

EVIDENCIAS EN PEDIATRÍA

Toma de decisiones clínicas basadas en las mejores pruebas científicas

www.evidenciasenpediatria.es

Fundamentos de medicina basada en la evidencia

Estudios ecológicos en Epidemiología

Ortega Páez E¹, Ochoa Sangrador C²

¹UGC Maracena. Distrito Granada-Metropolitano. Granada (España).

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

Correspondencia: Eduardo Ortega Páez, edortegap@gmail.com

Palabras clave en inglés: epidemiologic studies.

Palabras clave en español: estudios epidemiológicos.

Fecha de recepción: 8 de diciembre de 2015 • Fecha de aceptación: 10 de diciembre 2015

Fecha de publicación del artículo: 16 de diciembre de 2015

Evid Pediatr. 2015;11:69.

CÓMO CITAR ESTE ARTÍCULO

Ortega Páez E, Ochoa Sangrador C. Estudios ecológicos en Epidemiología. Evid Pediatr. 2015;11:69.

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: [artículo está disponible en: http://www.evidenciasenpediatria.es/EnlaceArticulo?ref=2015;11:69](http://www.evidenciasenpediatria.es/EnlaceArticulo?ref=2015;11:69)

©2005-15 • ISSN: 1885-7388

Estudios ecológicos en Epidemiología

Ortega Páez E¹, Ochoa Sangrador C²

¹UGC Maracena. Distrito Granada-Metropolitano. Granada (España).

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

Correspondencia: Eduardo Ortega Páez, edortegap@gmail.com

INTRODUCCIÓN

Los estudios ecológicos forman parte de los diseños observacionales en Epidemiología, que se distinguen del resto porque la población de estudio no son sujetos individuales, sino sujetos agrupados, normalmente grupos de personas (países, provincias, escuelas, etc.); esto hace que el nivel de inferencia de sus estimaciones sea también agregado. Clásicamente se han utilizado como generadores de hipótesis ya que son sencillos de realizar, rápidos, poco costosos y su análisis relativamente sencillo, o para la comprobación del impacto de la implantación de medidas de salud, por ejemplo, campañas vacunales. Habitualmente son estudios descriptivos y de temporalidad histórica, pero pueden ser analíticos y concurrentes, aunque la medición agrupada de exposición y efecto les confiere un alto riesgo de presentar sesgos. Por ello se les ha considerado tradicionalmente como estudios de bajo nivel de evidencia. En los últimos tiempos han resurgido, debido, en gran parte, al desarrollo de novedosas técnicas estadísticas que han mejorado su utilidad.

TIPOS DE VARIABLES Y ANÁLISIS¹

Los estudios ecológicos tienen peculiaridades en cuanto al tipo de variables que se miden y al tipo de análisis descriptivo e inferencial.

Tipos de variables

Clásicamente, las variables ecológicas se clasifican en variables agregadas, ambientales o globales. Las variables agregadas suelen ser medidas que resumen observaciones individuales, normalmente medias (por ejemplo, edad media al inicio de consumo de alcohol) o proporciones (por ejemplo, tasa de mortalidad neonatal/1000 recién nacidos vivos). Las medidas ambientales, son características físicas propias de un lugar determinado, que suelen tener su paralelismo a nivel individual (por ejemplo, niveles de polución atmosférica) o atributos de grupos, organizaciones o lugares, que no tienen una equivalencia a nivel individual (por ejemplo, densidad de población, tipo de sistema sanitario).

Niveles de análisis

En general, el análisis se realiza en un solo nivel, bien sea agregado (datos de grupos de sujetos) o individual (datos de

sujetos). Si los datos han sido recogidos en distintos niveles se deben transformar en un solo nivel. Lo usual es transformar los datos individuales en agregados, por ejemplo calculando las medias o proporciones. Pero también se puede realizar en sentido contrario, desagregando las variables; por ejemplo, en un estudio para medir la radiación individual después de la catástrofe de Chernóbil, la exposición de cada individuo se puede calcular por la distancia a la que estuviera del lugar de la explosión. Por último, se puede realizar un análisis en los dos niveles, mediante técnicas estadísticas multinivel jerarquizado, cuya explicación excede del objetivo de esta revisión².

Niveles de inferencia

El nivel de inferencia depende de nuestros objetivos. Supongamos que queremos hacer una inferencia sobre efectos de un factor de riesgo a nivel individual, entonces la inferencia será individual; por ejemplo, un estudio en el que se quiere relacionar el consumo de alimentos preprocesados (estimado a partir de una encuesta telefónica) con la incidencia de cáncer de colon (medido a partir de registro de tumores). Ahora supongamos que estamos interesados en conocer la efectividad de un programa vacunal, entonces nuestras medidas e inferencias se harán de forma agregada a partir de las tasas de coberturas vacunales de una población. Por último, podemos medir un factor de exposición a nivel individual y agrupado; por ejemplo, contamos con medidas de higiene a nivel poblacional (por ejemplo, grado de potabilización de aguas) y a nivel individual (por ejemplo, frecuencia de lavado de manos); en este caso estaríamos en una inferencia contextual.

TIPOS DE ESTUDIOS³

Clásicamente se han diferenciado los estudios ecológicos según dos tipos de dimensiones: el método de exposición y el de agrupación. En cuanto al método de exposición, nos podemos encontrar dos situaciones: si la variable primaria de exposición no se mide o solo se hace parcialmente se denominan exploratorios, en el caso contrario analíticos. Según el método de agrupación se pueden clasificar en múltiples, cuando se seleccionan varias zonas; de tendencia temporal, si existe medición a lo largo del tiempo, y mixtos, cuando se combinan ambos métodos. Las dos dimensiones a su vez se pueden combinar dando lugar a diseños más complejos como veremos a continuación.

Estudios de grupos múltiples

- **Exploratorios.** No existe medición de la variable de exposición (independiente) pero sí existe medición de la característica o enfermedad de estudio (variable dependiente). Un ejemplo consistiría en la comparación de las tasas ajustadas de cáncer oral entre varias zonas geográficas en el mismo periodo de tiempo. Este tipo de estudios es útil para generar hipótesis etiológicas que después deben ser comprobadas con estudios de mayor nivel de evidencia. Un ejemplo podría ser el estudio realizado en EE. UU. sobre tasas de mortalidad de cáncer ajustadas por edad y por zonas geográficas, donde se encontró que los cánceres orales eran más frecuentes entre las mujeres en las zonas del sureste, donde se sabe que es habitual que estas masquen tabaco; esto llevó a la hipótesis de que el cáncer oral podría estar relacionado con el tabaco¹. Podemos encontrarnos con dos problemas en el análisis, el primero es que poblaciones más pequeñas muestren tasas más extremas y con mayor variabilidad, y en segundo lugar que poblaciones muy cercanas muestren tasas muy parecidas (autocorrelación). Existen técnicas estadísticas específicas para analizar estas situaciones^{4,5}.
- **Analíticos.** Es el diseño más frecuente. Se trata de valorar la asociación entre el nivel medio de exposición o prevalencia de un factor y la tasa de enfermedad entre varios grupos. Un ejemplo sería describir las tasas de hospitalización (variable dependiente) en menores de un año en los distritos municipales de la ciudad de Madrid y analizar su asociación con variables socioeconómicas (variables independientes)⁶.

Estudios de tendencia temporal

- **Exploratorios.** Comparan las tasas de enfermedad a lo largo del tiempo en una misma zona geográfica. Un ejemplo sería, el análisis de la evolución de la utilización de antibióticos en la población pediátrica, medido como consumo anual medio de antibióticos, durante los últimos cinco años en la Comunidad Autónoma de Castilla y León⁷. El análisis estadístico se realiza por modelos de edad-periodo-cohorte, en el que mediante gráficos se trata de modelar de forma independiente los efectos de la edad, del tiempo y de la cohorte en los resultados. La utilización de modelos autorregresivos de series temporales (ARIMA) nos puede servir para realizar modelos predictivos de tendencias futuras entre la exposición y las tasas de enfermedad.
- **Analíticos.** Evalúa la asociación ecológica entre el cambio en el nivel medio o prevalencia de exposición y el cambio en la tasa de la enfermedad a lo largo del tiempo en una población geográficamente definida. Pueden existir dos inconvenientes: puede haber cambios en el tiempo de los criterios diagnósticos y de clasificación de las enfermedades, que distorsionen las tasas de enfermedad, y

pueden ser necesarios seguimientos muy prolongados si el tiempo de latencia entre la exposición del factor de riesgo y la detección de la enfermedad es muy largo. Como ejemplo mencionaremos el estudio realizado en México en los años 1993-1995 sobre la relación entre la contaminación ambiental y la mortalidad infantil⁸.

Estudios mixtos

Estos estudios comparan las tendencias temporales de varias áreas geográficas. El análisis estadístico es similar. Un ejemplo sería el estudio que realizaron Aceituno Madera *et al.* para relacionar la prevalencia de melanoma, la altitud y la radiación solar en varias zonas geográficas de la provincia de Granada durante el periodo de 1982-2007⁹.

ANÁLISIS ESTADÍSTICO

En muchos trabajos sobre estudios ecológicos, el análisis de los datos entre los factores de riesgo de exposición y la tasa de enfermedad se realiza por medio de pruebas estadísticas de asociación. Lo más común es medir la existencia de correlación lineal entre dos variables cuantitativas o de asociación entre variables cualitativas.

Mientras que los estudios observacionales con sujetos individuales podemos estimar directamente medidas de efecto (diferencias de riesgos), de riesgo (riesgo relativo, *odds ratio*) o de impacto (fracción atribuible), en los estudios ecológicos esto no es posible porque al usar medidas agrupadas desconocemos cuanto sujetos expuestos o no expuestos tienen la enfermedad.

Para estimar el efecto en los estudios ecológicos lo más habitual es recurrir a modelos de regresión entre la tasa de enfermedad (variable dependiente) y los factores exposición o prevalencia (variables independientes). El modelo más frecuentemente utilizado es el de **regresión lineal**, ya que las variables habitualmente son continuas, pudiéndose realizar modelos predictivos según la siguiente ecuación: $Y = b_0 + b_1X$, donde Y sería la tasa de enfermedad, X el factor de riesgo, b_0 sería la constante del modelo y b_1 el cambio medio de la tasa de enfermedad por unidad de exposición (X)^{1,3}. Con este modelo aditivo se asume una estimación del efecto biológico a nivel individual^{11,12}. De la ecuación anterior podemos calcular las tasas de enfermedad en los totalmente expuestos ($X = 1$) y en los totalmente no expuestos ($X = 0$), y a partir de ellas la diferencia de tasas, la razón de tasas, la fracción atribuible en los expuestos y la fracción atribuible poblacional¹⁰.

Para la estimación de los coeficientes de la ecuación de regresión se tienen en cuenta los valores medios de exposición; sin embargo, para calcular las tasas nosotros recurrimos a los valores extremos (0 y 1), lo que en ocasiones da lugar a tasas desproporcionadas o incluso negativas. Una posible solución es emplear **modelos multiplicativos log-lineales** o modelos

de **regresión de Poisson** si la variable objeto de estudio es muy poco frecuente o es discreta. Por último, cuando se miden variables en distintos niveles se pueden emplear **análisis multinivel** donde se pueden separar los efectos a nivel individual y poblacional y evaluar sus interacciones².

SESGOS

Los estudios ecológicos, por su especial diseño en el que la unidad de análisis es agregada, están sujetos a numerosos problemas metodológicos o sesgos que pueden desvirtuar la relación entre la exposición y la tasa de enfermedad de manera importante. Algunos de los sesgos presentes en estos estudios tienen peculiaridades propias que merece la pena comentar.

- **Sesgo ecológico**^{3,12}. Es el sesgo más frecuente e importante y la principal limitación de estos estudios. Ocurre cuando la medida agrupada elegida no mide de forma adecuada el efecto biológico a nivel individual, siendo errónea entonces la inferencia a nivel individual de una asociación encontrada a nivel ecológico (falacia ecológica). Su origen está en heterogeneidad de los niveles de exposición dentro de los grupos y en la existencia de diferencias de clasificación o de efecto entre los grupos.
- **Sesgo de confusión**. Clásicamente en los estudios a nivel individual se produce confusión cuando la variable de exposición estudiada no solo está relacionada con la enfermedad sino también con otro factor de riesgo de la enfermedad. En los estudios ecológicos, puede existir confusión si el factor de riesgo a nivel ecológico se comporta de forma análoga como en los estudios individuales, pero también aunque este no tenga relación con la exposición; sin embargo, factores confundentes únicamente a nivel individual puede que no produzcan confusión a nivel grupal¹³. **El control de la confusión**¹ se puede realizar de dos maneras. La primera, introduciendo las posibles variables confundentes en el modelo como covariables y realizar un análisis multivariante, lo que complica sobremanera el cálculo del efecto. La segunda posibilidad es estandarizar (ajustar) las tasas de enfermedad por las variables de confusión y realizar una regresión con las tasas ajustadas; para ello es imprescindible que todas las variables incluidas en el modelo se ajusten también a la misma variable de confusión y que deben conocerse las distribuciones conjuntas (covarianzas) de las variables, hecho que pocas veces está disponible en estos estudios. Por último, hay que decir que la aplicación de estas técnicas y otras más recientes como el análisis multinivel¹⁴, no garantiza el control de la confusión, ya que muchas veces el origen está en características no conocidas de la distribución de los datos entre los grupos.

VENTAJAS Y LIMITACIONES

Algunos de estos aspectos ya se han adelantado con anterioridad. Resumiendo, podemos decir:

Ventajas

- Son estudios baratos, rápidos porque utilizan datos de fuentes secundarias, generalmente disponibles. Relativamente sencillos en su análisis, aunque ya hemos detallado que a veces esto no se cumple.
- Son ideales para el estudio de nuevas hipótesis de trabajo.
- Son los únicos disponibles cuando no es posible medir la exposición de forma individual, por ejemplo, en estudios ambientales.
- Son muy útiles cuando la medida de efecto solo se puede medir a nivel ecológico. Por ejemplo, el impacto de una campaña de vacunación.

Limitaciones

- La mayor limitación es la realización de inferencias causales a nivel individual (falacia ecológica).
- A menudo no es posible controlar la confusión, por la existencia de factores propios de la distribución de las variables que rara vez están disponibles.
- La variabilidad de exposición dentro de una misma área geográfica hace que los estimadores de efecto de la enfermedad pierdan valor.
- La habitual ausencia de direccionalidad de estos estudios hace difícil asegurar la gradación temporal entre la exposición y enfermedad.
- Es frecuente que exista colinealidad entre las variables independientes, sobre todo en estudios de grupos múltiples, lo que daría a un error en la estimación de los coeficientes.

En la Tabla I se puede ver una comparación entre los estudios ecológicos e individuales.

TABLA 1. COMPARACIÓN DE LOS ESTUDIOS ECOLÓGICOS E INDIVIDUALES

	Estudios ecológicos	Estudios individuales
Variables	Agregadas	Individuales
Tiempo de realización	Rápidos	Lentos
Economía	Baratos	Caros
Tipos	Descriptivos, analíticos	Descriptivos, analíticos
Inferencia	Un solo nivel: agrupado	Nivel individual
Análisis e interpretación	Relativamente sencillos	Sencillos
Utilidad	Faltan datos individuales Programas de salud Variables actual nivel grupal	Presentes datos individuales
Sesgos	(++++)	(++)
Tipos de sesgos	Ecológico/colinealidad/ambigüedad temporal	Información/confusión
Control sesgos	Difícil/no posible	Fácil/posible
Análisis	Efecto: posible/difícil	Posible/fácil
Inferencia efecto	No siempre es posible	Posible/fácil

BIBLIOGRAFÍA

- Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3.ª edición. Filadelfia: Lippincott Williams & Wilkins; 2008.
- Diez-Roux AV. Multilevel analysis in public health research. *Annu Rev Public Health*. 2000;21:171-92.
- Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. *Annu Rev Public Health*. 1995;16:61-81.
- Cressie N. Regional mapping of incidence rates using spatial Bayesian models. *Med Care*. 1993;31:YS60-5.
- Clayton DG, Bernardinelli L, Montomoli C. Spatial correlation in ecological analysis. *Int J Epidemiol*. 1993;22:193-202.
- Sarría Santamera A, Franco Vidal A, Redondo Martín S, García De Dueñas Geli L, Rodríguez González A. Hospitalizaciones en menores de un año en la ciudad de Madrid y su relación con el nivel social y la mortalidad infantil. *An Esp Pediatr*. 2002;57:220-6.
- Vázquez Fernández ME, Luquero Alcalde FJ, Pastor García E, Bachiller Luque MR, Vázquez Fernández MJ, Eiros Bouza JM. Análisis del consumo de antibióticos en la población pediátrica de Castilla y León durante el periodo 2001 a 2005. *An Pediatr (Barc)*. 2007;67:11-7.
- Loomis D, Castillejos M, Gold DR, McDonnell W, Borja-Aburto VH. Air pollution and infant mortality in Mexico City. *Epidemiol*. 1999;10:118-23.
- Aceituno-Madera P, Buendía-Eisman A, Olmo FJ, Jiménez-Moleón JJ, Serrano-Ortega S. Melanoma, altitude, and UV-B radiation. *Actas Dermosifiliogr*. 2011;102:199-205.
- Morgenstern H. Uses of ecologic analysis in epidemiologic research. *Am J Public Health*. 1982;72:1336-44.
- Piantadosi S. Invited commentary: ecologic biases. *Am J Epidemiol*. 1994;139:761-4.
- Greenland S, Robins J. Invited commentary: ecologic studies—biases, misconceptions, and counterexamples. *Am J Epidemiol*. 1994;139:747-60.
- Greenland S, Morgenstern H. Ecological bias, confounding, and effect modification. *Int J Epidemiol*. 1989;18:269-74.
- Blakely TA, Woodward AJ. Ecological effects in multi-level studies. *J Epidemiol Community Health*. 2000;54:367-74.