

L. García Marcos<sup>1</sup>, J. Guillén Pérez<sup>2</sup>,  
G. Orejas Rodríguez-Arango<sup>3</sup>

*An Esp Pediatr 1999;50:519-524.*

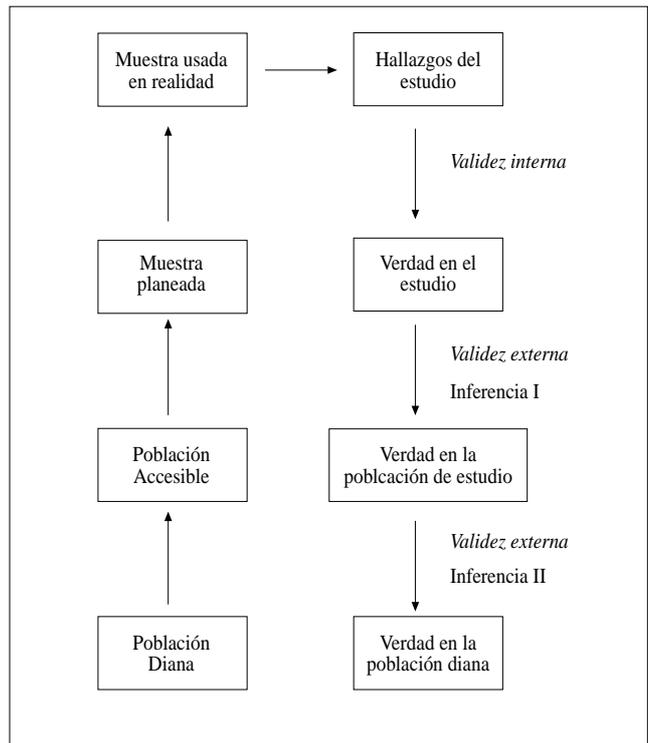
Un propósito fundamental de los estudios científicos en las ciencias de la salud es poder hacer generalizables los resultados. Esta generalización supone dos presunciones: la primera, que no puede trabajarse habitualmente con toda la población que se pretende estudiar y que hay que acudir, por tanto, a una muestra; y la segunda, consecuencia de la primera, que el diseño del estudio es lo suficientemente adecuado como para que los resultados obtenidos en la muestra sean un reflejo de lo que ocurre en toda la población. Dicho esto último de otra forma, que desde una muestra se puedan inferir los resultados a una población. Sin embargo, antes de inferir los resultados de una muestra a la población hay que asegurarse de que los resultados de la muestra están libres de error. El diseño de cualquier estudio debe intentar no incurrir en esos errores internos ni en aquellos que hicieran imposible la mencionada generalización o inferencia.

**Validez de la investigación. Validez interna y validez externa**

La validez de una investigación es la propiedad que ésta tiene de ser veraz o de haber obtenido resultados “verdaderos”, es decir, adecuados con la realidad. Clásicamente se distingue entre validez interna y validez externa. La validez interna hace referencia a la validez de los resultados obtenidos en la muestra estudiada (ausencia de errores internos). Este tipo de validez podría plantearse como: ¿son verdaderos los resultados que he obtenido en la muestra estudiada? o, por el contrario, ¿se han detectado errores (como, por ejemplo, un aparato mal calibrado, o un alejamiento sustancial del diseño inicial el estudio) que hacen que los resultados obtenidos no reflejen fielmente la realidad de la muestra diseñada?.

La validez externa es la capacidad de hacer generalizables los resultados obtenidos en la muestra hacia la población general. Puesto en forma de pregunta: estos resultados obtenidos en mi muestra ¿son un reflejo fiel de lo que ocurre en la población general? o, lo que es lo mismo, si hubiera estudiado toda la población, ¿hubiera obtenido los mismos resultados que en esta muestra? En la figura 1 se ilustran estos conceptos.

**Epidemiología y metodología aplicada a la pediatría (V): Sesgos**



**Figura 1.** Representación gráfica de los conceptos de validez externa, validez interna e inferencia.

**Concepto de error aleatorio y de error sistemático**

Los errores cometidos en el curso de una investigación y que interfieren en la validez del mismo pueden ser de dos tipos: errores aleatorios y errores sistemáticos (llamados comúnmente sesgos). Un error aleatorio, como su nombre indica, es un error debido al azar, y, como consecuencia, no evitable, aunque se debe contar con él, y las pruebas estadísticas así lo hacen. Por el contrario, el sesgo es un error sistemático, es decir, un error que se produce de forma constante y siempre en el mismo sentido (sistemático) en alguno o varios pasos del estudio. Suele tratarse, por tanto, de un error en el diseño de la investigación, que podría haberse evitado.

Por ejemplo, si la prevalencia verdadera de asma en España en la edad de 13-14 años es del 10%, en una muestra bien di-

<sup>1</sup>U.D. Pediatría. Universidad de Murcia y Unidad de Investigación de Cartagena.  
<sup>2</sup>U.D. Medicina Preventiva y Salud Pública. Universidad de Murcia.  
<sup>3</sup>Clinamat-Medycsa. Madrid.  
Correspondencia. Prof. Luis García Marcos. Unidad de Investigación. Pza. San Agustín, 3. 30201 Cartagena.

señada se podría encontrar un porcentaje del 12% o del 13%, o incluso aún más alejado del porcentaje real. Estas diferencias podrían deberse únicamente al azar y son consecuencia de que al trabajar con una muestra, ésta se encuentra aleatoriamente "inclinada" hacia uno u otro lado del porcentaje de la población. Las probabilidades de que el porcentaje hallado en la muestra sea lo más parecido al de la población aumentan a medida que el tamaño de la muestra se aproxima al de la población. Es decir, al aumentar el tamaño de la muestra disminuimos la probabilidad de error aleatorio, o lo que es lo mismo, aumentamos la precisión del estudio.

Sin embargo, no sólo el azar puede apartar a un estudio de la realidad. En el caso anterior, puede que estemos intentando medir la prevalencia del asma con un instrumento poco adecuado -una encuesta mal diseñada, por ejemplo- y que nuestros resultados no estén reflejando la realidad de la enfermedad asmática: podríamos estar estimando como asma lo que no lo es. En tal caso, no estamos estudiando lo que queremos estudiar y nos encontramos ante un sesgo, y por mucho que aumentemos el tamaño de la muestra no nos acercaremos al 10% que es la prevalencia real de asma. La manera de evitar estos errores sistemáticos es mejorar la exactitud del estudio (propiedad de medir lo que realmente se quiere medir), diseñándolo concienzudamente y utilizando los aparatos, medios y personas adecuados.

## Precisión

La precisión es la capacidad de un instrumento de medición (un aparato, una encuesta) para repetir el mismo valor cada vez que se mide lo mismo. Es evidente que determinados instrumentos de medición son intrínsecamente más precisos que otros: un tallímetro es más preciso que la encuesta de Graffar (método para determinar el nivel socioeconómico de una familia). La fiabilidad, reproducibilidad y coherencia de la medición son prácticamente sinónimos de precisión de la misma.

La precisión es una propiedad contraria a la de la variabilidad. Una manera de evitar errores aleatorios en los estudios científicos es reducir al máximo la variabilidad. La variabilidad debida al azar -como ya se ha dicho- puede combatirse aumentando el tamaño de la muestra. Sin embargo, en la mayoría de las ocasiones, bien por condicionamientos de presupuesto o de tiempo, se llega a un compromiso, aceptando la desventaja de una cierta variabilidad en aras de la gran ventaja de trabajar con una muestra reducida. Independientemente del azar, hay otras situaciones en las que se puede generar variabilidad y que pueden ser evitadas o reducidas.

## Fuentes de variabilidad en un estudio:

**Variabilidad debida al observador.** Se refiere a la diferente apreciación por parte de quien realiza la medición del fenómeno que está midiendo. Esta variabilidad puede depender desde la inflexión de la voz a quien le pasa una entrevista hasta los ruidos de fondo que pueden molestar a su oído para apreciar el momento de registrar la presión arterial. Cuando son varios los ob-

servadores, aunque lo que se esté midiendo sea lo mismo, cada uno de ellos puede no estar viendo exactamente lo mismo.

**Variabilidad debida al individuo.** Es la derivada al sujeto que se estudia, debida a cosas tales como el ritmo circadiano, cambio de humor, situación de nerviosismo o tranquilidad, etc.

**Variabilidad debida al instrumento de medición.** Cualquier instrumento de medición tiene unas fluctuaciones inherentes: por ejemplo, al igual que el esfigmomanómetro de mercurio es más preciso que el de reloj, una determinada encuesta puede ser más precisa para medir el agotamiento profesional que otra.

**Variabilidad debido a errores en el registro de los datos** o a su introducción en la base de datos electrónica. El cuidado en el registro de los datos en el momento en que éstos se producen, su automatización en la medida de lo posible, y la doble introducción en el ordenador y posterior comparación, son formas de evitar este error, bastante más frecuente de lo que a primera vista pudiera creerse.

Además de estas fuentes de variación más o menos evitables o reducibles, está la variación debida al registro equivocado de un dato.

## Determinación de la precisión:

**Desviación estándar y coeficiente de variación.** Como ya se ha comentado, la precisión es contraria a la variabilidad. Una forma de conocer la precisión de una variable es saber su variabilidad. En variables continuas (aquéllas en las que hay un número infinito de valores entre dos cualquiera), esta variabilidad puede medirse mediante la desviación estándar (DS) de una serie de mediciones repetidas. Cuando se trata de comparar la variabilidad de diferentes variables entre sí puede usarse el coeficiente de variación (CV): al dividir la desviación estándar por la media y multiplicar el resultado por 100 se obtiene una medida de variación en relación con el tamaño de la media que hace comparables dos variables, aunque éstas tengan unidades de medición muy distintas. Por ejemplo, en una encuesta sobre la actitud de las mujeres en relación a la alimentación de sus hijos durante el primer año de vida realizada en el Área de Salud de Cartagena sobre un total de 615 madres se determinó que la el peso de su primer hijo al nacer fué de  $3340,2 \pm 881,8$  gr, mientras que los días de lactancia materna exclusiva del niño fueron  $128,9 \pm 146,6$ . La DS del peso es mucho mayor, pero al ser variables muy distintas no podemos compararlas de esta manera. Si hallamos el CV de ambas variables podremos cuantificar la variabilidad y compararla entre ellas. La variabilidad de los días de lactancia es algo más del cuádruple de la del peso (CV 113,7% frente a 26,4%).

**Coherencia de resultados emparejados.** Estos datos emparejados se refieren a tres posibilidades:

**Coherencia examen-reexamen.** Es el grado de concordancia entre dos mediciones (antes y después) en un grupo de individuos. No debe dejarse pasar mucho tiempo entre las dos mediciones ya que si es así es posible que las diferencias que puedan encontrarse puedan tener una explicación distinta de la del azar.

**Coherencia interna.** Es la concordancia entre dos variables que miden la misma característica.

*Coherencia interobservador e intraobservador.* La primera se refiere a la relación de los valores obtenidos por dos o más observadores distintos en una misma muestra de individuos; la segunda es la repetibilidad de los valores obtenidos por un mismo observador cuando mide al mismo individuo varias veces.

Esta coherencia entre medidas emparejadas puede medirse con un simple coeficiente de correlación si se trata de variables continuas o con tablas de contingencia 2X2, teniendo en cuenta la concordancia debida al azar, cuando son variables dicotómicas (coeficiente kappa).

### Modos de mejorar la precisión:

La mejor forma de aquilatar al máximo la precisión es hacer lo más uniforme posible todo el método del estudio. Para ello, pueden tenerse en cuenta los siguientes consejos:

**Estandarizar los métodos de medición.** Fundamentalmente dejar muy claro, y por escrito, como llevar a cabo todo el proceso de la medición de manera que cualquier observador, aún no familiarizado con el estudio, pudiera llevarlo a cabo con facilidad. Para ello es muy interesante elaborar un manual de instrucciones lo más exhaustivo posible, más completo cuando más complejo sea el estudio.

**Entrenar a los observadores.** Es evidente que en la mayoría de los estudios se requiere la participación de diversos observadores. En la medida que su entrenamiento haga más uniforme su comportamiento en las mediciones, más precisas serán éstas.

**Refinamiento de los instrumentos de medición.** Al igual que deben prepararse y construirse los instrumentos electrónicos y mecánicos de manera que su variabilidad sea mínima, las encuestas y entrevistas deben ser redactadas de manera que puedan entenderse con absoluta claridad, teniendo siempre en cuenta a quien van dirigidas, evitando al máximo las ambigüedades: deberían estar siempre validadas.

**Automatización de los instrumentos.** Si se reduce al máximo la manipulación humana se evitará variabilidad. En determinadas mediciones esto puede hacerse con instrumentos automáticos. El equivalente en el caso de las encuestas es la auto-complimentación.

**Repetición.** El error aleatorio de la medición disminuye si se efectúan varias mediciones y se emplea la media.

### Exactitud

La exactitud de una variable es el grado con la que dicha variable mide lo que realmente quiere medir. La exactitud es fundamental para poder inferir correctamente los resultados del estudio tanto a la muestra (validez interna) como a la población diana (validez externa). Aunque los intentos de mejorar la exactitud suelen conducir a la mejora de la precisión, precisión y exactitud son conceptos distintos y no relacionados (cabe la precisión sin exactitud y viceversa).

### Fuentes de inexactitud

Así como la precisión es lo contrario de la variabilidad aleatoria (error aleatorio), la exactitud es lo contrario al error sis-

temático (sesgo). Los mismos errores comentados en el apartado referido a la precisión tienen aquí sus homólogos:

**Sesgo debido al observador.** Es la alteración constante y en el mismo sentido realizada por un observador, de manera consciente o inconsciente. Por ejemplo, el conocimiento de que un determinado grupo de pacientes está siendo tratado con una medicación antihipertensiva puede hacer que, inconscientemente, se tienda a infravalorar las cifras de la presión arterial. Los ensayos clínicos doble ciego, en los que ni el paciente ni el evaluador conocen el tipo de tratamiento, evitan este sesgo. En estudios de otro tipo, como los de cohortes, en los que se valora un factor de riesgo de padecer una enfermedad, la decisión sobre si un determinado individuo la padece o no debe ser tomada sin conocer sin conocer si dicho individuo está sujeto al factor de riesgo.

**Sesgo debido al individuo.** Las personas que participan en el estudio también pueden producir desviaciones sistemáticas de las mediciones. Se trata generalmente de *sesgos de selección*, ya que el error se comete en el proceso de selección de los individuos que se incluirán en la muestra.

### Estudios de Cohortes

Cuando son estudios de cohortes, puede que el grupo al que se va a seguir (cohorte) no refleje la misma distribución de variables tales como edad, sexo, etc. que la población general, debido, por ejemplo, a que algunos de los individuos originalmente seleccionados se niegan a participar.

Debido a que los estudios de cohortes son fundamentalmente de exposición a un determinado factor de riesgo de padecimiento de una enfermedad, otro de los posibles sesgos viene determinado por la desigualdad a esa exposición entre la cohorte y la población general. Por ejemplo, si se pretende estudiar la aparición de asma en una cohorte de recién nacidos y su relación con la contaminación atmosférica, cabe la posibilidad de que las madres que viven en una zona expuesta a la polución estén más sensibilizadas y tiendan a incluir más fácilmente a sus hijos en el estudio que aquéllas que no están sensibilizadas con el problema. De esta forma nos encontraríamos con una proporción mayor de niños expuestos en la cohorte que en la población general. Se trata de una forma de "efecto voluntario" que se comentará más adelante.

Otro tipo de sesgo de selección en este tipo de estudios ocurre cuando la enfermedad que es objeto de estudio tiene no sólo el factor de riesgo que se pretende determinar, sino que existen otros. Si la distribución de esos otros factores de riesgo distintos del objeto del estudio no es la misma en la cohorte que en la población general, los resultados del estudio no podrán generalizarse correctamente. Si en el ejemplo anterior, la distribución de los antecedentes familiares de alergia no son superponibles en la cohorte y en la población general, no se podrá saber con certeza qué papel juega la contaminación ambiental, ya que los resultados pueden ser consecuencia de la distinta distribución de la alergia. Para anular el efecto de los antecedentes familiares de atopía se requiere que la distribución de éstos entre

Tabla I Ejemplo de cómo se evita el sesgo de Neyman en un estudio de cohortes (seguimiento desde el nacimiento) sobre la influencia de la radiación nuclear en la aparición de malformaciones

	Vivos con malformación	Muertos con malformación	Vivos sin malformación
Expuestos a radiación	10	90	900
No expuestos a radiación	15	5	980

los expuestos y no expuestos sea la misma (que a su vez debe ser la de la población general).

Además de los sesgos de selección, en los estudios de cohortes -en los que se sigue a los individuos durante bastante tiempo- pueden sufrir sesgos debidos a las pérdidas en la participación. Si el número de pérdidas es equivalente en el grupo de expuestos y en el de no expuestos, no habrá problemas, siempre que la tasa de enfermedad sea la misma en los perdidos que en el grupo de donde proceden. Si no es así, entonces las cohortes se desequilibrarán y la validez interna se resentirá. Supongamos que en nuestra cohorte de recién nacidos, unos expuestos a contaminación y otros no, las pérdidas que se producen en el grupo no expuesto son sobre todo por niños que no acaban siendo asmáticos. Esto nos dejará en el grupo una tasa de asma menor a la real, y, por lo tanto, cuando se quiera comparar con los expuestos, el grupo de no expuestos estará artificialmente “más sano”. Una mayor prevalencia de asma en los expuestos podría deberse, desde luego, a la contaminación ambiental, pero también a la pérdida de asmáticos en el grupo de no expuestos. Como no podemos saber de antemano qué niños serán asmáticos no podemos saber si las tasas de asma en los perdidos son o no equivalentes al grupo del que proceden. Por ello, este problema es de difícil solución: en todo caso, debe intentarse que el número de pérdidas sea el menor.

### Estudios de Casos y Controles

Cuando los estudios son de casos y controles (enfermos y sanos en los que se investigan sus factores de riesgo en el pasado) los sesgos debidos al individuo son también diversos, alguno de ellos con nombre propio.

*Sesgo de prevalencia o incidencia (Neyman).* Se produce cuando se estudia una determinada patología que produce muertes precoces y en el momento del inicio del estudio esos individuos muertos ya no pueden incluirse en el grupo de los casos. Supongamos que queremos realizar un estudio caso-control sobre la relación de un factor de riesgo como la radiación nuclear sobre la prevalencia de malformaciones congénitas, en niños cuyo embarazo transcurrió inmediatamente tras un accidente nuclear, ocurrido 10 años atrás. Si la situación real de 1000 niños expuestos y 1000 niños no expuestos hubiera sido la de la ta-

Tabla II Ejemplo de sesgo de Neyman en un estudio de casos y controles cuando una buena parte de los casos ha muerto al iniciarse el estudio

	Vivos con malformación	Vivos sin malformación
Expuestos a radiación	10	900
No expuestos a radiación	15	980

bla I, algo que se hubiera puesto de manifiesto en un estudio de cohortes, lo que hubiéramos encontrado en un estudio caso-control hubiera sido lo que refleja la tabla II. Este sesgo, en este caso concreto, nos induce a pensar que el accidente nuclear “protegió” de las malformaciones, ya que en lo no expuestos a radiación la prevalencia de éstas (0,15%) fué mayor que en los expuestos (0,11%).

*Sesgo de la tasa de ingresos (falacia de Berkson).* Este sesgo fué ilustrado por Berkson en 1946 cuando evaluaba un estudio de casos y controles realizado años atrás que concluía que la tuberculosis protegía frente al cáncer: la prevalencia de tuberculosis entre los pacientes diagnosticados de cáncer era menor que entre los enfermos ingresados, no diagnosticados de cáncer. El problema radicaba en que las probabilidades de que un enfermo con las dos patologías estuviera ingresado era mucho menor que las de cualquiera de las enfermedades por separado. Dicho de otra forma: los pacientes con las dos enfermedades tienen más probabilidades de estar muertos y -por tanto- menos de estar ingresados en el momento del estudio. Es decir, las tasas de ingreso eran distintas para las patologías unidas y para las patologías por separado. La mejor manera de evitar este sesgo es conseguir un grupo control adecuado. En este caso hubiera sido suficiente elegir los controles de entre la población sana: hubiera debido compararse la prevalencia de cáncer entre los tuberculosos con la prevalencia de cáncer en la población general no tuberculosa. En resumen, debe tenerse cuidado al utilizar controles hospitalarios.

*Sesgo de los no respondedores o efecto voluntario* Se da cuando los pacientes o bien se presentan voluntarios para un estudio, o bien rehusan entrar en él. Por ejemplo, los estudios realizados en USA sobre la efectividad de la vacuna Salk frente a la poliomielitis utilizaron dos diseños diferentes. Algunas zonas utilizaron la vacuna frente a un placebo de forma aleatoria; pero otras, rehusaron este método y aceptaron que se ofreciera la vacunación a los niños de 2º grado y que los de 1º y 3º actuaran como controles. Los investigadores encontraron que las familias que aceptaron entrar en el estudio no aleatorio tenían un mejor nivel cultural y económico que las que no participaban, y, además, los niños faltaban más a clase. Es evidente que la ausencia a las clases de los niños con más probabilidades de

ser vacunados afectaron negativamente a los buenos resultados de la vacuna (ya que la cobertura vacunal hubiera sido más alta en los niños de 2º si los de mayor nivel hubieran asistido más a clase).

**Sesgo de pertenencia** Constituye un problema cuando se estudian grupos preexistentes, que el investigador no puede formar a su criterio. Se da cuando uno de los grupos se relaciona con la variable de estudio. Por ejemplo, cuando se estudian los efectos de la contaminación ambiental sobre la prevalencia de asma en Cartagena, la pertenencia del grupo de niños expuestos a contaminación se relaciona con un hábito de fumar en la madre considerablemente superior en su frecuencia a la de la zona no contaminada. La posible diferencia de la prevalencia o gravedad del asma en la zona contaminada no puede achacarse únicamente a la contaminación ambiental, ya que ésta se encuentra ligada a una mayor exposición al humo del tabaco, que es, de por sí, un factor de riesgo.

**Sesgo del trabajador sano** Es parecido al anterior y se puso en evidencia cuando trabajadores sometidos a ambientes peligrosos parecieron tener una tasa de supervivencia superior que la población general. La explicación no es que el ambiente peligroso proteja a los trabajadores, sino que la buena salud es un requisito previo para trabajar en esos ambientes, cosa que no ocurre en la población general.

### Ensayos Clínicos

En el caso de los ensayos clínicos, en los que el grupo activo y de tratamiento pueden escogerse evitando los sesgos de los estudios de cohortes y de casos y controles, pueden producirse también sesgos. El más característico es el *sesgo de procedimiento de selección*, que se produce cuando la asignación a los grupos de tratamiento se hace en base a algunas características de los pacientes, lo que puede provocar que los grupos acaben no siendo realmente iguales. Puede producirse cuando se compara un tratamiento quirúrgico con uno médico: es posible que los pacientes del tratamiento quirúrgico requieran un mejor estado general para su intervención y que, inconscientemente se deje para tratamiento médico a los pacientes que no pueden ser intervenidos. Está claro que los pacientes intervenidos están menos afectados por la enfermedad, por lo que no sería de extrañar que el tratamiento quirúrgico resultara superior al médico.

**Sesgos debidos al instrumento, procedimiento y/o seguimiento.** El *sesgo de medición* se debe fundamentalmente al mal funcionamiento del instrumento de medición, que desvía los resultados siempre en el mismo sentido. Ese instrumento puede ser tanto un aparato como una encuesta mal diseñada y no suficientemente bien validada.

El *sesgo de procedimiento* se produce cuando los grupos de sujetos no son tratados de la misma manera. Por ejemplo, los pacientes del grupo tratado -en caso de los ensayos clínicos- o expuestos al factor de riesgo -cuando son cohortes o casos-control- pueden ser seguidos con un mayor interés y una mayor profundidad, por lo que puede influirse en el sentido de que sean más obedientes al tratamiento -ensayos- o de que se investigue me-

jor el factor de riesgo, encontrándolo más frecuentemente. Someter a los grupos exactamente a los mismos procedimientos y examinar las variables de estudio con los mismos criterios son las mejores formas de evitar este sesgo.

**Sesgo de recuerdo.** Se produce cuando uno de los grupos tiene una mayor facilidad para recordar algo que se le pregunta. En un estudio de casos y controles es más fácil que los casos recuerden mejor si han estado sometidos al factor de riesgo. Volviendo al ejemplo del asma y la contaminación, si se hiciera un estudio de casos -asmáticos- y controles -no asmáticos- sobre la exposición pasada a picos de contaminación, es muy probable que los asmáticos recuerden mejor esos picos, ya que pudieron afectarles más que a los sanos.

**Sesgo del instrumento poco sensible.** Los instrumentos empleados para evaluar al individuo pueden no ser los adecuados para medir las variables del estudio, o pueden estar mal calibrados. Es evidente, por ejemplo, que en un estudio de nutrición no puede emplearse cualquier tallímetro, ni cualquier báscula, y que los que se empleen deben calibrarse periódicamente.

**Sesgo de detección.** Se da cuando se introduce una nueva técnica para el diagnóstico de una enfermedad, capaz de detectarla en un estadio anterior al que lo hacían los instrumentos clásicos: la detección precoz puede implicar una mayor supervivencia independientemente del tratamiento al que los enfermos sean sometidos.

**Sesgo de cumplimiento.** Se provoca cuando el paciente encuentra más agradable seguir un tratamiento que otro. Este sesgo debe tenerse muy en cuenta cuando en la evaluación de tratamientos antiasmáticos -por ejemplo- se compara la medicación oral con la inhalada: la comodidad de la administración o de la posología pueden determinar un mayor o menor grado de cumplimiento y, como consecuencia, un efecto terapéutico mayor o menor.

**Sesgo de migración.** Cuando un factor de riesgo de una enfermedad es conocido desde hace tiempo, los enfermos son enseñados a evitarlo, de manera que, en vez de encontrarse una relación factor de riesgo-enfermedad, lo que se halla es lo contrario, debido a la "migración" de los afectados. Por ejemplo, en los dos últimos estudios sobre la prevalencia de alergia realizados en Cartagena hemos encontrado que ésta era más frecuente entre los individuos que no tenían animales de compañía. Está claro que la interpretación correcta no es que los animales de compañía protejan de la alergia, sino que los alérgicos se deshacen de estas mascotas y evitan que sus hijos se expongan a este riesgo.

**Sesgos de evaluación.** Aún habiendo sido capaces de evitar todos los sesgos anteriores, aún cabe la posibilidad de encontrar algún error en la fase de evaluación de los resultados. Se trata del *sesgo de confusión* producido por un factor de confusión. Un factor de confusión es un factor que acompaña al factor de riesgo y actúa él mismo como factor de riesgo independiente, es decir, no es un factor intermedio en la cadena causal entre el factor de riesgo en estudio y la enfermedad. Estos factores de confusión pueden evitarse en el diseño del estudio, pero también

pueden tenerse en cuenta y amortiguar sus efectos por medio de determinados análisis estadísticos.

#### Medición de la exactitud

La forma de medir la exactitud de los resultados arrojados por un determinado instrumento es comparar esas mediciones con las realizadas por otro de reconocida exactitud o “estándar de oro”. Por ejemplo, si existe una encuesta muy complicada pero perfectamente aceptada como ajustada para medir el nivel socioeconómico y queremos utilizar una nueva, más sencilla, se debería hacer antes un estudio para comparar ambos métodos y si el novedoso resulta exacto (se correlaciona bien con el “estándar de oro”, ya aceptado) utilizarla sólo entonces. Esta comparación con un estándar de oro es clásica entre dos métodos diagnósticos, cuando uno de ellos es el que define la enfermedad.

#### Modos de mejorar la exactitud

Además de las formas comentadas en el apartado de la precisión (si mejoran la variabilidad aleatoria también mejoran la variabilidad sistemática), la mejor forma de mejorar la exactitud es no incurrir en los errores sistemáticos que se han comentado anteriormente.

#### Bibliografía

- 1 Gálvez Vargas R, Delgado Rodríguez M, Bueno Cabanillas A: Estudio de cohortes. En: G. Piédrola Gil y cols. eds. Medicina Preventiva y Salud Pública 9ª ed. Barcelona: Masson-Salvat 1994; pp 105-115.
- 2 Gálvez Vargas R, Delgado Rodríguez M: Estudio de casos y controles. En G. Piédrola Gil y cols. eds. Medicina Preventiva y Salud Pública 9ª ed. Barcelona: Masson-Salvat 1994; pp 105-115.
- 3 Hulley SB, Newman TB, Cummings SR: Puesta en marcha: anatomía y fisiología de la investigación. En: SB Hulley y SR Cummings eds. Diseño de la investigación clínica: un enfoque epidemiológico.