

Los estudios epidemiológicos analíticos como parte de la investigación de una epidemia

Victor Manuel Cardenas, Dora Rafaela Ramírez y Gloria Inés Suárez Rangel

Resumen

Al responder a una epidemia, el profesional epidemiólogo (a) de campo, establece mediante la epidemiología descriptiva hechos que sugieren una o mas hipótesis sobre factores de riesgo del ambiente (incluyendo las instituciones, sus políticas y otros determinantes sociales), el huésped o el agente y que requieren ser probadas. Las hipótesis deben formularse de manera que la expresión incluya el diseño del estudio y tenga una referencia a la hipótesis de nulidad, es decir, explícitamente diga como se va probar la hipótesis. La evidencia que se requiere implica recolectar sistemáticamente la información sobre las exposiciones de interés y el estado de salud mediante estudios epidemiológicos analíticos (por ejemplo, ecológicos, de cohortes o de casos y controles). Discutimos cuando está indicado escoger un tipo de estudio sobre otro y las ventajas y desventajas de los estudios de cohorte y de casos y controles.

Teniendo en cuenta la premisa de que dos epidemiólogos nunca seguirán exactamente el mismo camino de investigación, como menciona Michael B. Gregg en su libro *Epidemiología de Campo* [1], surge la pregunta de cuándo realizar un estudio epidemiológico, y cuándo utilizar un diseño u otro, ya sea de casos y controles, o de cohortes en una investigación de epidemia. Antes de contestar la pregunta, es necesario repasar los pasos en una investigación epidemiológica de una epidemia o brote. Más adecuadamente, la indagación o investigación en terreno de una epidemia o brote es parte de la respuesta de emergencia a un brote epidémico o epidemia. Empezaremos introduciendo el concepto, como nos lo enseñó el Dr. James S. Koopman en 1994 clase del curso de métodos intermedios en

Bogotá, que hay dos etapas en la indagación, una descriptiva, que a propósito de este artículo se llamará preliminar, fundamental para la indagación y que, aunque preliminar incluye aspectos de intervención en la respuesta, y la segunda fase, que es la analítica, que en este artículo se denominará definitiva y que puede o no confirmar los hallazgos de la fase preliminar. Los pasos de la indagación se recapitulan en el cuadro 1.

Cuadro 1. Pasos en la indagación sistemática como parte de la respuesta de salud pública a un brote epidémico

Fase Preliminar

1. Preparar la investigación luego de recibir la notificación o rumor

(equipo, papeles, responsabilidades, invitación, contactos, revisar la literatura, cuestionarios previamente usados, computadoras, suministros de recolección de muestras, transporte, alojamientos, viáticos)

2. Determinar la existencia del brote

(notificaciones pasadas contra la actual, o actual contra esperado, rasgos inusuales o nunca antes vistos)

3. Establecer el diagnóstico y una definición de trabajo de los casos

(con base en datos clínicos de entrevistar y explorar de preferencia personalmente, y pruebas de laboratorio, si no hay un diagnóstico *per se*, o al menos describir un síndrome, por ejemplo “síndrome febril indiferenciado”, adicione criterios epidemiológicos-variables tiempo, lugar, persona si es adecuado)

4. Buscar y enumerar o contar los casos por vigilancia pasiva o activa

(casos conocidos, identificar casos desconocidos, búsqueda entre contactos, en poblaciones a riesgo, sitios centinela, o una muestra aleatoria para estimar el total si son demasiados)

5. Elaborar y mantener actualizado un listado de casos y recolectar datos de exposiciones y factores de riesgo para orientarlos en persona, lugar y tiempo y calcular las tasas de ataque relacionando los casos con la población de riesgo

(epidemiología descriptiva, utilizar denominadores para estimar riesgos, estandarizar tasas de ataque y de mortalidad)

6. Formular hipótesis sobre el agente, el huésped y el ambiente, incluyendo aquellos inherentes a las instituciones y la sociedad.

(enunciar el diseño en la declaración de nulidad, por ejemplo, “las tasas de ataque en adultos fueron semejantes a las de niños”)

7. Considerar si las hipótesis están apoyadas en los hechos establecidos por la epidemiología descriptiva

(un ejercicio de inferencia, lógica e interpretación de los resultados del análisis descriptivo)

8. Establecer medidas de prevención y control

(generalmente se inician desde que se ha hecho el diagnóstico o aún antes con base en sospechas y que se refinan conforme se avanza en la indagación)

9. Planear, ejecutar y evaluar un plan de comunicación en salud pública

(informes inmediatos, diarios, semanales, notificación a los niveles regionales, nacionales o internacionales de acuerdo con procedimientos acordados)

10. Establecer vigilancia

(para evaluar el efecto de las medidas de control y detectar olas secundarias o subsecuentes)

Fase definitiva

1. Determinar la necesidad y factibilidad de indagaciones más sistemáticas

(considerar qué se conoce y desconoce para asignar las prioridades)

2. Determinar las estrategias epidemiológicas para probar de la manera más eficiente las hipótesis

(Estudios epidemiológicos, -ecológicos, cohortes y casos y controles, trazar productos que podrían haber vehiculado al agente, estudios ambientales incluyendo vectores, entre otros)

3. Recolectar y analizar la información de preferencia en terreno

4. Contrastar los resultados del o los estudios con los hechos establecidos por la epidemiología descriptiva

5. Reafirmar, ajustar o rectificar las medidas de control y prevención y la estrategia de comunicación de acuerdo con los nuevos resultados

En la práctica de la salud pública, en la primera fase al responder y estudiar una epidemia, el o la epidemiólogo(a) 1) preparara el trabajo de campo (antecedentes, invitaciones, definición de los papeles de los miembros del equipo, la documentación, entre otros); 2) determina la existencia de una epidemia, 3) obtiene un diagnóstico así sea solamente sindromático de la condición o enfermedad epidémica, desarrollando y buscando consenso en una definición de trabajo de casos a la que le agregará elementos de las variables de tiempo, lugar y persona en tanto sea adecuado; 4) busca o enumera a todos o la mayoría de los casos, o una muestra aleatoria de los casos si son demasiados, como ocurre a menudo en epidemias de enfermedades respiratorias o entéricas, (como edad y sexo), mediante vigilancia pasiva o activa, incluso por encuestas en hogares o sitios de trabajo y otros, 5)

elabora una lista de los casos con fechas y horas de inicio de síntomas, lugares de residencia y sitio de trabajo o escuela u otros, de ocupación y otros factores de riesgo relevantes para hacer la epidemiología descriptiva. 6) desarrolla hipótesis de trabajo sobre las exposiciones que pudieran haber dado lugar a la epidemia, sobre factores del agente, huésped y ambiente; 7) considera si una hipótesis tiene o no base en los hallazgos de la epidemiología descriptiva; 8) establecer o recomendar las medidas de control de la epidemia a las que hubiera lugar; 9) establece, ejecuta y evalúa un plan de comunicación que incluye las notificaciones a los distintos niveles del sistema de vigilancia y los programas de control, como con los medios y 10) establece y evalúa una vigilancia al menos ampliada que le permita evaluar las recomendaciones que se implementaron y anticipe el resurgimiento de la enfermedad en la población. Estos pasos se toman no necesariamente en ese orden. A veces, por ejemplo, si se trata de una condición que puede ser fatal, como la intoxicación botulínica de alimentos, con tan solo la sospecha de botulismo, el practicante debe indagar cómo procurar la antitoxina ya que toma tiempo conseguirla, e instaurar medidas que prevengan el consumo de alimento remanente que pudiera estar contaminado con la toxina. De haberse equivocado el diagnóstico clínico lo peor que puede pasar, en este ejemplo, es se tengan que cancelar los pedidos de toxina y liberar el alimento sospechoso.

El descubrimiento de todos los casos es muy importante, sobre todo cuando la frecuencia relativa, medida digamos por la tasa de ataque, es baja (<10%), y a veces requiere el ampliar la vigilancia pasiva e incluso establecer una vigilancia activa o el hacer una indagación en viviendas u otros sitios como escuelas o sitios de trabajo para obtener un cuadro completo de lo que ha ocurrido o está ocurriendo en la comunidad. El obtener una descripción exacta de lo que ocurre en la comunidad, representa el fundamento en que se basa todo el ejercicio científico de la disciplina epidemiológica. Si no hay una buena caracterización epidemiológica se estarán dando “palos de ciego”. Es de lo más importante el contar con una buena descripción de la ocurrencia de la enfermedad o condición de interés, por variables de persona, lugar y tiempo.

Los datos descriptivos pueden ser analizados con base en frecuencias absolutas y proporciones, pero es mejor relacionar los casos a la población a riesgo, utilizando una medida de frecuencia como la tasa de ataque ($\frac{\text{casos nuevos}}{\text{Población a riesgo}} \times k$) que es una proporción, k es una constante, generalmente 100, y la tasa de ataque es un nombre aplicado en epidemiología de terreno a la incidencia acumulada o riesgo, generalmente durante el estudio de un brote. La población a riesgo es aquella población de la que provienen los casos, sin reducirla solamente a los expuestos, sino a los potencialmente expuestos. Por ejemplo, en un brote de intoxicación alimentaria en una cafetería en una escuela, la población a riesgo son las personas que asistieron a la cafetería durante el periodo de riesgo. Si se sospecha digamos que el periodo de incubación es de 24 horas, se puede inferir el periodo de riesgo usando el rango y la mediana a partir de las fechas y horas de inicio. Como es sabido, si se conoce el periodo de incubación y los signos y síntomas y su duración se puede inferir el agente, aún antes de que los resultados de laboratorio estén disponibles. Sin embargo, la confirmación diagnóstica, aunque no sea esencial para el control, como discutiremos adelante, es importante. Las hipótesis sobre la etiología se postulan con base en los datos clínicos y el comportamiento epidemiológico de la enfermedad o condición bajo estudio.

El o la epidemióloga debe obtener una descripción de los signos y síntomas, conocer la etiología de estar disponible (por ejemplo, intoxicación alimentaria por *Salmonella enteritidis*), su duración, gravedad medida por la letalidad u otra manera, la tasa de ataque y distribución por persona lugar y tiempo, su modo de ocurrencia, incluyendo factores ambientales que contribuyeron a la epidemia para completar la investigación de esta. Siguiendo el ejemplo del brote de la cafetería de la escuela, será importante la descripción de los hallazgos de estudios de laboratorio de muestras de alimentos servidos, así como y la forma en que se prepararon los alimentos, describiendo los puntos críticos en que se pudo haber producido alguna contaminación. Todos estos elementos forman parte de esta indagación preliminar.

La mayoría de las veces, la investigación epidemiológica termina ahí, por varias razones. La primera razón es indudablemente que, en el contexto de la práctica de salud pública, hay otras demandas que hacen impertinente hacer énfasis en la investigación a expensas del uso de los recursos, el más importante de los cuales es el tiempo que el mismo epidemiólogo(a) y el equipo pueda dedicar a la investigación. Los Dres. Richard Goodman, Jay Buehler y Jeff Koplan, de los Centros para el Control y Prevención de las Enfermedades en 1990 [2] sintetizaron esta situación en el cuadro tetracórico o 2x2 siguiente (Cuadro 2).

Cuadro 2. Énfasis relativo de los esfuerzos en la investigación o control durante la respuesta a una epidemia de acuerdo al conocimiento de la fuente, modo de transmisión o agente causal

Agente u otros factores de riesgo críticos prevenibles	Fuente/Modo de ocurrencia o transmisión	
	Conocidos	Desconocidos
Conocidos	+++ Control	+Control
	+Investigación	+++Investigación
Desconocidos	+++Control	+Control
	+++Investigación	+++Investigación

+++ énfasis mayor + énfasis mínimo

Fuente: Referencia 2.

Como parte de la investigación de un brote epidémico o epidemia, ya sea que el énfasis esté en la investigación o no, se considera importante que se incluya la recolección de datos ambientales y otros que den contexto a los hallazgos de la epidemiología descriptiva y los datos clínicos y de laboratorio. Asimismo, debe dejarse establecido algún nivel de vigilancia, dado que a menudo el(la) epidemiólogo(a) puede ser sorprendido por el hecho de que la epidemia transite de un periodo de transmisión u ocurrencia baja, a una epidemia con olas secundarias o ulteriores a menudo aún mayores que la inicial.

Como parte de la descripción uno compara tasas de ataque por características utilizando los mejores denominadores disponibles. Tal comparación involucra necesariamente una hipótesis subyacente, es decir que hay diferencias en las tasas por características de lugar, persona y tiempo. La prueba de estas hipótesis que, a veces desdeñosamente se dice que es “muy descriptiva”, por no incluir diseños con los que uno está más familiarizado en los textos comunes de epidemiología, consiste en la comparación de experiencias de grupos de personas, o de la ocurrencia de la epidemia en el tiempo. Ambos implican estudios ecológicos, que fueron descritos por el Dr. Hal Morgenstern en 1982 [3] como la comparación de la ocurrencia de enfermedades entre características relacionadas con exposiciones de interés que son medidas a nivel de grupos de individuos para hacer inferencias sobre la relación entre exposiciones y la ocurrencia de enfermedades.

Por ejemplo, la comparación de tasas de COVID-19 entre Hispanos o Latinos en El Paso, Texas con las tasas correspondientes en personas no Hispanas o Latinos de esa comunidad, con razones de tasas de morbilidad y mortalidad de 1.6 y 2.0, respectivamente son un ejemplo de tal análisis de la experiencia de grupos. En el mismo estudio de la epidemia, los autores hicieron un estudio ecológico de series de tiempo que según Morgenstern son aquellos en que "un cambio de exposición, como el inicio de un programa de intervención, compara la pendiente en la tendencia de la enfermedad antes y después de la intervención" [3]. Los autores determinamos el efecto aparente de las órdenes del cierre de escuelas y negocios en la ocurrencia subsecuente de los casos notificado de COVID-19 por día en el condado de El Paso, Texas, hallando una reducción de 3.8% por día después de la intervención [4]. Los autores reconocieron entre las limitaciones de ambos hallazgos que pueden haber jugado un papel otras variables subyacentes en la diferencia por etnicidad o en la dinámica de la epidemia. Por ejemplo, los Hispanos pueden haber estado más expuestos que los no-Hispanos, por razones relacionadas a su bajo ingreso y nivel socioeconómico, o por ocupaciones de más riesgo o tener más hogares con mujeres como las únicas personas en la vivienda con empleo remunerado. En el segundo hallazgo sobre el efecto de las ordenes de cierre en la notificación diaria de casos, los

autores reconocieron que al mismo tiempo que se dieron las ordenes, hubo mayor uso de mascarillas, higiene de manos y se evitaron reuniones.

El estudio de la epidemia a veces continúa en lo que se considera la fase definitiva de la indagación por ejemplo para verificar hallazgos como los anteriormente señalados. Esta fase gira alrededor de hipótesis de trabajo, que pueden referirse digamos a la etiología del brote, o el mecanismo de transmisión o en términos más generales de su ocurrencia, o a factores prevenibles del huésped o ambiente, incluyendo las políticas de salud, y otros determinantes del estado de salud de las poblaciones. Si existen razones para evaluar hipótesis de trabajo, éstas surgen a menudo de la fase cualitativa de la indagación epidemiológica. Estas se basan a menudo en pistas generadas por las historias contadas por los casos, o de patrones de exposiciones que tienen los casos como se puede ver en el listado de casos. Un hallazgo de la experiencia de grupos puede generar hipótesis a comprobarse a nivel individual. Los autores del estudio COVID-19 en El Paso, Texas, ya referido, tuvieron la oportunidad de probar si era la etnicidad Hispana/Latina o el nivel socioeconómico o la ocupación los factores relacionados con el exceso observado en las tasas. Utilizando datos de una muestra de individuos encontraron que la etnicidad independientemente del nivel socioeconómico medido por escolaridad y la ocupación, y ajustada por edad y sexo, incrementó el riesgo de infección por SARS-CoV-2 (Cociente de Suertes [CS] de 6.6 (u OR -Odds Ratio en inglés) [4].

A menudo son varias hipótesis, a veces en competencia entre sí, las que surgen en la etapa preliminar de la indagación epidemiológica como parte de la respuesta al brote o epidemia. Las mejores hipótesis deben enfocarse en *factores prevenibles* del ambiente, huésped o agente, por ejemplo, el vehículo, si se sospecha alguno con base en la epidemiología descriptiva que lo hubo. Un brote de fuente común puntual en el tiempo suele tener un solo pico del que se puede inferir el punto de exposición substrayendo el valor de la mediana, más que el de la media, del periodo de incubación del pico de la fecha/hora de inicio de síntomas.

Una hipótesis de trabajo debe enunciar cómo se probará tal hipótesis. Una hipótesis es un enunciado o declaración de la relación entre una exposición de interés y la ocurrencia de la enfermedad, mientras que una pregunta de investigación es el enunciado de esa relación en forma interrogativa. Es necesario saber que la prueba estadística es sobre hipótesis nulas o de nulidad, es decir de que no existen diferencias según la medida de asociación que vamos a emplear. Las medidas de asociación como la razón o diferencia de tasas o riesgos, o la razón de probabilidades o cociente de suertes, tiene valores nulos o de nulidad (por ejemplo, razón de riesgos =1).

Supongamos que, en la epidemia en la escuela descrita antes, el epidemiólogo encontró en las muestras de heces y el remanente de alimentos servidos *S. enteritidis*, un patógeno que a menudo se encuentra en huevos y carne de pollo. Digamos que, para efecto de esta presentación, la parte cualitativa de la indagación -obtenida por entrevistas con varios trabajadores afectados y la administración de la fábrica, se supuso que, entre los diferentes alimentos y platillos servidos para el almuerzo en la escuela, la mayonesa casera podría haber sido el alimento que vehiculó al patógeno. La parte descriptiva añadió que la tasa de ataque fue del 32%. Como hay una población de riesgo que es numerable y que experimentó un riesgo alto (es decir, >10%) [5], el(la) epidemiólogo(a) piensa en un estudio de cohorte y escribe en sus notas la hipótesis alternativa: los asistentes a la cafetería de la escuela que consumieron mayonesa casera tuvieron una tasa de ataque de gastroenteritis mayor que los que no consumieron mayonesa casera (es decir, la razón de riesgos (RR) y fue estadísticamente diferente de 1, el valor de nulidad ($H_0: RR = 1$). Así la hipótesis enuncia un estudio de cohortes como parte del estudio del brote. En el estudio al que nos referimos, la razón de tasas de salmonelosis entre quienes consumieron mayonesa y quienes no la consumieron fue de 31.8.

El estudio de casos y controles es adecuado en situaciones en que la ocurrencia es menor del 10% y no hay una población a riesgo fácilmente numerable. Una circunstancia en la que funcionaría tal diseño, utilizando el ejemplo de la salmonelosis, para seguir en el tema, si hubiera por ejemplo cepas de salmonela con cierto patrón molecular que digamos fueron descubiertos en un laboratorio que sirve

diferentes jurisdicciones de salud pública. Por ejemplo, durante el estudio de un brote de *Salmonella enteritidis typhimurium* que ocurrió en el 2004 en nueve estados de EEUU, la historia de los casos sugirió que carne de res molida de una marca particular podría estar implicada. Para poner a prueba esta hipótesis, los investigadores ensamblaron una serie de 31 pacientes de los que se aislaron cepas *S. typhimurium* indistinguibles entre sí por electroforesis de campo pulsado en gel, y los controles fueron seleccionados entre personas seleccionadas mediante una encuesta telefónica de dígitos seleccionados al azar, que apareados por frecuencia por edad con los casos y para ser considerados controles no debían haber enfermado en los últimos siete días [6]. La hipótesis nula fue que el CS o la razón de probabilidades pareado era igual a 1. Si se hallara que medido por las suertes (lo que en inglés se llaman odds), y su cociente (CS) mayor que 1 para el consumo de carne molida de esa marca, se podría inferir que el riesgo de contraer *S. typhimurium* entre quienes consumieron carne de res molida de esa marca particular fue mayor que entre quienes no la consumieron, es decir, si el CS o razón de probabilidades fuesen estadísticamente diferente del valor de nulidad ($H_0: CS = 1$). El análisis pareado de los datos acerca del consumo de res molida de esa marca en el brote en cuestión arrojó un estimado del CS de 12.7.

Para relatar otro estudio de casos y controles como parte de una investigación de una epidemia, en 1986 estudiamos el último brote epidémico de poliomiélitis en México. La epidemiología descriptiva incluyendo los datos de antecedentes vacunales de los casos, si se distinguía el número de dosis recibidas antes de enfermarse de las dosis recibidas después de enfermarse, mostraban que 79.7% de los casos habían recibido menos de tres dosis y que los casos se conglomeraban en comunidades rurales. El análisis de una encuesta de cobertura de vacunación por conglomerados mostro una alta cobertura en el estado, pero en localidades con baja cobertura de vacunación de acuerdo con la edad pediátrica estaba correlacionada con el nivel de cobertura alcanzado en la nueva estrategia llamada Días Nacionales de Vacunación (DNV). Decidimos realizar un estudio a nivel individual de casos de poliomiélitis y controles seleccionados al azar entre los niños listados como elegibles para asistir a los

DNV (2 por caso) y entre los vecinos de los casos (2 por caso). Se compararon los resultados entre los dos grupos para evaluar si había diferencias en los estimados del CS según el tipo de caso, es decir para identificar algún sesgo de selección. Los controles fueron pareados por edad (± 3 meses en menores de 3 años y ± 6 meses en los niños de 3-5 años). Los resultados del análisis pareado se muestran en el cuadro 3, donde hubo 5 conjuntos pareados, en que los controles estaban sin vacunar y los casos si estaban vacunados, mientras que en 27 instancias los casos estaban sin vacunar mientras que los controles sí estaban vacunados. La razón de estos conjuntos con estatus discordantes, el CS o razón de probabilidades fue 5/27 o 0.19, del que se calcula la eficacia vacunal $1 - \text{CS}$ o $(1 - 0.19=0.81)$ del 81% [7]. No hubo diferencias en el estimado del CS por tipo de control.

Cuadro 3. Comparación pareada de los 18 casos de poliomielitis y sus 63 controles a la vacunación antipoliomielítica, Sinaloa, México, 1984-1986

Estatus vacunal de los casos	Estatus vacunal de los controles		Total
	Vacunados (3+ dosis)	No Vacunados	
Vacunados (3+ dosis)	15	5	20
No Vacunados	27	16	43
Total	42	21	63

Fuente: Referencia 7

Los resultados de los análisis de los estudios de cohortes y casos y controles descritos antes deben ser contrastados con los hallazgos de la epidemiología descriptiva. Deberían ser consistentes en cuanto a explicar la mayoría de los casos si se trata de un brote de fuente común. Además, los hallazgos pueden refinar la búsqueda de detalles en la preparación de los alimentos, o en general la cadena de eventos que llevo a la exposición. Por ejemplo, en el caso de brotes por alimentos, puede encontrarse asociaciones débiles con otros alimentos, y que sugiera que hubo una contaminación cruzada digamos por el uso de los mismos utensilios. También el hallazgo de exposiciones que

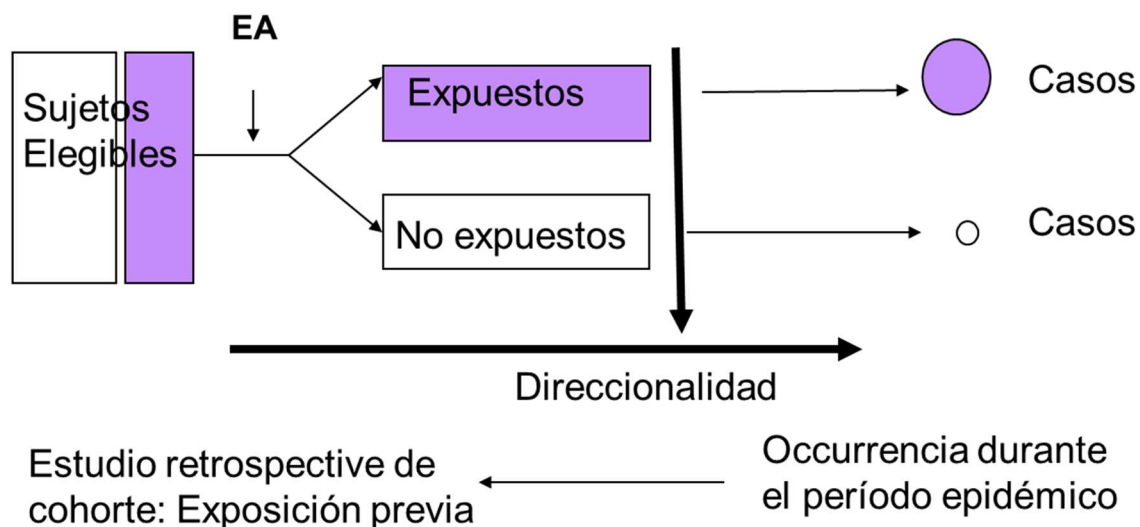
aparentan ser “protectoras” (esto es, por tener un RR o CS <1) podría explicarse porque quienes comieron esos alimentos “protectores” evitaron los alimentos implicados, lo cual puede mirarse en los datos disponibles.

El lector se preguntará si acaso hay equivalencias entre estos estudios, el de casos y controles y el de cohortes, ya que aparentan ser semejantes. Aunque los resultados son equivalentes, nada más lejos de la realidad. El estudio de cohorte es por muchas razones más informativo que el de casos y controles. Uno puede entender una cohorte como una población que comparte una experiencia en común. De hecho, el nombre se toma de la palabra *cohors* en Latín que significa unidad de soldados, dado que en el ejército romano una cohorte eran 600-800 soldados, es decir, seis centurias (100-160 soldados o centuriones o más por centuria) ordenadas en años de servicio o señoría, siendo la primera cohorte la más experimentada. El equivalente moderno de las cohortes romanas en los ejércitos modernos es un batallón. Las cohortes romanas a menudo tenían un estandarte que las identificaba. De manera que es correcto definir en términos de la demografía y la epidemiología a una cohorte como una población definida por compartir una característica o exposición en común. En un estudio de enfermedades crónicas por ejemplo las exposiciones de interés como el hábito tabáquico permiten definir cohortes de fumadores y no fumadores, o puede uno pensar en una cohorte de trabajadores expuestos a fibras de asbesto. En un estudio de cohorte retrospectiva de un brote, digamos de enfermedades transmitidas por alimentos, hay cohortes de personas que consumieron y de las que no consumieron los alimentos indagados. Para cada una de estas exposiciones se elaboran mediciones de la exposición con base en cuestionarios generalmente, y se sigue la ocurrencia de la enfermedad u otro evento en el tiempo. También puede uno pensar en seguir la experiencia de mortalidad en el tiempo de cohortes de personas nacidas en 1900, 1910, 1920, ...y así sucesivamente hasta el año 2000, digamos, para ver su experiencia, por ejemplo, como víctimas de homicidios en un estudio ecológico. De manera que el estudio de cohorte es central en la indagación epidemiológica. Lo distintivo del estudio epidemiológico de cohorte es que las personas son seleccionadas a un estudio

con base en su exposición, no con base en que tengan la enfermedad (ser caso) o no (control). La informatividad del estudio de cohortes radica en que uno puede medir directamente la ocurrencia de la enfermedad según la exposición.

El estudio de cohortes es parecido a un experimento, solamente que la asignación a cada exposición no depende del investigador. Para la mayoría de las exposiciones prevenibles o modificables, es raro que se tenga oportunidad de hacer un experimento. Digamos, por ejemplo, al estudiar el papel de la dieta usual en la salud, por decir en la diabetes o el cáncer, no hay una posibilidad realista o ética de que un investigador asigne a las personas a llevar o no una dieta saludable por años, entonces el estudio epidemiológico de cohortes puede ser pensado como el remedio más cercano al experimento aleatorio, ya que en vida real las trayectorias de vida llevan a la elección de estilos de vida. Del mismo modo, en estudios de brotes epidémicos o epidemias, no hay forma realista o ética de hacer tales experimentos, y sin embargo, ocurren de manera natural. Esta relación entre el diseño del estudio de cohorte retrospectiva que a menudo se hace en un estudio de epidemia y el experimento o ensayo aleatorio (EA), que cuando se trata de enfermos, como se realiza en la investigación en el contexto de la práctica clínica, se denominan ensayos clínicos aleatorizados, se presenta en la figura 1.

Figura 1. Relación entre los ensayos aleatorios y el estudio epidemiológico de cohorte retrospectiva



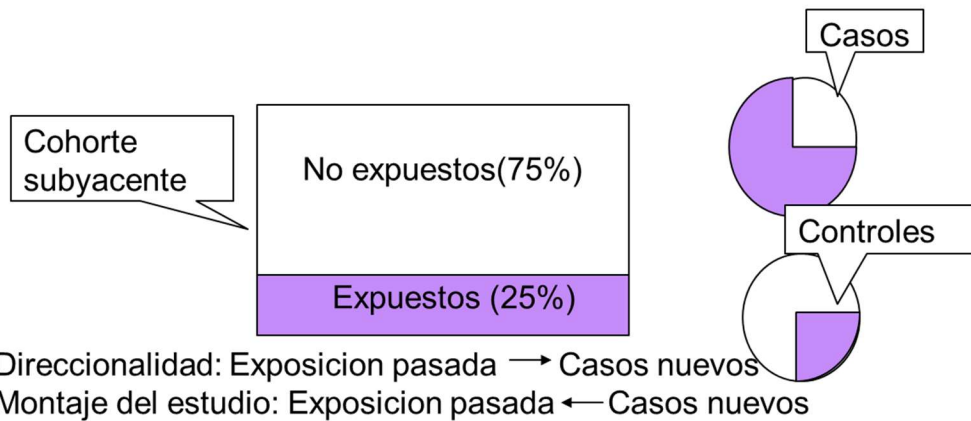
Como se muestra en la figura 1, en el experimento aleatorizado el investigador asigna a los sujetos elegibles a tener o no la exposición, y luego seguirlos en el tiempo para medir la ocurrencia de casos en ambos grupos. Se cuenta con una base poblacional, se puede medir directamente y comparar el riesgo o la incidencia entre los grupos según su exposición. Ahora véase bien que, cuando se hace un experimento necesariamente la direccionalidad es prospectiva, los sujetos elegibles (generalmente libres de la enfermedad) son asignados a la exposición, y la ocurrencia en el tiempo es medida prospectivamente. El estudio de cohorte es también siempre prospectivo en su naturaleza misma, porque si uno lo ve bien, las personas que eran elegibles desarrollaron la enfermedad luego de estar expuestas. Por ejemplo, en el estudio del brote en la cafetería escolar, todos los niños y el personal que fue a almorzar no había tenido salmonelosis antes, -si hubiera alguien con esta condición antes de haber comido en la cafetería, puede uno excluir del denominador también. El estudio de cohorte retrospectivo es aquel en que, al momento que se empieza el estudio, ya han ocurrido todos o casi todos los casos y por supuesto la exposición. Sin embargo, uno puede saber qué ocurrió primero. Es el caso del estudio del brote de salmonelosis en la cafetería escolar: se notificaron los casos, se identificaron algunos más, y luego se encontró que en común tenían haber almorzado en la cafetería, y de ahí se decidió encuestar a todos los que fueron a la cafetería en el periodo a riesgo para ver que artículos del menú comieron todos los asistentes y comparar las tasas de ataque por alimentos consumidos. El estudio de cohorte retrospectivo es el más comúnmente utilizado en la epidemiología de campo.

Ahora pensemos qué es el estudio de casos y controles en este mismo contexto. El estudio de casos y controles es una contribución original de la epidemiología a los métodos de investigación y veamos por qué es tan valioso: es aquel estudio epidemiológico en el que se empieza por seleccionar individuos afectados por un evento o condición o enfermedad de interés, quienes forman una serie (llamada serie de casos) y se compara usando el cociente de suertes de exposición o así llamada también la razón de probabilidades con individuos que, perteneciendo a la cohorte subyacente de la que provienen los

casos, están libres del evento, condición o enfermedad de interés. Los controles son una muestra de aquellos miembros de la cohorte subyacente libres de la enfermedad, y que, de haber tenido tal evento, condición o enfermedad, serían parte de la serie de casos. Ellos forman la serie control, es decir, son los controles.

La figura 2 muestra la relación entre la cohorte y el estudio de casos y controles. Nótese que en la cohorte entera solamente el 25% están expuestos. Si los controles son una muestra representativa de la cohorte subyacente de la que provienen los casos, la frecuencia de exposición entre ellos debería reflejar la frecuencia en la cohorte, es decir 25%. Los casos son los individuos afectados que provienen de la cohorte y la frecuencia de exposición entre esta serie dependerá de la relación que existe en la naturaleza entre la exposición y el riesgo de la enfermedad, evento o condición estudiadas. Nótese que, en el ejemplo, el 75% de los casos estaban expuestos. La medida de asociación que se utiliza en los estudios de casos y controles utiliza la comparación de prevalencias de exposición entre las dos series para derivar la probabilidad de haber desarrollado la enfermedad dado que se estuviera expuesto y la probabilidad de haberla desarrollado dado que no se esté expuesto, es decir la razón de riesgos o tasas, y no es simplemente una comparación de prevalencias de exposición. En nuestro ejemplo no es $75\%/25\%$ o 3, sino $(0.75/0.25)$, o la suerte de exposición entre los casos dividido entre $(0.25/0.75)$, o la suerte de exposición entre los controles, o $3/0.33$ o 9.1, el cociente de suertes o razón de probabilidades u odds ratio en inglés, que aproxima a la razón de riesgos o tasas.

Figura 2. Relación entre el estudio de cohorte y el estudio de casos y controles



Nótese que tanto el estudio de cohortes como el estudio de casos y controles, la direccionalidad o temporalidad, el que la exposición preceda o no a la ocurrencia del evento, condición o enfermedad, queda bien establecida por una indagación correcta que precise la línea de tiempo en que ocurren tanto la exposición como la ocurrencia del evento, condición o enfermedad de interés. Lo genial del estudio de casos y controles es que se parte de ensamblar las series de casos y sus controles para después indagar la exposición pasada, sin necesidad de esperar la ocurrencia de los casos, que es problemático en enfermedades crónicas de latencia prolongada o bien enfermedades transmisibles también de incubación prolongada. Cuando se estudian brotes epidémicos, sin embargo, como el estudio de cohorte es generalmente retrospectivo, no hay tal ventaja, en ahorro de la espera a la ocurrencia de los casos, solamente en términos de costos de recolección de la información entre decenas o cientos de personas expuestas y no expuestas. Sin embargo, el ahorro de hacer un estudio de casos y controles en tales circunstancias viene a costa de no poder calcular directamente el riesgo u otras medidas de ocurrencia.

Es claro que una ventaja del estudio de casos y controles sobre el estudio de cohortes es, no obstante, que no hace falta una enumeración de todos o una muestra de expuestos y no expuestos. Sin embargo, el estudio de casos y controles no permite estudiar más que una enfermedad, condición o evento a la vez, mientras que el estudio de cohortes permite estudiar diferentes enfermedades, condiciones o eventos. A su vez, en general, el estudio de casos y controles permite explorar más exposiciones que

el estudio de cohortes. Este último se limita generalmente, por razones logísticas, a unas cuantas exposiciones.

Deseamos terminar esta recapitulación, con algunos comentarios sobre malentendidos comunes que observamos sobre estos dos tipos de estudios.

El primero es en relación con llamarle equivocadamente a los estudios de brotes epidémicos de cohorte retrospectivos, “estudios transversales”. Los estudios de cohorte retrospectivos parecieran transversales si no se toma en cuenta que, durante las investigaciones de epidemias sabemos cuándo empezó la enfermedad o episodios de enfermedad entre los afectados y también tenemos claridad de cuál fue la línea de tiempo de la exposición. Estos estudios son diferentes de los estudios transversales basados en encuestas de salud, en los que se hace una indagación en una muestra de la población y se miden simultáneamente la ocurrencia de enfermedad y la(s) exposición(es), a través de preguntas mayormente estandarizadas para encuestas. En las encuestas de salud, por naturaleza de los períodos de referencia de las preguntas, generalmente es imposible conocer que ocurrió primero, si la exposición o la enfermedad o condición de interés. En los estudios transversales se identifican casos existentes no necesariamente nuevos, es decir son prevalentes, por lo que las medidas de asociación son generalmente las razones de prevalencia, no las razones de riesgo o incidencia.

El segundo, es la malinterpretación de los estudios de casos y controles como una moda saludable o deseable en la investigación de epidemias, particularmente adecuado para la investigación de epidemias al investigar varias hipótesis, una especie de “expedición para ver que se pesca” y para afrontar las dificultades encontradas al identificar una población a riesgo numerable [8]. A menudo la falta de una aplicación adecuada de los métodos epidemiológicos en la descripción puede llevar a no identificar una población a riesgo, o, aunque se identifique, a obviar su enumeración por la aparente facilidad de sustituir un método por el otro. Una de las consecuencias de esta confusión es que se haga un estudio de casos y controles, seleccionando como controles miembros aparentemente no

afectados de una población experimentando un alto nivel de riesgo, que resulte en un abultamiento de los estimados de la asociación ya que los cocientes de suerte suelen sobreestimar las razones de riesgo o de tasas cuando la enfermedad es común (esto es >10%).

Referencias

1. Gregg, MB. *Field Epidemiology*. 3rd ed. New York, Oxford University Press, 2008.
2. Goodman RA, Buehler JW, Koplan JP. The epidemiologic field investigation: science and judgment in public health practice. *Am J Epidemiol* 1990;132:9–16. Citado por Dicker RC, Coronado F, Koo D, Parrish RG. Principles of Epidemiology in Public Health Practice, 3rd ed. Atlanta: CDC, 2011. Disponible en <https://www.cdc.gov/csels/dsepd/ss1978/>
3. Morgenstern H. Use of ecologic analysis in epidemiologic research. *Am J Pub Hlth* 1982; 72; 1336-1344. doi:10.2105/ajph.72.12.1336
4. Cárdenas VM, Cardenas IC, Ocaranza HI, Castillo R, Ortega-Madani A, Barrera, S, Urrea C, Mora A. Observations on the interaction of older age and diabetes on COVID-19 mortality, and of non-pharmaceutical interventions on daily reported cases in El Paso, Texas, 2020. *American Journal of Field Epidemiology*, 1(1), 41–52.
5. World Health Organization. Foodborne disease outbreaks: guidelines for investigation and control. Geneva: World Health Organization 2008. <https://apps.who.int/iris/handle/10665/43771>.
6. Centers for Disease Control and Prevention (CDC). Multistate outbreak of Salmonella typhimurium infections associated with eating ground beef--United States, 2004. *MMWR Morb Mortal Wkly Rep*. 2006;55(7):180-182.
7. Cárdenas-Ayala VM, Vilchis-Licón H, Stetler HC, et al. Risk factors for the persistence of wild poliovirus transmission in Sinaloa, Mexico, 1984-1986. *Bull Pan Am Health Organ*. 1988;22(3):227-239.

8. Dwyer DM, Strickler H, Goodman RA, Armenian HK. Use of case-control studies in outbreak investigations. *Epidemiol Rev.* 1994;16(1):109-123. doi:10.1093/oxfordjournals.epirev.a036137.