

DISEÑOS DE ESTUDIOS

ESTUDIOS DE CORTE TRANSVERSAL

Introducción	2
Ventajas	3
Limitaciones y problemas.....	4
El Estudio de la asociación entre exposición y enfermedad	5
Elementos a considerar en el cálculo y selección de muestras.....	5
Modalidad de muestreo:.....	5
Prevalencia estimada del fenómeno:.....	5
Los factores de confusión y modificadores del efecto	7
Análisis	8

INTRODUCCIÓN

Uno de los diseños de investigación más sencillos y frecuentemente utilizados es el que se utiliza en los denominados estudios de prevalencia o de corte transversal. De **332** artículos originales publicados en el [New England Journal of Medicine](#) en dos años, un **39 %** correspondían a estudios transversales. De estos, el 20% describían **nuevas enfermedades** o variantes de enfermedades conocidas. Un 26 % estudiaban **métodos diagnósticos** o diferentes **estadios** de las enfermedades, 54 % estudiaban el **curso** de diferentes enfermedades.

Un **estudio transversal** o **estudio de prevalencia** es un estudio observacional, descriptivo, en el que en un único momento temporal, se mide a la vez la prevalencia de la exposición y del efecto en una muestra poblacional, es decir, permiten estimar la magnitud y distribución de una enfermedad o condición en un momento dado.

También se denominan:

- Estudio de corte.
- Cross-sectional studies.
- Cross sectional surveys.
- Estudio vertical.

Una de las características principales de estos estudios es el **tipo de muestreo** utilizado, denominado por FLEISS como *muestreo tipo I* (FLEISS, 1981, Cap. 5). En primer lugar se selecciona un determinado número de individuos de una población que se desea estudiar, sin conocer previamente quienes presentan o no el fenómeno objeto de estudio. Una vez obtenida la información que nos interesa -parámetros biológicos, características o conductas-, se buscará extrapolar los resultados obtenidos a la población de partida.

En un estudio transversal, los sujetos que entrarán en la investigación se seleccionan frecuentemente a partir de listas en la que están registrados los individuos de la población que se desea estudiar. Esta lista puede ser, por ejemplo, el censo o la guía telefónica, o cualquier otra fuente de información independiente del fenómeno objeto del estudio (es decir, que el hecho de presentar el fenómeno de interés no modifique la probabilidad de estar o no estar en la lista).

En un estudio transversal típico, la información que se obtiene de los participantes corresponde a las características que tienen en el momento de realizar el estudio. Aunque se pueden introducir matices y mejoras sobre este diseño típico (fundamentalmente, obtener información histórica sobre estados anteriores), en esta discusión se asumirá que la información recogida se refiere al momento en que se realiza el estudio.

Un estudio transversal ofrece una visión de lo que ocurre en un momento dado en una población determinada, lo que proporciona una **fotografía instantánea de la realidad**. Suele ser realizado dentro de un período relativamente corto de tiempo (ejemplo, prevalencia de infección nosocomial durante un mes o una semana).

VENTAJAS

Los estudios transversales son útiles para estimar la prevalencia de una enfermedad. Aunque normalmente proporcionan estimaciones de la prevalencia en grupos específicos de la población –por ejemplo, pacientes internados en terapia intensiva de un hospital–, si la población de estudio se obtiene mediante un muestreo aleatorio, proporcionan prevalencias representativas de la población.

La prevalencia de una determinada enfermedad se calcula dividiendo el número de casos que presentan la enfermedad por el número total de personas estudiadas en la muestra. Esto requiere una definición exacta y precisa de la enfermedad. Mediante la estratificación de la población en estudio de acuerdo con valores de otras variables, se pueden estimar las prevalencias de la enfermedad en distintos estratos. Esta descripción de los datos, aunque simple, es de gran utilidad. Es más, diferencias considerables entre las prevalencias de los estratos, sugieren que la variable que define el estrato puede estar asociada con la aparición de la enfermedad. Identificar factores asociados a la aparición de la enfermedad constituye el objetivo principal de la investigación epidemiológica. Por esta razón, en el resto del capítulo se resalta la utilidad de los estudios de prevalencia para determinar factores asociados a la aparición de la enfermedad.

Los estudios transversales tienen la ventaja sobre los de cohorte y los de casos y controles, de permitir estimar tanto la **probabilidad de presentar la enfermedad** como la de **presentar la exposición**.

En un estudio transversal puede estimarse la probabilidad de que una persona expuesta a un factor de riesgo presente la enfermedad en un momento dado, de forma similar a un estudio de cohorte (aunque las posibilidades de inferencia causal son mucho menores ya que, como se verá más adelante, a menudo es difícil determinar si la exposición es una causa o una consecuencia de la enfermedad). Adicionalmente, en un estudio transversal puede calcularse, al igual que en un estudio de casos y controles, la probabilidad de que una persona que está enferma presente, o haya presentado, la exposición estudiada.

En un diseño de casos y controles se estudia una sola enfermedad, mientras que en uno transversal, al igual que sucede en los diseños de cohortes, se pueden estudiar múltiples enfermedades, con el consiguiente ahorro de tiempo y dinero.

Los estudios transversales tienen la ventaja sobre los estudios de cohorte de poder realizarse en un corto período de tiempo. Por lo tanto, un diseño transversal suele ser bastante más rápido y económico que uno de cohorte que, a menudo, requiere un largo período de seguimiento para que un número suficiente de participantes desarrollen la enfermedad (con el problema implícito de la pérdida de participantes o perdidos de vista).

LIMITACIONES Y PROBLEMAS

La principal limitación de este tipo de estudios deriva de que **miden la prevalencia y estudian casos prevalentes** que, como tales, **pueden no ser representativos de todos los casos de enfermedad que se han producido en la comunidad.**

- Una **primera limitación** surge como consecuencia de que **en este tipo de estudios se investiga lo que ocurre en una población en un momento dado.** NO podemos establecer en ellos la secuencia temporal de las diferentes variables estudiadas. Por lo tanto, la exposición que se recoge es la que presenta el sujeto en el momento de realizar el estudio, y no es posible establecer si un determinado factor estaba presente antes de aparecer la enfermedad (pudiendo ser, por tanto, un factor de riesgo) o si apareció después de la misma (pudiendo ser una consecuencia de ella). Por lo tanto, puede no ser posible diferenciar entre factores de riesgo y consecuencias de la enfermedad, o, en otras palabras, no es posible determinar si la exposición precede o es una consecuencia de la enfermedad.

Éste es el problema que encontraron CONTI y COL. (1986), al estudiar la relación entre capacidad vital e hipertensión. En un análisis transversal encontraron una relación inversa entre capacidad vital y la hipertensión, lo que sugería que los enfermos con una capacidad vital baja estaban más predispuestos a desarrollar hipertensión arterial. Es decir, una capacidad vital baja, sería un factor de riesgo de hipertensión. Esta asociación desaparecería al realizar un análisis longitudinal de los datos: la disminución de la capacidad vital, en lugar de ser un factor de riesgo de hipertensión, se mostraba como un efecto de la misma. Esta dificultad para diferenciar entre causa y efecto, se agudiza, obviamente al estudiar enfermedades que se desarrollan de forma lenta y/o cuando la exposición es de larga duración.

- Un **segundo problema** de los estudios que investigan casos prevalentes, es decir, individuos que presentan la enfermedad en el momento de realizar el estudio, es la **dificultad de separar los factores de riesgo (causas) de los factores pronósticos que influyen en la duración de la enfermedad.** Los casos prevalentes o han sobrevivido lo suficiente como para ser considerados enfermos en el momento de realizar el estudio, o su enfermedad ha sido lo suficientemente prolongada como para estar enfermos en ese mismo momento. Si existen una serie de factores que, por ejemplo, incrementan la duración de la enfermedad, y éstos aparecen de forma selectiva en un determinado grupo de la población, este grupo quedará sobre representado con respecto al conjunto de enfermos; por tanto, las características que lo diferencian de la población general aparecerán como asociadas a la enfermedad. El mismo fenómeno, pero en sentido inverso, ocurre con los factores que disminuyen la duración de la enfermedad, ya sea por una mayor letalidad o una mayor velocidad de curación. En este caso, todas las características del grupo poblacional sin dichos factores, aparecerían en el estudio como posibles factores de riesgo. Por lo tanto, las estimaciones de la prevalencia (y como consecuencia, las medidas de asociación que se construyan con ellas) pueden estar sesgadas, al excluir los enfermos que fallecen o se recuperan de forma rápida.

Un ejemplo de esta situación es el estudio transversal en donde se observó asociación entre el antígeno HLA2 y la leucemia, en el cual los autores concluyeron que este podría ser un factor predisponente de presentar la enfermedad. En un estudio posterior (de incidencia) lo que se vio era que los pacientes que presentaban el HLA2 sobrevivían más tiempo. Todas las consideraciones anteriores se resumen en que los factores de riesgo asociados a los casos

prevalentes pueden ser distintos a los asociados a los casos incidentes de enfermedad, lo que puede inducir a conclusiones erróneas sobre cuáles son los determinantes de la enfermedad. Este problema puede aparecer en todo estudio epidemiológico que trabaje con casos prevalentes en lugar de incidentes.

EL ESTUDIO DE LA ASOCIACIÓN ENTRE EXPOSICIÓN Y ENFERMEDAD

Considerando los problemas enumerados anteriormente, existen algunos aspectos que han de ser tenidos en cuenta si se desea evaluar la relación entre exposición y enfermedad.

- En primer lugar, es fundamental que los estudios de prevalencia tengan un **tamaño de muestra adecuado**, no sólo para estimar la prevalencia con un grado de precisión suficiente, sino también para establecer asociaciones entre exposición y enfermedad con la suficiente potencia estadística entre subgrupos de interés.
- En segundo lugar, es fundamental «**desenredar**» **los factores de confusión** para establecer correctamente la asociación entre exposición y enfermedad. Los métodos para «eliminar» el efecto de los factores de confusión no son específicos de los estudios transversales y, por tanto, se incluye aquí sólo una breve descripción.
- Finalmente, incluso más importante que ajustar por el efecto de los factores de confusión es **determinar las interacciones**, es decir, que la asociación entre exposición y enfermedad sea distinta para distintos valores de otra variable denominada, por tal razón, **factor modificador de efecto**.

Elementos a considerar en el cálculo y selección de muestras

Modalidad de muestreo:

En el cálculo del tamaño muestral de un estudio de prevalencia, debe considerarse la posible distribución del fenómeno a estudiar en la población objetivo. De no observarse una distribución particular de la variable en la población, puede realizarse un **muestreo aleatorio simple**. Si por el contrario se advierte que el atributo a medir guarda relación con alguna característica poblacional, debe considerarse la posibilidad de realizar un muestreo **aleatorio estratificado**, respetando la conformación poblacional de los estratos de interés epidemiológico (estratos tales como sexo, grupos de edad, condición socioeconómica, etnia) o considerando grupos como un todo (clusters).

Prevalencia estimada del fenómeno:

Como ya se mencionó, debe recurrirse al empleo de algún **estimador o parámetro del atributo a medir en la población**. Si no se cuenta con este antecedente, puede recurrirse a su estimación sobre la base de una exploración poblacional en pequeña escala o bien utilizar estimaciones en grupos seleccionados en los cuales el problema puede haber sido medido (por ejemplo, grupos de alto riesgo de contraer Hepatitis B en el caso de prevalencia de dicha enfermedad).

También puede recurrirse a la literatura en búsqueda de algún estimador, considerando especialmente poblaciones de características similares. Esta información corresponderá a la prevalencia estimada.

El tamaño adecuado de la muestra para una encuesta relativa a la población está determinado en gran medida por tres factores:

- 1- prevalencia estimada de la variable considerada (en este caso, la malnutrición crónica);
- 2- nivel deseado de fiabilidad; y
- 3- margen de error aceptable.

El **tamaño de la muestra** para un diseño de encuesta basado en una **muestra aleatoria simple**, puede calcularse mediante la siguiente fórmula.

$$N = \frac{t^2 \times p(1-p)}{m^2}$$

Donde:

n = tamaño de la muestra requerido

t = nivel de fiabilidad de 95% (valor estándar de 1,96)

p = prevalencia estimada

m = margen de error de 5% (valor estándar de 0,05)

Ejemplo

En el proyecto de AL HAOUZ en Marruecos, se ha calculado que cerca del 30% (0,3) de los niños de la zona del proyecto padecen de malnutrición crónica. Este dato se basa en estadísticas nacionales sobre malnutrición en las zonas rurales. Utilizando los valores estándar indicados anteriormente se efectúa el cálculo siguiente:

$$N = \frac{1.96^2 \times 0.3(1-0.3)}{0.05^2}$$

$$N = \frac{3.8416 \times 0.21}{0.0025}$$

$$n = \frac{0.8068}{0.0025}$$

$$n = 322.72 \sim \mathbf{323}$$

Imprevistos

El tamaño de la muestra se aumenta en un 5% para hacer frente a imprevistos como la ausencia de respuesta o errores de registro.

LOS FACTORES DE CONFUSIÓN Y MODIFICADORES DEL EFECTO

Otro error que puede cometerse al estudiar la asociación entre un factor de exposición y una enfermedad es el de establecer que la asociación -positiva o negativa- existe cuando en realidad ésta es debida, en parte o en su totalidad, a un **tercer elemento o factor de confusión**. Este fenómeno es extremadamente importante, ya que al estudiar el riesgo de enfermedad asociado a una exposición, éste puede aparecer aumentado o disminuido por la existencia de dicho factor de confusión.

No entra en los objetivos de esta unidad discutir este complejo fenómeno. Simplemente les recordamos que este problema surge cuando se estudia la relación entre una determinada exposición y una enfermedad y existe otra exposición (que no es consecuencia de la exposición estudiada) que está asociada con la enfermedad y con la exposición estudiada.

- Por tanto, un factor de confusión es un factor de riesgo de la enfermedad, es decir, la prevalencia de la enfermedad es más alta en el grupo que lo presenta que en el que no lo presenta.
- En segundo lugar, el factor de confusión está asociado con la exposición estudiada, es decir, la prevalencia del factor de confusión en el grupo expuesto es diferente a la del grupo de los no expuestos.
- Finalmente, el factor de confusión no debe ser una consecuencia de la exposición estudiada. En la práctica, todo factor de riesgo para una determinada enfermedad, que se distribuye de forma desigual en los expuestos y en los no expuestos, puede ser un factor de confusión de la asociación que se esté estudiando. Un ejemplo de factor de confusión frecuente es la edad.

Otro **factor de confusión habitual**, que aparece al estudiar la relación entre contaminantes y enfermedad, es el tabaco. El tabaco es un factor de riesgo para muchas enfermedades relacionadas con los contaminantes y suele estar asociado con la exposición.

Los factores de confusión pueden ser **controlados**, bien en el momento de diseñar el estudio (mediante el apareamiento según los principales factores de confusión), o bien en el momento del análisis de los datos.

A menudo no es posible realizar un apareamiento en los estudios transversales debido al método de muestreo utilizado. Si, por ejemplo, la edad y el sexo son las variables de confusión, el apareamiento se realizaría escogiendo para cada persona expuesta una no expuesta del mismo sexo y de la misma edad.

En los estudios transversales este apareamiento no puede realizarse, ya que se selecciona a un determinado número de personas sin conocer previamente sus características. En el momento del análisis, los instrumentos más utilizados para ajustar por los factores de confusión son la estratificación y la regresión. Estos métodos se presentan más adelante en el apartado de análisis de los datos.

Otro fenómeno de gran importancia en el momento de analizar la relación entre exposición y enfermedad es la **interacción**, es decir, que una variable modifique el efecto de la exposición de interés.

Un investigador puede afirmar que existe una fuerte relación entre una exposición y una enfermedad, cuando en realidad ésta se debe a la existencia de un fenómeno de interacción.

La interacción aparece cuando la asociación entre exposición y enfermedad tiene diferentes valores según los niveles de una tercera variable. En este caso, no se puede hablar de una única relación entre exposición y enfermedad, sino que ésta varía en función de los valores de otra variable.

Un ejemplo de modificación del efecto es el encontrado por Shapiro y col. (Shapiro M, Munoz A, Tager IB, Schoenbaum SC, Polk BF. Risk factors for infection at the operative site after abdominal or vaginal hysterectomy. N Engl J Med. 1982 Dec 30;307(27):1661-6.) al estudiar los factores de riesgo asociados a infección quirúrgica en mujeres a las que se les había realizado una histerectomía. Se estudió la asociación entre infección quirúrgica y una serie de variables como la edad, profilaxis antibiótica y el tipo de intervención. Al estudiar la relación entre la administración de antibióticos (exposición) y la prevalencia de la infección quirúrgica (enfermedad) los autores encontraron que la relación no era única, sino que variaba en función de la duración de la intervención.

Los autores hubiesen cometido un error si hubiesen presentado una única relación entre administración de antibióticos e infección, sin tener en cuenta la duración de la intervención. Sin tener en cuenta esta interacción, las mujeres a las que no se les administraba antibióticos tenían una probabilidad de presentar una infección de la herida 1.88 veces mayor que las que recibían antibióticos. Sin embargo, esta probabilidad era mucho mayor (6.23) si las intervenciones eran de corta duración (75 minutos) y mucho menor si duraban más de 165 minutos (1.28). Si los autores no hubieran estudiado el efecto de la duración de la intervención simplemente habrían afirmado que la no administración de antibióticos aumentaba la probabilidad de infección en un 88 %, y no se hubiese puesto en evidencia que el beneficio era mucho mayor en intervenciones de corta duración.

ANÁLISIS

Habitualmente los datos de un estudio de prevalencia se presentan en forma de tabla 2 X 2 (Cuadro 11.3).

	Enfermedad +	Enfermedad -	Total
Exposición +	a	b	a+b = n_{x+}
Exposición -	c	d	c+d = n_{x-}
Total	a + c = n_{E+}	b+d = n_{E-}	a+b+c+d = n

Cuadro 11.3. Tabla 2 x 2 que resume los datos de un estudio transversal.

Esta tabla se asienta sobre la base de que se ha estudiado una muestra de tamaño **n**, en la que **a+b** individuos están expuestos, es decir, presentan un posible factor de riesgo o exposición, y **c+d** no lo están. Del grupo de sujetos expuestos, **a** corresponde a individuos que presentan la enfermedad, y entre los no expuestos **c** representa a los enfermos.

A fin de cuantificar la frecuencia de la enfermedad encontrada en el estudio (Cuadro 11.3) se utiliza la prevalencia (**P**) que es la probabilidad de presentar la enfermedad. La prevalencia de la enfermedad en el total de la población estudiada se estima mediante la fórmula **p = (a+c) / n**. La prevalencia en expuestos se estima mediante **P_{x+} = a/(a+b)** y la prevalencia en no expuestos mediante **P_{x-} = c/(c+d)**. Lo que hicimos fue una lectura horizontal de los datos.

También podría plantearse que de los $a+c$ enfermos, a presenta el factor de riesgo y que de los $b+d$ no enfermos, b lo presenta (lectura vertical de los datos). En este caso la medida de frecuencia que se utiliza es la prevalencia de la exposición en los $a+c$ enfermos del Cuadro 11.3 $P_{E+}(X) = a/(a+c)$ y la prevalencia en no enfermos $P_{E-}(X) = b/(b+d)$.

Otra medida de frecuencia que se utiliza son los «odds» (denominados por algunos autores como «prevalence odds» [KLEINBAUM Y COLS., 1982, pág. 118] y traducido en este capítulo como «oportunidad» [MARTÍN MORENO, 1990]).

Mientras que la **prevalencia** es la probabilidad de que un individuo de la población presente la enfermedad en un momento dado, la **oportunidad (odds)** es la razón entre la probabilidad de presentar la enfermedad (P), y la probabilidad de no presentarla ($1 - P$).

De la misma forma que, como se ha visto, se puede calcular la prevalencia de la enfermedad y la prevalencia de la exposición, se pueden también estimar la oportunidad de enfermedad y la oportunidad de exposición.

Las oportunidades así calculadas y la prevalencia son medidas relacionadas y equivalentes. Específicamente, a mayor prevalencia mayor oportunidad, y a cada valor de la prevalencia corresponde un único valor de la oportunidad. Por ejemplo, a prevalencias de 10, 50 y 90 %, le corresponden oportunidades de 0,11, 1 y 9, respectivamente. Nótese que mientras la prevalencia está comprendida entre cero y uno, la oportunidad puede tomar valores entre cero e infinito.

En concreto, la relación entre la oportunidad de enfermedad y la prevalencia (P), de la enfermedad, viene dada por:

$$\text{Odds} = \frac{\text{Prevalencia}}{1 - \text{Prevalencia}}$$

Cuanto más grande sea la prevalencia de la enfermedad, mayor será la diferencia entre la oportunidad y la prevalencia, mientras que para prevalencias pequeñas, la oportunidad y la prevalencia serán prácticamente iguales (por ejemplo, para una prevalencia de 5%, la oportunidad es de $0.05/(1-0.05) = 5.3\%$).

Otra medida relacionada y equivalente a la prevalencia es la **transformación logarítmica de la oportunidad**. Aunque no tiene una interpretación intuitiva y directa, el logaritmo de la oportunidad tiene características matemáticas útiles. Por ejemplo, el logaritmo de la oportunidad cubre todos los valores desde menos infinito hasta infinito.

Así, a prevalencias de 10.50 y 90%, corresponden logaritmos de oportunidades de -2.20, 0 y 2.20, respectivamente. Esta característica de no estar acotado hace que el logaritmo de la oportunidad sea la medida que se utiliza en los modelos de regresión para la determinación de factores de riesgo asociados a la prevalencia de la enfermedad de interés (MUÑOZ Y COLS., 1993, pág. 965).

La prevalencia calculada a partir de los datos de un estudio epidemiológico es sólo una estimación puntual de la prevalencia que existe en la población. Esta estimación está basada en una muestra de la misma. Como consecuencia de la variabilidad biológica, esta estimación puntual

difícilmente será idéntica a la real de la población. Por este motivo, es frecuente que el investigador proporcione un intervalo que muy probablemente contenga la verdadera prevalencia de la población ("intervalo de confianza de la prevalencia"). La prevalencia es una proporción, por lo que puede calcularse su intervalo de confianza mediante la utilización de los métodos descritos para las proporciones.

Recuerden que el intervalo de confianza del 95% de una proporción P se calcula mediante la fórmula (ROSNER, 1995, págs. 174-176):

$$P_I = P - 1.96 \sqrt{\frac{P \cdot (1 - P)}{n}} \quad (1)$$

$$P_s = P + 1.96 \sqrt{\frac{P \cdot (1 - P)}{n}} \quad (2)$$

Reemplazando en la fórmula (1) P por el valor de la prevalencia y n por el tamaño de la muestra que se ha utilizado para estimar esta prevalencia, se obtiene el límite inferior (P_I) del intervalo. Para obtener el límite superior (P_s), se realizan los mismos cálculos con la ecuación (2). Dado que el método presentado asume una aproximación normal de la distribución binominal, sólo será válido para muestras grandes.