

## CÓMO PRESENTAR LOS RESULTADOS DE LOS ESTUDIOS EPIDEMIOLÓGICOS<sup>1</sup>

Alexander M. Walker<sup>2</sup>

### Introducción

A menudo, la investigación epidemiológica produce resultados difíciles de sintetizar. Si consideramos que la epidemiología es, en líneas generales, el estudio de los factores determinantes de la salud en las poblaciones humanas, la dificultad principal en esa tarea surge de la distribución no aleatoria de las exposiciones en la población general, y de los errores de medición en los sujetos de estudio. En los ensayos clínicos la aleatorización anticipa y hasta cierto punto define los análisis que presuntamente seguirán, y permite la valoración cuantitativa de las distorsiones que pueden haberse introducido por causas desconocidas o mal medidas de la entidad sometida a estudio. Por el contrario, cuando la exposición puede ir naturalmente aparejada a las fuerzas que producen o previenen la enfermedad, y cuando la calidad de la observación puede resultar modificada por esas mismas fuerzas, tanto en la investigación como en el informe subsiguiente ha de hacerse especial hincapié en ir más allá del formalismo de la comparación previamente planificada.

Si durante la planificación, el diseño y el análisis de los estudios epidemiológicos se considera el informe que habrá de seguir, podrá conseguirse posteriormente una presentación más lúcida de los datos. El lector tendrá a la postre que sentir que ha sido capaz de deshacer la madeja de hallazgos casuales, sesgos e intención tan frecuentes en la investigación epidemiológica. Deberá recibir la suficiente información para juzgar la calidad de la recogida de datos, y la pertinencia de dichos datos para las circunstancias y eventos que son el objeto real del estudio. El lector debe ser guiado a través de una minúscula fracción de todas las relaciones que pueden observarse en los datos, y aun así ha de resultar convencido de que ha obtenido la información esencial. El propósito de las observaciones siguientes es señalar cómo es posible informar

---

<sup>1</sup> Este artículo, solicitado al autor por el Editor de *American Journal of Public Health* apareció originalmente con el título "Reporting the results of epidemiological studies" (*Am J Public Health*; 1986; 76(5): 556-558), con la siguiente *Nota del Editor*: "La forma en que se han indicado los resultados en números recientes de la revista ha provocado preocupación en algunos lectores. Por dicha razón hemos solicitado este trabajo al Dr. Walker y hemos pedido al Dr. Joseph Fleiss que presente su punto de vista distinto en el artículo que sigue a continuación. [...] La presente traducción se publica con autorización de dicha revista y del autor. © American Public Health Association.

<sup>2</sup> Departamento de Epidemiología, Escuela de Salud Pública de Harvard. Dirección postal: Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, Estados Unidos de América.

claramente de los resultados si se presta la debida atención a los métodos de investigación y a los aspectos importantes que subyacen a la realización y notificación de un estudio. Estas ideas constituyen una reflexión realizada a partir de lo que, a mi juicio, han sido informes bien presentados que he observado en el trabajo de otras personas y ocasionalmente en el mío propio.

## Contexto

Un informe debe comenzar por exponer las relaciones que se pretende estudiar y los motivos para considerarlas de interés. Quizá la investigación se suscitó a partir de observaciones de laboratorio, pero más a menudo se trata de series de casos clínicos, estudios de correlación, o análisis epidemiológicos formales que sugieren la presencia de una relación exposición-enfermedad que tiene importancia científica o sanitaria. Como los detalles de exposición y definición de la enfermedad a menudo varían de unas poblaciones a otras, puede ser conveniente repetir casi sin más un protocolo anterior. En cualquier caso, deben hacerse constar las deducciones verificables de otros estudios, exponiéndolas clara y brevemente como hipótesis positivas. Es de poco interés dar cuenta de hallazgos previos que no sean verificables, es decir, refutables.

Un aspecto importante del contexto del estudio, que no tiene nada que ver con sus antecedentes científicos, son sus características operacionales o “logísticas”. Hay que decir si el trabajo se basa en datos recogidos *ad hoc*, si es parte de una serie de estudios llevados a cabo en la misma población, o si es un producto colateral de un estudio de mayor alcance, una investigación de objetivos múltiples o un sistema de vigilancia. En gran parte, el significado de las expresiones trilladas que aparecen en los informes (por ejemplo, “entrevistadores especialmente capacitados”), dependerá del contexto; por ello es imprescindible ofrecer datos, incluso circunstanciales, que indiquen hasta qué punto los resultados deducidos de la investigación pueden suponerse reflejo de la realidad del universo observado. Si los resultados han reproducido relaciones bien conocidas, estas deben notificarse de la manera más concisa posible.

## Sujetos del estudio y población de partida

En todos los informes de descubrimientos epidemiológicos dignos de consideración pueden encontrarse tasas, fracciones y funciones de estas dos medidas elementales. Cada una de ellas exige la definición de una población de partida en cuyo ámbito se hacen las observaciones. La población de partida puede especificarse con detalle o puede resultar implícitamente definida en el proceso de selección de los sujetos de estudio. En cualquier caso, debe ser descrita en los aspectos que puedan tener relación con la frecuencia de la enfermedad y la factibilidad de la recogida de datos. Las características conocidas antes de iniciar el estudio pueden describirse en la sección de métodos, junto con una especificación que en principio ha de permitir determinar en cualquier momento si una persona dada pertenece o no a la población en estudio.

A veces, el intento de descripción de la población de partida correspondiente a un procedimiento de selección de casos echa luz sobre alguna debilidad del estudio. La población de partida de los casos admitidos a un solo hospital es un ejemplo citado a menudo. Puede que una cierta zona sea cubierta por varios hospitales y la predisposición de los pacientes a acudir a un centro u otro puede estar relacionada de manera no cuantificada con factores que tienen interés directo o colateral en un estudio etiológico, por ejemplo, el grupo étnico o el nivel de ingresos. Es mala señal que el investigador sea incapaz de especificar su población de partida en términos adecuados, sean descriptivos u operativos; es presagio de otras vaguedades en el resto del diseño, el análisis y la interpretación.

En la mayor parte de los estudios de cohortes y en muchos estudios de casos y controles, el diseño exige identificar y enumerar todos los miembros de la población de partida, pero en muchos estudios de casos y controles solo se estudia realmente una muestra que representa la población de partida o la población-tiempo expuesta a riesgo (el grupo control). En tal situación, el método de muestreo debe ser descrito con detalle suficiente para que el lector entienda su relación con la población de partida y pueda juzgar si es plausible que el procedimiento de selección haya producido un grupo control adecuado, que refleje fielmente la distribución en la población de partida de las características estudiadas. Cuando la selección de controles se realiza mediante muestreo directo de un grupo definido según criterios prácticos, las frecuencias de respuesta y las características de los que no responden tienen interés. A veces el grupo control es seleccionado en un proceso de varias etapas que puede estar solo parcialmente controlado por el investigador. Un control enfermo puede ser "seleccionado" de la población de partida por su enfermedad, luego por el investigador de entre todos los individuos similarmente enfermos, luego por sí mismo al dar su conformidad para participar en el estudio. Cada etapa debe ser examinada para detectar posibles dependencias con respecto al estado de exposición o no exposición a un cierto factor de riesgo.

Si los controles son seleccionados sin estratificación previa, tras apareamiento basado en criterios generales (por ejemplo, edad y sexo), o ultraespecíficos (como los usados en un estudio de casos y controles emparejados según vecindario), el principio común que rige y que debe ser evidente al lector es el siguiente: en el esquema de muestreo, los controles dan información no sesgada sobre la población de la que proceden los casos. Una manera útil de describir el grupo control es presentar el proceso de selección de controles como homólogo del proceso de selección de casos: cada uno define una población de partida. Las dos poblaciones de partida deben ser idénticas con respecto a los determinantes de exposición. El procedimiento más simple para garantizar que las características son idénticas es hacer que en ambos procesos la población de partida sea la misma. Si el investigador ni siquiera tiene una idea clara de cuál es la población de procedencia de los casos, lo más fácil es que ni él ni el lector puedan juzgar si los controles son o no adecuados.

## **Recogida de datos**

Al describir el origen de los datos deben darse detalles de quién recoge y proporciona la información, cómo se registran los datos y la ruta por la cual la información inicial llega a adoptar la forma en la que finalmente es analizada. Hay que hacer constar cualquier proceso de control de calidad y los métodos para detectar respuestas obviamente equivocadas o incoherentes. Cuando los métodos utilizados son habituales, hay que ser breve; cuando son novedosos, hay que ser amplio (y prudente). Cuando el detalle de los datos de los que se dispone pone limitaciones previas a las preguntas que pueden hacerse, hay que hacerlo constar.

## Resultados

### *Indicar la sustancia de los datos*

Nunca se ha hecho una observación epidemiológica importante que no haya podido ser presentada claramente en unos pocos cuadros o tablas de datos numéricos en bruto y algunas estadísticas descriptivas sencillas. Los cuadros de casos y poblaciones en riesgo o de recuentos de casos y controles clasificados según la exposición al factor de riesgo sirven a un doble propósito: transmitir el mensaje clave de un conjunto de observaciones y la incertidumbre que puede seguirse de la cortedad de los números. A menudo una simple tabla  $2 \times 2$  (tetracórica) o  $2 \times k$  ilustra claramente un resultado; a veces se requiere estratificar en relación a algún factor de confusión importante; pocas veces es necesario algún procedimiento de mayor complejidad.

Presentar los datos con sencillez no significa que mediante un análisis ramplón se llegue a la selección de unas pocas tablas clave. Los factores que pueden confundir o influir sobre un resultado deben ser examinados mediante estratificación y cálculo adecuado de valores descriptivos, o mediante su inclusión en un modelo estadístico. Los programas de análisis estadístico para computadoras están ya tan difundidos que las técnicas de análisis de múltiples variables, antes reservadas para la última etapa del análisis, ahora se utilizan para hacer una primera criba en conjuntos amplios de factores que pueden interactuar y producir confusión. Cuando esta práctica se realiza responsablemente, la revisión sistemática por el analista de los cambios en las estimaciones paramétricas, la matriz de covarianza y las medidas de bondad del ajuste reemplaza el repaso de las tablas de valores numéricos para "hacerse una idea" de la variabilidad y la interrelación de los datos.

Si el analista ha profundizado en las relaciones entre las variables estudiadas usando montones de tablas o docenas de ecuaciones de regresión, tendrá una comprensión de los datos que no podrá ser transmitida con propiedad bajo las limitaciones habituales de la publicación en revistas. En este caso, debe seleccionar los temas clave que quiere presentar. El lector ha de entender la estrategia empleada para analizar los datos, pero no hay razón alguna para que haya de seguir paso a paso el proceso por el que pasó el investigador. Una práctica muy útil y cada vez más común es presentar las tablas numéricas más sencillas en las que se observa el efecto, junto con las estimaciones basadas en el análisis más completo que se pueda hacer.

### *Certidumbre de las estimaciones*

Los valores  $P$  pueden ser útiles cuando no se dispone de una estimación directa del efecto o dicha estimación es difícilmente interpretable, como ocurre a veces con los términos de grado superior en los modelos estadísticos utilizados para analizar bases de datos muy ricas. Sin embargo, los valores  $P$  no deben presentarse aislados o con una sola estimación puntual, y mucho menos en la forma degradada de una calificación de "significativo" o "no significativo". Los epidemiólogos estudian y estiman la magnitud de las relaciones biológicas, y la dicotomía producida por una afirmación infundamentada de significación suele estar fuera de lugar. Los intervalos de confianza proporcionan estimaciones del espectro de verdaderas relaciones compati-

bles con un conjunto dado de observaciones. De manera que permiten conciliar resultados aparentemente contradictorios y generalmente (ya que suelen ser más amplios de lo deseado) introducen una advertencia apropiada de precaución en la interpretación de resultados "evidentes".

Ni los valores  $P$  ni los intervalos de confianza dan una visión acabada de la incertidumbre inherente al análisis de los resultados epidemiológicos. En este contexto, la distinción entre datos observacionales y experimentales es que el analista reemplaza el acto físico de la aleatorización con una hipótesis de trabajo respecto a la naturaleza de las variables no medidas. Tanto el intervalo de confianza como el valor  $P$  tienen definiciones operacionales sencillas en los ensayos clínicos, en los que el mecanismo estocástico (o de "aleatorización" o asignación al azar) asigna determinantes del resultado mediante un proceso comprensible. En un estudio observacional la hipótesis es que los determinantes no medidos se distribuyen entre los grupos de comparación al azar; las técnicas apropiadas al análisis de fenómenos verdaderamente probabilísticos se aplican para evaluar la posible contribución del "azar" a los resultados de la investigación. La afirmación de que los determinantes no medidos están distribuidos de manera arbitraria, condicionalmente a los factores medidos, no es verificable. Pero debe comentarse si es o no plausible en la sección de discusión del informe, bajo el apartado general de factores de confusión no controlados.

### *Datos perdidos o faltantes*

Ciertos datos resultan incompletos incluso en investigaciones en las que los sujetos participan adecuadamente en el estudio. Los individuos que responden cuestionarios dan a veces respuestas irrelevantes, incluso a preguntas perfectamente planteadas, y los registros y archivos suelen estar incompletos. La frecuencia con la que faltan datos por cualquier razón es una pieza importante de información sobre la calidad de un estudio y debe ser presentada explícitamente. Un supuesto común que permite eliminar sin más del análisis los casos en los que faltan datos (o, a veces, estimar cómo habrían sido los datos faltantes si hubieran podido conseguirse) es que la pérdida de datos es un evento arbitrario, sin relación con los verdaderos valores de las magnitudes investigadas. El supuesto es a veces obviamente falso, como cuando un valor falta porque está fuera del intervalo de características registrables; más a menudo el error es más sutil, como cuando una variable crucial resulta censurada en función de un predictor de riesgo. Por ejemplo, la verificación histológica de un tumor difícil de diagnosticar puede haber sido muy deficiente en los muy viejos, o inusualmente exacta en los ricos. En estos casos será erróneo cualquier análisis de la relación entre tipo de tumor y algún correlato de la edad o la clase social. Incluir la categoría de "desconocidos" en cada tabla y considerarla en los procedimientos de variables múltiples como categoría específica de riesgo o enfermedad puede tener cierta utilidad. Con ello no siempre se previenen distorsiones importantes pero, si las hay, su presencia puede ser puesta de manifiesto por la asociación entre la categoría de "desconocidos" y la enfermedad o los factores de riesgo. Cuando se han perdido o se desconocen muchas respuestas, hay que hacer constar alguna consideración de su impacto, ya sea formalmente en el análisis o informalmente en la discusión de los resultados.

### *Hipótesis múltiples*

Los resultados imprevistos son frecuentes en los estudios en los que se investiga un gran número de factores. En los límites impuestos por la capacidad de los sujetos para proporcionar respuestas válidas, obtener tanta información como

sea posible de entrevistas costosas es un objetivo meritorio, pero presenta diversos problemas, sobre todo cuando el número de casos no es muy grande. Las principales dificultades impuestas por los datos “demasiado ricos” son las comparaciones múltiples, el análisis de subgrupos y la invalidación de la representatividad de los controles. Cada una de estas dificultades exige una atención especial al presentar los datos.

Tal como aparece habitualmente, el problema de las comparaciones múltiples radica en que, cuando se somete a prueba un gran número de hipótesis independientes, se espera obtener resultados estadísticamente significativos en la proporción de veces especificada por el error tipo I de la prueba de hipótesis. Una recomendación frecuente es reducir el tamaño de la región de rechazo, como compensación. El remedio propuesto pone de manifiesto un aspecto lamentable de que dependamos de los valores  $P$ , ya que nos incapacita para la detección de cualquier efecto, en la medida en que el nivel de significación requerido se acerca a cero según crece el número de hipótesis. Más grave es el hecho de que, a menudo, la sugerencia no es aplicable de manera coherente, ya que el número de hipótesis independientes que podrían ser sometidas a prueba en un conjunto de observaciones profusamente interrelacionadas puede no depender exclusivamente de los datos disponibles; las hipótesis que podrían ser razonablemente sometidas a prueba difieren en función de información externa al estudio. ¿Debe desecharse un resultado interesante porque el investigador sometió a prueba una hipótesis que a nuestro juicio es absurda? Lo mejor es presentar los resultados inesperados y sus intervalos de confianza no ajustados con un comentario apropiado que identifique dichas hipótesis como no consideradas al comienzo del estudio y, si es posible, someter a prueba con los datos disponibles implicaciones ulteriores de las nuevas hipótesis. La interpretación de resultados no esperados depende en gran medida de criterios externos de verosimilitud biológica y coherencia con otros resultados.

El análisis de subgrupos es una variante del problema de las comparaciones múltiples, en el que una sola hipótesis se multiplica por investigaciones separadas en muchas subpoblaciones. Empeñarse en conseguir significación en pruebas de heterogeneidad del efecto sobre los subgrupos como prerequisite del análisis de subgrupos protege al investigador de la distracción debida a variaciones espurias de poca monta, pero al precio de una incapacidad casi total para reconocer la variabilidad que tiene sus raíces en la población sometida a estudio. Excepto cuando hay pocos estratos y estos son muy numerosos, las pruebas de heterogeneidad tienen escasa potencia en comparación a otras muchas opciones interesantes. Por ello, observaciones relevantes ajenas al estudio pueden ser cruciales para decidir si tomar en serio o no un efecto observado en un subgrupo. Por lo tanto, hay que presentarlas junto con los resultados.

A menudo se adaptan series de controles no seleccionados por un procedimiento aleatorio a partir de una población bien definida, para utilizarlas en estudios específicos. Por ejemplo, es posible escoger como controles personas hospitalizadas, seleccionadas entre pacientes con diagnósticos considerados no relacionados con el consumo de alcohol o tabaco, en un estudio para valorar los efectos de tales exposiciones. Un grupo control de ese tipo permite estimaciones válidas del uso de alcohol y tabaco, pero genera a la vez estimaciones muy sesgadas de la prevalencia de otros há-

bitos relacionados con los diagnósticos usados para especificar los controles. Una forma de reducir el riesgo de error en esa situación es definir los diagnósticos de los controles por inclusión (no por exclusión) e indicar las frecuencias de exposición en las categorías de los controles. No obstante, en general es prudente limitar los análisis exploratorios mediante estudios de casos y controles a estudios en los que el proceso de selección de controles a partir de la población general tiene unos pocos pasos bien definidos y cuantificables, como ocurre cuando se usan los típicos métodos de muestreo de las encuestas con el fin de seleccionar controles para grupos de casos extraídos de la población general.

## **Implicaciones**

El ímpetu de los estudios epidemiológicos puede proceder de diversas disciplinas y las ramificaciones de una observación pueden extenderse igualmente a muchos campos, pero es raro que la investigación epidemiológica por sí misma sea suficientemente detallada para justificar una discusión extensa de los mecanismos de acción propuestos. La discusión debe situar claramente los resultados en el contexto de otros trabajos epidemiológicos relevantes, mostrando paralelismos donde sea posible y poniendo de manifiesto puntos de conflicto aparente con estudios previos. Lo más frecuente es que el conocimiento científico nuevo surja de resultados que no concuerdan con hipótesis previas y no a partir de resultados confirmatorios, de manera que los desacuerdos deben ser explorados con tanta atención como merezcan los datos.

## **Agradecimiento**

Muchas de las ideas aquí presentadas son fruto de comentarios útiles y extensos de los revisores y de colegas de la Escuela de Salud Pública de Harvard, el Organismo Internacional para la Investigación sobre el Cáncer y Epidemiological Resources, Inc. Durante la preparación de este trabajo tuve el respaldo del Organismo Internacional para la Investigación sobre el Cáncer, como miembro de plantilla, y luego conté con una beca de la Fundación Mellon.