

BIBLIOGRAFÍA

Asen KE; Tomson P. Intervención familiar. Guía práctica para los profesionales de la salud. 1ª ed. Barcelona: Paidós Ibérica SA; 1997.

Carrillo JE, Green AR, Betancourt JR. Cross-cultural primary care: A patient-based approach. *Ann Intern Med* 1999; 130:829-834.

Cassell EJ, Talking With Patients, volume I: The theory of doctor-patient communication (Cambridge, Mass: The MIT Press; 1985) citado en Laine C, Davidoff F Patient-centered medicine: a professional evolution. *JAMA* 1996; 275:152-156.

Emanuel EJ, Emanuel LL. Four models of the doctor-patient relationship. *JAMA* 1992; 267:2221-2226.

Kassirer JP. Adding insult to injury: usurping patients' prerogatives. *N Engl J Med* 1983;308:898-901.

Kleinman A, Eisenberg L, Good B. Culture, illness, and care. *Ann Intern Med* 1978; 88:251-258.

Laine C, Davidoff F. Patient-centered medicine: a professional evolution. *JAMA* 1996; 275:152-156.

Roa R, Ruiz Morosini ML, Siede JA. Medicina Familiar: hacia un nuevo modelo de atención de la salud. 1ª ed. Buenos Aires: Akadia; 1997.

Slack W. The patient's right to decide. *Lancet* 1977;2(8031):240.

CAPÍTULO

6

MEDICINA BASADA EN LA EVIDENCIA

FERNANDO RUBINSTEIN, SERGIO TERRASA Y KARIN KOPITOWSKI

6-1. ASPECTOS GENERALES DE LA MEDICINA BASADA EN LA EVIDENCIA

FERNANDO RUBINSTEIN

6-2. OBSTÁCULOS PARA LA PRÁCTICA DE LA MEDICINA BASADA EN LA EVIDENCIA

SERGIO TERRASA

6-3. LA FORMULACIÓN DE LA PREGUNTA

KARIN KOPITOWSKI

6-4. ESTIMACIÓN DEL RIESGO EN LA PRÁCTICA CLÍNICA

FERNANDO RUBINSTEIN

6-5. INTRODUCCIÓN AL DISEÑO DE LOS ESTUDIOS DE INVESTIGACIÓN

FERNANDO RUBINSTEIN

6-6. EVALUACIÓN DE UN ARTÍCULO DE INVESTIGACIÓN

FERNANDO RUBINSTEIN

6-1. ASPECTOS GENERALES DE LA MEDICINA BASADA EN LA EVIDENCIA

FERNANDO RUBINSTEIN

INTRODUCCIÓN

La medicina basada en la evidencia es, esencialmente, una estrategia de aprendizaje y formación continua que intenta llenar la brecha entre la investigación y la práctica clínica cotidiana. Intenta brindar los elementos necesarios para poder hacer una interpretación crítica de la información publicada e integrarla efectivamente a

la experiencia propia y así garantizar una atención médica de buena calidad.

DIFICULTADES PARA MANTENERSE ACTUALIZADO

Cualquiera sea nuestra especialidad, los médicos que trabajamos en el área de la atención médica primaria, tenemos la responsabilidad de

mantenernos actualizados. Esta tarea se ha vuelto más complicada a causa de los avances científicos y los aportes que la literatura médica nos brinda diariamente.

Explosión del conocimiento

En épocas pasadas, el limitado conocimiento sobre la biología humana y la escasez de estudios de investigación clínica determinaban que los avances científicos fueran publicados con mucha menor frecuencia y, por lo tanto, “estar al día” era considerablemente más fácil. Los viejos médicos de familia podían sentirse relativamente tranquilos (con fundamento o sin él) de que el cuerpo de sus conocimientos se mantenía estable, o que, por lo menos, les era posible seguir trabajando con un buen nivel de calidad a través de sus años de profesión. Sin embargo, en los últimos cuarenta años el progreso de las ciencias básicas hizo posible comprender muchos fenómenos biológicos y el importante desarrollo de los métodos epidemiológicos determinó el punto de partida de muchos estudios clínicos y su aplicación directa a la población, en general, y a los pacientes, en particular. Se comenzaron a conocer factores asociados al desarrollo de algunas enfermedades, a medir la ocurrencia de los eventos, a diseñar tratamientos y a evaluar sus resultados.

Para cualquier médico de nuestro tiempo es absolutamente imposible mantenerse al día con la información que aparece cada día ya que según lo calculado por Sackett, sería necesario leer alrededor de 20 artículos diarios, los 365 días del año.

Importancia de priorizar la información de buena calidad

El número de publicaciones se multiplicó en tal medida que hoy existen miles si sumamos todo el espectro de las especialidades médicas. Por otra parte, la permanente presión que se ejerce sobre los investigadores de los centros académicos para publicar más, y así garantizar su promoción y su supervivencia mediante la obtención de los fondos que financien los distintos proyectos de investigación, lleva frecuentemente a la publicación de un sinnúmero de estudios de calidad objetable, existiendo además un constante bombardeo de información médica. Por este motivo, necesitamos establecer prioridades y adquirir nociones básicas de revisión de la literatura para priorizar qué vamos a leer en el escaso tiempo que tenemos.

Cuanto más escasos son los recursos de un país, mayor es la necesidad de planificar sobre la base de prioridades que deben indefectiblemente fundarse en evidencias sólidas. Es importante que, durante nuestra formación profesional, recibamos entrenamiento formal sobre metodología de la investigación y sobre estrategias de revisión crítica de la literatura.

Lamentablemente también es verdad que en la mayoría de las circunstancias tampoco existe el tiempo necesario para buscar información (o la infraestructura necesaria para obtenerla) que dé respuesta a la infinidad de preguntas que surgen en la práctica clínica diaria sobre los problemas particulares que enfrentamos cuando estamos con un paciente. Esto es muy importante ya que, en un gran número de ocasiones, seguimos basando nuestras conductas y decisiones médicas en conocimientos incorporados durante nuestra formación que ya no están actualizados o en la sobreinterpretación de experiencias sobre pacientes individuales que pueden carecer de la necesaria objetividad y representatividad.

ESTRATEGIAS PARA APLICAR LA EVIDENCIA DISPONIBLE A LA PRÁCTICA CLÍNICA COTIDIANA

Cuando hablamos de la mejor evidencia disponible, nos referimos a la que proviene de estudios relevantes de investigación clínica que, una vez evaluados, se consideren apropiados para responder las preguntas de la práctica clínica diaria.

La estrategia utilizada como punto de partida de ese aprendizaje se basa en la solución de problemas (aprendizaje basado en el problema) y fue introducida en los años sesenta en la escuela de medicina de la universidad de McMaster en Canadá. Especialmente para nuestro medio, más acostumbrado a la enseñanza estructurada, este tipo de metodología puede resultar menos habitual pero, sin lugar a dudas, ofrece una oportunidad insustituible de participar como sujeto activo en el proceso de aprendizaje.

La formulación de la pregunta

A los médicos nos resulta familiar el concepto de empezar nuestro razonamiento clínico con una pregunta, ya que surgen naturalmente a partir del contacto con nuestros pacientes. El gran desafío reside en organizar estas preguntas de tal manera que puedan responderse usando la estrategia de solución de problemas. Así, en lugar de preguntarnos si valdrá la pena agregar anticoagulantes a un paciente con una miocardiopatía dilatada, nos deberíamos preguntar concretamente cuál es el riesgo de eventos embólicos en un grupo de pacientes similar al nuestro (pronóstico), cuanto se modifica con el tratamiento y si el beneficio de esa intervención supera el riesgo de complicaciones (terapéutica).

La mayor dificultad radica en nuestra habilidad para ser lo más específicos posible en la pregunta, definiendo claramente el tipo de paciente, el problema, la intervención y el resultado a evaluar, como veremos más adelante.

La búsqueda de la información

Una vez esbozadas las preguntas, el paso siguiente consiste en buscar las respuestas. Más allá de las fuentes habituales de información a las que diariamente recurrimos (otros colegas, textos, etc.), existen grandes bases de datos electrónicas que se van haciendo cada vez más accesibles a través de bibliotecas, colegios médicos, instituciones, etc. o a través de Internet. Las bases de datos bibliográficas, de las cuales la más conocida y difundida es MEDLINE (véase **cap. 7, “Recursos médicos en Internet”**) permiten identificar un gran porcentaje de las publicaciones actuales utilizando estrategias de búsqueda que van desde la más sensible (identificando la mayor cantidad de artículos relacionados con el tema) a la más específica (limitando la búsqueda a través de criterios que filtran y reducen el número de trabajos a revisar). Estas estrategias de búsqueda se pueden aprender e incorporar con facilidad para utilizar en forma rutinaria. No debemos olvidar a LILACS para buscar en publicaciones de Latinoamérica y del Caribe.

Otras bases de datos disponibles en Internet son la “Base de Revisiones Sistemáticas de Cochrane” (Cochrane Database of Systematic Reviews) ahora también en español y revistas como *EVIDENCIA en la práctica ambulatoria* que incluye información sobre estudios evaluados críticamente y comentados (publicaciones secundarias) lo que ahorra tiempo al lector brindando una fuente objetiva y confiable de actualización.

Revisión crítica de la bibliografía

Cuanto más clara y específica sea la pregunta, menor será la cantidad de estudios que deberemos evaluar para decidir cuál o cuáles son los apropiados para responderla. Como resultado inicial de la búsqueda, uno tiene a su disposición los resúmenes de los estudios seleccionados y una breve lectura general nos indicará los estudios más adecuados para analizar en detalle.

Una gran cantidad de trabajos publicados, ya sea por su falta de rigor metodológico o por su escasa relevancia, no son útiles para obtener respuestas. Por ello, la etapa de la revisión o evaluación crítica de los estudios es crucial para decidir si la información que brindan es confiable, válida y clínicamente útil como para ser aplicada en nuestra práctica.

No es necesario ser un epidemiólogo o tener experiencia en investigación para poder realizar una excelente revisión. La estrategia

consiste en poder responder una serie de preguntas estructuradas sobre los distintos tipos de estudios, ya sean de diagnóstico, pronóstico, tratamiento, calidad de atención médica, análisis económico, revisiones sistemáticas, guías de práctica clínica, etc., que toman en consideración aspectos metodológicos para evaluar la validez interna del estudio, sus resultados y la posibilidad de generalización de las conclusiones (validez externa). Las preguntas son en general muy claras pero se necesita una base mínima de conocimiento para poder aplicarlas correctamente a los distintos tipos de trabajos. Estas guías han sido publicadas en la literatura y son objeto de continua revisión y refinamiento. Obviamente, la experiencia se va adquiriendo a medida que uno realiza este proceso una y otra vez. Una vez aprendidos los pasos, la mayoría de los trabajos se podrán evaluar rápidamente.

La evidencia obtenida

Una vez identificada la información como válida, relevante y aplicable, podrá utilizarse directamente en la toma de decisiones sobre un paciente. Además podrá tomarse en cuenta para desarrollar guías de práctica, para enriquecer la discusión grupal sobre pacientes individuales, para fundamentar nuestras opiniones a la hora de decidir políticas de salud en una organización y para implementar programas de control de calidad.

JUSTA MEDIDA DE LA MEDICINA BASADA EN LA EVIDENCIA

Críticas a la medicina basada en la evidencia

Muchas voces críticas se han levantado destacando que la Medicina Basada en la Evidencia (MBE) no es nada nuevo y que podría ser una innovación peligrosa llevada a cabo por arrogantes con el fin de servir al control de costos y cercenar la libertad de la práctica asistencial (Sacket, 1996). Las distintas limitaciones derivadas de la metodología y del diseño de los estudios de investigación podrían afectar la práctica clínica cotidiana basada en la experiencia, en el paciente individual y en los mecanismos fisiopatológicos.

Dificultades para aplicar la evidencia disponible a la práctica cotidiana

El diseño metodológico de los estudios muchas veces impone estrictas reglas que terminan limitando, en aras de un estudio impecable, su aplicabilidad clínica. Este conflicto entre la validez interna y la externa cuenta con innumerables ejemplos en la literatura en los que las poblaciones estudiadas no incluyen ciertos subgrupos como ancianos, mujeres y grupos étnicos o raciales a los que, de todas maneras, se intenta extrapolar los hallazgos.

El grado de evidencia aportado por cada tipo de diseño guarda una estrecha relación con la pregunta que intenta contestar y con los métodos utilizados en el desarrollo del estudio particular. El ensayo clínico controlado y aleatorizado (ECCA) es el paradigma dentro de los distintos diseños, ya que es el único que garantiza la distribución balanceada de los factores confundidores entre los distintos grupos, tanto de los conocidos como los desconocidos, a través de la asignación aleatoria de la intervención. Esto es cierto y especialmente útil cuando queremos establecer una relación causal inequívoca entre la intervención y el evento (ej. si un determinado tratamiento modifica la morbilidad de una enfermedad o si una determinada

estrategia de rastreo modifica su historia natural). Sin embargo, estos estudios pueden tener limitaciones éticas o económicas y su validez externa puede verse restringida por el tipo de población seleccionada.

No todas las decisiones clínicas pueden estar basadas en la evidencia

La práctica clínica nos enfrenta con innumerables situaciones en las que se toman decisiones, siendo la evidencia disponible para sustentarlas controvertida, o simplemente inexistente. También es cierto que muchas veces, lo que es blanco o negro en el papel se torna gris al confrontarlo con la práctica y estas brechas de “áreas grises” serían mayores si buscáramos evidencia proveniente de ECCA para la mayoría de nuestras decisiones cotidianas.

Los estudios de investigación observacionales o los ECCA no están diseñados para responder preguntas sobre el tratamiento de pacientes individuales. Estos ofrecen una estimación de la eficacia promedio de una intervención al ser comparada con otra o del rango probable de la magnitud de asociación de un factor con un evento determinado. Este “resultado en el paciente promedio” es, para muchos, difícil de trasladar a su práctica porque sus pacientes no son similares a los reportados o la intervención no es reproducible. Es justamente en este punto donde la estrategia global de la MBE adquiere la mayor importancia ya que la evidencia externa obtenida puede informar pero nunca reemplazar el juicio clínico individual y es ese criterio el que finalmente debe prevalecer para decidir cómo integrar esa información en la toma de decisiones con el paciente.

Utilizar apropiadamente la evidencia disponible nos permite actualizar nuestro conocimiento y trasladarlo a la práctica cuando es aplicable. También permite conocer la indicación (o no indicación) de distintos tratamientos o intervenciones diagnósticas ante determinadas condiciones.

Los potenciales abusos que bajo el nombre de la MBE se puedan cometer nos lleva a alertar sobre la interpretación que puedan hacer sobre esta estrategia los profesionales de la salud, así como las autoridades gerencadoras o financiadoras de los distintos sistemas; tanto públicos como privados. Por tal motivo, los profesionales no deben adoptar dogmáticamente los principios de la MBE como única fuente de información e inspiración en detrimento de otros elementos fundamentales que, armónicamente integrados, conforman la práctica clínica. Los principios deben integrarse con la propia experiencia, el conocimiento de cada realidad sanitaria y la relación médico paciente. Respecto de las autoridades gerencadoras o financiadoras de los distintos sistemas de salud, debe evitarse el establecimiento de políticas rígidas de control y reconocimiento económico basadas en la adopción de reglas o guías inflexibles que requieran una justificación que la literatura médica, hoy por hoy, no es capaz de brindar.

CONCLUSIONES

Es absolutamente cierto que la evidencia disponible puede no adaptarse a nuestras necesidades con los pacientes individuales ya que muchas de las condiciones presentes en los estudios de investigación no son replicables en la práctica cotidiana; pero esto no debe impedir que, aun reconociendo sus limitaciones, utilicemos esta estrategia como una poderosa herramienta de enseñanza y aprendizaje.

6-2. OBSTÁCULOS PARA LA PRÁCTICA DE LA MEDICINA BASADA EN LA EVIDENCIA

SERGIO TERRASA

INTRODUCCIÓN

Como estrategia de aprendizaje y educación médica continua, la medicina basada en la evidencia intenta obtener en el menor tiempo

y con el menor esfuerzo posible respuestas confiables a preguntas específicas. Una de cada dos preguntas que surgen durante la práctica clínica podría ser potencialmente respondida a través de medicina basada en la evidencia (MBE). Sin embargo, durante nuestra práctica

intentamos responder nuestras dudas principalmente formulándoselas a otros colegas en un pasillo o en forma de una interconsulta formal o bien buscando en algún libro de texto.

¿CUÁNDO VALE LA PENA UTILIZAR LA ESTRATEGIA DE MEDICINA BASADA EN LA EVIDENCIA?

La estrategia de MBE para responder una pregunta clínica es necesaria cuando se cumplen los siguientes tres requisitos: 1) clara formulación de la pregunta; 2) alta probabilidad de encontrar evidencia de buena calidad para responderla; 3) potencial impacto en los cuidados del paciente.

OBSTÁCULOS DESCRITOS PARA LA PRÁCTICA DE LA MEDICINA BASADA EN LA EVIDENCIA

Muchas veces el profesional ignora lo que no sabe o lo niega por incomodidad, falta de tiempo o vergüenza. Por ejemplo, en un estudio realizado por Ely y cols., se encontró que los médicos de familia reportan 3,2 preguntas por cada diez consultas, cuyas dos terceras partes ni siquiera se formulan; mientras que de las que sí se formulan, el 80% se responden.

Otro obstáculo importante es la falta de tiempo, causada por la celeridad con la que nos vemos obligados a atender a nuestros pacientes, en ocasiones. Y cuando existe buena disponibilidad, puede resultar más rápido y práctico realizar una interconsulta a otro colega con mayor experiencia en el tema que llevar a cabo los distintos pasos que conducen a hallar la respuesta respaldada por la mejor evidencia disponible (alternativas en realidad no excluyentes sino complementarias). Si a esto le sumamos la aprensión a asumir el tratamiento de problemas de salud que no manejamos con soltura, la repetición de este proceso consolida nuestra falta de habilidad para el manejo de este tipo de situaciones clínicas que terminamos prefiriendo derivar antes que aprender a manejar.

Además muchas veces existe incertidumbre sobre dónde buscar información, hay dificultades para interaccionar con los recursos bibliográficos y algunas preguntas que surgen en el consultorio no han sido formuladas por los estudios disponibles.

Sin embargo, suele haber dificultades para convertir una pregunta clínica referida a *un individuo particular* (ej. qué prueba diagnóstica le debo solicitar a este paciente o qué medicamento le debo indicar) a una pregunta referida a una *población de pacientes parecidos al mío*. El no cumplir satisfactoriamente este proceso dificulta el hallazgo de una respuesta con las herramientas disponibles ya que la información existente suele ser poblacional y no individual.

Asimismo, el profesional suele ser desconfiado sobre la información que ha hallado, por ejemplo, por carecer de los recursos para evaluarla en forma crítica o porque la información encontrada puede no ser la más adecuada (ej. no estar dirigida a médicos) o existir dificultades para sintetizarla y concluir una respuesta.

CONCLUSIÓN

Volvemos a insistir en que la medicina basada en la evidencia nunca puede reemplazar a la experiencia y a la destreza clínica, pero sí debe complementar estas habilidades. Dicho de otro modo, si no se sabe interrogar y revisar al paciente, si no se tiene la experiencia para poder clasificarlo en las categorías clínicas que permiten pronosticar cómo evolucionará sin realizarle intervención alguna o administrándole algún tratamiento o bien qué probabilidad existe de que quede mejor categorizado luego de realizarle tal o cual prueba diagnóstica, no habrá pregunta clínica lo suficientemente bien construida ni búsqueda bibliográfica que pueda dar una respuesta útil. Por otro lado, si carecemos de la capacidad para evaluar objetivamente la evidencia disponible para avalar una decisión o una recomendación, corremos el peligro de caer en prácticas predominantemente sesgadas por experiencias personales.

6-3. LA FORMULACIÓN DE LA PREGUNTA

KARIN KOPITOWSKI

INTRODUCCIÓN

La medicina basada en la evidencia (MBE) es una estrategia que tiene por objetivo optimizar los procesos para que puedan obtenerse en el menor tiempo y esfuerzo posible las respuestas más confiables a preguntas específicas.

El cuidado de la salud requiere un continuo aprendizaje. Se trata de un proceso en el que se generan preguntas y respuestas. En efecto, al atender pacientes, permanentemente debemos encontrar respuestas a preguntas como estas: *¿cuál es el diagnóstico? ¿cuál es la causa de lo que le sucede? ¿cuál es su pronóstico? ¿cuán beneficioso o riesgoso es cierto tratamiento? ¿qué es lo que desea el paciente?*

En estas circunstancias, cuando encontramos la respuesta válida, seguimos adelante. Por suerte, la mayoría de las preguntas pueden responderse gracias a nuestra experiencia clínica, la educación médica recibida, los libros, las publicaciones científicas y/o preguntando a un colega.

Sin embargo, cuando nos acostumbramos a no obtener respuestas, solemos comenzar a olvidarnos de hacernos más preguntas. Además, debido a la universal escasez del tiempo, los médicos queremos y necesitamos que el hallazgo de la respuesta no demande más que unos pocos minutos. Por eso acudimos principalmente a otros colegas, a los libros y, escasamente, a las fuentes de información electrónicas. Por lo tanto, luego de formularnos las preguntas, sería ideal poder tener un acceso rápido a fuentes confiables de información.

METODOLOGÍA PROPUESTA

Primero: reconocer lo que uno no sabe

Generalmente, cuando el personal del equipo de salud adopta la MBE como estrategia para actualizarse y realizar educación continua, pareciera que las principales barreras con las que se topa son la búsqueda de información en bases de datos complejas, el acceso a los trabajos originales, su interpretación crítica y la necesidad de adquirir herramientas que permitan cierto análisis cuantitativo.

Sin embargo, la dificultad de los médicos y otros profesionales de la salud de reconocer y aceptar los déficit en el conocimiento constituye una barrera importante. Saber que no se sabe y reconocerlo no es una tarea sencilla.

Si bien vamos a dedicar este capítulo a la confección adecuada de la pregunta que se genera ante un déficit de conocimiento ("déficit global" o "grieta") que habitualmente se reconoce como el primer paso en la práctica de la MBE, consideramos que el primer y real paso es reconocer que algo no se sabe.

Segundo: determinar el tipo de déficit de conocimiento (grieta o el déficit global)

Asumamos que estamos preparados para reconocer aquello que desconocemos. A nuestras deficiencias de información importantes y amplias las llamaremos déficit globales y a las más puntuales, grietas.

Cuando hablamos de un “déficit global”, nos referimos a que sabemos muy poco sobre ese tema; dicho de otro modo: se trata de una pregunta amplia y de orden general. Supongamos que recibimos a un individuo con una enfermedad extraña (de esas que solo recordamos porque alguna vez las hemos marcado en alguna opción de un examen de respuestas múltiples) o que no es tan rara pero que hace mucho que no la vemos en el consultorio. En este caso nuestro déficit en el conocimiento es muy general y, probablemente, nos veremos obligados a consultar de qué se trata en algún libro de texto de los clásicos.

Pero supongamos ahora que nos consulta un paciente por una patología o condición que conocemos bien y manejamos frecuentemente pero, ya sea porque se trata de un tema que se actualiza rápidamente (ej. diabetes tipo II, virus de la inmunodeficiencia adquirida) o tenemos una duda puntual acerca del problema, estamos ante una grieta del conocimiento. En este caso, es probable que los libros no puedan ayudarnos.

Es decir que al diferenciar si estamos ante una grieta o un “déficit global” ya podemos ir imaginando el mejor sitio para encontrar la respuesta.

Ejemplo: una mujer de 50 años llega a su consultorio con diagnóstico previo de colitis ulcerosa de siete años de evolución. La controla un gastroenterólogo que le indica colonoscopias periódicas. Actualmente se encuentra asintomática y sin medicación.

El médico, quien hace tiempo no ve pacientes con colitis ulcerosa decide leer el capítulo correspondiente del libro de medicina interna ya que considera insuficiente su información y comprensión global del tema, lo que más le interesa es obtener información con buen grado de decantación (no la última novedad). Esto sería entonces una pregunta *general*. Si el médico estuviera avezado en las características de la enfermedad, pero tuviera dudas respecto de la frecuencia con que debe realizarse la colonoscopia, estaría ante una grieta del conocimiento y es posible que el mejor lugar donde buscar la respuesta no sea un libro de texto.

Ejemplo: una paciente de 20 años le consulta a un residente de primer año que está atendiendo en la guardia. Al examinarla, Uld. diagnóstica varicela. El residente no recuerda el período de incubación, las características distintivas de la erupción y las complicaciones de la enfermedad. Estamos frente a un “déficit global” del conocimiento.

Uld., que ha visto muchas varicelas en su vida médica, se pregunta si el aciclovir acelerará la recuperación.

Usted tiene una grieta.

Es muy probable que el residente pueda resolver su problema consultando un libro de texto (pregunta general). Es más, no solo lo podrá resolver sino que esa es la fuente recomendada para lo que busca ya que el período de incubación y las características de la erupción son atributos decantados que, en general, no se desactualizan.

En cambio su pregunta seguramente requerirá alguna estrategia de búsqueda más sofisticada. A este tipo de preguntas se las denomina *específicas*.

Para obtener información global y decantada sobre un tema la mejor y más accesible fuente de información suele ser un libro de texto actualizado.

Tercer paso: formular la pregunta

La pregunta clínica es el vehículo que permite diseñar la estrategia adecuada de búsqueda de la respuesta. De la precisión con que la formulemos dependerá la facilidad y rapidez con que hallemos lo que necesitamos. En otras palabras, para poder encontrar el mejor

camino es imprescindible saber primero nuestro destino de llegada. ¿Cómo se formula la pregunta? Cuando reconocemos que ignoramos algo, y ese desconocimiento es puntual y específico (grieta del conocimiento), está recomendado transformar ese escenario clínico en una pregunta bien construida. El cumplimiento de este primer paso (que no suele ser sencillo) permite hallar la respuesta con la mayor precisión y eficiencia posible. Con el ejercicio de la práctica de la MBE este primer paso tiende a automatizarse, sin embargo, recomendamos explicitar la pregunta, especialmente a quienes se inician en esta disciplina.

Las preguntas pueden estar referidas a la terapéutica, pronóstico, etiología, diagnóstico, costo efectividad, etc. Cualquiera fuese la categoría, debe respetar una anatomía que, como veremos, no es caprichosa pues contemplar los componentes que se enumeran a continuación:

- El paciente, la población o el problema al que nos referimos.
- La intervención o el factor de exposición que nos interesa.
- Comparación con otra intervención o exposición (si corresponde).
- Resultado de interés.

Paciente o población

El paciente o la población deben definirse claramente y con precisión. Sin embargo, no conviene estrechar en exceso la definición porque corremos el riesgo de fracasar en la búsqueda: nunca encontraremos un estudio de investigación con pacientes idénticos al que tenemos en frente.

Ejemplo: usted atiende a un paciente diabético, hipertenso, dislipémico con el antecedente de un infarto agudo de miocardio (IAM) anterior extenso ocurrido hace dos años. Presenta microalbuminuria positiva en el análisis de orina y tiene microaneurismas en el fondo de ojo. Usted quiere saber si el control estricto de la tensión arterial retrasará el progreso hacia la insuficiencia renal.

En rigor de verdad, la definición de la primera parte de la pregunta sería: paciente diabético tipo II infartado, con hipercolesterolemia, hipertensión, microalbuminuria y retinopatía. Esta sería la definición precisa, pero estrecha tanto el espectro de pacientes diabéticos, que probablemente no encontremos ningún ensayo que haya incluido pacientes con estas características. Por lo tanto, puede ser suficiente definir al paciente como diabético tipo II con complicaciones.

Intervención o factor de exposición y resultados

La intervención o el factor de exposición también deben definirse claramente ya que queremos saber si el tratamiento con drogas produce tal o cual efecto o nos interesa una familia de drogas o nos interesa una droga en particular; o tal vez queremos saber si dosis altas o bajas de un mismo fármaco producen efectos similares; o bien la intervención de interés es el ejercicio físico o una dieta. Del mismo modo, debemos definir si nos interesa comparar una intervención con otra.

Ejemplo: nos interesa saber si los antagonistas de receptores de angiotensina reducen la mortalidad en la insuficiencia cardíaca al menos tan eficazmente como los inhibidores de la enzima convertidora de angiotensina (comparación entre dos familias de drogas) o si el bupropion aumenta la tasa de cese tabáquico en mayor proporción que el placebo (comparación contra placebo).

Tal vez no nos interesa comparar sino que queremos saber si una exposición se asocia a algún resultado (etiología).

Ejemplo: queremos saber si los adultos (población) que usan teléfono celular (exposición) tienen más riesgo de desarrollar tumores cerebra-

les (resultado). Sin embargo, aunque no tan evidente, también hay una comparación subyacente (los que usan teléfono celular versus los que no lo usan).

Es crucial que el resultado de interés esté claramente definido, siendo un error muy frecuente preguntarse si tal o cual droga es mejor que otra porque puede llevar a errores graves de interpretación.

Ejemplo: si nos interesa saber si los inhibidores de la enzima convertidora de angiotensina bajan la presión en mayor medida que las tiazidas, en esta situación el resultado es: ¿bajan más la presión? En cambio, si queremos evaluar si a igual descenso de presión las tiazidas son más económicas, el resultado aquí es ¿cuánto dinero es necesario gastar por cada mmHg de descenso de la tensión arterial? Pero pudiera ser que sólo nos interese si disminuyen en forma equivalente el riesgo de accidente cerebrovascular (ACV) mayor, siendo en este caso el resultado: prevención de ACV mayor. Más aún, quizás a alguien le interese si algún fármaco es más seguro que otro y frente a este interrogante el resultado serían los efectos adversos graves.

Si bien todos estos resultados podrían agruparse, las implicancias de cada respuesta para el tratamiento de los pacientes es muy diferente.

Veamos con ejemplo otro error frecuente en relación a la definición del resultado de interés al formular la pregunta.

Ejemplo: usted atiende a un paciente con lumbalgia aguda que no presenta signos de alarma (p. ej., pérdida de peso, fiebre, antecedentes de cáncer conocido, inicio luego de los 50 años) que hagan sospechar una enfermedad grave o sistémica (ej. osteomielitis, metástasis óseas, etc.) y se pregunta acerca de la eficacia de los relajantes musculares.

Usted puede sentirse tentado a preguntar si en pacientes con lumbalgia aguda (población), los relajantes musculares (intervención) sirven (resultado).

Sin embargo, volvemos al mismo punto ya que preguntarse si una intervención sirve no es una buena manera de comenzar, ya que este interrogante puede estar referido al alivio del dolor, a la disminución del ausentismo laboral, al aumento de la satisfacción de los pacientes, etc.

Cuando queremos saber si una intervención es útil en algún sentido, debemos formular cómo desearíamos que esa utilidad se refleje en la vida real. Podría ser que los relajantes musculares hicieran que el paciente pase de referir un dolor de seis puntos sobre diez de intensidad a referir un dolor de cinco sobre diez. En ese caso, ¿significa que sirven? Por lo tanto, es fundamental precisar qué nos importa como médicos cuando queremos saber si algo sirve. Un modo de pregunta bien formulada con relación al caso de la lumbalgia es el que sigue:

Ejemplo: ¿en pacientes con lumbalgia aguda (población), los relajantes musculares (intervención) disminuyen la intensidad y/o duración del dolor de manera clínicamente relevante (resultado)?

Veremos varios escenarios clínicos y algunas de las preguntas que pueden formularse:

Ejemplo: usted atiende a un varón de 50 años que ha sufrido un infarto agudo de miocardio (IAM) inferior sin complicaciones hace tres meses. Al revisar las indicaciones, observa que el paciente recibe betabloqueantes, aspirina, atorvastatina y enalapril. El paciente no era hipertenso y tiene una tensión arterial de 110/70. Usted se pregunta acerca de la pertinencia de la indicación del enalapril: ¿en pacientes infartados (población) y comparado con placebo, el enalapril (intervención y comparación) disminuye el riesgo de reinfarcto, ACV o muerte (resultados de interés)?

Usted acaba de formularse una pregunta de terapéutica.

En la misma consulta, el paciente relata síntomas de depresión, y usted se pregunta acerca del pronóstico de los pacientes infartados que además están deprimidos: ¿en los pacientes infartados (población), la depresión (exposición) es un factor de riesgo para el reinfarcto o la muerte (resultados de interés)?

Usted se ha formulado una pregunta sobre pronóstico.

Veamos un ejemplo sobre un método de diagnóstico:

Ejemplo: usted se pregunta cuán buena es la densitometría ósea (DMO) para el diagnóstico de osteoporosis. Es habitual escuchar la pregunta con el formato siguiente: ¿sirve la densitometría ósea?

El lector debe haber advertido que esta no es una buena manera de formularla debido a que detrás de esta interrogación pueden, en realidad, coexistir diferentes inquietudes cuya respuesta deberá buscarse de manera diferente ya que el tipo de estudio ideal para responder cada una de esas inquietudes también es diferente.

¿Qué quiere saber usted en realidad?, ¿le interesan las características operativas de la prueba diagnóstica con relación a la biopsia de hueso para el diagnóstico de osteoporosis?, ¿o le interesa saber cuán bien predice una fractura futura una densitometría ósea anormal?, ¿o en qué grado reduce la probabilidad de una fractura futura una densitometría normal? ¿O tal vez usted ya sabe las respuestas a estas preguntas y lo que en realidad quiere saber es si un programa de rastreo aplicado en la población general reduce el riesgo de fracturas?

Con estos ejemplos acerca de preguntas que surgen con relación a una prueba diagnóstica se hace evidente que la anatomía debe ser respetada. Veamos ejemplos de preguntas bien confeccionadas:

Ejemplos: ¿en las mujeres postmenopáusicas (población) la presencia de una densitometría patológica (exposición/factor de riesgo/test anormal) predice riesgo de fractura futura (resultado)?

¿En las mujeres postmenopáusicas, la densitometría (exposición) es capaz de diagnosticar con buena especificidad y sensibilidad la presencia de osteoporosis (resultado)?; o bien, ¿en las mujeres postmenopáusicas (población) la densitometría de rastreo (intervención) disminuye el riesgo de fracturas?

En ocasiones, al reflexionar sobre pruebas diagnósticas, ya conocemos las respuestas a todas estas preguntas y lo que en realidad nos interesa saber es si esta estrategia es sustentable desde el punto de vista económico.

Ejemplo: ¿en las mujeres postmenopáusicas (población) la realización rutinaria de densitometría (intervención) tiene un costo por fractura prevenida que el sistema puede financiar (resultado)?

CONCLUSIÓN

Las preguntas y respuestas deben estar integradas al proceso de cuidado de los pacientes. Vale la pena aclarar que muchas veces es difícil reconocer que uno no tiene la respuesta, y esta es muchas veces la peor barrera para el aprendizaje. Por ello, el desafío central consiste en seguir haciéndonos preguntas.

6-4. ESTIMACIÓN DEL RIESGO EN LA PRÁCTICA CLÍNICA

FERNANDO RUBINSTEIN

CONCEPTO DE RIESGO

El riesgo es la probabilidad de ocurrencia de un evento en una población definida a lo largo de un período determinado (ej. riesgo anual de accidente cerebrovascular en los pacientes hipertensos leves no tratados).

En la literatura se utiliza el concepto de riesgo para brindar tanto información cualitativa como cuantitativa. Si, por ejemplo, queremos saber si un tratamiento es eficaz, necesitamos conocer cuál es el riesgo de sufrir el evento en una población no tratada y cuánto se modifica en aquellos que reciben la intervención. Esto nos va a permitir establecer si son o no diferentes (cualitativa) y cuán diferentes son al compararlos (cuantitativa).

Medidas de asociación comúnmente utilizadas en la literatura

Utilizaremos para explicar estos conceptos el siguiente ejemplo obtenido de la literatura. El **cuadro 6-4-1** muestra la cantidad de accidentes cerebrovasculares (ACV) en hipertensos moderados a severos tratados y no tratados seguidos durante 60 meses.

Riesgo absoluto

El riesgo absoluto (RA) es la incidencia (casos nuevos detectados) del evento de interés en cada grupo.

$$RA = \frac{\text{Número de individuos que tienen un evento nuevo en un lapso determinado}}{\text{Número total de individuos que podrían tenerlo (susceptibles) en ese mismo lapso}}$$

En nuestro ejemplo:

$$RA \text{ en pacientes tratados} = 1.800/15.000 = 0,12$$

$$RA \text{ en pacientes no tratados} = 3.000/15.000 = 0,20$$

Esto significa que el 20% (0,20) de los pacientes hipertensos moderados a severos estudiados que no recibieron tratamiento antihipertensivo desarrollaron ACV en los 60 meses siguientes, mientras que el 12% (0,12) de los que sí lo recibieron tuvo un ACV.

Riesgo relativo

El riesgo relativo (RR) responde a la siguiente pregunta: ¿cuántas veces más probable (o menos probable en estudios de tratamientos efectivos) es la ocurrencia del evento en los expuestos al factor en estudio comparado con el grupo control?

$$RR = \frac{\text{Riesgo absoluto del evento en los expuestos}}{\text{Riesgo absoluto en los no expuestos (grupo control)}}$$

Si el riesgo de sufrir el evento es igual en ambos grupos, el riesgo relativo es igual a 1 (no hay diferencia), si es mayor en el grupo expuesto es mayor a 1 y si es menor, es menor a 1.

En nuestro ejemplo:

$$RR = RA \text{ en tratados} / RA \text{ en no tratados} = 0,12/0,20 = 0,6$$

En este caso, al ser el RR menor que 1 significa que el efecto del tratamiento es protector. Es decir que por cada ACV registrado en el grupo control se registraron 0,6 en el grupo tratado (o, hablando en porcentajes, por cada 100 ACV en el grupo control se registraron 60 en el tratado).

Veamos ahora cuánto protege el tratamiento.

Reducción del riesgo relativo

La reducción del riesgo relativo (RRR) es una de las formas más comunes de reportar los resultados de los estudios terapéuticos en la literatura. Reporta el porcentaje de reducción relativa del evento en el grupo tratado con respecto al basal.

$$RRR = (1 - RR) \times 100 \text{ (Se mide en porcentaje de reducción).}$$

O

$$RRR = \frac{(\text{Riesgo en no expuestos o grupo control} - \text{Riesgo en expuestos})}{\text{Riesgo en no expuestos o grupo control}} \times 100$$

En nuestro ejemplo:

$$\text{La RRR} = (1 - 0,6) \times 100 = 40\%.$$

O

$$RRR = \frac{(0,20 - 0,12)}{0,20} \times 100 = 40\%$$

De esta forma la conclusión del estudio será: “en los pacientes con hipertensión moderada a severa que recibieron tratamiento durante cinco años se redujo en un 40% la incidencia (o el riesgo) de ACV”.

Riesgo o beneficio atribuible o reducción del riesgo absoluto o diferencia de riesgo

La reducción del riesgo absoluto (RRA) es una forma poco frecuente de expresar resultados en la literatura pero de gran aplicación para la práctica clínica. Nos responde la pregunta: ¿cuál es, en términos absolutos, la reducción del riesgo de sufrir el evento que se le atribuye al tratamiento?, (o el aumento del riesgo, en caso de tratarse de un factor nocivo).

Se obtiene simplemente calculando la diferencia de los riesgos absolutos en una y otra población.

$$RRA = RA \text{ en no expuestos} - RA \text{ en expuestos}$$

En nuestro ejemplo:

$$RRA = \text{Riesgo en no tratados} - \text{Riesgo en tratados}$$

$$RRA = 0,20 - 0,12 = 0,08 = 8\%$$

La reducción del riesgo en este estudio significa que de cada 20 ACV que ocurren en esta población de hipertensos moderados a severos, 8 pueden ser prevenidos con el tratamiento antihipertensivo o, lo que es lo mismo, el beneficio atribuible al tratamiento es del 8%.

CUADRO 6-4-1. RIESGO DE ACV EN HIPERTENSOS MODERADOS Y SEVEROS CON Y SIN TRATAMIENTO HIPERTENSIVO LUEGO DE UN SEGUIMIENTO DE 60 MESES

	ACV	No ACV	Total
Tratamiento	1.800	13.200	15.000
Placebo	3.000	12.000	15.000

Número de pacientes necesario para tratar

El número de pacientes necesario para tratar (NNT) equivale al número de pacientes que necesitamos tratar para evitar un evento. También con el NNT podemos calcular el número de pacientes que se necesitan tratar para producir algún efecto adverso. De alguna manera, el NNT es una medida que resume el impacto producido por una intervención cuando uno se pregunta: ¿a cuántos pacientes hay que someter al tratamiento para obtener un beneficio o para que uno de ellos se beneficie? Cuando una intervención es costosa, invasiva o está acompañada de muchos efectos colaterales que comprometen la calidad de vida, habrá que evaluarla muy cuidadosamente para justificar su implementación. Al igual que el riesgo atribuible, el NNT incorpora el concepto de cuán frecuente es el evento en la población que no recibe la intervención.

$NNT = \text{la inversa del RRA} = 1/RRA$.

En nuestro ejemplo:

$NNT = 1/0,08 = 12,5$

Es decir que se necesita tratar a 12.5 pacientes hipertensos moderados a severos para prevenir un ACV o, lo que es lo mismo, de cada 100 pacientes tratados se van a prevenir 8 ACV.

Frecuentemente nos quedamos con la idea de que un evento se produce tantas más veces en un grupo que en otro sin ponernos a pensar cuán frecuente es. Cuando se trata de un evento poco frecuente, o de bajo riesgo, las medidas de asociación habitualmente publicadas (RR y RRR) no nos van a dar una completa idea de la magnitud del problema. El riesgo relativo (RR) de un evento asociado a un determinado factor de riesgo puede ser de dos o tres, pero si es poco frecuente, puede significar un aumento desde dos eventos cada 10.000 personas hasta cuatro a seis eventos cada 10.000, siendo en términos absolutos poco importante desde el punto de vista clínico.

El RR y la RRR brindan información sobre el incremento del riesgo (en el caso de un factor de riesgo) o el beneficio (en el caso de un tratamiento) solo en términos comparativos entre uno y otro grupo en estudio. Para tener una idea más completa del impacto global de estos riesgos, necesitamos incorporar información sobre la población no expuesta y su riesgo basal de sufrir el evento en estudio, así como también los posibles efectos adversos. Esto lo logramos calculando el riesgo absoluto (RA) en unos y otros, su diferencia (RRA) y la recíproca de esa diferencia o número necesario para tratar (NNT), lo que nos va a dar una idea de la magnitud de la intervención necesaria para evitar un evento. Estos conceptos, que en general no se reportan específicamente en los trabajos (aunque sí los datos para calcularlos), proveen información clave para evaluar el impacto en la población y no solo la comparación entre dos estrategias o dos grupos con diferente exposición a un factor determinado.

Veamos cómo se trasladan estos conceptos a las implicancias clínicas del siguiente estudio comparándolo con el ejemplo anterior.

El **cuadro 6-4-2** muestra el riesgo de ACV en hipertensos leves con y sin tratamiento hipertensivo luego de un seguimiento de 60 meses.

$RA \text{ en expuestos} = 135/15.000 = 0,009$

$RA \text{ en no expuestos} = 225/15.000 = 0,015$

$RR = 0,009/0,015 = 0,6$

$RRR = (1 - 0,6) \times 100 = 40\%$

$RRA = 0,015 - 0,009 = 0,006 = 0,6\%$

$NNT = 1/0,006 = 167$

Tanto el RR como la RRR en pacientes hipertensos leves tratados son idénticos a aquellos estimados para la población de hipertensos más severos. ¿Esto significa que el beneficio del tratamiento es el mismo?

CUADRO 6-4-2. RIESGO DE ACV EN HIPERTENSOS LEVES CON Y SIN TRATAMIENTO HIPERTENSIVO LUEGO DE UN SEGUIMIENTO DE 60 MESES

	ACV	No ACV	Total
Tratamiento	135	14.865	15.000
Placebo	225	14.775	15.000

Como beneficio atribuible al tratamiento en esta población, notamos que de 15 ACV que se producen cada 1000 hipertensos leves no tratados durante cinco años, 6 ACV pueden evitarse con el tratamiento. El NNT es sustancialmente mayor que en los hipertensos severos, pues se necesita tratar a 167 pacientes durante cinco años para evitar 1 solo ACV. Esto simplemente se explica porque el evento en estudio (ACV) es mucho más frecuente en los hipertensos severos que en leves, por lo tanto, aunque las medidas relativas sean iguales, el RRA y el NNT evidencian el diferente impacto del tratamiento. De todas maneras, siendo la hipertensión leve un hallazgo tan frecuente en la población adulta (aproximadamente el 25 a 30% de la población), el número de eventos evitados por el tratamiento es fundamental y justifica plenamente su indicación.

Estos conceptos son muy importantes ya que constantemente se nos trata de “convencer” acerca de la utilidad de nuevos descubrimientos.

Ejemplo: la droga X mostró disminuir a la mitad la incidencia del evento Y.

En este caso la reducción del riesgo relativo (RRR) fue del 50%, pero lo que más importa es la prevalencia del evento Y, la reducción del riesgo absoluto (RRA) o el beneficio atribuible al tratamiento y el impacto clínico de utilizar esa nueva droga X.

Intervalo de confianza

El intervalo de confianza (IC) es una medida de la dispersión o la variabilidad de las estimaciones que se hacen en los estudios epidemiológicos. Toda medida estadística tiene una medida de dispersión asociada ya que las estimaciones de la realidad no son nunca exactamente iguales (los tiros al blanco de un mismo tirador). Las medidas de riesgo como el RR, la RRR, el RA, etc. deben interpretarse como parámetros estimados para la muestra en estudio. Idealmente, las muestras, para ser representativas de la población general, no deben presentar sesgos, es decir que sus características no deben ser diferentes a las de la población de la cual provienen (**fig. 6-4-1**). Sin embargo, aun utilizando una perfecta técnica de selección, una muestra de la realidad no es nunca idéntica a otra, y esta variabilidad se debe al azar. Si tomáramos 100 muestras distintas de la misma población, obtendríamos en la mayoría estimaciones de riesgo (parámetro p) muy similares pero no exactamente iguales a los de la población general (parámetro P).

Estos 100 valores de p definirán un rango dentro del cual se encuentra el verdadero parámetro P que se denomina IC.

Si repitiéramos el estudio con 100 muestras diferentes, obtendríamos un rango de p que, en el caso del IC del 95% (el más usado), incluirá el verdadero parámetro de la población (P) 95 de cada 100 veces. Existe hasta un 5% de probabilidad de que el verdadero valor de P se encuentre por fuera de este rango y esto lo consideraremos error por azar. Como vimos antes al explicar error por azar, si la probabilidad de que esa estimación sea debida al azar es igual o mayor al 5% (p mayor al 0,05), la consideraremos **no significativa** estadísticamente.

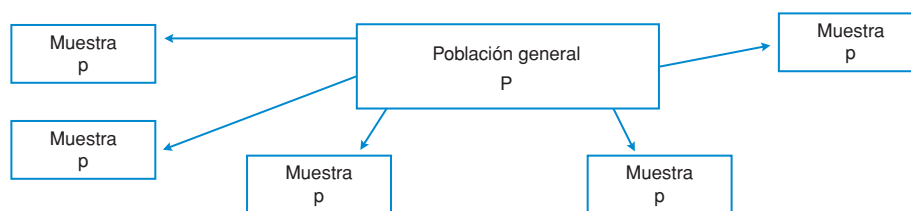


FIG. 6-4-1. Las características de las muestras representativas deben ser similares a las de la población general (P) de la cual provienen.

La precisión de un IC, es decir, la mayor o menor dispersión alrededor de un valor estimado está directamente relacionada con el número de participantes (n) en el estudio y el número de eventos (n).

Cuanto mayor es el “ n ” de un estudio más estable será la estimación del riesgo.

Ejemplo: si el RR de cáncer de mama (CM) en mujeres usuadoras de estrógenos es reportado como 1,3 (IC95%: 1,15-1,45), esto significa que la verdadera estimación de riesgo (parámetro P) estará entre 1,15 y 1,45 en 95 de cada 100 estudios similares. Hay un IC que define que el riesgo es el 95% de las veces mayor que 1 por lo que, aunque sea de 1,15 o de 1,45, siempre hay mayor riesgo.

Si cuando hablamos de RR, el IC incluye al 1, esto quiere decir que la asociación NO es significativa ya que incluye el valor de $RR = 1$, es decir riesgos iguales en ambos grupos.

En el ejemplo anterior si el RR de CM en las usuadoras de estrógenos fuera reportado como igual a 1.3 (IC95%: 0,80 - 1,45), esto significa que para esa muestra, el riesgo de CM en madres fue de 1,3 pero que, dado que el IC incluye la unidad y su rango va de 0,80 a 1,45, la estimación de riesgo en otras muestras puede ser desde un 20% menos a un 45% más, lo cual no permite establecer una estimación de riesgo estable.

Por consiguiente, conocer el intervalo de confianza nos permite tener una idea mucho más clara de una medida de riesgo porque nos brinda la idea cuantitativa de la variabilidad de esa estimación en la realidad.

ODDS RATIO

La razón de posibilidades u *odds ratio* (OR) es otra medida de efecto comúnmente usada en estudios clínicos y puede encontrarse en la literatura en español como Razón de Producto Cruzado. Tiene relación matemática con el riesgo relativo y se utiliza para reportar la chance relativa (*odds* relativo u *odds ratio*) que tiene un paciente del grupo expuesto de tener el evento, comparado con un paciente del grupo control, o la frecuencia relativa de exposición entre casos y controles.

En el ejemplo, de los 15.000 pacientes tratados, 1800 tienen el evento mientras que 13.200 no. En el grupo placebo, la relación es de 3.000 a 12.000.

El *odds* (chance) en el grupo expuesto es $a/b = (1800/13200) = 0,136$, el *odds* en el grupo control es $c/d = 0,25$. El *odds ratio* es $(a/b) / (c/d)$ o su equivalente $(ad/bc) = 0,136/0,25 = 0,54$, similar al RR estimado de 0,6. La interpretación es similar al RR y cuando los eventos en estudio son infrecuentes, se la se utiliza como una buena aproximación al RR. Siempre se reporta como el **número de eventos esperables en los expuestos por cada evento esperable en los no expuestos**.

¿Cuál es la ventaja de estimar OR si el RR es más sencillo de expresar? En muchos estudios clínicos –por ejemplo estudios de casos y controles o de corte transversal– no contamos con la información necesaria para estimar incidencia de eventos. Por lo tanto, no podemos estimar riesgos. La estimación del OR se puede realizar cualquiera sea el diseño del estudio, y permite obtener una medida de efecto que cuantifique la magnitud de la asociación entre la exposición y el evento.

El OR también se utiliza como medida para resumir una revisión sistemática u otros estudios que utilizan modelos matemáticos multi-variables como la regresión logística.

VARIABLES CONFUNDIDORAS O FACTORES CONFUNDIDORES

A veces se puede encontrar una asociación entre el factor de exposición (o variable de interés) y el evento en estudio sin que en realidad exista relación alguna entre ambas variables. Esto es producido por la presencia de una tercera variable llamada “confundidor” que está asociada tanto con el factor de exposición como con el evento en estudio en forma independiente.

Ejemplo: si analizamos en la población general la asociación entre el uso de anteojos como factor de riesgo de infarto agudo de miocardio (IAM), encontraremos que quienes usan anteojos tienen más riesgo de sufrir un IAM que quienes no los usan. Sin embargo, en los individuos mayores de 65 años el riesgo de IAM es igual para los usuarios y para los no usuarios y lo mismo ocurre con los menores de 65 años.

El **cuadro 6-4-3** nos muestra en forma gráfica este ejemplo.

Si estimamos el riesgo relativo de IAM entre quienes usan anteojos y quienes no, vemos que de 200 pacientes que usan anteojos, 34 tuvieron un IAM, mientras que de 200 pacientes que no usan anteojos, solo 16 tuvieron un IAM. Si consideramos el riesgo relativo (RR): el riesgo absoluto del evento (IAM) en los expuestos (los pacientes que usan anteojos) dividido el riesgo absoluto en los no expuestos (los que no los usan) es 2,12 (34 de 200 dividido 16 de 200) o expresado de otro modo:

Riesgo absoluto en los que usan anteojos	$34/200 = 0,17$
Riesgo absoluto en los que no usan anteojos	$16/200 = 0,08$
Riesgo relativo	$0,17/0,08 = 2,12$ (más del doble)

El **cuadro 6-4-4** nos muestra lo que ocurre cuando consideramos la edad de los pacientes y analizamos la influencia del factor de riesgo en función de la edad (mayores o menores de 65 años).

CUADRO 6-4-3. ASOCIACIÓN ENTRE EL RIESGO DE INFARTO AGUDO DE MIOCARDIO Y EL USO DE ANTEOJOS

	Pacientes con IAM	Pacientes sin IAM	Total de pacientes
Pacientes que usan anteojos	34	166	200
Pacientes que no usan anteojos	16	184	200
Total de pacientes	50	350	400

En las personas mayores de 65 años el riesgo de IAM es igual para las que usan o no usan anteojos. El riesgo relativo es igual a 1 (32 de 160 dividido 8 de 40).

Riesgo absoluto en los que usan anteojos	$32/160 = 0,2$
Riesgo absoluto en los que no usan anteojos	$8/40 = 0,2$
Riesgo relativo	$0,2/0,2 = 1$

En las personas menores de 65 años, el riesgo de IAM es igual entre los que usan y los que no usan anteojos, siendo, por lo tanto, el riesgo relativo igual a 1 (2 de 40 dividido 8 de 160):

Riesgo absoluto en los que usan anteojos	$2/40 = 0,05$
Riesgo absoluto en los que no usan anteojos	$8/160 = 0,05$
Riesgo relativo	$0,05/0,05 = 1$

Como vemos en el ejemplo anterior, la asociación entre usar anteojos (factor de riesgo) y riesgo de IAM (evento) desaparece cuando tomamos en cuenta la edad de las personas. Cuando estratificamos por el factor “edad”, en mayores o menores de 65 años, la asociación entre el uso de anteojos y el IAM desaparece.

En este ejemplo la edad está actuando como factor confundidor de la asociación entre el uso de anteojos y el IAM.

Lo que ocurre es que el factor edad está independientemente relacionado tanto con el uso de anteojos (los mayores de 65 años los usan más) como con la aparición de IAM (su riesgo aumenta con la edad), y eso confunde la asociación entre el uso de anteojos y el riesgo de IAM porque la hace aparecer como una asociación real cuando en realidad no existe.

Este es un claro ejemplo de la influencia que pueden tener los factores confundidores cuando uno no los toma en cuenta en el análisis de los datos o, lo que es peor, cuando no se los conoce. Por esta razón, cuanto menor sea la presencia o influencia de confundidores en un estudio, más creíbles y válidas serán sus conclusiones.

TIPOS DE ERRORES

Error aleatorio (random error)

Se denomina error aleatorio a la variabilidad de las mediciones o estimaciones debidas completamente al azar. No tienen direccionalidad, es decir, no modifican los resultados sistemáticamente hacia uno u otro lado. Podemos manejar el error por azar (llamado también error tipo I o α) a través del uso de la estadística definiendo cuál será el límite que toleraremos. Arbitrariamente se ha definido que, para

considerarse estadísticamente significativa, esta probabilidad deberá ser menor al 5% o, lo que es lo mismo, la famosa $p < 0,05$.

El nivel de α (alfa) o la “p” se refiere a la probabilidad de que la asociación entre dos variables o la diferencia entre las muestras encontrada en el estudio sea debida al azar. Se acepta por convención que si la probabilidad es mayor o igual a 0,05 (5%) las asociaciones o diferencias halladas en un estudio no podrán ser aceptadas como estadísticamente significativas, es decir, no se consideran verdaderas.

Error sistemático (sesgo o bias)

El error sistemático es la desviación sistemática (y por lo tanto direccional) de los resultados o estimaciones de su valor correcto o real debido a fallas del diseño (en la selección de la muestra o en los métodos de recolección de la información) o fallas del análisis (falla en el control de posibles confundidores). Los sesgos de diseño no se pueden corregir una vez cometidos (no se pueden manipular estadísticamente) mientras que los de análisis son corregibles con el uso apropiado de la estadística.

Los sesgos de información se refieren a las fuentes y formas de recolección de los datos que puedan llevar a error. Por ejemplo, la revisión de historias clínicas que pueden estar incompletas, encuestas con preguntas cuyas respuestas se remontan al pasado con la consiguiente dificultad para recordar, utilización de familiares o personas muy próximas en lugar de los participantes, etc. Por otra parte, los sesgos de selección se refieren a la metodología utilizada en la selección de los participantes del estudio.

Ejemplo: si quisiéramos estudiar la relación entre número de parejas y el riesgo de embarazo en la adolescencia, la selección de los participantes debe tener en cuenta la población de donde provienen, es decir, no podríamos elegir la muestra de una comunidad severamente religiosa porque es probable que la asociación sea sistemáticamente subestimada ya que el factor de riesgo (número de parejas) no será representativo de la población general.

Como ejemplo gráfico diremos que si tenemos un tirador experto de arco y flecha que da siempre en el blanco, las pequeñas variaciones entre tiro y tiro serían consideradas error por azar del mismo tirador. Si algún tiro fuera diferente y la probabilidad de que esta diferencia sea por azar fuera menor al 5%, la consideraríamos estadísticamente significativa y, por ende, como proveniente de otro tirador.

Si el tirador tuviera constantemente viento desde un costado (sesgo), esto podría desviar el tiro (estimación) sistemáticamente hacia el otro, obteniéndose entonces valores diferentes de los “reales” a causa de la presencia de ese factor modificador.

CUADRO 6-4-4. ASOCIACIÓN ENTRE EL RIESGO DE INFARTO AGUDO DE MIOCARDIO Y EL USO DE ANTEOJOS, ESTRATIFICADO EN FUNCIÓN DE LA EDAD DE LOS PACIENTES

	Mayores de 65 años			Menores de 65 años		
	Pacientes con IAM	Pacientes sin IAM	Total de pacientes mayores de 65 años	Pacientes con IAM	Pacientes sin IAM	Total de pacientes menores de 65 años
Pacientes que usan anteojos	32	128	160	2	38	40
Pacientes que no usan anteojos	8	32	40	8	152	160
Total de pacientes	40	160	200	10	190	200

VALIDEZ

Validez interna

La validez interna se refiere a la estructura interna del estudio. Los resultados y su interpretación son coherentes con los métodos utilizados y fueron los adecuados para responder la/s pregunta/s formulada/s.

Si el trabajo tiene validez interna, está bien hecho y, por lo tanto, carece de sesgos que, de estar presentes, puedan invalidar completamente el estudio.

Validez externa

La validez externa se refiere a si las conclusiones del estudio (asumiendo que tiene validez interna) son generalizables a otras poblaciones.

Si el trabajo tiene validez externa, significa que además de estar bien hecho, las conclusiones se pueden extrapolar a otras poblaciones, como por ejemplo, nuestra práctica clínica cotidiana.

6-5. INTRODUCCIÓN AL DISEÑO DE LOS ESTUDIOS DE INVESTIGACIÓN

FERNANDO RUBINSTEIN

ESTUDIOS DESCRIPTIVOS

Los estudios de investigación se clasifican en descriptivos y analíticos o explicativos (fig. 6-5-1).

Los estudios descriptivos solo describen características o frecuencias. No intentan la búsqueda de asociaciones causales entre variables. Estudios de este tipo son el reporte de casos, los diagnósticos más comunes, la descripción de frecuencias, etc. (p. ej., de 1236 pa-

cientes que consultaron en este período, un 23% lo hizo por ansiedad, 37% por hipertensión, etc.).

Si bien este tipo de estudio no intenta explicar las causas del fenómeno a estudiar, tiene una importancia fundamental para conocer nuestra realidad, que muchas veces difiere de la descrita en otros medios. Son el primer paso necesario para planificar cualquier estudio o intervención posterior porque nos brinda la información básica del entorno en el que desarrollamos nuestra práctica.

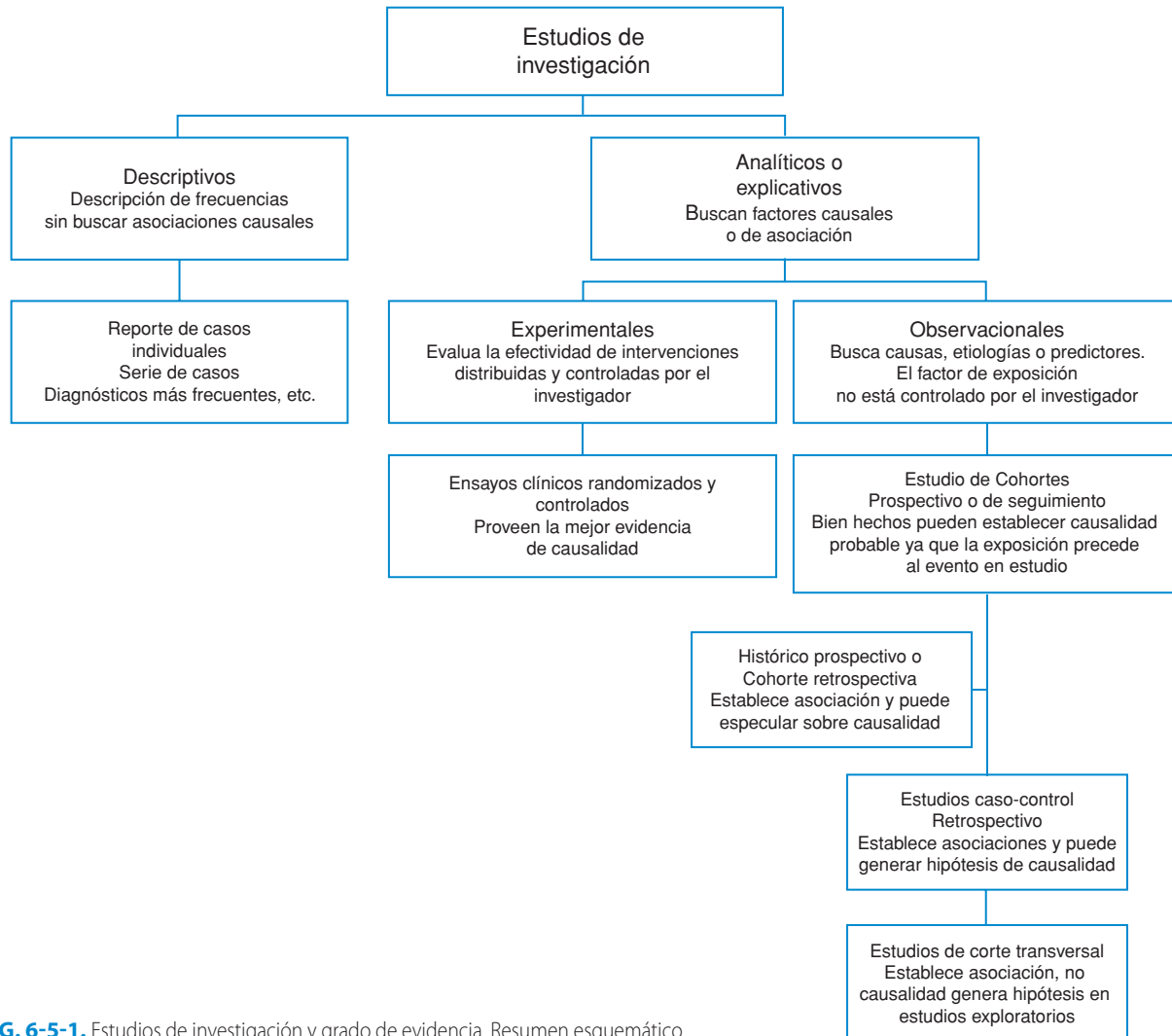


FIG. 6-5-1. Estudios de investigación y grado de evidencia. Resumen esquemático.

ESTUDIOS ANALÍTICOS O EXPLICATIVOS

Los estudios explicativos intentan explicar lo que ocurre y, por lo tanto, buscan la asociación entre las variables (en general alguna en particular) y el evento de interés. Pueden ser experimentales u observacionales.

Los distintos tipos de estudios analíticos se valen de diferentes estrategias para la selección de la muestra en estudio y la distribución del llamado factor de exposición o de riesgo. Utilizan diferentes métodos para recolectar la información y manejar las variables llamadas “confundidores” (que pueden confundir parcial o totalmente la asociación entre una variable de interés y el evento en estudio). Estas características hacen que cada uno de ellos tenga pros y contras desde el punto de vista metodológico y sus conclusiones deban ser analizadas a la luz del grado de evidencia que aportan.

Estudios experimentales

Los estudios experimentales son analíticos. En ellos el investigador decide y maneja las condiciones del estudio (selección de pacientes, distribución de la intervención a evaluar, seguimiento y manejo clínico, etc.) con el propósito de lograr una comparación no sesgada de los grupos evaluados.

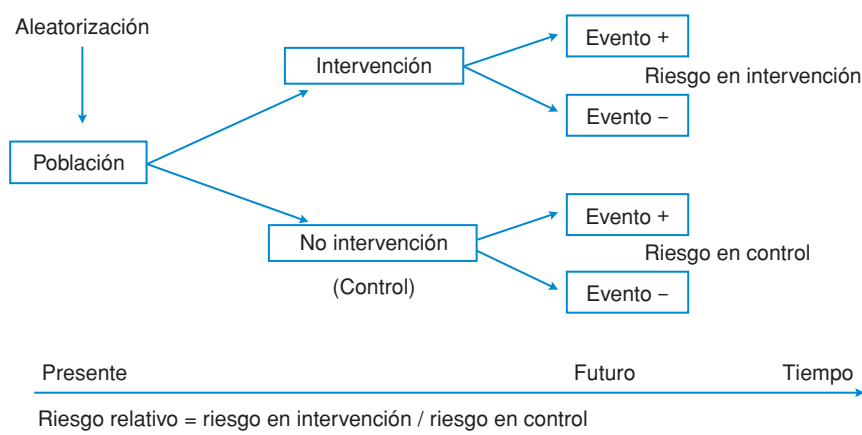
En este sentido el investigador realiza un experimento clínico ya que controla las variables que va a analizar. El grupo que recibe la intervención (droga X, programa educativo, maniobra de prevención, etc.) se llama grupo experimental y el grupo de referencia contra el cual se comparan los resultados es el grupo control. Es siempre prospectivo, es decir, toda la información se recolecta hacia adelante en el

tiempo, evitando los problemas suscitados por la búsqueda de datos o eventos que ocurrieron en el pasado. Para evitar sesgos en la evaluación idealmente será “doble ciego”: ni el participante ni quien lo evalúa sabe a qué grupo pertenece. El estudio experimental más utilizado es el ensayo clínico aleatorizado y controlado (ECRC), el cual se esquematiza en la **figura 6-5-2**.

- Randomización o aleatorización: la característica más saliente de los ECCA es, justamente, el control sobre la distribución del factor que se quiere estudiar, es decir, la intervención. Randomizar significa que todos los participantes del estudio deben tener la misma chance de estar en el grupo experimental o en el control (por ejemplo, a través de un sorteo o una lista de números elegidos completamente al azar). De esta manera se evita que la presencia de otras variables (confundidores) puedan explicar las eventuales diferencias entre los grupos ya que la exposición al factor en estudio no tiene relación con absolutamente ninguna otra variable en juego, como edad, sexo, educación, etc. Como concepto, si una randomización está bien hecha, todos los posibles confundidores (tanto los conocidos como los desconocidos) quedarán distribuidos en forma equivalente en ambos grupos a comparar y, por lo tanto, cualquier diferencia hallada será atribuida a la intervención que se desea evaluar. Por este motivo fundamental, los ECRC proveen la evidencia más fuerte de la asociación causal entre el factor de exposición y el evento en estudio.

Algunas características de los ECCA se mencionan en el **cuadro 6-5-1**. Cabe recordar que para todo ECRC hay que contar con el con-

FIG. 6-5-2. Estudios experimentales. La distribución de la investigación en estudio está en manos del investigador y, en general, es asignada en forma aleatoria.



CUADRO 6-5-1. VENTAJAS Y DESVENTAJAS DE LOS ESTUDIOS EXPERIMENTALES

VENTAJAS	DESVENTAJAS
<p>Brindan la mejor evidencia de la relación causa-efecto.</p> <p>Evitan sesgos por confundidores ya que asignan la intervención al azar.</p>	<p>Muy caros y relativamente complejos.</p> <p>Limitados por barreras éticas.</p> <p>Incluyen población autoseleccionada.</p>

Potenciales problemas de los estudios experimentales

- Población en estudio: los criterios de inclusión y exclusión deben ser claros y aceptables ya que pueden afectar la validez externa del estudio.
- Intervención: debe estar definida claramente, ser reproducible y práctica. Todos los participantes deben tener la misma chance de recibirla (aleatorización).
- Adherencia al tratamiento asignado: la diferente adherencia individual al tratamiento asignado puede aumentar o disminuir la efectividad en forma espúria.
- Pérdidas y contaminación de los grupos: el análisis debe ser adecuado al grupo asignado en la aleatorización ("intención de tratar"). Debe informar qué pasó con los que abandonaron el estudio.

sentimiento de los individuos que participan. Esto, sin dudas, produce un sesgo de selección importante que puede disminuir la validez externa del trabajo ya que no todos los individuos elegidos para el estudio van a querer ser parte y solo lo hacen quienes voluntariamente acordaron al firmar el consentimiento.

Al ser un estudio prospectivo (hacia adelante), existe el riesgo de “perder” participantes ya que algunos se arrepienten y no desean seguir porque tienen efectos adversos, otros simplemente dejan de venir a los controles o no contestan las encuestas, etc. Si esta pérdida es importante (mayor al del 30 o el 40 %), se debe tener información sobre los que abandonaron (“drop outs”), de lo contrario los resultados del estudio pueden estar fatalmente sesgados. La contaminación de los grupos ocurre cuando aquellos pacientes que fueron asignados a una determinada estrategia terminan recibiendo la del otro grupo.

El tipo de intervención debe ser claramente definida, práctica y reproducible. La eficacia se evalúa en función del impacto que produce la intervención únicamente en quienes la reciben en condiciones ideales (total adherencia y cooperación con la recomendación médica), mientras que la efectividad se establece evaluando el impacto que tiene la recomendación del tratamiento incluyendo en el análisis tanto a aquellos que lo toman como también a los que por diferentes razones no siguieron la recomendación. Esto significa que, por ejemplo, un tratamiento puede producir mejoría clínica en aquellos que lo toman y entonces será eficaz pero no será efectivo si la mayoría de los pacientes en quienes está indicado o recomendado no lo toman por razones de costo, efectos adversos, dificultades operativas, etc. Ejemplo de esto en la vida real es el DCCT (*Diabetes Control and Complications Trial*) que demostró claramente la eficacia del tratamiento insulínico intensivo para retardar y disminuir el progreso de complicaciones en diabéticos insulino dependientes. No obstante, dadas las dificultades operativas y de costo para poder recomendar globalmente varias inyecciones diarias de insulina corriente o una bomba sumado a los controles diarios con hemoglucotest hace que esta intervención sea eficaz pero poco efectiva.

A los fines de generalizar los resultados (validez externa) la evaluación de un ECCA debe basarse en la efectividad y no solo en la eficacia de la intervención.

Por último, aun cuando las dificultades operativas, económicas y éticas que caracterizan a este tipo de estudios hacen que en muchos casos no se puedan realizar, cuando tenemos que evaluar un tratamiento, no deberíamos alejarnos de los estudios experimentales, ya que representan la evidencia más contundente para establecer una asociación causal entre el factor de exposición y el evento de interés.

Estudios observacionales

Los estudios observacionales son analíticos en los cuales, a diferencia de los experimentales, el investigador se limita a observar lo que ocurre en los grupos en función de la exposición no controlada a un factor de riesgo. En rigor de verdad, la mayoría de las preguntas sobre factores de riesgo no pueden ser evaluadas con estudios experimentales aunque sean los más rigurosos desde el punto de vista científico.

Por lo tanto, los estudios observacionales son muchas veces la única alternativa factible para responder preguntas sobre riesgo en epidemiología.

Sin embargo, este tipo de estudios está sujeto a muchos potenciales sesgos ya que el hecho de que una persona esté expuesta o no a un factor de riesgo puede no ser independiente de otras características que determinen esa exposición y, si estas estuvieran a su vez relacionadas con el evento en estudio, actuarían como variables confundidoras. Por eso, además de la rigurosidad del diseño, el mayor desafío de los estudios observacionales es controlar la presencia de confundidores lo más estrictamente posible.

Estudios de cohortes o de seguimiento

Los estudios de cohorte son analíticos observacionales. Una cohorte es un grupo de personas que tienen en común alguna característica particular. En un estudio de cohortes, el investigador observa dos cohortes libres del evento de interés al comienzo del seguimiento (ej. expuestos y no expuestos a determinado factor de riesgo) a lo largo de un período y registra los eventos nuevos que ocurren en uno y otro grupo para luego poder compararlos (fig. 6-5-3).

El tiempo de seguimiento debe ser lo suficientemente largo como para que el evento pueda manifestarse; por lo tanto, no son estudios apropiados para estudiar eventos poco comunes ya que se requeriría un tiempo inaceptable de seguimiento.

Como todo estudio prospectivo, se deben extremar las medidas para obtener información completa sobre la mayoría de los participantes ya que la pérdida de datos puede modificar significativamente los resultados, especialmente cuando hay mayor pérdida en uno de los dos grupos.

Un ejemplo clásico de este tipo de diseño es el famoso estudio de Framingham que comenzó en 1949 con el objeto de identificar factores asociados al riesgo de enfermedad coronaria. El 50% de la población de entre 30 y 59 años de ese pueblo de los EE.UU. fue examinada y se determinó que 5 127 de ellos no tenían ninguna evidencia de enfermedad coronaria. Estas personas fueron examinadas cada dos años

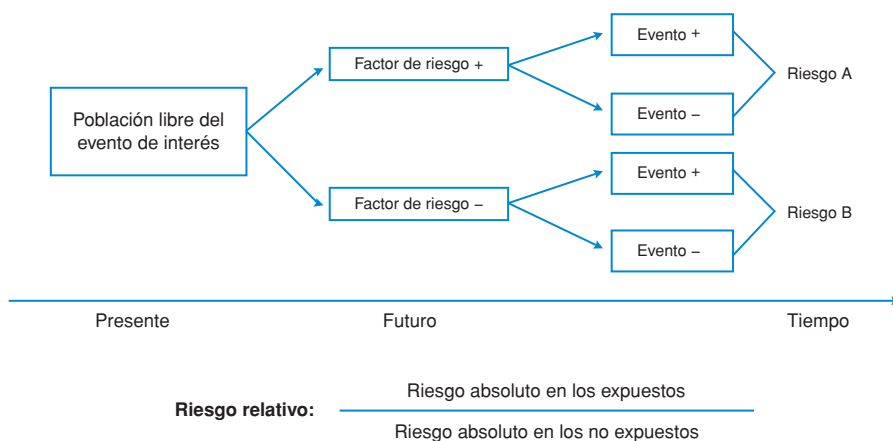


FIG. 6-5-3. Estudios de seguimiento (de cohortes o *follow up*). Estudia la ocurrencia del evento (incidencia) en función de la exposición al riesgo y el tiempo. Recolección prospectiva de la información.

CUADRO 6-5-2. VENTAJAS Y DESVENTAJAS DE LOS ESTUDIOS DE COHORTE O SEGUIMIENTO

VENTAJAS	DESVENTAJAS
Establecen una secuencia temporal y miden incidencia. Mejor control en la selección de la población en estudio. Identifican confiablemente los eventos en estudio. Pueden evaluar exposiciones poco frecuentes.	Caros y prolongados. Es difícil seguir a los grupos. Sesgo de vigilancia (búsqueda más exhaustiva del evento en los expuestos). Distribución no balanceada de variables confundidoras. No son prácticos para evaluar eventos poco prevalentes.

registrándose todos los eventos coronarios identificados a lo largo de casi 30 años. Este estudio permitió establecer que aquellos que tenían uno o más de los hoy clásicos factores de riesgo coronarios tenían un riesgo significativamente mayor de padecer eventos coronarios.

La mayoría de los grandes estudios epidemiológicos son estudios de cohortes. En el **cuadro 6-5-2** se mencionan algunas ventajas y desventajas de este diseño. Como principal virtud, los estudios de cohortes que evalúan el riesgo de contraer enfermedades en función de la exposición a un determinado factor son los únicos posibles ya que no es ético ni apropiado evaluar tales riesgos con diseños experimentales.

Si bien es cierto que solo los experimentos (ECRC) pueden establecer fehacientemente la relación causal entre un factor y el evento de interés, los estudios de cohortes bien hechos permiten una fuerte presunción de causalidad. Por ser prospectivos miden la ocurrencia de eventos nuevos en la población susceptible (incidencia) y pueden definir los criterios para identificarlos en forma confiable.

Son particularmente útiles para estudiar la exposición a factores poco comunes. Además del costo y la dificultad de su coordinación, existe un potencial sesgo particular de los estudios prospectivos que es la búsqueda más exhaustiva del evento en el grupo de los expuestos ya que se corre el riesgo de subregistrar los eventos en los no expuestos. Este sesgo de vigilancia puede ocurrir cuando los evaluadores no están ciegos con respecto al grupo al que pertenece el paciente (expuesto o no) y esto puede atentar contra los hallazgos del estudio.

Estudios de casos y controles

Los estudios de casos y controles son analíticos observacionales. Comenzaremos a explicar su diseño mediante un ejemplo.

Ejemplo: un estudio evaluó la relación existente entre el consumo de café como factor de riesgo y la ocurrencia de cáncer de páncreas (CAPN), necesitamos diseñar un estudio cuyas características permi-

tan comparar dos grupos, controlar todos los posibles confundidores y reclutar un número suficiente de participantes en cada grupo para poder compararlos estadísticamente. El CAPN no es un evento común en la población, motivo por el cual, si diseñáramos un estudio de cohortes, necesitaríamos reclutar un número inmenso de participantes y seguirlos por un lapso extremadamente prolongado para poder lograr identificar un número de eventos suficiente como para poder comparar a aquellos que consumen café (expuestos) con los no consumidores (no expuestos). Una alternativa posible entonces es diseñar un estudio que busque hacia atrás en el tiempo y compare la frecuencia del factor de riesgo consumo de café en un grupo de casos (pacientes con CAPN) y en uno de controles (sin CAPN).

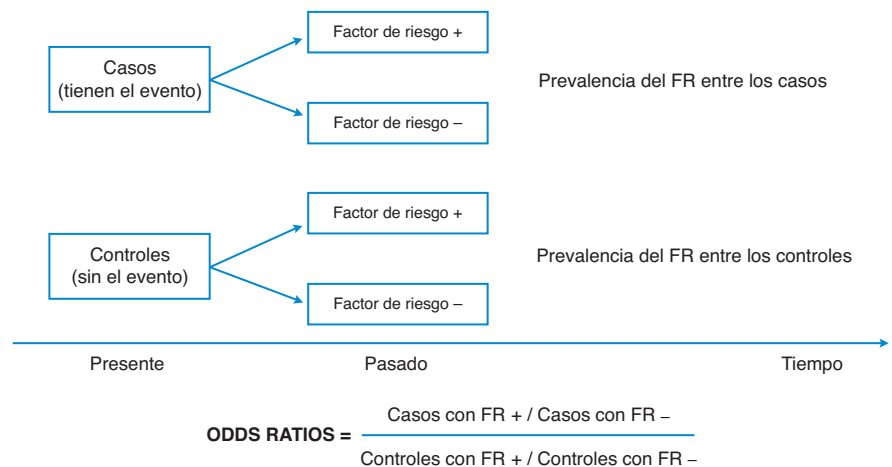
Esta es la estructura de un estudio de casos y controles (**fig. 6-5-4**)

A diferencia de los estudios de cohortes en los que el investigador selecciona la población en función de la exposición o no a un factor de riesgo determinado y evalúa la ocurrencia del evento, la selección de los grupos en los estudios de casos y controles se realiza en función de la ocurrencia del evento en función del pasado (retrospectivo) para buscar si la exposición al factor de riesgo fue diferente en uno y otro grupo. Se parte de la selección de individuos que hayan sufrido el evento de interés y un grupo control que no lo haya sufrido pero que sea comparable en todas o al menos la mayoría de las características al grupo de casos.

Como primer concepto importante, los estudios de casos y controles son especialmente útiles para indagar eventos poco comunes. Otra gran ventaja es su costo y rapidez ya que no requieren seguimiento.

Dado que evalúan retrospectivamente la distribución del factor de riesgo entre los casos y los controles, no pueden calcular el riesgo de ocurrencia del evento en el tiempo como sí pueden hacerlo los de cohortes. El odds ratio o razón de probabilidad es un cociente que estima la diferencia en la distribución del factor de riesgo entre los casos y los controles e informa cuánto más prevalente es el factor de riesgo

FIG. 6-5-4. Estudios de casos y controles. Estudian en forma retrospectiva la distribución del factor de riesgo entre casos y controles.



en una población comparada a la otra. Es una medida de asociación que, en la mayoría de los casos, puede funcionar como una aceptable aproximación a una medida de riesgo relativo.

Los problemas típicos a los que están sujetos los estudios de casos y controles (cuadro 6-5-3) son los sesgos de selección y de información y, al igual que el resto de los estudios observacionales, la presencia de factores confundidores que deben ser controlados en el análisis.

Para graficar el sesgo de selección volvamos al ejemplo del estudio de café y CAPN.

Los participantes de un grupo control adecuado deben ser lo más parecido posible a los casos excepto por el evento. Por eso los autores de este trabajo buscaron los pacientes del grupo control en el mismo lugar de donde habían reclutado los casos; los servicios de gastroenterología de 11 hospitales del área de Boston. Luego de analizar los datos, concluyeron que el consumo de café era mucho más prevalente entre los pacientes con CANP que entre los controles y la asociación era altamente significativa. Luego de revisar los criterios de selección del grupo control, se vio que los pacientes gastroenterológicos (grupo control) tenían un consumo de café muy inferior al de la población general y, por lo tanto, al compararlos con los pacientes con CAPN, estos aparecían como “grandes consumidores”. ¿Cuál había sido el problema? El grupo control no era representativo de la población general, ya que por ser pacientes con dolencias del aparato digestivo consumían poco café.

De este ejemplo se desprende que una de las condiciones fundamentales para elegir un adecuado grupo control, además de que puedan ser comparables con los casos, radica en que la distribución del factor de riesgo entre los controles sea representativa de la distribución de ese factor en la población general.

Estudios de corte transversal o de prevalencia

Los estudios de corte transversal o de prevalencia están entre los estudios analíticos observacionales más comúnmente reportados en la literatura médica. Es un diseño relativamente accesible en cuanto

al costo y no requiere seguimiento (fig. 6-5-5). Los datos sobre los factores de riesgo posibles y los eventos son recolectados al mismo tiempo y, aunque pueden establecer asociaciones entre variables, no pueden establecer ninguna presunción de causalidad. Estos estudios son muy útiles para generar hipótesis de asociaciones pero son especialmente utilizados en mayor o menor escala para determinar la prevalencia de distintos problemas en la población.

Un ejemplo de este tipo de estudio es una evaluación del número de personas que suman más de diez puntos en una escala de depresión a través de una encuesta administrada a una muestra de todos los pacientes que concurran a la consulta en el próximo mes. Esto nos dará una estimación de la prevalencia de depresión en ese grupo y además podemos recolectar información sobre otras variables de interés y relacionarlas con la depresión para explorar si existe alguna asociación. Algunas características de los estudios de corte transversal se mencionan en el cuadro 6-5-4.

Estos estudios son muy útiles para describir la realidad rápidamente y así permitir conocer y priorizar los problemas con los que lidiamos a diario en nuestros lugares de trabajo, siendo además un excelente punto de partida para generar hipótesis de trabajos futuros.

Como ocurre en otros estudios, una de las características más importantes para determinar la prevalencia de un evento es la selección apropiada de la muestra, ya que debe ser representativa de la población a la que se quiere extrapolar esos datos. Cualquier sesgo de selección limitará los resultados.

A modo de conclusión podemos decir que los estudios observacionales brindan información muy valiosa y ofrecen diferentes grados de evidencia de la asociación entre variables. De ellos, el estudio de cohortes es el que con mayor fuerza la establece y en muchos casos es considerado suficiente para explicar causalidad, especialmente en aquellos casos en que una randomización del factor de riesgo no se puede realizar por limitaciones de índole ética (p. ej., tabaco y cáncer de pulmón). Si bien es cierto que todos los estudios observacionales

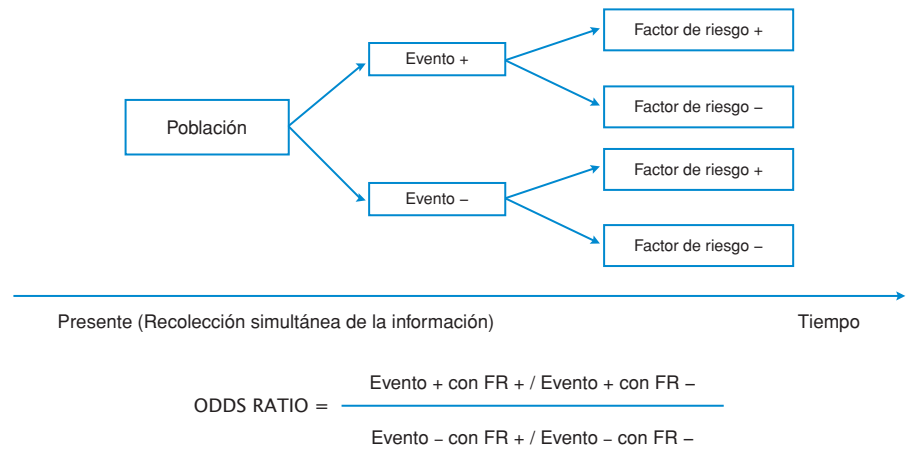
CUADRO 6-5-3. VENTAJAS Y DESVENTAJAS DE LOS ESTUDIOS DE CASO-CONTROL

VENTAJAS	DESVENTAJAS
Efficientes (baratos y rápidos). Establecen posibles causas a través de la asociación entre el factor de riesgo y el evento de interés. Muy útil para el estudio de eventos de baja prevalencia.	Susceptibles a muchos sesgos (p. ej., revisión de historias clínicas). Sesgo del recuerdo: los casos y los controles pueden recordar selectivamente determinada exposición (p. ej., las pacientes que sufrieron una trombosis venosa profunda es más probable que recuerden que habían tomado anticoagulantes que las que no la sufrieron). Problemas para la definición de los casos y selección de una adecuada población control. Distribución desequilibrada de muchas variables que actúan como confundidores.

CUADRO 6-5-4. VENTAJAS Y DESVENTAJAS DE LOS ESTUDIOS DE CORTE TRANSVERSAL

VENTAJAS	DESVENTAJAS
Determinan la prevalencia de un problema en un momento dado. Exploran la asociación de diferentes variables con un evento de interés. Generan hipótesis de asociación causal. Son rápidos y baratos.	La información sobre los factores de riesgo y los eventos es simultánea, lo que impide interpretar la relación causa-efecto. Dificultad para llegar a conclusiones generalizables si la muestra no es representativa de la población a la que se quiere trasladar las conclusiones. Susceptibilidad a todo tipo de sesgos (selección, información y confundidores) y difícil control. Sesgo de incidencia y prevalencia (p. ej., como los sujetos expuestos al factor de riesgo que sufren el evento fallecen antes de poder ser seleccionados, una proporción menor de expuestos con la enfermedad va a ser observada, subestimando la asociación).

FIG. 6-5-5. Estudios de corte transversal (*cross-sectional*). Estudian la asociación de eventos y factores de riesgo en un momento determinado o en un período muy corto (estudios de prevalencia).



pueden tener muchos sesgos, éstos pueden prevenirse en el diseño original y/o controlarse eficazmente en el análisis de los datos.

Aun el mejor estudio de investigación está lejos de ser perfecto y siempre se podrán encontrar posibles problemas en su diseño y/o análisis. El concepto fundamental es poder identificar la presencia de sesgos que nos impidan aceptar las conclusiones del estudio como válidas (falta de validez interna) o reconocer características del trabajo que limiten su generalización (falta de validez externa).

GRADO DE EVIDENCIA

Dentro del concepto de evidencia está implícito el grado de confiabilidad de la información, que depende en gran parte de los métodos utilizados para obtenerla. En medicina, la información se difunde diariamente en grandes cantidades y de nosotros depende saber evaluarla con responsabilidad para decidir si es útil, relevante y válida como para incorporar o no a nuestra práctica diaria.

Ya vimos que existen distintas fuentes de información y que cada una nos brinda distinto grado de comodidad, accesibilidad y confiabilidad. También vimos que los estudios de investigación ofrecen diferentes enfoques para contestar preguntas y que proveen distinto grado de evidencia para establecer la verdadera relación entre los factores en estudio. Es por esto que se decidió clasificar el grado de evidencia que respalda los conocimientos actuales en función de la estrategia utilizada para obtenerlos.

Calidad de la evidencia

Según del tipo de estudio del que proviene, la evidencia se clasifica en tres grados:

Buena

Se dispone de evidencia consistente obtenida de estudios bien diseñados y conducidos en poblaciones representativas con valoración directa de resultados.

Aceptable

La evidencia es suficiente pero limitada en cuanto al número, validez interna, externa o consistencia de los estudios; o bien valoración indirecta de resultados.

Pobre

La evidencia es insuficiente en cuanto al número, validez interna de los estudios, brechas en la cadena de evidencia o falta de información sobre importantes resultados en salud.

Fuerza de la recomendación según la calidad de la evidencia que la sustenta

Existe una relación entre la calidad de la evidencia, la factibilidad de llevarla adelante y el grado de recomendación de una práctica según la Fuerza de Tareas Preventivas de los EE.UU. (*US Preventive Service Task Force*) (**cuadro 6-5-5**).

En cuanto al grado de recomendación de las prácticas médicas, una gran cantidad de las que hoy se realizan rutinariamente no cuentan con la evidencia científica necesaria como para recomendarlas (grado C o I) y que muchas otras se siguen realizando aun cuando ya existe evidencia suficiente como para dejar de hacerlo (grado D).

CUADRO 6-5-5. GRADO O FUERZA DE LA RECOMENDACIÓN SEGÚN LA CALIDAD DE LA EVIDENCIA QUE LA SUSTENTA

Tipo	Dirección	Evidencia	Resultados	Beneficio neto
A	Se recomienda firmemente	Buena	Mejoran	++++
B	Se recomienda hacer	Aceptable	Mejoran	++
C	No se recomienda (opcional)	Aceptable	Pueden mejorar	0
D	Se recomienda no hacer	Aceptable	No mejoran	0 o -
I	No se puede recomendar	Insuficiente	Indeterminados	

METANÁLISIS

El metanálisis es una revisión sistemática y cuantitativa de distintos ensayos clínicos controlados cuya intención es resumir la evidencia existente sobre un tema o pregunta en particular. La calidad del metanálisis dependerá fundamentalmente de los criterios de rigurosidad científica utilizados en la selección de los estudios y el peso relativo que en función de estos criterios se le otorgue a cada uno para la estimación del resultado final promedio.

Tomar en cuenta las diferencias de diseño entre los estudios (meta-regresión) y, aunque muy poco probable, tener la posibilidad de analizar los datos individuales de los participantes de cada estudio

(meganálisis) aumenta notablemente el peso de la recomendación surgida de este tipo de análisis.

Otras técnicas de síntesis en fase de desarrollo son las síntesis de diseños cruzados que combinan los resultados de estudios experimentales con los de estudios observacionales, principalmente de bases de datos, en la evaluación de tecnología médica.

La Agencia de Evaluación de Tecnología Médica de Cataluña ha propuesto en un artículo reciente un nuevo e interesante esquema de gradación de la evidencia científica que coloca al metanálisis bien diseñado por encima de la evidencia resultante de un solo ensayo clínico controlado y randomizado.

6-6. EVALUACIÓN DE UN ARTÍCULO DE INVESTIGACIÓN

FERNANDO RUBINSTEIN

Finalmente, después de explicar brevemente algunos importantes conceptos epidemiológicos, vamos a verlos aplicados en la evaluación de los distintos tipos de trabajos de investigación que aparecen en la literatura.

Existen muchas razones para leer artículos médicos, pero la más importante es el objetivo que tenemos al comenzar la búsqueda. Es muy distinto ir “a ver qué hay de nuevo” durante algún rato libre que buscar información para responder una pregunta específica.

Por eso, es importante entender la estructura en la cual se basan los distintos trabajos y tener cierta cuota de escepticismo o al menos espíritu crítico al evaluarlo, para no ser solo pasivos consumidores sino activos interpretadores de la información.

Antes de extendernos en la descripción de la estructura básica de un artículo y de los criterios explícitos para evaluar cada tipo de artículo, queremos advertir al lector de cuatro falacias en las que puede caer si lee una publicación científica en forma desprevenida (**cuadro 6-6-1**).

ESTRUCTURA BÁSICA DE UN ARTÍCULO

Título

Debe ser interesante y útil para responder nuestra pregunta.

Resumen

Debe incluir los siguientes puntos: 1) clara formulación de la hipótesis a investigar; 2) Diseño del estudio; 3) Información sobre los métodos utilizados para recolectar la información; 4) información sobre población utilizada en el estudio; 5) resultados y la conclusión de los autores.

Los resúmenes (*abstracts*) condensan en general lo mejor del estudio y, obviamente, tratan de presentar los conceptos más destacables. No obstante, lo que debe decidirse si vale la pena seguir adelante con la lectura del artículo después de haber leído un resumen interesante es el análisis detallado del diseño del estudio (véase más adelante materiales

y métodos). También, la lectura del resumen nos va a permitir descartar rápidamente los artículos que no reúnen las características mínimas de interés y, por ende, no invertiremos más tiempo en su lectura.

Si este análisis previo es satisfactorio, el siguiente paso será evaluar la validez interna (presencia de posibles sesgos) y en qué medida nos puede ayudar en nuestra práctica (validez externa).

Introducción

En general provee información sobre el tópico y los motivos que llevan a los autores a plantear el trabajo. Muchas veces presentan controversias o una extensa revisión de la literatura, lo que la hace mucho más interesante.

Materiales y métodos

La sección “materiales y métodos”, habitualmente presentada en letra chica y tradicionalmente la parte más aburrida, es una de las secciones esenciales del trabajo ya que informa sobre la población estudiada, el diseño utilizado, los métodos de recolección de la información y las estrategias de análisis de los datos. Solo leyendo los métodos uno puede evaluar la presencia de sesgos que pueden comprometer la validez interna del estudio. Tienta no leerla, para avanzar hacia secciones más entretenidas como los resultados o la discusión. Sin embargo, a pesar de parecer una pérdida de tiempo, en realidad significa lo contrario, ya que de la apreciación crítica de esta sección puede surgir la conclusión de que no vale la pena continuar adelante con la lectura del artículo.

Resultados y discusión

Usualmente presentados como texto y gráficos o tablas, deben ser claros y coherentes. El resultado primario del trabajo tiene que surgir claramente. Es importante que el lector no se vea inmerso en múltiples análisis de subgrupos que no hayan sido planificados y presentados como hipótesis al comienzo del trabajo. La discusión debe poner los resultados del trabajo en perspectiva e integrarlos al cuerpo de conocimientos actuales sobre el tema destacando las coincidencias o diferencias existentes. En lo posible no debe ser extremadamente especulativa aunque debe tener una clara interpretación de los resultados por parte de los autores.

EVALUACIÓN DE LOS DISTINTOS TIPOS DE TRABAJOS

Se describe a continuación una serie de preguntas que debe hacerse quien se encuentre frente a un artículo científico.

Recordamos que para en los estudios de prevención o tratamiento, no nos debemos alejar de los diseños experimentales como los ECRC (**cuadro 6-6-2**). Respecto de los artículos sobre estudios diagnósticos, el **cuadro 6-6-3** muestra el ejemplo de la evaluación de la *ergometría* (PEG) en el diagnóstico de enfermedad coronaria (EC).

CUADRO 6-6-1. ERRORES EN LOS QUE SE PUEDE CAER DURANTE LA LECTURA CIENTÍFICA

- Si está publicado en la literatura, entonces es **bueno**.
- Si es estadísticamente significativo, entonces es **verdad** (o si no es estadísticamente significativo, entonces **no es verdad**).
- Si es estadísticamente significativo, entonces es **importante**.
- Si es estadísticamente significativo, entonces es **clínicamente útil**.

CUADRO 6-6-2. PREGUNTAS PARA LA EVALUACIÓN DE UN ARTÍCULO DE PREVENCIÓN O TRATAMIENTO**Validez**

- **¿Los resultados son válidos?** Este aspecto se refiere a la validez interna del estudio (si fue bien diseñado y analizado, sin sesgos identificables).
- **¿La asignación de los pacientes a uno y otro grupo fue randomizada?** Significa si todos los participantes del estudio tuvieron la misma chance de pertenecer a uno u otro grupo. Esto asegura una distribución balanceada de las poblaciones a estudiar evitando factores “confundidores”.
- **¿Se tomaron en cuenta todos los pacientes que entraron en el estudio para el análisis y las conclusiones?** ¿Cómo fue el seguimiento? (porcentaje de los participantes que tuvo seguimiento completo). ¿Qué pasó con los que se fueron del estudio? ¿Cómo se analizaron los grupos?
- **¿El estudio fue “ciego”?** ¿El personal que participó en el análisis de los resultados conocía la asignación de cada paciente? ¿Los participantes del estudio sabían a qué grupo pertenecían?
- **¿Se describen todos los resultados y efectos adversos?**
- **¿Cuál es la magnitud del efecto?** (del tratamiento o de la intervención). ¿Se presentan todos los elementos para calcular no solo el riesgo relativo (RR) y el absoluto, y la RRR, sino también el riesgo atribuible y el NNT?

Aplicabilidad

- **¿Los beneficios de la intervención superan los costos y/o los potenciales efectos adversos?**
- **¿Los pacientes son similares a los míos?** ¿Cuán diferentes a los míos son? ¿Cuáles son sus preferencias preocupaciones y expectativas?
- **¿Los resultados son aplicables en mi práctica clínica?** Estos aspectos se refieren a la validez externa del estudio para poder generalizar los resultados a mi práctica.

CUADRO 6-6-3. EVALUACIÓN DE UN ARTÍCULO DE TRATAMIENTO A TRAVÉS DEL EJEMPLO DE LA EVALUACIÓN DE LA ERGOMETRÍA (PEG) EN EL DIAGNÓSTICO DE ENFERMEDAD CORONARIA (EC)**Validez**

- **¿Existió una comparación “ciega” e independiente con un test de referencia (*gold standard*)?** Esto significa que tanto el test de referencia, en este caso coronariografía (CGF), como el que se está evaluando (PEG) fue realizado en TODOS los participantes en forma independiente sin que aquellos que los interpretaron conocieran previamente los resultados del otro test.
- **¿La muestra incluyó un espectro apropiado de pacientes en quienes uno ordenaría el test diagnóstico?** No solo aquellos muy probablemente enfermos o no enfermos, ya que la prevalencia de la enfermedad en la población utilizada modifica los valores predictivos del test (véase cap. 11, “Interpretación de los estudios diagnósticos”). Si realizamos la PEG en un grupo de hombres fumadores mayores de 60 años y otro grupo de estudiantes de educación física de 20 años la prevalencia de enfermedad va a variar significativamente y con ella los valores predictivos del test en estudio. Es de esperar que el espectro de pacientes en el estudio sea similar a los pacientes en quienes uno, en la práctica, ordenaría el test.
- **¿El resultado del test en estudio tuvo alguna influencia en la decisión de ordenar el test de referencia?** En el ejemplo, si el resultado de la PEG influyera en la decisión de indicar o no la CGF en lugar de realizar ambos estudios en forma ciega. Si se decidiera hacer CGF solo en los pacientes con PEG positiva, esto solo permitiría evaluar el valor predictivo positivo de la PEG. Al no tener los datos de la CGF de aquellos con PEG negativa, no podremos estimar la tasa de falsos y verdaderos negativos y, por lo tanto, no vamos a poder estimar la sensibilidad ni la especificidad del test. Por otra parte, la interpretación del test de referencia debe hacerse sin conocimiento del resultado del test que se está evaluando.
- **¿Fueron descriptos claramente los métodos del test como para poder replicar el estudio?**

Aplicabilidad

- **¿Se discute cómo se modifican las probabilidades posttest o se incluyen los datos para calcularla?** Coeficientes de probabilidad o sensibilidad y especificidad para calcularlos.
- **¿Los resultados van a modificar mi manejo posterior del paciente?**
- **¿Los pacientes realmente se benefician con este test?**
- **¿Los pacientes incluidos en el estudio son similares a los míos?**
- **¿Los resultados son aplicables en mi práctica?**

La mayoría de los artículos que intentan evaluar la causalidad buscan asociar una exposición (factor de riesgo o intervención) y un evento (**cuadro 6-6-4**).

Los **cuadros 6-6-5** y **6-6-6** brindan información de ayuda en la evaluación de artículos de pronóstico y revisiones.

COMENTARIO FINAL

La conducta de los médicos se basa en sus conocimientos y creencias sobre qué acciones son apropiadas para responder a un problema par-

ticular. Este conocimiento, a su vez, puede ser moldeado por factores como sus experiencias previas, el consejo de colegas respetados y lo que pueda haber leído en diversas publicaciones que proveen información que puede ser o no relevante, válida, y/o útil para responder el problema.

Los criterios explícitos de interpretación crítica de la literatura se utilizan para decidir qué tipo de información buscar y para evaluar la validez y la relevancia de lo que uno encuentra.

La evaluación de síntomas, signos o estudios de laboratorio para el estudio de un problema, la decisión de realizar maniobras de rastreo

CUADRO 6-6-4. EVALUACIÓN DE UN ARTÍCULO DE CAUSALIDAD**Validez**

- **¿Los grupos comparados son similares excepto por el factor en estudio?**
- **¿Cuál fue el diseño del trabajo?** Ensayo clínico controlado, estudio de cohortes, caso control, etc.
- **¿Se mencionan técnicas para controlar posibles “confundidores”?**
- **¿Cuán fuerte es la asociación entre la exposición y el evento y cuál es el impacto en la población?** Magnitud de la asociación: RR, RRR, RA, etc.
- **¿Las conclusiones coinciden con las de otros trabajos?** Es importante saber si en la literatura hay otros trabajos que concuerden o avalen las conclusiones del que estamos evaluando ya que esto les otorga mayor fuerza.
- **¿Existe una relación temporal?** ¿El factor causal propuesto precede al evento en estudio? Esto se asume en todo trabajo prospectivo pero no en los retrospectivos o de corte transversal. ¿Existe un efecto dosis-respuesta? Si la mayor exposición al factor causal propuesto aumenta la incidencia del evento (o disminuye si se trata de un factor “protector”). Los tres últimos puntos brindan fuerza a la posibilidad de causalidad.

CUADRO 6-6-5. EVALUACIÓN DE ARTÍCULOS DE PRONÓSTICO**Validez**

- **¿Se incluyó una muestra representativa en un momento común (usualmente precoz) de la enfermedad?**
- **¿Fue el seguimiento completo y por un período suficiente?**
- **¿Se aplicaron criterios objetivos para medir resultados de una manera ciega a la exposición?**
- **¿Se identificaron distintos subgrupos con diferente pronóstico?** ¿Se ajustó por otros factores pronósticos? ¿Se validó el resultado en otro grupo de pacientes?
- **¿Cuán probables fueron los resultados del seguimiento?** (cantidad de eventos).
- **¿Cuán precisas son las estimaciones de pronóstico?** Cuanto mayor sea el intervalo de confianza, menos precisas lo serán.

Aplicabilidad

- **¿Los pacientes del estudio eran similares a los míos?**
- **¿Es el impacto de la evidencia clínicamente importante como para influir en las conclusiones sobre qué decirle u ofrecerle al paciente?**

o la elección de una alternativa terapéutica representan conductas clínicas cotidianas que requieren decisiones racionales. Los conceptos discutidos brevemente en este capítulo intentan ser una guía que asista al médico en su tarea diaria de ver pacientes brindando el sustento epidemiológico necesario para la adecuada evaluación de la evidencia y la toma de decisiones prácticas.

De ninguna manera queremos decir que el médico debe ser esclavo de los ensayos controlados o conocer los números exactos de los

CUADRO 6-6-6. EVALUACIÓN DE ARTÍCULOS DE REVISIÓN**Validez**

- **¿La revisión es “sistemática” (metanálisis o revisiones sistemáticas sin síntesis cuantitativa de datos) o no sistemática?** En las revisiones no sistemáticas el o los autores manifiestan una opinión basada en los trabajos que citan sin que exista una estrategia explícita de selección de trabajos. Las llamadas revisiones sistemáticas deben incluir los criterios de selección de los trabajos originales y la metodología utilizada en la evaluación de cada uno de ellos.
- **¿Trata un tema en general o responde a una pregunta específica?** Las revisiones sistemáticas en general responden preguntas concretas mientras que las no sistemáticas actualizan un tema en particular.
- **¿Se utilizan criterios explícitos para la selección de los trabajos originales?** Por ejemplo, solamente ensayos clínicos controlados o estudios prospectivos, etc.
- **¿La revisión incluye todos los trabajos importantes sobre el tema en cuestión?** Esto es importante para no sesgar las recomendaciones surgidas de la revisión.
- **¿Los resultados de los estudios incluidos son similares entre sí?** ¿Se reportan los resultados de todos los estudios relevantes y la validez individual de cada uno de ellos?

Aplicabilidad

- **¿Los resultados son aplicables en mi práctica?**

riesgos involucrados en cada decisión clínica, sino que es conveniente que esté familiarizado con conceptos que le van a facilitar la tarea de discriminar qué información de la que llega a sus manos o le cuentan “de pasillo” tiene fundamento, vale la pena leer y va a ayudar a su práctica.

Lograr el balance adecuado entre la experiencia, el conocimiento y el criterio clínico es, sin lugar a dudas, el desafío más importante y a la vez más estimulante de nuestra actividad como médicos de atención primaria.

RESUMEN FINAL

- La MBE es una herramienta poderosa que permite que el médico y otros profesionales de la salud se actualicen para tomar decisiones clínicas en el cuidado de los pacientes basadas en la mejor evidencia disponible.
- Sin embargo, no todas las decisiones clínicas pueden y deben ser basadas en la evidencia.
- Resulta un desafío lograr que la conducta médica se base en la experiencia, el conocimiento, el criterio clínico y también teniendo en cuenta las preferencias individuales de los pacientes.

BIBLIOGRAFÍA

- Bligh J. Problem based, small group learning. *BMJ* 1995;311:342-343.
- Bucher HC, et al. Influence of method of reporting study results on decision of physicians to prescribe drugs to lower cholesterol concentration. *BMJ*. 1994;309:761-764.
- Ciaponni A. Guía de práctica clínica: Tercer reporte de la USPSTF. Evidencia en atención primaria. 2001;4(5):153-158.
- Davidoff F, et al. Evidence based medicine. *BMJ* 1995;310:1085-1086.
- Ely J, Osheroff J, Ebell M. Obstacles to answering doctors questions about patient care with evidence: qualitative study. *BMJ* 2002;324: 710-713.

- Ely JW, Burch RJ, Vinson DC. The information needs of family physicians: case-specific clinical questions. *J Fam Pract* 1992;35: 265-269.
- Endarterectomy for asymptomatic carotid artery stenosis. Executive Committee for the Asymptomatic Carotid Atherosclerosis Study. *JAMA* 1995;273: 1421-1428.
- Evidence-based medicine. A new approach to the teaching of medicine. Evidence-Based Medicine Working Group. *JAMA* 1992;268:2420-2425.
- Feinstein A, Horowitz R. Problems in the "Evidence" of "Evidence Based Medicine". *Am J Med* 1997;103:529-535.
- Gehlbach S. Interpreting the medical literature. Practical epidemiology for clinicians. 2nd ed. New York, NY: Mc Millan; 1988.
- Guyatt GH, et al. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1993;270:2598-2601.
- Guyatt GH, et al. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA* 1994;271:59-63.
- Hennekens C, et al. Epidemiology in medicine. 1st ed. Boston/Toronto: Little, Brown & Co. 1987.
- Jaeschke R, et al. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA* 1994;271:703-707.
- Jaeschke R, et al. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1994;271:389-39.
- Kistin C. Patient health literacy and the practice of evidence-based medicine. Editorial. *Evid Based Med ebmed-2012-100712* Published Online First: 22 May 2012.
- Jovell A, et al. Evaluación de la evidencia científica. *Med Clin (Barc)*. 1995; 105:740-743.
- Levine M, et al. Users' guides to the medical literature. IV. How to use an article about harm. Evidence-Based Medicine Working Group. *JAMA*. 1994;271:1615-1619.
- MacMahon B, et al. Coffee and cancer of the pancreas. *N Eng J Med* 1981; 304:630-633.
- McKibbin, K, Ann MLS, Richardson W. Finding answers to well-built questions. *ACP Journal Club* 1999; 4: 164-167.
- Naylor D. Grey zones of clinical practice: some limits to evidence based medicine. *Lancet* 1995;345:840-842.
- Oxman AD, et al. Users' guides to the medical literature. I. How to get started. The Evidence-Based Medicine Working Group. *JAMA* 1993;270:2093-2095.
- Oxman AD, et al. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. *JAMA* 1994;272:1367-1371.
- Packer M, et al. The effect of carvedilol on morbidity and mortality in patients with chronic heart failure. *N Eng J Med* 1996;334:1349-1355.
- Richardson WS, Wilson MC, Nishikawa J. The well built clinical question: a key to evidence-based decisions. *ACP Journal Club* 1995;123:A12.
- Rosenberg W, Donald A. Evidence based medicine: an approach to clinical-problem solving. *BMJ* 1995;310:1122-1126.
- Rubinstein F. Editorial: Evidencia en atención primaria: ¿es necesaria? Evidencia en atención primaria. 1997;0(0):1-2.
- Rubinstein F. Editorial: Medicina Basada en la Evidencia: su justa medida. Evidencia en atención primaria. 1998;1(2):2-3.
- Rubinstein F. Editorial: Práctica clínica y evidencia disponible: ¿cuál es la estrategia? Evidencia en atención primaria. 1998;1(1):1-2.
- Rubinstein F. Práctica clínica basada en la evidencia. En Rubinstein E, Zárate M, Carrete P, Deprati M, editores. PROFAM: Herramientas en Medicina Familiar y Ambulatoria. 3ra ed. Buenos Aires: Fundación MF "Para el Desarrollo de la Medicina Familiar y la Atención Primaria de la salud"; 2006. p. 97-114.
- Sackett D, et al. Clinical epidemiology: a basic science for clinical medicine. 2nd ed. Boston/Toronto: Little, Brown & Co.; 1991.
- Sackett DL, et al. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312:71-72.
- Shepherd J, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med* 1995;333:1301-1307.
- The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. *N Eng J Med* 1993;329:977-986.