

Cirugía y epidemiología

Dr. José Miguel Avilán Rovira

Individuo de Número

La epidemiología, después de haber superado, desde hace ya largo tiempo, la clásica definición de “estudio de las epidemias”, se concibe en la actualidad como el análisis de todas las variables relacionadas con la ocurrencia de la salud y la enfermedad en las poblaciones. De acuerdo a su propósito y dominio, se considera como la ciencia que trata de explicar las causas del proceso salud-enfermedad y de generar conocimientos para mejorar los patrones poblacionales de salud y bienestar.

Ya Greenwood (1) precisó hace algún tiempo, que se trata del “estudio de la enfermedad, cualquier enfermedad, como un fenómeno de masa”

Como se trata de cualquier enfermedad, pues incluye no sólo aquellas tratadas médicamente, abarca, por supuesto, aquellas con tratamiento preferiblemente quirúrgico. Y comprende igualmente el estudio de los resultados y complicaciones posoperatorias, a corto, mediano y largo plazo.

Gran parte del esfuerzo de los investigadores médicos para comprender la etiología del proceso salud-enfermedad y desarrollar tratamientos adecuados, se dedica al estudio de las relaciones entre eventos o características, las llamadas “variables de predicción” con sus efectos o “variables-respuesta”.

Existen dos formas básicas de estudiar estas relaciones: la observación y la experimentación. En epidemiología se recurre a estudios observacionales o experimentales, de acuerdo al tipo de problema a resolver. La diferencia está en que en el primer

caso, el investigador se limita a observar, y establece las mejores condiciones posibles de observación; en el segundo, el investigador interviene y asigna la exposición al objeto de estudio. Kleinbaum distingue la subcategoría cuasi-experimental, cuya característica es no cumplir con la asignación aleatoria de la exposición (2).

Tradicionalmente, los estudios observacionales son de dos tipos: descriptivos y analíticos. En los primeros se determina la incidencia, prevalencia y mortalidad de la enfermedad en grupos de población, de acuerdo a características básicas, como persona, lugar y tiempo. En los segundos, se plantean interrogantes específicas que se intentan responder mediante estudios de riguroso diseño y análisis de los datos recogidos en la investigación. En la práctica, muchos estudios son mixtos: descriptivos y analíticos.

La causalidad múltiple del proceso salud-enfermedad es la concepción dominante en la epidemiología actual, expresada desde la sexta década de este siglo por la conocida metáfora de la “red causal” (3). A pesar de los logros hasta ahora alcanzados con estos supuestos, algunos estudiosos consideran que tiene serias limitaciones derivadas de un enfoque biomédico predominante, con demasiado énfasis en el estudio individual de los factores de riesgo.

Entre ellos, Mervyn Susser, en un provocativo ensayo que estimula profundamente a la reflexión, distingue tres niveles de estudios epidemiológicos: el molecular, el individual y el social, que propone englobar en una nueva concepción que denomina “socioepidemiología” (4).

El debate está abierto y en la actualidad existen más interrogantes que respuestas en relación a las

estrategias metodológicas que puedan aplicarse para estudiar el proceso salud-enfermedad a los niveles micro, individual y macro, de modo secuencial o en su conjunto

La discusión de estas concepciones está más allá del objetivo de este informe, donde solamente pretendemos ilustrar con ejemplos, el uso de los diseños epidemiológicos más comunes en las enfermedades quirúrgicas y sus complicaciones.

Estos ejemplos los hemos reunido intencionalmente, es decir, de ninguna manera se intentó realizar su selección en forma aleatoria, pues la finalidad es describir la aplicación de la epidemiología por parte de los cirujanos en la investigación de las variables que intervienen en la ocurrencia y evolución de las afecciones que tratan y no determinar la frecuencia con la cual se aplica determinado diseño.

La fuente de los ejemplos son estudios publicados en revistas, cuyos autores y demás datos incluimos en las referencias, con respeto de las normas vigentes.

Como podrá observarse, los ejemplos corresponden a investigaciones realizadas desde la tercera década de este siglo, por lo cual al evaluarlos, debemos considerar los problemas que se trataron de solucionar y los resultados obtenidos, de acuerdo al nivel de los conocimientos para la época. De ninguna manera se pretendió actualizar los problemas de investigación con los cuales ilustramos los distintos tipos de diseños epidemiológicos utilizados, lo cual no sólo está fuera del alcance de este informe, sino que lo volvería interminable.

Tal como era de suponer, la mayoría de los ejemplos localizados corresponden al nivel intermedio de la investigación epidemiológica, es decir, en general su propósito fue la identificación de factores de riesgo a nivel individual.

Así, por ejemplo, no encontramos en las revistas de cirugía consultadas, estudios clasificables al nivel de microepidemiología, como los que se están realizando para identificar biomarcadores producto de la influencia de contaminantes atmosféricos en zonas altamente industrializadas —como podrían ser los hidrocarburos policíclicos aromáticos, para citar al menos un ejemplo— donde las tasas de cáncer del pulmón son más elevadas que en zonas menos contaminadas (5).

En el nivel macro, sin embargo, existen ejemplos que muestran el análisis de la influencia de variables tales como el nivel social en la “incidencia”

(para utilizar el mismo término utilizado por los autores) de las amigdalectomías o las cesáreas, o la variabilidad de las tasas de ejecución de algunas intervenciones quirúrgicas, que algunos investigadores tratan de explicar por diferencias en la organización y el pago de los servicios médicos, entre otras determinantes, del tipo de sistema de salud.

Estudios descriptivos. Debido a la especificidad de muchos procedimientos quirúrgicos, dentro de este tipo de investigaciones, además de las que describen las variables que explican la incidencia, prevalencia y mortalidad de las enfermedades propias de la especialidad y de sus complicaciones, una de las características de la descripción epidemiológica en cirugía, es la identificación de las variables que determinan la mayor o menor frecuencia de las intervenciones más conocidas, su evolución en el tiempo o sus diferencias entre áreas geográficas o niveles sociales.

En otras palabras, se sustituye en estos estudios la descripción de la enfermedad por la de su procedimiento quirúrgico.

Es importante señalar que, entre nosotros, el interés por tratar de explicar la frecuencia con la cual se ejecutan ciertos procedimientos quirúrgicos, no es reciente. En efecto, Salvador Córdoba, en un trabajo publicado en 1928, sobre apendicitis, refiere su inquietud por su aparente alta frecuencia en la época (6). Textualmente dice: “es tal la frecuencia observada y comprobada operatoriamente de la apendicitis aguda en Caracas en estos últimos años, que la primera idea que viene a la mente es la de investigar si es ciertamente real o aparente el aumento notado y tratar de conocer la causa de esta ocurrencia...”.

Con el objetivo de responder a esta pregunta, dirigió una carta-circular, en forma de cuestionario, a los cirujanos residentes en el país para conocer su opinión, con la poca fortuna de ser atendido solamente por cinco de sus colegas, entre ellos, José Izquierdo, Fermín Díaz y Hérmán de las Casas.

En consecuencia, Córdoba sólo pudo limitarse a una conjetura: “la apendicitis ha existido siempre en Venezuela con la misma frecuencia y las mismas modalidades clínicas encontradas actualmente... (pero) los médicos de hoy hacen con más seguridad y facilidad el diagnóstico” (6).

Otro ejemplo, que ilustra el interés por describir

las características del conjunto de intervenciones por una determinada enfermedad, es el de Blanco Gásperi (7), quien analizó “sus primeras cien operaciones del tiroides”, trabajo con el cual se incorporó a la Academia. Además de la distribución por género, edad, tipo de enfermedad tiroidea, procedencia y mortalidad, el autor trata de evaluar la calidad de la información presentada, advirtiendo al comienzo de su exposición, en la forma diáfana que lo caracterizaba, textualmente: “Es muy difícil en nuestro medio, cuando no se ha comenzado a hacer un trabajo desde un principio con un objeto determinado, como el caso presente mío en que comencé a hacer operaciones sobre el tiroides hace catorce años y luego resuelvo ahora hacer un trabajo de conjunto y me encuentro que mi estadística y la del Hospital Vargas, que es donde están hechas la mayor parte de mis operaciones, no coinciden totalmente...”.

Dentro de los ejemplos de este tipo de estudio, tenemos el trabajo ya clásico de J. Alison Glover, sobre “incidencia” de la amigdalectomía en Inglaterra y Gales, en las primeras tres décadas de este siglo (8). De acuerdo a las cifras analizadas, mientras en 1923 la tasa anual de amigdalectomías fue de 0,9%, alcanzó un poco más del doble en 1931. Con base a estos datos, se estimó que para esa época, cerca del 20% de los escolares, antes de cumplir sus 14 años, habían sido sometidos a la intervención. De acuerdo al autor, la incidencia era similar en Estados Unidos. Se intervenían más varones que hembras, lo cual parecía extraño, pues se disponía de información sobre la mayor incidencia de amigdalitis y de reumatismo agudo en las hembras que en los varones. Las edades pico para la intervención eran las de 5 a 7 años, lo cual coincide con la hipertrofia temporal de las amígdalas, que podría tener dos explicaciones: fisiológica, por el desarrollo de la cavidad bucal propia de este período; e inmunológica, por respuesta al cambio ambiental del hogar a la escuela y a la sepsis resultante de las caries de la primera dentición. No pudo determinarse ninguna relación entre las tasas de intervención por localidades y los supuestos factores que podrían producir sus diferencias, por lo cual la única explicación posible sería la variación de la opinión médica para indicar la operación. No se encontró tampoco una explicación razonable para que la tasa de intervención fuese hasta tres veces más alta en las clases acomodadas. De acuerdo al autor, cuanto más afortunado es un niño, mayor su oportunidad para la amigdalectomía. A pesar de que en algunas

áreas ocurrió un descenso brusco de la incidencia de las tasas de intervención, no se observaron consecuencias insatisfactorias: no se registró un aumento de las otitis medias o de los defectos auditivos. En su discusión, el autor concluye que resulta innegable que la amigdalectomía tiene sus indicaciones en ciertos casos, con resultados beneficiosos, por lo cual es posible que se intente realizar en casos donde su aplicación sea dudosa y no la ameriten, por hipertrofia temporal, fisiológica o inmunológica.

Cita para finalizar la observación del Comité de Epidemias Escolares, del Consejo de Investigación Médica, según la cual “no puede negarse la existencia de una tendencia para practicar la operación como un ritual profiláctico rutinario sin razones ni resultados particulares” (8).

Después del descenso de la amigdalectomía registrado a partir de 1932, en Inglaterra y Gales, Glover observó un repunte en 1937.

En el estudio de Bakwin (9) se presentan datos de Estados Unidos e Inglaterra, que indican que la amigdalectomía y adenoidectomía, eran realizadas tan frecuentemente en la década de los años cincuenta, como en el pasado. Para el autor, la continuación de tal práctica sigue siendo un enigma, en especial en una época en la que se supone que se utilicen tratamientos y procedimientos quirúrgicos solamente después de ser verificados por rigurosos ensayos clínicos y quirúrgicos.

En algunos de estos estudios descriptivos, se intenta relacionar las tasas diferenciales de distintos tipos de intervenciones quirúrgicas con algunas variables de predicción.

Así, por ejemplo, en un análisis de los registros de la Asociación Cruz Azul de Kansas, se demuestran variaciones de las tasas regionales de procedimientos quirúrgicos. Estas variaciones no solamente son características de las amigdalectomías y adenoidectomías, sino en procedimientos con indicaciones más precisas, tales como apendicectomías y herniorrafias (10). Se consideran los factores que podrían producir tales variaciones entre regiones y se demuestra que podrían explicarse por el número de camas y cirujanos disponibles.

El autor opina que una variación de 3 a 4 veces entre los procedimientos quirúrgicos comunes, parece difícil de explicar por diferencias de la incidencia o prevalencia de las enfermedades en la población. Parece más reconfortante asumir que las

camas hospitalarias se instalan y los cirujanos se agrupan en áreas que tienen el mayor número de personas predispuestas a la apendicitis o a las enfermedades de la vesícula biliar, entre otras enfermedades quirúrgicas. Agrega que es difícil descartar la impresión que la llamada "cirugía ritualista", no se extiende a otras operaciones además de la circuncisión y la amigdalectomía.

Otros estudios sobre el mismo tema presentan resultados similares. En Rochester, New York, en 23 áreas hospitalarias, con una tasa promedio de 38,3 apendicectomías por 10 000 habitantes, en algunas áreas asciende a 83,2 por 10 000. En una de las áreas con una tasa de 63 por 10 000, la mitad de las operaciones fueron realizadas por 2 de los 16 cirujanos, con "privilegios quirúrgicos", ninguno de los cuales atiende referencias. No pudo demostrarse asociación alguna entre bajas tasas por apendicectomías y altas tasas de mortalidad por apendicitis (11).

En este estudio los autores concluyen diciendo que los resultados podrían interpretarse como el apoyo de la Ley de Parkinson: la admisión de pacientes a cirugía se expande hasta llenar las camas, los pabellones quirúrgicos y el tiempo de los cirujanos.

Otro ejemplo, es el estudio del Consejo Hospitalario de Michigan, en 1953, sobre 1 647 pacientes que egresaron con el diagnóstico de apendicitis primaria, sobre un total de 65 000 pacientes atendidos (12). De acuerdo a la distribución por edad y género, la enfermedad es más común en la adolescencia, con un pico marcado en las hembras, a los 16 años. Este aumento está precedido de una meseta entre los 9 y 13 años que no se registra en los varones. La distribución por género fue similar: 827 hembras y 820 varones. Cuando se analiza la distribución de los casos que se confirmaron, por el resultado del examen del espécimen quirúrgico (hembras 282, varones 456), se nota que el pico marcado de las adolescentes desaparece y se aprecia un predominio real de apendicitis confirmada en los varones. En la distribución de las apendicitis no confirmadas (hembras 545, varones 364), reaparece el pico entre las hembras. Como dicen los autores, parecería de acuerdo a estos datos, "que en la adolescencia los varones padecen la enfermedad, pero las operadas son las hembras". También en este estudio, cuando se distribuyen las apendicitis confirmadas como agudas, por el examen del espécimen quirúrgico, de acuerdo a los hospitales,

el porcentaje varía entre 68% y 18%. De acuerdo a los cirujanos que realizaron la intervención, el porcentaje de confirmación varió entre 0 y 87%. Los autores aclaran que no se pretende que el cirujano siempre logre el diagnóstico correcto, pero los datos parecen demostrar que las variaciones registradas son de tal magnitud que revelan verdaderas diferencias en la práctica médica.

En otros estudios se intenta encontrar apoyo de algunas hipótesis, tales como la relación entre el volumen quirúrgico hospitalario y la mortalidad (13). Basados en los datos de registros hospitalarios obtenidos por la Comisión de Actividades Profesionales y Hospitalarias, de Ann Arbor, Michigan, entre 1974-75, para 12 procedimientos quirúrgicos de complejidad variada, en 1498 hospitales, para encontrar apoyo a la hipótesis antes mencionada, se obtuvo los siguientes resultados:

- a. Para cirugía de corazón abierto, cirugía vascular, resección transuretral de la próstata y derivación coronaria, la mortalidad decrece con el número mayor de operaciones. Los hospitales con 200 o más de estas operaciones anuales, registraron tasas de mortalidad ajustadas, 25% a 40% más bajas que los hospitales con menores volúmenes;
- b. Para otros procedimientos, como colectomías, cirugía del tracto biliar, reemplazo pélvico, resección e injerto del aneurisma aórtico abdominal, vagotomía u otra piloroplastia para úlcera, la mortalidad se estabiliza a volúmenes bajos. Por ejemplo, hospitales que practican de 50 a 100 reemplazos de cadera, alcanza una mortalidad para este procedimiento casi tan baja, como los hospitales que practican 200 o más;
- c. Algunos procedimientos, como la colecistotomía, no muestra ninguna relación entre volumen y mortalidad. Para los autores, los resultados podrían reflejar el efecto del volumen o la experiencia, sin descartar las referencias de pacientes seleccionados a instituciones con mejores resultados obtenidos en el pasado. Sin importar la explicación, estos datos parecen apoyar el valor de la regionalización para ciertas operaciones.

Dos ejemplos sobre la frecuencia de cesáreas en dos ciudades del Brasil, intentan mostrar la relación con factores culturales y sociales (14,15). De acuerdo a los datos revisados por los autores, el 15% de cesáreas registrado en 1970, se duplicó para 1980,

aun cuando en ese período el Seguro Social implementó la política de igual remuneración al médico para partos quirúrgicos y no quirúrgicos. Se observaron sin embargo, las siguientes diferencias: la incidencia promedio fue muy baja entre los pacientes que no pagan (13%) y la más alta entre los pacientes privados (50%). Los pacientes asegurados ocuparon una posición intermedia. Las variaciones las atribuyen los autores a diferentes patrones de atención por parte de los médicos a las pacientes de acuerdo a su clase social. Reflejan las características de la medicina clasista prevalente en el Brasil de la época. Las expectativas del médico y su paciente, en término de los procedimientos a utilizar, varía de acuerdo a la clase social de ambos.

La misma variabilidad y tendencia al aumento, se observa en las estadísticas de Agüero (comunicación personal), para años más recientes: en la Maternidad "Concepción Palacios" el porcentaje de cesáreas aumentó de 15,6 en 1991 a 18,4 en 1994. En la Clínica "Santa Ana", del Instituto Venezolano de los Seguros Sociales (IVSS), el porcentaje aumentó para los mismos años de 27,7 a 29,4. Finalmente, en el Centro Médico de Caracas el porcentaje osciló para los mismos años, entre 35,1 y 43,2.

La tendencia al aumento del porcentaje de cesáreas, por las razones que sean, se incrementará aún más en el futuro, de verificarse los resultados publicados recientemente sobre sus efectos protectores en los hijos de madres HIV positivas (16).

Aun sin medicación específica, las madres infectadas con HIV, podrían reducir a la mitad el riesgo de transmitir el virus a sus descendientes y hasta el 87%, si son tratadas. La vía del parto disminuye en forma importante el tiempo de exposición del feto a la sangre materna.

No podemos omitir en esta ya larga enumeración de descripciones epidemiológicas, el estudio de Doyle (17) sobre histerectomías injustificadas. Con el fin de conocer los criterios que indicaron la intervención y la confirmación posterior, el autor analizó los datos de 7 252 pacientes operadas en 35 hospitales de Estados Unidos en 1948. En el 60,7% el diagnóstico clínico se confirmó en el posoperatorio o podría considerarse correctamente indicada. Por consiguiente, en el grupo restante la indicación de la intervención podría criticarse: en un 2,3% el diagnóstico definitivo la contraindicaba; en un 3% podría justificarse, pero no en la extensión con que se practicó; en un 12,5% no se determinaron lesiones

ni micro ni macroscópicas y en el 21,5% restante lo indicado debió ser más observación o un tratamiento más conservador que el practicado. Afortunadamente, la mortalidad operatoria fue baja (0,04%). La justificación de la histerectomía varió entre el 95% y el 16%, entre los hospitales estudiados.

En algunos de estos estudios se hace énfasis en la comparación de tasas quirúrgicas entre países (18) o de la letalidad entre instituciones hospitalarias docentes y no docentes (19).

Con los egresos hospitalarios de Inglaterra y Gales en 1967 y del Canadá en 1968, donde figuran los procedimientos quirúrgicos y los diagnósticos de los pacientes dados de alta, se calcularon tasas ajustadas por edad y género, con el fin de compararlas por tipo de intervención y la mortalidad asociada. En general se encontró que las tasas del Canadá resultaron 1,8 veces más altas para los varones y 1,6 veces más altas para las hembras, que las de Inglaterra y Gales. De las variadas explicaciones que ofrece el autor, considera que las más importantes serían la prevalencia de las enfermedades, estilos de tratamiento, personal quirúrgico e incentivos financieros.

Si la prevalencia de las enfermedades fuese la misma y la cirugía practicada en Inglaterra y Gales fuese insuficiente, lo esperado sería que las afecciones potencialmente fatales que se tratan quirúrgicamente, presentasen tasas de mortalidad más elevadas en este país que en el Canadá. Los datos no confirmaron esta hipótesis. Las tasas específicas de mortalidad por edad para cáncer cervical, uterino y de la mama fueron similares, a pesar de las diferencias en tasas quirúrgicas. Para las enfermedades de la vesícula biliar, la tasa de mortalidad en los varones y hembras de mayor edad fue dos veces más alta en Canadá, a pesar de que la tasa de colecistectomías fue cinco veces más elevada. Como dicen los autores, parte del exceso de la mortalidad puede atribuirse a la mayor práctica quirúrgica.

Otras diferencias en las tasas quirúrgicas podrían explicarse por la mayor o menor utilización de técnicas conservadoras o radicales. Si las tasas de cáncer de la mama son comparables en ambos países, las diferencias en tasas de mastectomía radical probablemente reflejan distintas preferencias en modos de tratamiento, que mayor frecuencia de la enfermedad. Es interesante señalar, que las tasas de apendicectomía en los dos países fueron casi idénticas, lo que sugiere que la enfermedad tiene la

misma incidencia, los mismos criterios diagnósticos y terapéuticos. La disponibilidad de camas y de personal quirúrgico, en ambos países, es otro factor a considerar. Para la época, en Inglaterra y Gales se disponía de un 30% menos de camas para casos agudos, lo cual limitaba el número de admisiones y en consecuencia las tasas quirúrgicas. Mientras en este país casi toda la cirugía es realizada por especialistas, en el Canadá pueden practicarla los médicos generales. El análisis de los datos disponibles de la forma de remuneración en programas de atención médica prepagada, en ambos países, parece influir en las diferencias de algunas de las tasas. Mientras en Inglaterra y Gales, los cirujanos reciben un salario fijo, en Canadá es posible recibir pagos por servicios realizados en algunas instituciones. Una comparación de las tasas quirúrgicas entre dos grupos de una misma localidad y de características demográficas similares, mostró menores tasas en el grupo con el programa médico prepagado que en el atendido por médicos pagados por servicios realizados. La tasa de amigdalectomía, por ejemplo, fue un poco menos de la mitad en el primer grupo.

Los autores reconocen que las causas para estas diferencias deben continuarse estudiando para poder responder muchas interrogantes. Dado que para muchos procedimientos quirúrgicos, en especial los electivos, la tasa canadiense es más elevada que la inglesa, cabría preguntarse: ¿es que la primera es muy alta o la segunda muy baja? Como las afecciones para las cuales pueden aplicarse procedimientos electivos, por lo regular no amenazan la vida, proponen que la pregunta se responda con ensayos quirúrgicos controlados, con pacientes asignados aleatoriamente a los tratamientos a comparar. Recomiendan que estas investigaciones deben ser realizadas por cirujanos pues son ellos los que mejor comprenden las dificultades de los problemas quirúrgicos.

Entre 1956 y 1959, con los registros hospitalarios analizados en el Servicio Nacional de Salud del Reino Unido, se constató la diferencia en tasas de letalidad por diferentes procedimientos quirúrgicos, entre hospitales docentes y no docentes, observada a comienzos de la década. Así por ejemplo, la tasa de letalidad, entre los primeros y los segundos, fueron: 6,1 y 9,7 para hernia abdominal con obstrucción; 2,8 y 4,3 para apendicitis con peritonitis; 8,1 y 10 para úlcera péptica perforada; 9,4 y 13 para hiperplasia de la próstata; 2,7 y 3,4 para fracturas

del cráneo, respectivamente.

Las diferencias podrían explicarse porque los hospitales docentes ofrecen un tratamiento superior, pues por lo regular tienen mayor cantidad de médicos consultantes y otro personal. No parece haberse demostrado desventajas sociales entre los pacientes atendidos en los dos tipos de hospitales. Los autores recomendaron más estudio de este problema, pues permaneció durante el tiempo de observación.

Estudios analíticos. El cirujano, como todo médico, está interesado en conocer las relaciones etiológicas de las enfermedades, lesiones o complicaciones quirúrgicas, con la finalidad de prevenirlas. Este tipo de investigación no puede realizarse experimentalmente por razones éticas, pues sería insensato exponer o no a seres humanos, aleatoriamente, a los supuestos agentes etiológicos, para determinar la ocurrencia o no de sus consecuencias lesivas.

Por consiguiente, en epidemiología se recurre a los llamados estudios de observación de personas expuestas o no a los agentes etiológicos de forma espontánea o natural, no deliberada o intencionada por parte del investigador. La exposición ocurre en estos individuos debido a sus hábitos, sus patrones alimentarios, sometimiento a procedimientos terapéuticos médicos o quirúrgicos, la influencia de contaminantes atmosféricos o tóxicos ocupacionales, entre otros muchos que sería largo enumerar, la cual no se produce en el vacío, sino que en gran parte está condicionada, por la posición del grupo al cual pertenecen dentro del contexto social.

Dentro de estos estudios analíticos se distinguen fundamentalmente dos: los de cohortes y los de casos y controles, ejemplos de los cuales describiremos a continuación.

Estudios de cohortes. De estos estudios de observación el que más se asemeja a un experimento es el estudio de cohortes. Una vez que el supuesto factor etiológico o de riesgo, se identifica o no, en grupos de individuos, estos son seguidos en el tiempo con el fin de determinar si la enfermedad, lesión o complicación aparece en consecuencia.

La tasa calculada de ocurrencia nos indica el riesgo de cada grupo. La importancia relativa de las dos tasas de riesgo, se obtiene mediante su diferencia

(riesgo atribuible) o su cociente (razón de riesgo o riesgo relativo).

Este diseño epidemiológico es el más comúnmente utilizado en el seguimiento de pacientes intervenidos por tumores malignos de cualquier localización, observando su recaída, complicaciones y tiempo de sobrevida, de acuerdo a los factores de riesgo del enfermo, tipo y estadio del tumor y modalidades del tratamiento coadyuvante, entre otras características relacionadas de interés.

Así, por ejemplo, en el trabajo de Ravelo Celis y Ravelo Pagés, en el cual analizan la sobrevida de 102 casos de cáncer de mama, en 100 pacientes, dos de ellos bilaterales, se estudian las características de mayor interés de la historia clínica y familiar, en especial, los factores pronósticos que se consideran más importantes: tamaño del tumor, positividad axilar, histología de riesgo, estado menstrual, receptores hormonales y tratamientos adyuvantes (20).

De esta aplicación específica del seguimiento del cáncer intervenido, pueden encontrarse múltiples ejemplos. Sin embargo, citaremos uno más, en el que se compara la evolución del cáncer de la mama en el varón con el de la hembra (21). Los investigadores justifican su estudio pues los informes revisados de observaciones anteriores han sido conflictivos. En este estudio de cohortes, realizado retrospectivamente, pues se disponía de todos los datos de los pacientes durante 20 años, se parearon 41 varones con 123 hembras con cáncer de la mama, por edad, sitio y grado del tumor primario y estado de los nódulos linfáticos axilares. No se pudo demostrar diferencias en 72 meses de seguimiento, en el intervalo libre de enfermedad y la sobrevivencia entre los pacientes estudiados. De acuerdo a los autores, para el manejo del cáncer de la mama en varones son válidas las mismas recomendaciones conocidas para el cáncer de mama en las hembras.

Otro ejemplo de este modelo de estudio es la relación entre recurrencia y la respuesta de fase aguda, posterior a la cirugía, en pacientes aparentemente curados de cáncer colorrectal (22).

Aproximadamente, 70% de los pacientes sometidos a resección por cáncer del colon o el recto, que recurren, lo hacen dentro de los dos primeros años de la cirugía. Es un hecho conocido que la progresión o recurrencia de estos tumores, se asocian a una respuesta de proteínas de fase aguda, de acuerdo a resultados de estudios retrospectivos.

En este estudio se siguieron 36 pacientes, hasta 24 meses después de haber sido sometidos a cirugía curativa colorrectal. Al inicio del seguimiento, ningún paciente presentó evidencia clínica de recurrencia tumoral, infección u otra condición inflamatoria. En todos los pacientes se realizaron mediciones de antígeno carcinoembrionario, proteína reactiva-C (PRC), albúmina, número de plaquetas, cuenta y diferencial de leucocitos. Los exámenes periódicos se verificaron cada 3 meses hasta completar los dos años de observación, cuando se examinaron radiológicamente para diagnosticar recurrencia.

Según los resultados, de los 21 pacientes con valores de PRC inferiores a 5 mg/L sólo 2 recurrieron localmente, mientras que en los 15 pacientes restantes con valores de PRC iguales o superiores a 5 mg/l, recurrieron 11, siete de ellos a distancia y cuatro, localmente. La concentración de antígeno carcinoembrionario, se mantuvo dentro de límites normales (60 U/L). Evidencia adicional de que un PRC de 5 mg/L o más indicaba una respuesta de fase aguda, fue el aumento del número de leucocitos y la alta proporción de neutrófilos.

Los autores concluyen que hubo una tasa de mayor recurrencia de cáncer colorrectal en los pacientes con una respuesta de fase aguda que en los pacientes sin esa respuesta con significancia estadística.

A este tipo de estudio epidemiológico pertenece el intento por apoyar con datos, el antiguo y conocido supuesto del efecto del retardo en el diagnóstico de la apendicitis aguda en la aparición de complicaciones y mayor morbilidad (23). Como es sabido, el retardo antes de la admisión al hospital, se considera responsabilidad del paciente y el posterior a la admisión, al cirujano. Se estudiaron 486 pacientes admitidos entre 1980 y 1992. El retardo del paciente en acudir al médico y el retardo del cirujano entre la admisión al hospital y la intervención, se relacionaron con el estadio de la apendicitis y las complicaciones posoperatorias. El retardo promedio para las apendicitis clasificadas como agudas simples fue de 1,7 días, en comparación con 2,3 días en las gangrenosas y perforadas. La diferencia entre los promedios (14,4 horas), es estadísticamente significativa. El promedio del retardo por parte del cirujano fue de 13,6 horas en la apendicitis aguda simple, versus 14,5 horas en la apendicitis avanzada. Esta diferencia de 54 minutos

no tiene significancia estadística. De acuerdo a los autores, la decisión de observar los pacientes para clarificar el diagnóstico está justificada y no influye negativamente en el resultado. En cuanto a las complicaciones posoperatorias, sólo un 10% de las apendicitis agudas simples se complicaron, contra el 20% de las avanzadas. Esta diferencia resulta, además, estadísticamente significativa. Estos datos parecen comprobar que el retardo del paciente en ser visto por el médico, afecta adversamente el estadio de la enfermedad en apendicitis aguda y se asocia a un incremento de las complicaciones infecciosas y la estancia hospitalaria: 8 días promedio en la aguda simple, contra 11,5 días en la avanzada. En este estudio es crucial el interrogatorio de los pacientes para la determinación de la duración de los síntomas antes de la admisión al hospital. De la misma manera lo es, el registro del tiempo entre la admisión, la hora del examen por parte del médico y la hora de la intervención. En su discusión, los autores no mencionan ninguna duda en relación a la recolección de estos datos. Al no disponer de datos adicionales y dados los resultados presentados, podríamos concluir que es probable que la información suministrada por los pacientes y los cirujanos, esté ajustada a la verdad.

Otro ejemplo del mismo tipo de estudio, en otra aplicación de la cirugía, es el de los determinantes de mortalidad en niños con quemaduras masivas (24). Consiste en el seguimiento de 103 menores de 17 años, con quemaduras de no menos del 80% de la superficie corporal y 70% o más con pérdida completa de las capas de la piel (tercer grado), tratados en una institución especializada con excisión temprana masiva e injertos. Mediante análisis univariado, corroborado por regresión logística, se identificaron, además de los conocidos factores pronósticos —menor edad, porcentaje de superficie corporal, grado de la quemadura, sepsis— el tiempo para iniciar la hidratación (no su cantidad) y la dependencia de un ventilador para controlar la falla respiratoria. Mientras los sobrevivientes recibieron fluido intravenoso en más o menos media hora de ser admitidos al centro, los que fallecieron la recibieron a las dos horas, en el promedio. A su vez, los requerimientos de ventilación durante los primeros 28 días del tratamiento, fueron cuatro veces menos en los sobrevivientes, que en quienes fallecieron. Se hace referencia en este estudio al seguimiento y evaluación de las secuelas psicosociales de menores quemados con 80% o más de su superficie corporal,

al primer año, 5 y 10 años después de la ocurrencia de la lesión. Sus conclusiones, basadas en evaluaciones estandarizadas de su adaptación psicológica y su habilidad funcional, son prometedoras, pues evidencian, hasta la fecha de la comunicación, que estos pacientes pueden alcanzar una calidad de vida que ellos consideran satisfactoria.

Queremos resaltar, que así como en este estudio se menciona la evaluación de la calidad de vida de los intervenidos, es muy frecuente encontrar que este importante aspecto de los pacientes preocupa cada vez más a los cirujanos, sin limitarse a comparar únicamente complicaciones posoperatorias, mayor o menor estancia hospitalaria, recurrencia o mortalidad. Así, por ejemplo, en el Foro quirúrgico del Colegio Americano de Cirujanos, celebrado en Orlando, en octubre de 1998, en el 7% de los 350 trabajos presentados, se analizaron resultados relacionados con la calidad de vida de los pacientes intervenidos (25). Por ejemplo, en un estudio comparativo de la evolución entre pacientes con tumores benignos o malignos, sometidos al procedimiento de Whipple y un grupo control sometido a colecistectomía, pareado por edad, género, fecha de la intervención (dentro de 1 año) y la presencia de enfermedades serias, en el Hospital General de Toronto (26), el objetivo principal fue comparar la calidad de vida entre los pacientes sometidos a los dos procedimientos. Como la colecistectomía es realizada como parte del procedimiento de Whipple, es posible evaluar el efecto de la pancreático-duodenectomía. Esta experiencia estuvo justificada por la percepción de que este procedimiento está asociado a complicaciones, sintomatología gastrointestinal y deficiencia exo y endocrina pancreática, lo cual afecta la calidad de vida de estos pacientes.

Todos los pacientes respondieron seis cuestionarios estandarizados para evaluar su estimación de "utilidad" o calidad de vida percibida. Las utilidades pueden variar desde 0, que significa muerte, hasta 1, que significa vida con salud normal. Todas las evaluaciones fueron realizadas por una enfermera investigadora con la esperanza de evitar el deseo de todo paciente de agrandar a su cirujano. Los investigadores están claros, sin embargo, que el sesgo introducido por la percepción del paciente es muy difícil de evitar. En general, las utilidades promedio, en cada cuestionario, no difieren entre los dos grupos de pacientes, y la calidad de vida es percibida como normal en aquellos pacientes sin

recurrencia de la enfermedad. La muestra fue estimada previamente como de 19 pacientes por grupo, para identificar una diferencia mínima de 0,15 entre las utilidades promedio. El estudio se realizó con 25 en cada grupo.

En relación a la percepción del paciente, es importante considerar la opinión de los urólogos del llamado proyecto de Maine, que trata de evaluar los resultados de las alternativas para las enfermedades de la próstata: observación, prostatectomía abierta o transuretral (27). Como la prostatectomía por lo regular reduce la esperanza de vida, los beneficios de la cirugía derivan de la mejoría de la sintomatología asociada. Por consiguiente, para evaluar los resultados es esencial identificar la percepción de los pacientes a ser intervenidos y respetar su escogencia.

Estudios de casos y controles. En contraposición a los estudios de cohortes —que en general se asocian con el seguimiento prospectivo— en otros problemas, el diseño de investigación que resulta más apropiado es indagar hacia el pasado de los pacientes. Es la evaluación comparativa de los antecedentes de los pacientes con una enfermedad, con los de los controles, con otra o sin ninguna enfermedad.

Identificados los casos y sus controles, el investigador obtiene su historia para conocer la presencia o ausencia de un factor de riesgo o su nivel. La intención es verificar si el factor ocurre con más frecuencia en los casos que en los controles. Si ésta es la situación, el investigador concluye que existe una asociación entre el factor de riesgo y la enfermedad.

Debido a que por lo regular, los datos de los pacientes y sus controles, se obtienen de su vida anterior a la presencia o no de la enfermedad, se les llamó estudios retrospectivos, término que se utiliza cada vez menos.

En el área de la cirugía este tipo de estudio se utiliza con frecuencia para determinar los factores de riesgo asociados a complicaciones posoperatorias. Entre ellos encontramos el realizado en Japón para establecer los factores que predisponen a la neumonía posoperatoria (28).

Es conocido que las neumonías ocupan el segundo lugar, entre las infecciones intrahospitalarias, después de las del tracto urinario, la

mitad de ellas en pacientes quirúrgicos.

De acuerdo a la revisión de los antecedentes de este problema, los autores afirman que a pesar de los avances en cirugía y anestesia, en los últimos 40 años se ha observado una variación significativa en la ocurrencia de la neumonía posoperatoria. Constituye una causa importante de morbilidad y mortalidad en pacientes quirúrgicos, dado que en la actualidad se intervienen más pacientes de mayor edad.

Muchos de los factores que predisponen a la neumonía posoperatoria han sido identificados, pero no los relacionados con la evolución de la atelectasia hacia la neumonía. Este fue el objetivo principal del estudio. Para definir atelectasia se adoptaron los criterios de Windsor: moderado o marcado colapso pulmonar mediante radiografía del tórax, temperatura de al menos 38°C y antecedentes de terapia agresiva torácica, incluida broncoscopia. Para el diagnóstico de neumonía se consideró la presencia de un infiltrado en la radiografía, temperatura de al menos 38°C y esputos purulentos. En todos los pacientes con estas características se completó un análisis de los factores que pudieran contribuir a la infección pulmonar. Los factores analizados fueron edad, género, antecedentes tabáquicos, preexistencia de enfermedades respiratorias, concentración de albúmina y colinesterasa a la admisión, porcentaje del volumen de espiración forzada predicho en 1 segundo (%FEV₁), porcentaje de capacidad vital predicha (%VC), utilización de adyuvantes inhaladores preoperatorios, duración de la anestesia, pérdida de sangre durante la intervención, volumen de fluidos administrados y sedantes epidurales aplicados después de la intervención. De la revisión de 2 969 historias de pacientes intervenidos entre enero de 1980 y diciembre de 1992, por cirugía mayor abdominal, se encontró neumonía en 45 (1,5%), de las cuales 16 habían presentado atelectasia previa. Además se encontraron 44 pacientes con atelectasia severa, ninguno de los cuales se complicó con neumonía. El tipo de cirugía al que fueron sometidos ambos grupos de pacientes fue similar estadísticamente. Los factores se analizaron con métodos estadísticos univariados y posteriormente mediante análisis discriminante multivariado. Las variables que resultaron estadísticamente significativas, fueron: edad, albúmina, colinesterasa, VC%, terapia inhalatoria y pérdida de sangre. La primera y las dos últimas permanecieron estadísticamente significativas con el análisis multivariado. Es bien conocido

que los pacientes mayores son más susceptibles a las infecciones pulmonares después de la cirugía, pues los niveles de inmunoglobulina decrecen con la edad y presentan una mayor colonización del tracto superior con organismos Gram negativos. Lo mismo ocurre con los pacientes con menores concentraciones de proteínas plasmáticas. La pérdida de más de 1 200 ml de sangre —que fue el promedio observado en los pacientes que desarrollaron neumonía— conduce a transfusión con mayor cantidad de sangre y existen evidencias de que las transfusiones homólogas son inmunosupresivas. Esto podría explicarse por haberse constatado una disminución de linfocitos circulantes posterior a ellas. Las inhalaciones previas administradas a pacientes con enfermedades respiratorias preexistentes, especialmente nebulizaciones, son un factor predisponente bien conocido de las neumonías, lo que tal vez podría evitarse con el uso de equipo debidamente esterilizado, además de racionalizar su aplicación. Los autores consideran que una proporción apreciable de neumonías posoperatorias podrían prevenirse.

Los aspectos críticos de los estudios de casos y controles que merecen destacarse son: 1°. los criterios utilizados para definir los casos y los factores analizados; 2°. la selección de los controles y su comparabilidad con los casos; 3°. la precisión de las historias de exposición a los factores de riesgo.

Consideramos que los criterios para distinguir los casos de neumonía de los de atelectasia (controles en este estudio), se cumplieron en todos los pacientes. En realidad aquí no hubo selección de controles por parte de los investigadores: las atelectasias simples ocurren cuando no están presentes los factores de riesgo que predisponen a la neumonía. Los factores de riesgo estudiados, en especial los que resultaron estadísticamente diferentes entre ambos grupos, son relativamente fáciles de definir y registrar. Como el análisis es a posteriori, es probable que las imprecisiones de registro hayan afectado por igual ambos grupos de pacientes, pues se ignoraba su evolución posterior al momento de su recolección.

En este ejemplo, tanto los casos como los controles son pacientes hospitalizados, debido a la naturaleza misma del problema investigado. Esto representa ciertas ventajas y desventajas. La primera de las ventajas es que tanto los casos como los controles son fáciles de identificar, lo que minimiza costo y esfuerzo para reunir el número requerido.

En segundo lugar, como se trata de hospitalizados, es más probable que estén más conscientes de sus antecedentes que los controles sanos. Se reduce así la posibilidad de introducir sesgos por recuerdo. En tercer lugar, al usar pacientes hospitalizados, se presume que los factores que pudieran influir en su admisión a este hospital, no deben diferir entre los grupos a comparar. Finalmente, los pacientes hospitalizados por lo general son más cooperadores que los controles sanos, por lo cual se minimiza el sesgo por no respuesta.

La mayor desventaja de utilizar controles hospitalizados es que por definición están enfermos y, por consiguiente, se diferencian de los sujetos sanos en un número de variables asociadas a la condición de enfermedad y de estar hospitalizados. En general, se ha observado que los hospitalizados tienden a fumar más, consumir más bebidas alcohólicas, utilizar más contraceptivos orales y beber menos café. Es así como una de las explicaciones de no encontrar relación con hábitos tabáquicos en el estudio, pudiera ser debido a que tanto los casos como los controles fumaban con la misma frecuencia.

Otro ejemplo que ilustra la utilización del estudio de casos y controles, es el realizado como complemento de una experiencia controlada en el programa de atención médica permanente, de la fundación Kaiser en California, EE.UU. (29). Esta experiencia realizada en cerca de 5 000 miembros adscritos al plan de salud de la fundación, residentes en Oakland, Berkeley y San Francisco, se inició en 1964, con la finalidad de evaluar los beneficios de someterse a exámenes médicos periódicos. El estudio duró 16 años y los resultados demostraron, que el grupo que cumplió con los exámenes, se había beneficiado, aunque no tanto como era de esperarse. Las afecciones detectadas mediante el examen periódico fueron la hipertensión y el tumor maligno del intestino grueso. Sin embargo, la diferencia en mortalidad fue pequeña y no estadísticamente significativa. En síntesis, si bien los resultados parecen haber sido favorables a los exámenes periódicos para evaluar la salud, no dejan de ser controversiales. No existen dudas en cuanto a la prevención de la hipertensión, pero los resultados no permitieron concluir claramente respecto a la sigmoidoscopia en su efecto en la reducción de la mortalidad por cáncer colorrectal, puesto que el grupo de estudio se sometió a este examen casi con la misma frecuencia que el grupo control. Fue por

ello que la fundación planeó y ejecutó un estudio de casos y controles para evaluar la eficacia de la sigmoidoscopia periódica, en la reducción de la mortalidad por cáncer del colon y el recto.

Este estudio fue posible, pues para la fecha de su inicio la población adscrita alcanzaba cerca de 2 millones de personas, por lo que se estimaba que ocurriera un número apreciable de muertes por cáncer colorrectal (los casos) y se pudiera generar un número apreciable de controles, pareados por las variables de interés. Asimismo, los registros permitían conocer si habían sido sometidos o no a la sigmoidoscopia, si habían desarrollado o no cáncer y su posterior evolución (30). Existían pues todas las facilidades para realizar el estudio de casos y controles basado en los registros disponibles.

El estudio demostró que las personas que murieron de cáncer colorrectal habían sido sometidas a un número muchísimo menor de sigmoidoscopias que los controles. De acuerdo a los resultados, se constató una reducción del 59% del riesgo de morir por cáncer colorrectal, entre los sometidos a la sigmoidoscopia. Esta reducción, por supuesto, estuvo confinada a los cánceres localizados en aquellas áreas que pueden alcanzarse por el examen; los localizados en las regiones superiores del colon no se previnieron. Otro de los hallazgos fue que no se encontraron diferencias de eficacia en la realización del procedimiento de tamizaje cada 10 años, en comparación con la frecuencia generalmente recomendada de efectuarla cada 3 ó 5 años.

Estudios experimentales. Según la tradición científica, los beneficios de un procedimiento quirúrgico se investigan como un experimento. Después de haber sido examinados y constatada la elegibilidad y la presencia de la enfermedad, lesión o complicación a ser intervenida, los sujetos a conformar la experiencia, se asignan aleatoriamente al nuevo procedimiento o al viejo ya conocido. Si este no existe, se compara el resultado con una operación simulada o placebo (*sham operation*).

Al final de un período conveniente de observación, la ocurrencia de los síntomas y signos que precedieron la intervención, se enumeran y se expresan como tasas para los distintos grupos, así como otros indicadores de bienestar, tales como, recuperación de la salud, percepción por parte de los pacientes de su calidad de vida e incorporación a sus labores habituales.

Si se aprecia una diferencia importante entre las tasas, o entre los indicadores, estadísticamente significativa y no existe otra explicación plausible, el resultado podría atribuirse al procedimiento con la tasa más favorable.

Esta conducta, sin embargo, es relativamente reciente. Hacia comienzos de la sexta década de este siglo, se comentaba en unos de los editoriales del Journal de la Asociación Médica Americana, sobre la creciente convicción de que ciertos supuestos resultados de procedimientos terapéuticos médicos o quirúrgicos, pudiesen ciertamente ser el producto de la sugestión (31).

Se hacía referencia a que en “un cuidadoso estudio controlado, un hombre con angina pectoris, quien después de 4 minutos de un ejercicio normalizado presentaba dolor intolerable e inversión de sus ondas T, después de ser sometido a una intervención quirúrgica, fue capaz de ejercitarse por 10 minutos, sin dolor y sin inversión de sus ondas T”. Pero, en este caso, el paciente sólo había sufrido una operación simulada: incisión cutánea sobre las arterias mamarias internas. Beecher, citado por Colton (32), discutió los aspectos éticos implicados en esta conducta terapéutica, y presentó razonadamente argumentos en favor del estudio controlado, antes que continuar sometiendo pacientes a procedimientos dudosos. Refiere que la ligadura de las arterias mamarias internas, para el tratamiento de la angina pectoris, originaria de Italia desde 1939, se difundió rápidamente en Estados Unidos, donde por falta de control adecuado, se volvió bastante popular.

Shindell, del Departamento de Medicina Interna, Escuela de Medicina de Pittsburgh, cita la historia de esta intervención, como un ejemplo de la pregunta que debemos hacernos siempre: ¿de qué otra manera podemos explicarnos la diferencia encontrada? (33). La ligadura bilateral de las arterias mamarias internas se realizaba con la intención de mejorar los canales circulatorios colaterales de la región cardíaca. Los entusiastas de los resultados favorables de la operación publicaron informes de pacientes, quienes actuaron como sus propios controles. Se refería que en un período de seguimiento, de 2 a 6 meses, 34 de los 50 pacientes habían mejorado clínicamente.

Otros investigadores (34), refirieron que una mejoría inicial de sus 20 pacientes no duró más allá de dos meses. Beecher (32) observaba que en estos informes, podían notarse efectos tan nocivos como la muerte.

Los ensayos clínicos controlados —más propiamente, ensayos “quirúrgicos” controlados— no se hicieron esperar. En dos trabajos (35,36), con un número reducido de pacientes y con anestesia local, el cirujano recibía la orden de ligar las arterias o realizar una simple incisión de la piel, al momento de iniciar la intervención. Los médicos que evaluaban posteriormente al paciente desconocían si aquél había sido sometido a la ligadura o a la operación simulada. Los resultados de estas evaluaciones no mostraron diferencias en la evolución de los síntomas anginosos en ambos grupos de pacientes. Así se respondió a la pregunta: ¿Si existe una diferencia resultante de la operación, es debida a la operación por si misma o sería que las personas sometidas a la operación difieren de quienes no la sufrieron?

Otros ensayos están diseñados para responder otras preguntas, entre ellas: ¿Cómo explicar que unos pacientes respondan satisfactoriamente y otros no?

En un estudio, informado en 1957, (37) de fracturas del cuello femoral, se trataron 204 pacientes, 139 con instalación de prótesis y 65 con el tratamiento conservador. Del análisis de los resultados, mientras casi la mitad de los sometidos al tratamiento conservador (31 de 65), fallecieron en el posoperatorio, sólo aproximadamente la cuarta parte (36 de 139), falleció con el tratamiento quirúrgico.

Al discutir estas diferencias, lo autores advirtieron que al subdividir el grupo de pacientes con fracturas, de acuerdo al riesgo de sobrevivir la intervención, se constató que en ninguno de los dos grupos, el tratamiento quirúrgico presentó menor letalidad. Como correctamente señalaron los autores de esta investigación, en la cual la letalidad fue sólo una de las variables analizadas, las diferencias son debidas a características de los pacientes y no a los tratamientos empleados. Con la aplicación de prueba estadística a los resultados, se pudo mostrar que en ninguno de los grupos hubo diferente letalidad de acuerdo al tratamiento.

Hay muchos otros aspectos a considerar en la interpretación de los resultados de un ensayo quirúrgico controlado, pero concluimos la descripción de esta aplicación de la epidemiología en cirugía, con el llamado “análisis de los datos por intención a tratar”.

Este análisis debe realizarse, mediante compa-

ración de los resultados de los grupos incluidos en el ensayo, considerados en su totalidad. Esto quiere decir, que si se trata de dos grupos, se analicen los datos completos, tanto del grupo de ensayo quirúrgico, como los del grupo control. Este grupo control pudo haberse sometido a tratamiento médico o a otro procedimiento quirúrgico, que se utiliza como término de comparación.

Como es de todos conocido, no todos los pacientes seleccionados (preferiblemente al azar), para uno u otro grupo, reciben el tratamiento para el cual fueron asignados, ni lo cumplen de la misma forma. Es sabido, que a pesar de planear cuidadosamente un ensayo quirúrgico, pueden ocurrir desviaciones del protocolo de estudio. Por ejemplo, algunos pacientes aleatorizados para intervención laparoscópica o abierta, pudieran en realidad tener que someterse a la modalidad opuesta a la cual fueron asignados previamente, por circunstancias imprevistas. Así por ejemplo, en un estudio de colecistectomía por vía laparoscópica (38), sus autores reconocen en el informe que 4 pacientes (2,6%), se convirtieron a cirugía abierta. Como en este ejemplo, se trata de una sola serie de pacientes, no hay problemas en cuanto al tipo de análisis. En otro ejemplo, en el cual se comparan los resultados de apendicectomías por vía laparoscópica y abierta, los autores informan que 2 pacientes (7,1%), asignados a la primera de las dos vías, se convirtieron a cirugía abierta durante la intervención, sus datos permanecieron para el análisis, en el grupo de laparoscopia (39). Es decir, en el grupo al que de primera intención fueron ubicados. En consecuencia, una interrogante legítima es cuáles de los pacientes deben incluirse en el análisis para una evaluación válida y no sesgada de la eficacia y la seguridad del tratamiento.

Debemos recordar, para entendimiento del problema, que la aleatorización de los pacientes genera grupos comparables, pero no garantiza una distribución semejante de las variables básicas, tales como las características demográficas y otras relativas a la enfermedad, lesión o complicaciones del tratamiento. Esto no sólo en cuanto a las variables que podríamos identificar, medir y controlar con anterioridad, sino también de aquellas no medidas y de las cuales desconocemos su posible influencia.

Estos puntos son básicos para comprender el fundamento del análisis de acuerdo “a la intención a tratar”. Una vez que los grupos a comparar han sido

constituidos mediante aleatorización, es así como tenemos que analizarlos, sin importar el tratamiento que en realidad recibieron. En otras palabras, el análisis trata de respetar la “intención” de la distribución aleatoria.

Este tipo de análisis se diferencia del llamado “evaluable”, que se refiere a aquellos pacientes que completaron el estudio y de acuerdo a la intervención terapéutica que en realidad recibieron. Existen otros criterios para incluir pacientes para el análisis evaluable, que no describimos aquí por razones de espacio, pero en todo caso deben especificarse debidamente en el protocolo de estudio, con anticipación a los resultados. Para mayores detalles, remitimos al lector a la obra de Chow y Liu (40).

Para ilustrar el concepto, utilizaremos el ejemplo de Sackett y col. (41), con los resultados de la comparación entre tratamiento médico y quirúrgico de la estenosis bilateral de la carótida. El análisis evaluable se realizó con los pacientes que completaron el seguimiento: 79 sometidos a la intervención quirúrgica y 72 al tratamiento médico. De la comparación de las tasas de ataque isquémico transitorio, accidente cerebrovascular o muerte, se observó una reducción del 27% de dichos riesgos, a favor de la intervención quirúrgica ($P= 0,02$). Sin embargo, de una lectura cuidadosa del informe, se observa que en total 167 ingresaron al estudio, de los cuales 16 no completaron el seguimiento, porque sufrieron accidente cerebrovascular o murieron al iniciarse la hospitalización, después de haber sido distribuidos aleatoriamente en uno de los dos grupos a comparar: 15 a intervención quirúrgica y el restante a tratamiento médico. Si estos pacientes, a quienes originalmente se tenía “intención” de tratar, se incluyen en el análisis, pues se disponía de todos sus datos, se observó que la reducción del riesgo a sufrir las afecciones antes señaladas o fallecer, es solamente del 16%, siempre a favor de la intervención quirúrgica, pero la diferencia pierde significancia estadística ($P= 0,09$).

Como puede observarse en este ejemplo, el análisis evaluable es mucho más vulnerable a la inclusión o no de algunos de los pacientes con los cuales se inició el estudio. La conducta correcta a seguir es incluir todos los datos de los pacientes, para que puedan compararse los resultados del análisis evaluable y el de intención a tratar. Si se observan discrepancias entre los dos tipos de análisis, la conclusión debe ser el producto de una discusión debidamente razonada.

REFERENCIAS

1. Greenwood M. Epidemic and crowd diseases. London: Williams & Norgate; 1935.
2. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research. Belmont: Lifetime Learning Publications; 1982.
3. Krieger N. Epidemiology and the web of causation: Has anyone seen the spider? *Soc Sci Med* 1994;39(7):887-903
4. Susser M. Does risk factor epidemiology put epidemiology at risk? *J Epidemiol Health* 1998;52:608-611.
5. Perera FP. Uncovering new clues to cancer risk. *Sci Am* 1996;274(5):40-46.
6. Córdoba S. Apendicitis aguda en Venezuela, estudio clínico, tratamiento y técnica operatoria. *Gac Méd Caracas* 1928;35(16):242-249.
7. Blanco Gásperi P. Mi primer centenar de tiroidectomías. *Gac Méd Caracas* 1947;55(7/12):47-128.
8. Glover JA. The incidence of tonsillectomy in school children. *Proc Roy Soc Med* 1938;31:1219-1236.
9. Bakwin H. The tonsil-adenoidectomy enigma. *J Pediatr* 1958;52(3):339-361.
10. Lewis Ch E. Variations in the incidence of surgery. *N Engl J Med* 1969;281(16):880-884.
11. Lembcke PA. Measuring the quality of medical care through vital statistics based on hospital service areas. I. Comparative study of appendectomy rates. *Amer J Pub Health* 1952;42:276-286.
12. Eisele CW, Slee VN, Hoffmann RG. Can the practice of internal medicine be evaluated? *Ann Int Med* 1956;44(1):144-161.
13. Luft HS, Bunker JP, Enthoven AC. Surgical hospital volume and mortality. *N Engl J Med* 1979;301:1364-1369.
14. Yazlle-Rocha JS, Ortiz PC, Tai-Fung Y. The incidence of cesarean sections and remuneration for treatment during childbirth. *Caderno de Saúde Pública, (Rio de Janeiro)*, 1985;1(4):457-466.
15. Barros FC, Vaughan JP, Victora CG. Why so many cesarean sections? The need for a further policy change in Brazil. *Health Policy and Planning* 1986;1(1):19-29.
16. Riley L, Greene MF. Elective cesarean delivery to reduce the transmission of HIV [editorial]. *N Engl J Med* 1999;340(13):1032-1033.
17. Doyle JM. Unnecessary hysterectomies: Study of 6 248 operations in thirty-five hospitals during 1948. *JAMA* 1953;151(5):360-365.

18. Vayda E. A comparison of surgical rates in Canada and in England and Wales. *N Engl J Med* 1973;289(23):1224-1229.
19. Lipworth L, Lee JAH, Morris JN. Case-fatality in teaching and non-teaching hospitals, 1956-1959. *Medical Care* 1963;1:71-76.
20. Ravelo Celis JA, Ravelo-Pagés R. Tratamiento preservador del cáncer de la mama. Experiencia personal en 15 años. *Gac Méd Caracas* 1999;107(1):32-55.
21. Willsher PC, Leach H, Ellis IO, Bourke JB, Blamey RW, Robertson JFR. A comparison outcome of male breast cancer with female breast cancer. *Am J Surg* 1997;173:185-188.
22. McMillan DC, Wotherspoon HA, Fearon KCH, Cooke TG. A prospective study of tumor recurrence and the acute-phase response after apparently curative colorectal cancer surgery. *Am J Surg* 1995;170:319-322.
23. Eldar S, Nash E, Sabo E, Matter I, Kunin J, Mogilner JG, Abrahamson J. Delay of surgery in acute appendicitis. *Am J Surg* 1997;173:194-198.
24. Wolf SE, Rose JK, Desai MH, Mileski JP, Barrow RE, Herndon DN. Mortality determinants in massive pediatric burns. *Ann Surg* 1997;225(5):554-569.
25. Proceedings for the 53rd Annual Sessions of the Owen H. Wangenstein Surgical Forum, 84th Annual Clinical Congress, 1998, October 25-30, Orlando FL.
26. McLeod RS, Taylor BR, Oconnor BI, Greenberg GR, Jeejeebhoy KN, Royall D, Langer B. Quality of life, nutritional status, and gastrointestinal hormone profile following the Whipple procedure. *Am J Surg* 1995;169:179-185.
27. Wennberg JE, Mulley AG, Hanley D, Timothy RP, Fowler FJ. An assessment of prostatectomy for benign urinary tract obstruction: Geographic variations and the evaluation of medical care outcomes. *JAMA* 1988;259(20):3027-3030.
28. Fujita T, Sakurai K. Multivariate analysis of risk factors for postoperative pneumonia. *Am J Surg* 1995;169:304-307.
29. Friedman GD. *Primer of epidemiology*. 4^a edición. New York: McGraw-Hill, Inc.; 1994.
30. Selby JV, Friedman GD, Quesenberry CP, Weiss NS. A case-control study of sigmoidoscopy and mortality from colorectal cancer. *N Engl J Med* 1992;326:653-657.
31. Beecher HK. Ethics and experimental therapy (editorial). *JAMA* 1963;186(9):858-859.
32. Colton T. *Medical statistics*. Boston: Little Brown and Co.; 1974.
33. Shindell S. Hypotheses and conclusions. *JAMA* 1963;186(9):849-853.
34. Kitchell JR, Glover RP, Kyle RH. Bilateral internal mammary artery ligation for angina pectoris. *Am J Cardiol* 1958;1:46-50.
35. Diamond EG, Kittle CF, Crockett JE. Evaluation of internal mammary artery ligation and sham procedure in angina pectoris. *Circulation* 1958;18:712-716.
36. Cobb LA. Evaluation of internal mammary artery ligation by double blind technic. *N Engl J Med* 1959;200:1115-1118.
37. Murray DB, Young S. Fractures of femoral neck. Review of management and results. *Lancet* 1957;1:507-510.
38. Ayala LA, Souchon E, Belloso R, Henríquez L. La colecistectomía por video laparoscopia. *Video-Rev Cir* 1992;9(1):21-30.
39. Frazee RC, Roberts JW, Symmonds RE, Snyder SK, Hendricks JC, Smith RW, Custer MD, Harrison JB. A prospective randomized trial comparing open versus laparoscopic appendectomy. *Ann Surg* 1994;219(6):725-731.
40. Chow SC, Liu J. *Design and analysis of clinical trials*. New York: John Wiley and sons; 1998.
41. Sackett DL, Brian R, Tugwell P. *Clinical epidemiology*. Boston: Little, Brown and Co. 1985.