

EVIDENCIAS EN PEDIATRÍA

Toma de decisiones clínicas basadas en las mejores pruebas científicas

www.evidenciasenpediatria.es

Fundamentos de medicina basada en la evidencia

Estudios observacionales (I). Estudios transversales. Medidas de frecuencia. Técnicas de muestreo

Molina Arias M¹, Ochoa Sangrador C²

¹Servicio de Gastroenterología y Nutrición. Hospital Infantil Universitario La Paz. Madrid. (España).

²Servicio de Pediatría. Hospital Virgen de la Concha. Zamora. (España).

Correspondencia: Manuel Molina Arias, mma1961@gmail.com

Palabras clave en inglés: prevalence; sampling; incidence; cross-sectional studies

Palabras clave en español: prevalencia; muestreo; incidencia; estudios transversales

Fecha de recepción: 18 de octubre de 2013 • **Fecha de aceptación:** 21 de octubre de 2013

Fecha de publicación del artículo: 13 de noviembre de 2013

Evid Pediatr.2013;9:72.

CÓMO CITAR ESTE ARTÍCULO

Molina Arias M, Ochoa Sangrador C. Estudios observacionales (I). Estudios transversales. Medidas de frecuencia. Técnicas de muestreo. Evid Pediatr. 2013;9:72.

Para recibir Evidencias en Pediatría en su correo electrónico debe darse de alta en nuestro boletín de novedades en <http://www.evidenciasenpediatria.es>

Este artículo está disponible en: <http://www.evidenciasenpediatria.es/EnlaceArticulo?ref=2013;9:72>

©2005-13 • ISSN: 1885-7388

Estudios observacionales (I). Estudios transversales. Medidas de frecuencia. Técnicas de muestreo

Molina Arias M¹, Ochoa Sangrador C²

¹Servicio de Gastroenterología y Nutrición. Hospital Infantil Universitario La Paz. Madrid. (España).

²Servicio de Pediatría. Hospital Virgen de la Concha. Zamora. (España).

Correspondencia: Manuel Molina Arias, mma1961@gmail.com

Los estudios descriptivos son aquellos cuya única finalidad es describir los datos encontrados en una muestra de una población, pero sin intervención alguna por parte del investigador y sin tratar de establecer relaciones causales entre factores de riesgo y/o exposición y efectos observados.

Los estudios descriptivos más sencillos son las series de casos y los estudios transversales. Son estudios observacionales que carecen de direccionalidad. En general, ambos tratan de responder a preguntas del tipo ¿a cuántos?, ¿dónde? o ¿a quién?, por lo que su principal finalidad es la medición de la existencia de factores de riesgo o enfermedades en la muestra de población estudiada.

Existen varias formas de realizar estas mediciones, cada una de las cuales presenta una serie de peculiaridades que la hacen más adecuadas para un determinado diseño de estudio descriptivo.

MEDIDAS DE FRECUENCIA

La forma más sencilla de medir la presencia de enfermedad es contar el número de personas enfermas. Sin embargo, resulta una herramienta poco útil, ya que no proporciona ninguna estimación del tamaño relativo de la población de donde provienen. Por ello, se han ideado diversas formas de medir las enfermedades o los factores de riesgo o exposición que varían según lo que incluyamos en el denominador de la ecuación de cálculo. Las tres principales son proporción, razón y tasa.

Una **proporción** es un cociente en el que el numerador está incluido en el denominador. Un ejemplo sería el porcentaje de pacientes celíacos que existe en la población. El numerador está compuesto por los pacientes con enfermedad, que se cuentan también en el denominador como parte de la población general. En realidad, transmite el concepto de probabilidad, desde el punto de vista frecuentista, de que un fenómeno ocurra, por lo que puede tomar valores entre 0 y 1 (o entre 0% y 100%).

Por otra parte, una **razón** es un cociente en el que el numerador no se encuentra incluido en el denominador. Por ejemplo, los celíacos que siguen bien la dieta (el numerador) respecto a los que no la siguen (el denominador). Cuando el numerador de una razón expresa la probabilidad de que un suceso ocurra y el denominador la probabilidad de que no ocurra, este cociente recibe el nombre de *odds*. En realidad, proporción y *odds* son expresión del mismo fenómeno y podemos calcular una conociendo la otra, utilizando las siguientes fórmulas:

$$\begin{aligned} \text{Proporción} &= \text{Odds} / (\text{Odds} + 1) \\ \text{Odds} &= \text{Proporción} / (1 - \text{Proporción}) \end{aligned}$$

Finalmente, una **tasa** es una proporción que tiene en cuenta el paso del tiempo, por lo que relaciona el cambio de una magnitud por unidad de tiempo. Como ejemplo, el número de celíacos nuevos en una población durante un periodo de un año.

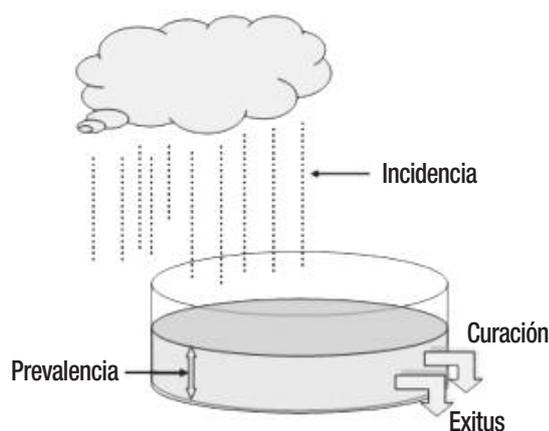
Una vez descritas las diferencias entre proporción, razón y tasa, pasaremos a definir brevemente las principales medidas de frecuencia de enfermedad: prevalencia e incidencia.

La **prevalencia** es el número de individuos de una población que presenta la enfermedad en un momento dado. Representa la situación en un momento fijo, pero no nos dice nada de lo que pasó antes o de lo que pasará en el futuro, ni cómo variará con el tiempo. Si lo pensamos un momento, los enfermos forman parte tanto del numerador como del denominador, por lo que la prevalencia es una proporción, a pesar del error frecuente de referirse a ella como tasa de prevalencia.

Por su parte, la **incidencia** es el número de casos que se producen en una población en un periodo de tiempo dado. En realidad, este término hace referencia a la incidencia acumulada, ya que podemos distinguir dos tipos de incidencia: incidencia acumulada y densidad de incidencia.

La **incidencia acumulada** es una proporción, ya que representa el número de enfermos (numerador) en

Figura 1. MEDIDAS DE FRECUENCIA: INCIDENCIA



una población general (denominador) que aparecen en un periodo de tiempo. Su valor radica en que estima el riesgo de desarrollar la enfermedad en ese periodo. Este valor representa el riesgo que tiene un individuo de enfermar siempre que, claro está, no muera por otras circunstancias, por lo que la incidencia acumulada puede no tener mucho sentido si el periodo de tiempo es demasiado largo. Este periodo será elegido arbitrariamente y dependerá de las características de la enfermedad que estamos midiendo.

Pero a veces ocurre que el periodo de observación no es el mismo para todos los sujetos, ya sea por incorporarse tarde o porque se pierden durante el seguimiento. En estos casos resulta útil la **densidad de incidencia**, que es el número de casos por unidad de tiempo-población. Este parámetro es una tasa, ya que incorpora la incidencia acumulada en el numerador y el paso del tiempo en el denominador, estimando que el riesgo de enfermar es el mismo durante todo el periodo. En este sentido, una tasa de un caso por cada 100 personas/año sería el resultado de encontrar un caso en 100 personas seguidas un año o en diez personas seguidas diez años. El periodo de seguimiento debe ajustarse según las características de la enfermedad.

SERIES DE CASOS

Una serie de casos es una descripción atemporal de un grupo reducido y altamente seleccionado de sujetos que tienen en común una misma característica.

Las series de casos pueden ser transversales y longitudinales. En los transversales se describen las características del grupo de personas en un momento dado del tiempo, mientras que en las longitudinales, también llamadas de seguimiento de una cohorte, se describe la evolución temporal de estas características observadas.

Este tipo de estudios puede ser útil para conocer la evolución natural o el cambio de lo que estamos estudiando después de una intervención o tratamiento, siempre teniendo en cuenta que, por definición, la intervención no se realizó de forma deliberada para llevar a cabo el estudio. Por ejemplo, puede describirse la presentación de un efecto secundario en un grupo de pacientes tratados con un fármaco, pero solo si tanto el tratamiento como el efecto adverso se han presentado sin la participación del investigador.

Estos estudios pueden servir para generar hipótesis que luego deben comprobarse con estudios experimentales. Sin embargo, tienen una serie de inconvenientes. En primer lugar, no sirven para calcular la frecuencia del efecto en la población, ya que la muestra no es representativa de la misma. Además, suelen estar impregnados de un componente subjetivo por parte del investigador, no permiten comparaciones con estudios similares y la presencia de los factores observados puede deberse al azar. Por último, son muy susceptibles al sesgo de publicación, de forma que existe tendencia a publicar únicamente las series con resultados favorables para el investigador. Por todos estos motivos, el grado de evidencia que aportan es pequeño.

ESTUDIOS TRANSVERSALES

Son estudios observacionales y descriptivos que carecen de direccionalidad. Su finalidad es estimar la frecuencia de una enfermedad o característica en una muestra de una población en un momento determinado, por lo que se conocen también como **estudios de prevalencia**. Para llevarlos a cabo, se define la población de estudio, de la que se extrae una muestra representativa en la que se mide la frecuencia de la variable que se quiere estudiar.

Existen diversas técnicas de muestreo que comentaremos más adelante. En cualquier caso, y sea cual sea la técnica de muestreo empleada, es preciso que se defina con precisión la población de referencia sobre la que se extrapolarán los resultados, la población susceptible de pertenecer al estudio y sus criterios de inclusión, así como la selección y definición de las variables que van a medirse en el estudio y las escalas de medida que van a utilizarse. Todos estos aspectos, junto con la calidad de los datos recogidos y el número y motivo de los datos perdidos, condicionarán la calidad del estudio y su validez.

En ocasiones, los estudios transversales se diseñan para estudiar la relación entre dos o más variables en una población en un momento dado. Reciben entonces el nombre de **estudios de asociación cruzada**.

De todas formas, aunque las características de la variable puedan cambiar a lo largo del tiempo, este tipo de diseños no permite establecer una clara relación temporal ni de causa-efecto.

Las ventajas de los estudios transversales, además de en su sencillez, radican en que se puede controlar bien la selección de los sujetos y las variables a medir, no existen pérdidas durante el seguimiento, no hay que esperar a que se desarrolle la enfermedad y son de realización más rápida y económica que otros estudios. Por todo esto, son útiles para generar hipótesis que, una vez más, habrá que comprobar posteriormente con estudios más complejos.

Entre sus inconvenientes, el fundamental es el ya comentado de no poder establecer claramente la secuencia de los acontecimientos, con lo que no pueden proporcionar relaciones de causa-efecto. Además, son poco útiles para enfermedades raras en la población de estudio o para características que puedan modificarse sensiblemente a lo largo del tiempo. Por último, están sujetos al riesgo potencial de todos los estudios observacionales.

TÉCNICAS DE MUESTREO

En muchas ocasiones, la técnica de muestreo más sencilla es recurrir a un **muestreo de conveniencia** o **no probabilístico**, en el que elegimos una muestra de sujetos que están fácilmente asequibles y disponibles. Por ejemplo, los pacientes que acuden a nuestra consulta durante un periodo de tiempo. El problema de este tipo de muestreos es que la muestra puede no ser representativa de la población de estudio, lo cual compromete la generalización de las conclusiones obtenidas a partir de nuestros resultados.

Por este motivo, es preferible recurrir a alguna de las técnicas de **muestreo probabilístico**, que tienen en común el hecho de que todos los sujetos del marco muestral (la población accesible para ser seleccionada) tienen una probabilidad conocida y distinta de cero de ser incluidos en la muestra. El más sencillo de los muestreos probabilísticos es el **muestreo aleatorio simple**, en el que todos los sujetos del marco muestral tienen la misma probabilidad de ser seleccionados, para lo cual debemos conocer previamente toda la

población. En este caso, la probabilidad de ser seleccionado es independiente de los sujetos, lo que asegura una mayor representatividad de la muestra y nos permite generalizar los resultados con mayor seguridad.

El muestreo aleatorio simple es, como hemos visto, una técnica equiprobabilística, pero en algunas ocasiones interesa que los elementos de muestreo del marco no tengan todos las mismas probabilidades de ser incluidos en la muestra, ya sea por características de la población o de determinadas variables de estudio.

En el **muestreo por cuotas** se eligen sujetos de forma que en la muestra se mantenga la misma proporción que en la población. Pensemos, por ejemplo, qué características, como el peso, el sexo o la raza, son importantes para el estudio. En estos casos, la selección se hace de forma que asegure la misma proporción dentro de la muestra que la que se observa en la población. El problema es que, al tener distinta probabilidad de ser seleccionado, podemos favorecer de forma inadvertida alguna otra característica, comprometiendo la representatividad de la muestra.

En otras ocasiones, la población puede estar dividida en estratos que interesa tener en cuenta a la hora de realizar el estudio. Pensemos en un muestreo entre distintas zonas geográficas, diferentes en tamaño o densidad de población. Si realizamos un muestreo aleatorio simple, puede ocurrir que las zonas más pequeñas queden poco representadas en la muestra. En estos casos puede realizarse un **muestreo estratificado**, que intenta asegurar que la muestra tenga la misma distribución que la población de donde procede. De esta forma, la estimación global estará ponderada según la fracción a la que pertenece cada estrato.

Por último, consideremos el caso de que los sujetos pertenezcan de forma natural a grupos excluyentes. Por ejemplo, niños de diferentes escuelas o pacientes de diferentes hospitales. En estos casos puede realizarse un **muestreo por conglomerados**. Se listan los conglomerados y se hace un muestreo aleatorio simple dentro de cada uno, en lugar de hacer directamente el muestreo aleatorio dentro de la población general. El problema de este tipo de muestreos es que obligan a tener en cuenta la variabilidad intra- e intergrupos, por lo que son menos eficientes desde el punto de vista estadístico y suelen precisar tamaños muestrales mayores.