. J Epidemiol Community Health. 2004;58(8):635-641. DOI: 10.1136/jech.2003.008466
11. Hernández-Ávila M, Garrido F, Salazar-Martínez E. Sesgos en estudios epidemiológicos. Salud Publica Mex. 2000;42(5):438-

446. Disponible en: <https://scielosp.org/pdf/spm/2000.v42n5/438-446/es>
12. Chavalarias D, Ioannidis JP. Science mapping analysis characterizes 235 biases in biomedical research. *J Clin Epidemiol*. 2010;63(11):1205-1215. DOI: 10.1016/j.jclinepi.2009.12.011
13. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Assessing bias in studies of prognostic factors. *Ann Intern Med*. 2013;158(4):280-286. DOI: 10.7326/0003-4819-158-4-201302190-00009
14. Zurita-Cruz JN, Márquez-González H, Miranda-Novales G, Villasís-Keever MÁ. Estudios experimentales: diseños de investigación para la evaluación de intervenciones en la clínica. *Rev Alerg Mex*. 2018;65(2):178-186. DOI: 10.29262/ram.v65i2.376
15. Jørgensen L, Paludan-Müller AS, Laursen DR, Savović J, Boutron I, Sterne JA, et al. Evaluation of the Cochrane tool for assessing risk of bias in randomized clinical trials: overview of published comments and analysis of user practice in Cochrane and non-Cochrane reviews. *Syst Rev*. 2016;5:80. DOI: 10.1186/s13643-016-0259-8
16. Hróbjartsson A, Boutron I, Turner L, Altman DG, Moher D, Cochrane Bias Methods Group. Assessing risk of bias in randomised clinical trials included in Cochrane Reviews: the why is easy, the how is a challenge. *Cochrane Database Syst Rev*. 2013;(4):ED000058. DOI: 10.1002/14651858.ED000058
17. Porta N, Bonet C, Cobo E. Discordance between reported intention-to-treat and per protocol analyses. *J Clin Epidemiol*. 2007;60(7):663-669. DOI: 10.1016/j.jclinepi.2006.09.013
18. Rendón-Macías ME, Villasís-Keever MÁ. Fases para determinar la utilidad clínica de las pruebas diagnósticas. *Rev Alerg Mex*. 2020;67(3):279-285. DOI: 0000-0002-8566-0811
19. Rendón-Macías ME, Valenzuela M, Villasís-Keever MÁ. Sesgos en los estudios de pruebas de diagnóstico: implicación en la estimación de la sensibilidad y especificidad. *Rev Alerg Mex*. 2020;67(2):165-173. DOI: 10.29262/ram.v67i2.771

ORCID

Jessie Nallely Zurita-Cruz, 0000-0003-1389-7964; Miguel Ángel Villasís-Keever, 0000-0002-8566-0811