

Práctica clínica basada en la evidencia

A)

INTRODUCCIÓN

En la práctica clínica diaria surgen infinidad de preguntas sobre los problemas particulares que enfrentamos cuando estamos con un paciente. ¿Cuál es el beneficio de usar betabloqueantes en este paciente con un infarto reciente? ¿Conviene iniciar tratamiento con drogas en este hombre de 50 años con 270 mg% de colesterol total y sin otros factores de riesgo? ¿Qué utilidad tiene la sangre oculta en materia fecal en la prevención secundaria del cáncer de colon? ¿Es necesario el suplemento de hierro en los lactantes sanos? ¿Debería indicarles ácido fólico a todas mis pacientes embarazadas?

La Práctica clínica basada en la evidencia es una estrategia de aprendizaje que intenta llenar la brecha existente entre la investigación y la práctica clínica corriente, brindando los elementos necesarios para poder interpretar críticamente la literatura y asegurar una buena calidad de atención médica. David L. Sackett, uno de los “fundadores” de esta disciplina que aplica directamente los conceptos de la epidemiología clínica a la práctica diaria, la describe como el proceso de convertir los problemas clínicos en preguntas, sobre la base de las cuales se busca la información, se evalúa el grado de evidencia que provee y se la utiliza como guía para una toma de decisiones adecuada.

Este capítulo no pretende ser un curso de metodología sino dar algunos elementos básicos para dedicar el poco tiempo que tenemos para la lectura a la información relevante y, de esta manera, poder traducir evidencia clínica en calidad de atención médica.

El médico de familia debe estar capacitado para interpretar críticamente la literatura médica y evaluar la evidencia disponible antes de tomar posición con respecto a una conducta determinada. Solo de esta forma podrá informar y aconsejar correctamente a sus pacientes.

B)

OBJETIVOS

- 1) Comprender la necesidad de la evaluación crítica de la literatura médica.
- 2) Adquirir conocimientos básicos de los distintos diseños de investigación clínica.
- 3) Conocer los distintos tipos de riesgo para interpretar los reportes de los estudios.
- 4) Incorporar una metodología mínima para la revisión de distintos tipos de artículos.

C)

CONTENIDOS

- 1) Fuentes de información
- 2) Estudios de investigación
- 3) Grado de evidencia
- 4) Concepto de riesgo
- 5) Evaluación de un artículo
- 6) Reflexiones sobre la práctica clínica basada en la evidencia



1

FUENTES DE INFORMACIÓN

Las preguntas que nos hacemos cotidianamente son formuladas tomando en cuenta el estado clínico del paciente (enfermo, sano, asintomático pero con factores de riesgo, etc.) y la posibilidad o no de producir mediante la acción o inacción médica un evento determinado, ya sea mejoría clínica, prevención primaria o secundaria, etc. La respuesta a estas preguntas nos conduce a decisiones que implican tratar o no tratar, con qué tratar, realizar o no un estudio diagnóstico, realizar o no una maniobra de rastreo (*screening*), realizar o no un consejo médico, etc.

Las fuentes de información que habitualmente utilizamos para responder a estas preguntas son básicamente cuatro:

a) Buscar la opinión de otro colega

La opinión de otro colega es, en general, un método rápido y conveniente. El riesgo que corremos es el de tomar una conducta sugerida por alguien a quien consideramos “experto” y que se basa exclusivamente en experiencias personales (“yo he visto que...”, “ojo, que una vez tuve un paciente que...”) que, si bien son interesantes, no dejan de ser anecdóticas y, en ocasiones, faltas de la necesaria objetividad.

b) Asistir a simposios o conferencias

Si estas opiniones son consensuadas por un grupo de colegas, en general estamos en mejores condiciones para responder a nuestra duda. Si, por ejemplo, asistimos a una conferencia sobre “Actualizaciones en el manejo de la enfermedad X”, no tenemos otra alternativa (a menos que conozcamos el tema con cierta profundidad) que confiar en la honestidad y capacidad académica del disertante para brindar una información objetiva sin sesgos de ningún tipo.

c) Leer la información de algún texto actualizado

Recurrir a un buen texto es un excelente método si está razonablemente actualizado, ya que el autor tiene que revisar exhaustivamente el tema y exponer todas las opiniones y controversias si las hubiera. Un buen texto es también una buena fuente para identificar la bibliografía original en la cual se basa.

d) Realizar una búsqueda bibliográfica

Este método es, sin lugar a dudas, el que parece más difícil (de hecho, es el más difícil), pero también es el que, a largo plazo, nos va a resultar más útil a la hora de tomar decisiones importantes para recomendar o informar a nuestros pacientes. De todas las formas de obtener la información, este método es el único que nos pone en contacto con todo o casi todo lo que hay escrito sobre el tema en cuestión, y permite que quien

realiza la búsqueda sea el que saque las conclusiones. Para realizar una búsqueda, no solo hay que familiarizarse con el modo de hacerla (aspecto que no está incluido en este capítulo) sino también con cómo interpretar lo que uno lee.

Veamos un ejemplo

Un médico residente de primer año de Medicina Familiar del Hospital Italiano vio por la mañana en el consultorio a una pareja que quería comenzar a utilizar anticonceptivos orales (ACO), pero la mujer tenía miedo de tener más riesgo de desarrollar cáncer de mama (CM) en el futuro. El residente se comprometió a asesorarlos y los citó para la semana siguiente. En el almuerzo, consulta con un amigo que se está formando en oncología clínica, quien le dice: "No les des ACO porque hay estudios que muestran un aumento del riesgo de CM". Este médico le dice eso porque acaba de leer un artículo en una revista de oncología donde alertan acerca del uso de ACO. Luego consulta con el jefe de residentes de Medicina Familiar, quien le dice: "Dales ACO y explícales que el riesgo es muy bajo". Luego, va a un seminario que organiza un laboratorio donde se promociona el nuevo ACO marca "NOOVUL" que no aumenta el riesgo de CM. Levemente confundido, decide realizar una búsqueda bibliográfica actualizada y encuentra datos nuevos recientemente publicados sobre el tema que describen un aumento leve del riesgo en un subgrupo de mujeres premenopáusicas. Finalmente, decide leer alguna información actualizada y objetiva sobre el tema y busca el capítulo de PROFAM de Anticoncepción, donde se muestra que hay mucha controversia en la literatura, que existiría un ligero riesgo y que este no parece ser de mucha importancia clínica dada la baja incidencia de la enfermedad en ese grupo etario. Con todo este bagaje de información nueva en la cabeza, debe asesorar a sus pacientes.

Como vemos, la búsqueda bibliográfica tiene la ventaja de que provee una gran cantidad de información actualizada en forma objetiva. La desventaja es que lleva tiempo encontrarla y que hay que saber interpretarla. Es absolutamente esencial tener en cuenta la responsabilidad que implica aconsejar a nuestros pacientes con datos objetivos y actualizados, sin sesgos de ningún tipo.

La evidencia de la que hablamos proviene, además del insustituible criterio clínico, de los estudios de investigación clínica que llegan a nosotros a través de las distintas publicaciones médicas. Sabemos positivamente que, para cualquier médico, hoy es imposible mantenerse al día con la información que aparece periódicamente. Por este motivo, necesitamos establecer prioridades y adquirir nociones básicas de revisión de la literatura para decidir qué vamos a leer en el escaso tiempo que tenemos y cuál es la mejor evidencia para responder a nuestras preguntas.

Un excelente recurso es recurrir a las publicaciones dedicadas a la revisión periódica de la literatura, tales como el *ACP Journal Club* (aparece cada 2 meses con el *Annals of Internal Medicine*), el *Journal Watch* o el recientemente aparecido *Evidence Based Medicine (EBM)*. Todas estas publicaciones brindan un buen resumen estándar de los trabajos, y un análisis clínico y metodológico comentado por alguno de los editores asociados. (Si Ud. está interesado en recibir el *ACP* o el *EBM*, debe solicitar información en la siguiente dirección: *American College of Physicians* Independence Mall West, Sixth Street at Race, Philadelphia, PA 19106 - 1572, E-mail: ebm@acpmhs.com).

En ese sentido, PROFAM funciona como un texto actualizado en el que la información es revisada por el autor del capítulo, quien marca las controversias existentes y refiere al lector a las publicaciones originales que sostienen los distintos puntos de vista, tratando de evitar los sesgos personales. No obstante, el texto está escrito por médicos de familia y tiene una perspectiva específica que hace que la información sea encarada y planteada desde la perspectiva de la Medicina Familiar.



Antes de explicar los diferentes estudios de investigación, es necesario familiarizarnos con los siguientes términos: confundidores, tipos de errores y validez.

Variables 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 ninguna relación 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.

Por ejemplo, si analizamos la asociación entre el uso de anteojos como factor de riesgo y el infarto agudo de miocardio (IAM) como evento en estudio:

		IAM	
		SÍ	NO
ANTEOJOS	SÍ	34	166
	NO	16	184
	50	350	400

Si hacemos las cuentas y estimamos el riesgo relativo (ver más adelante) de IAM entre los que usan anteojos y los que no, veremos que el riesgo es casi = 2 (34 de 200 dividido 16 de 200); es decir, que los que usan anteojos tienen el doble de riesgo de sufrir un IAM que los que no los usan.

¿Esto es cierto? ¿Cómo están compuestos los grupos en cuanto a edad?

		IAM	
		SÍ	NO
ANTEOJOS	SÍ	32	128
	NO	8	32
	40	160	200

> de 65 años

En los individuos mayores de 65 años, el riesgo de IAM es igual para usadores o no usadores:

Riesgo de IAM en usadores de anteojos = 32 IAM en 160 usadores = $32/160 = 0.2$ ó 20%

Riesgo de IAM en no usadores de anteojos = 8 IAM en 40 no usadores = $8/40 = 0.2$ ó 20%

Lo mismo ocurre con los menores de 65 años:

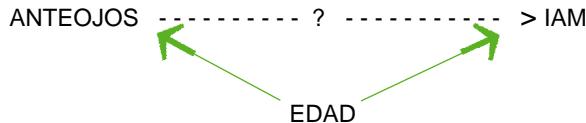
		IAM	
		SÍ	NO
ANTEOJOS	SÍ	2	38
	NO	8	152
		10	190
			200

< de 65 años

Riesgo de IAM en usadores = $2/40 = 0.05$ ó 5%, riesgo en no usadores = $8/160 = 0.05$ ó 5%

Al analizar la misma relación en cada grupo de edad, vemos que la asociación entre usar anteojos y la presencia de IAM desaparece cuando tomamos en cuenta la edad de los individuos, y que el riesgo de tener un IAM es el mismo en usadores y no usadores de anteojos en ambos grupos.

En este caso, la edad está funcionando como factor confundidor de la asociación entre el uso de anteojos y el IAM.



El factor edad está relacionado tanto con el uso de anteojos (los mayores de 65 años los usan más) como con la aparición de IAM (> riesgo de IAM según edad) independientemente, y eso confunde la asociación entre uso de anteojos y IAM, y 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 por azar (*random error*)

Se denomina **error por azar** a la variabilidad de las mediciones o estimaciones debidas completamente al azar. No tienen dirección, es decir, no modifican los resultados sistemáticamente hacia uno u otro lado. Podemos manejar el error por azar (llamado también error tipo I o alfa) a través del uso de la estadística definiendo cuál será su límite. Arbitriariamente, se ha definido que, para considerarse estadísticamente significativa, esta probabilidad deberá ser menor al 5% o, lo que es lo mismo, la famosa $p < 0.05$.

Este nivel de α (alfa) es 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 halladas en un estudio no podrán ser aceptadas como estadísticamente significativas, es decir, no se considerarán verdaderas.

Error sistemático (sesgo o *bias*)

Es la desviación sistemática (y, por lo tanto, direccional) de los resultados o estimaciones de su valor correcto o real, debida a fallas de diseño (en la selección de la muestra o en los métodos de recolección de la información) o 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, de encuestas con preguntas cuyas respuestas se remontan al pasado con la consiguiente dificultad para recordar, la utilización de familiares o personas muy próximas en lugar de los participantes, etc.

Los sesgos de selección se refieren a la metodología utilizada en la selección de los participantes del estudio. Por ejemplo, si quisieramos estudiar la relación entre número de parejas y 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: 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 costado, y se obtendrían entonces valores diferentes de los “reales” por la presencia de ese factor modificador.

Validez

Validez interna

Se refiere a la estructura interna del estudio. Los resultados y su interpretación son coherentes con los métodos utilizados y estos métodos fueron los adecuados para responder a 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, podrían invalidar completamente el estudio.

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, más allá de que esté bien hecho, las conclusiones se pueden extrapolar a otras poblaciones, por ejemplo, a nuestra práctica.

Clasificación de los estudios de investigación

Los estudios de investigación se clasifican en:

1) Descriptivos

Los estudios descriptivos solo describen características o frecuencias. No intentan la búsqueda de asociaciones causales entre variables.

Ejemplos de este tipo de estudio son el reporte de casos, los diagnósticos más comunes, la descripción de frecuencias, etc. (por ejemplo: “De 1236 pacientes que consultaron en este período, el 23% lo hizo por ansiedad, el 37% por hipertensión, etc.”).

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

2) Analíticos o explicativos

Los estudios analíticos 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.

a) Estudios experimentales (son estudios analíticos)

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

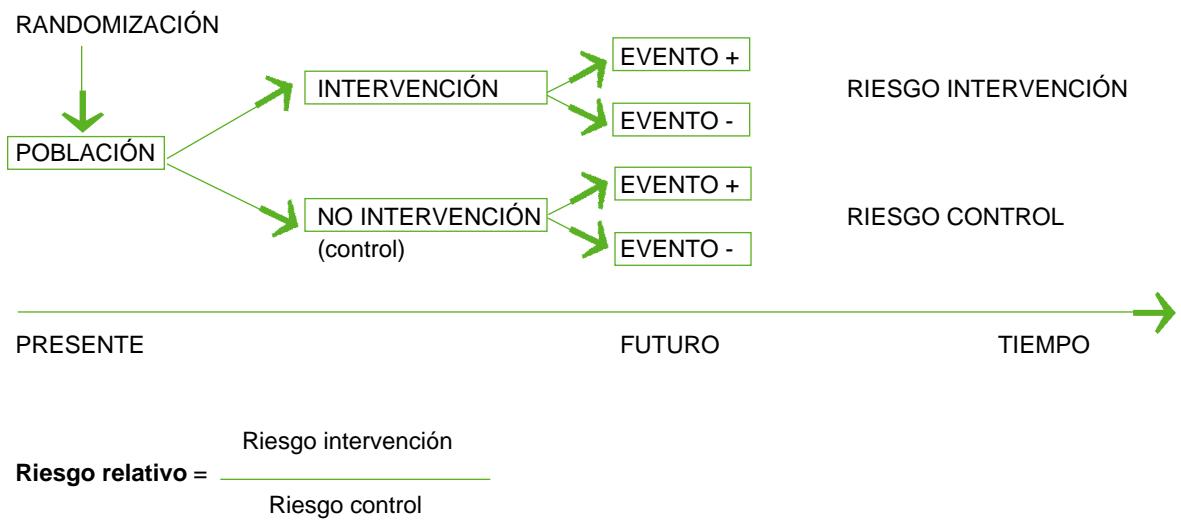
En este sentido, el investigador está realizando un **experimento clínico**, ya que controla las variables que va a analizar. El grupo que recibe la intervención (droga X, programa educativo, maniobra de prevención, etc.) se llama **grupo experimental** y el grupo de referencia contra el cual se comparan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

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”, es decir, que ni el participante ni quien lo evalúa sabe a qué grupo pertenece.

- El estudio experimental más utilizado es el ensayo clínico randomizado y controlado (ECRC).

Estudios Experimentales

La distribución de la intervención en estudio está en manos del investigador y, en general, es asignada en forma aleatoria.



La característica más saliente de los ECRC 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 posibilidad 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 los dos grupos que se comparan 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 ECRC se mencionan en la Tabla I. Hay que recordar que para todo ECRC es necesario contar con el **consentimiento** 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 formar parte y solo lo hacen quienes voluntariamente acordaron al firmar el consentimiento.

Tabla I.

Estudios experimentales

Ventajas	Desventajas
<ul style="list-style-type: none"> • Establecen la mejor evidencia de relación causa-efecto. • Evitan sesgos por confundidores. • Asignan de la intervención al azar. 	<ul style="list-style-type: none"> • En general, muy caros y relativamente complejos. • Limitados por barreras éticas. • Incluye población autoseleccionada.
Potenciales problemas de los estudios experimentales	
Población en estudio: ¿quién entra?	¿Criterios de inclusión y exclusión claros y aceptables? (Pueden afectar la validez externa del estudio.)
Intervención.	¿Definida claramente? ¿Reproducible? ¿Práctica? ¿Todos los participantes tienen la misma chance de recibirla?
Adherencia al tratamiento asignado.	Puede aumentar o disminuir la efectividad en forma espuria.
Drop outs y contaminación de los grupos.	Manejo adecuado en el análisis. ¿Intención de tratar? ¿Tratamiento recibido? Investigación de los sujetos que abandonan.

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 ir a los controles o no contestan las encuestas, etc. Si esta pérdida es importante (> del 30 ó 40%), se debe tener información sobre los que abandonaron (*drop out*); 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 para ser efectiva, aun cuando se demuestre su eficacia. 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 toma por razones de costo, efectos adversos, dificultades operativas, etc. Un 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, las dificultades de costo y operativas para poder recomendar globalmente varias inyecciones diarias de insulina corriente o una bomba, sumado a los controles diarios con hemoglucotest, hacen que esta intervención sea eficaz pero poco efectiva.

A los fines de generalizar los resultados (validez externa), la evaluación de un ECRC 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.

b) Estudios observacionales (son estudios analíticos)

En este tipo de estudios, a diferencia de los estudios 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. Si nos detenemos a pensar, la mayoría de las preguntas sobre factores de riesgo no pueden ser evaluadas con estudios experimentales, aunque estos sean más rigurosos desde el punto de vista científico.

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

Los estudios observacionales están sujetos 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.

b1) Estudios de cohortes o de seguimiento (son estudios 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 (por ejemplo, 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.

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

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. Aproximadamente el 50% de la población de entre 30 y 59 años de ese pueblo de los Estados Unidos fue examinada y se determinó que 5.127 de ellos no tenía ninguna evidencia de enfermedad coronaria. Estas personas fueron examinadas cada dos años, y se registraron 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.

Estudios de Seguimiento (cohorte o *follow up*)

Estudian la ocurrencia del evento (incidencia) en función de la exposición al riesgo y el tiempo. Recolección prospectiva de la información.



$$\text{Riesgo Relativo} = \frac{\text{Riesgo absoluto en los expuestos}}{\text{Riesgo absoluto en los no expuestos}}$$

La mayoría de los grandes estudios epidemiológicos son estudios de cohortes. En la Tabla II 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.

Tabla II.

Estudios de cohortes o *follow up*

Ventajas	Desventajas
<ul style="list-style-type: none"> • Establecen secuencia temporal. • Miden incidencia. • Estiman riesgos absolutos. • Mejor control en la selección de la población en estudio. • Identificación confiable de los eventos en estudio. • Pueden evaluar exposiciones poco frecuentes. 	<ul style="list-style-type: none"> • Insumen costo y tiempo. • Pérdida de seguimiento en 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.

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 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.

b2) Estudios caso-control (son estudios analíticos observacionales)

Si quisieramos evaluar la relación que existe entre el consumo de café como factor de riesgo y la ocurrencia de cáncer de páncreas (CAPN), necesitaríamos diseñar un estudio cuyas características permitieran comparar dos grupos, controlar todos los posibles confundidores y reclutar un número suficiente de participantes en cada grupo para poder compararlos estadísticamente.

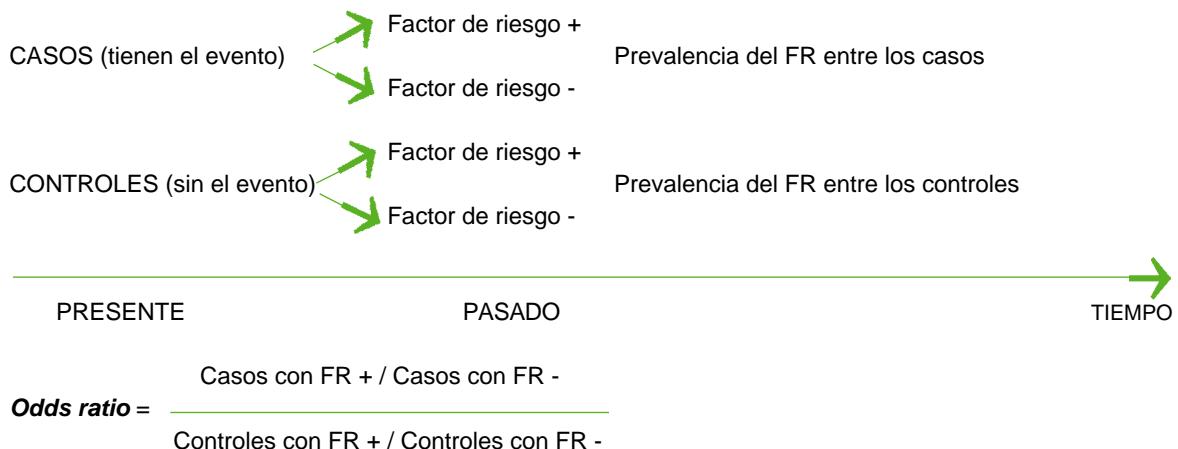
El CAPN no es un evento común en la población, motivo por el cual si diseñáramos un estudio de cohortes, necesitaríamos reclutar un número inmenso de participantes y seguirlos por un lapso extremadamente prolongado para lograr identificar un número de eventos suficiente como para poder comparar a aquellos 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 caso-control.

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 caso-control se realiza en función de la ocurrencia del evento y se va hacia el pasado (retrospectivo) a 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 de un grupo control que no lo haya sufrido, pero que sea comparable en todas o la mayoría de las características con el grupo de casos.

Estudios caso-control

Estudian en forma **retrospectiva** la distribución del factor de riesgo (FR) entre casos y controles.



Como primer concepto importante, los estudios caso-control son especialmente útiles para estudiar eventos poco comunes. Otras grandes ventajas son 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 (que ya tuvieron o no tuvieron el evento), estos estudios no pueden evaluar riesgo de ocurrencia del evento en el tiempo como sí pueden hacerlo los de cohortes. El *odds ratios* es un cociente que estima la diferencia en la distribución del factor de riesgo entre los casos y los controles, y nos informa cuánto más prevalente es

el factor de riesgo en una población comparada con 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.

Tabla III.

Estudios caso-control

Ventajas	Desventajas
<ul style="list-style-type: none">• Eficientes (costo y rapidez).• Establecen posibles causas a través de la asociación entre el FR y el evento de interés.• Muy útiles para el estudio de eventos de baja prevalencia.	<ul style="list-style-type: none">• Susceptibles a muchos sesgos.• Adecuación de la información por analizar (por ejemplo, revisión de historias clínicas).• Recuerdo selectivo y diferente a determinada exposición entre casos y controles (<i>recall bias</i>).• Problemas de definición de casos y selección de una población control adecuada.• Distribución desequilibrada de muchas variables que actúan como confundidores.

Los problemas típicos a los que están sujetos los estudios caso-control 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 parecidos posible a los casos, excepto por el evento. Por eso, los autores de este trabajo buscaron a los pacientes del grupo control en el mismo lugar en 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 claramente 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 casos (pacientes con CAPN) aparecían estos como “grandes consumidores”. ¿Cuál había sido el problema?

La condición fundamental para elegir un adecuado grupo control, además de que puedan ser comparables con los casos, es que la distribución del factor de riesgo entre los controles sea representativa de la distribución de ese factor en la población general.

b3) Estudios de corte transversal o de prevalencia (son estudios analíticos observacionales)

Están entre los estudios más comúnmente reportados en la literatura médica. Tienen un diseño relativamente accesible en cuanto al costo y que no requiere seguimiento. 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.

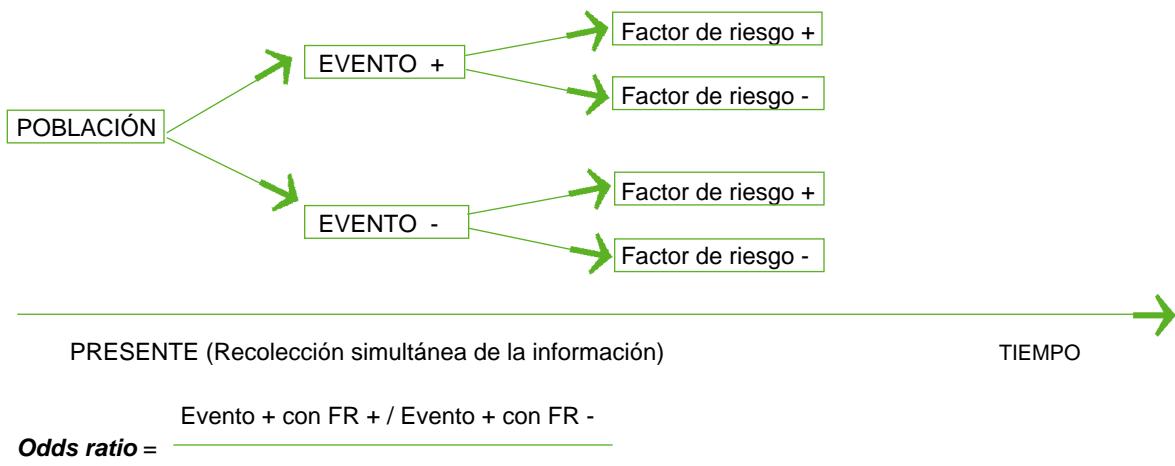
Un ejemplo de este tipo de estudio es una evaluación del número de personas que suman más de 10 puntos en una escala de depresión, a través de una encuesta administrada a una muestra de todos los pacientes que

concurran a la consulta en el mes siguiente. 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 la Tabla IV.

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).



Estos estudios son muy útiles para describir la realidad rápidamente y así permitir conocer y priorizar los problemas con los que lidiamos diariamente en nuestros lugares de trabajo. Además, son 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.

Tabla IV.

Estudios de corte transversal

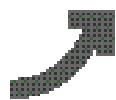
Ventajas	Desventajas
<ul style="list-style-type: none"> Determinan la prevalencia de un problema en un momento dado. Exploran la asociación de diferentes variables con un evento de interés. Generan hipótesis de asociación causal. Son rápidos y baratos. 	<ul style="list-style-type: none"> La información sobre factores de riesgo y eventos es simultánea, lo que impide interpretar la relación causa-efecto. Población en estudio. ¿Las conclusiones son generalizables? ¿De dónde se extrajo la muestra del estudio? Susceptibles a todo tipo de sesgos (selección, información y confundidores) difíciles de controlar. Sesgo de incidencia-prevalencia (por ejemplo: los sujetos expuestos al FR que sufren el evento se mueren antes de poder ser seleccionados, es decir, una proporción menor de expuestos con la enfermedad va a ser observada, subestimando la asociación).

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 muchas oportunidades, 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 (por ejemplo, tabaco y cáncer de pulmón). Si bien es cierto que todos los estudios observacionales pueden tener muchos sesgos, estos pueden ser previstos en el diseño original y/o controlados eficazmente en el análisis de los datos.

Estudios de investigación y grado de evidencia. Resumen esquemático



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).



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 responsablemente para decidir si es útil, relevante y válida como para ser incorporada 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, según la estrategia utilizada para obtenerlos.

Calidad de la evidencia

Evidencia obtenida a partir de al menos un ensayo clínico randomizado y controlado, bien diseñado.	Grado I
Evidencia obtenida a partir de ensayos controlados no randomizados, bien diseñados.	Grado II-1
Evidencia obtenida a partir de estudios de cohortes (a = prospectivos, y b = retrospectivos) o estudios caso-control (c) bien diseñados, realizados preferentemente en más de un centro o por más de un grupo de investigadores.	Grado II-2
Evidencia obtenida a partir de múltiples series comparadas en el tiempo con o sin grupo control. Incluye resultados contundentes producidos por experimentos no controlados (por ejemplo, introducción de la penicilina en los años cuarenta).	Grado II-3
Opiniones basadas en experiencias clínicas, estudios descriptivos o informes de comités de expertos.	Grado III

Grados de recomendación de la tercera edición de la Fuerza de Tareas Preventivas de los Estados Unidos (*United States Preventive Services Task Force*, o USPSTF, su sigla en inglés)

- A: Existe **buenas evidencia** científica para recomendar la práctica y la magnitud del beneficio es **sustancial**.
- B: Existe **buenas o suficientemente buenas evidencias** científicas para recomendar la práctica y la magnitud del beneficio **superan los riesgos** (beneficio sustancial o moderado).
- C: Existe **buenas o suficientemente buenas evidencias** para recomendar (o no) la práctica, pero la magnitud del beneficio es **pequeña**.
- D: Existe **buenas o suficientemente buenas evidencias** para NO recomendar la práctica y la magnitud del beneficio es **nula**.
- I: La evidencia disponible es **insuficiente** para concluir a favor o en contra de la práctica y la magnitud del beneficio no puede ser determinada.

En su tercera edición, publicada en abril de 2001, la USPSTF modificó los grados de recomendación para acercarlos más a la práctica real en el consultorio. En esta nueva edición se le ha incorporado al grado de evidencia la magnitud del beneficio de la intervención y se redefinen los grados de recomendación: se incorporó el grado I (insuficiente evidencia para recomendar o no la intervención) y se eliminó el grado E. Sin embargo, esta clasificación está en proceso de cambio y aún no se han incluido todas las patologías, encontrándose algunas de ellas con la clasificación anterior.

Relación entre el grado de recomendación de una práctica y los niveles de la calidad de la evidencia, según la USPSTF

Grado de Recomendación	Grado de Evidencia	
A	I	II-1
B	II-1	II-2
C	II-1	II-2
D	I	II-1
I		III

Veamos algunos ejemplos.

- *La recomendación de usar enalapril en pacientes con insuficiencia cardíaca es una recomendación de tipo A basada en evidencia grado I (ECRC de enalapril en mortalidad).*
- *La indicación de ecografía obstétrica de rutina en el tercer trimestre del embarazo es una recomendación grado C, ya que hay evidencia proveniente de estudios observacionales como para no recomendarla.*
- *El uso de flecainida en arritmias ventriculares postinfarto agudo de miocardio es recomendación grado D (contundente evidencia para NO usarla) basado en evidencia grado I (Estudio CAST, ECRC que demostró mayor mortalidad en el grupo que recibió la intervención).*
- *El rastreo de cáncer de piel realizando un examen corporal total de la piel es una recomendación I, ya que no está claro si es beneficioso o no para la detección temprana del melanoma, del cáncer basocelular o del epidermoide.*

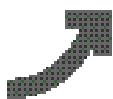
En cuanto al grado de recomendación de las prácticas médicas, es de hacer notar que una gran cantidad de las que hoy se realizan rutinariamente no cuentan con la evidencia científica necesaria como para recomendarlas (grado I) y que muchas otras se siguen realizando aun cuando ya existe evidencia suficiente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

Metaanálisis

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 metaanálisis dependerá fundamentalmente de los criterios de rigurosidad científica utilizados en la selección de los estudios analizados y del 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 estudios (metarregresión) y, aunque muy poco probable, tener la posibilidad de analizar los datos individuales de los participantes de cada estudio (megaanálisis) 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 e Investigación Médica de Cataluña (AATRM) ha propuesto un nuevo e interesante esquema de gradación de la evidencia científica, que coloca al metaanálisis bien diseñado por encima de la evidencia resultante de un solo ensayo clínico controlado y randomizado.



4

CONCEPTO DE RIESGO

El riesgo es la probabilidad de ocurrencia de un evento en una población definida a lo largo de un período determinado (por ejemplo, riesgo anual de accidente cerebrovascular en 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 los tratamientos (*cualitativa*) y cuán diferentes son al compararlos (*cuantitativa*).

Medidas de asociación comúnmente utilizadas en la literatura

Para explicar estos conceptos, utilizaremos el siguiente ejemplo obtenido de la literatura.

Riesgo de accidente cerebrovascular (ACV) en hipertensos moderados a severos seguidos durante 60 meses (Sackett, D. *Understanding clinical trials*. BMJ, 1994; vol 309:755 - 756).

Los resultados de este estudio son los siguientes:

	ACV	NO ACV	TOTAL
Tratamiento	1.800	13.200	15.000
Placebo	3.000	12.000	15.000

Riesgo Absoluto (RA): es la incidencia (casos nuevos detectados) del evento de interés en cada grupo.

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

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 desarrollaron ACV en los 60 meses siguientes si no recibían tratamiento antihipertensivo, mientras que, con tratamiento, el 12% (0.12) tuvo un ACV.

Riesgo Relativo (RR): responde 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 comparada con el grupo control?

Riesgo Absoluto del evento en los expuestos

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

Si el riesgo de sufrir el evento es igual en ambos grupos, el RR = 1 (no hay diferencia); si es mayor en el grupo expuesto, es > 1; y si es menor, es < 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 por cada 100 ACV en el control se registraron 60 en el tratado).

Veamos ahora cuánto protege el tratamiento.

Reducción del Riesgo Relativo (RRR): es una de las formas más comunes de reportar resultados de estudios terapéuticos en la literatura. Reporta el porcentaje de reducción relativa del evento en el grupo tratado.

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

En nuestro ejemplo:

La RRR = $(1 - 0.6) \times 100 = 40\%$. De esta forma, la conclusión del estudio será: “En los pacientes con hipertensión moderada a severa que recibieron tratamiento durante 5 años, se redujo en un 40% la incidencia (o el riesgo) de ACV”.

Riesgo o beneficio atribuible, Reducción del Riesgo Absoluto (RRA) o diferencia de riesgo: es una forma poco frecuente de expresar resultados en la literatura, pero de gran aplicación para la práctica clínica. Responde la pregunta siguiente: ¿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 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 = Riesgo en no tratados - 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%.

Número de pacientes necesario para tratar (NNT): es el número de pacientes que necesitamos tratar para evitar un evento. También, con el NNT podemos calcular el número de pacientes que se necesita 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?

Cuando una intervención es costosa, invasiva, está acompañada de muchos efectos colaterales que comprometen la calidad de vida, etc., 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 = 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 expuesto, sin ponernos a pensar en cuán frecuente es dicho evento. 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 2 ó 3, pero si este es poco frecuente eso puede significar un aumento de 2 c/10.000 a 4 ó 6 c/10.000, siendo 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 del 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 posibles efectos adversos. Logramos esto calculando el Riesgo Absoluto (RA) en unos y otros, su diferencia (RRA) y la recíproca de esa diferencia o Número Necesario para Tratar (NNT), elementos que nos van 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.

Riesgo de ACV en hipertensos leves seguidos durante 60 meses (Sackett, D. *Understanding clinical trials. BMJ, 1994; vol 309:755 - 756*)

	ACV	NO ACV	TOTAL
Tratamiento	135	14.865	15.000
Placebo	225	14.775	15.000

$$\text{RA en expuestos} = 135 / 15.000 = 0.009$$

$$\text{RA en no expuestos} = 225 / 15.000 = 0.015$$

$$\text{RR} = 0.009 / 0.015 = 0.6$$

$$\text{RRR} = (1 - 0.6) \times 100 = 40\%$$

$$\text{RRA} = 0.015 - 0.009 = 0.006 = 0.6\%$$

$$\text{NNT} = 1 / 0.006 = 167$$

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

Si miramos el beneficio atribuible al tratamiento en esta población, vemos que de 15 ACV que se producen cada 1.000 hipertensos leves no tratados en 5 años, 6 ACV pueden evitarse con el tratamiento. El NNT es sustancialmente mayor que en los hipertensos severos, necesitando tratar a 167 pacientes durante 5 años para evitar 1 solo ACV. Esto simplemente se explica porque el evento en estudio (ACV) es mucho más frecuente en hipertensos severos que en leves; por lo tanto, aunque las medidas relativas sean iguales, el RRA y el NNT nos dan idea del diferente impacto del tratamiento. De todas maneras, siendo la hipertensión leve un hallazgo tan frecuente en la población adulta (aproximadamente, el 25 al 30% de la población), el número de eventos evitados por el tratamiento es muy importante y justifica plenamente su indicación.

Estos conceptos son muy importantes, ya que constantemente nos tratan de “vender” descubrimientos muy “útiles”, por ejemplo diciéndonos: “Dr., recete la droga X que demostró disminuir a la mitad la incidencia del evento Y”. En este caso, nos están diciendo que la reducción del riesgo relativo (RRR) fue del 50%, pero lo que más nos debe importar 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 (IC)

El IC es una medida de dispersión o 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, para ser representativas de la población general, las muestras no deben presentar sesgos, es decir que sus características no deben ser diferentes de las de la población de la cual provienen. 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 las de la población general (parámetro P).



■ Estos 100 valores de p definen un rango. Ese rango dentro del cual se encuentra el verdadero parámetro P , 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% ($p > 0.05$), la consideraremos *no significativa* estadísticamente.

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.

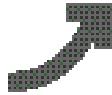
Ejemplo:

Si el RR de cáncer de mama (CM) en mujeres usadoras de estrógenos es reportado como = 1.3 IC 95% [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 este 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, si el RR de CM en las usadoras de estrógenos fuera reportado como = 1.3 IC 95% [0.80 - 1.45], esto significa que el RR (al pasar por 1) puede ser tanto mayor como menor en el 95% de las veces, por lo que las conclusiones del estudio no son válidas porque no son estadísticamente significativas.

■ Conocer el intervalo de confianza nos permite tener una idea mucho más clara de una medida de riesgo porque nos brinda la información cuantitativa de la variabilidad de esa estimación en la realidad.



Finalmente, después de ver brevemente algunos importantes conceptos epidemiológicos, vamos a verlos aplicados en la evaluación de 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 leer “para ver qué hay de nuevo” durante algún rato libre que buscar información para responder a una pregunta específica.

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

Estructura básica de un artículo

Título:

Debe ser interesante y útil para responder a nuestra pregunta.

Resumen:

Debe incluir los siguientes puntos:

- Clara formulación de la hipótesis por investigar.
- Diseño del estudio.
- Información sobre los métodos utilizados para recolectar la información.
- Información sobre la población utilizada en el estudio.
- Resultados y 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, si vale la pena seguir adelante después de leerlo, es el análisis detallado de su estructura lo que nos va a permitir sacar conclusiones válidas. 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 va a 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 su trabajo. Muchas veces, presentan controversias o una extensa revisión de la literatura, lo que la hace mucho más interesante.

Métodos:

Habitualmente presentado en letra chica, es, casi siempre, una de las partes más tediosas. Es una de las secciones esenciales del trabajo ya que informa sobre la población en estudio, 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.

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.

Existen mitos sobre la información que brindan los artículos:

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

Los datos extraídos de un artículo deben ser cuidadosamente evaluados para poder obtener conclusiones que modifiquen la práctica clínica del médico.

Evaluación de distintos tipos de trabajos: criterios explícitos

1) Evaluación de artículos sobre tratamiento o prevención

Recordar que para este tipo de estudios, que evalúan una estrategia de tratamiento o una maniobra de prevención, no nos debemos alejar de los diseños experimentales como los ECRC.

a) 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?** Alude a si todos los participantes del estudio tuvieron la misma posibilidad 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?

b) 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.

2) Evaluación de artículos sobre estudios diagnósticos

Tomaremos como ejemplo la evaluación de la ergometría (PEG) en el diagnóstico de enfermedad coronaria (EC).

a) Validez

- ¿Existió una comparación “ciega” e independiente con un test de referencia (*gold standard*)? Significa que tanto el *test* de referencia, en este caso coronariografía (CGF), como el que se está evaluando (PEG), fueron realizados 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? Es decir, no solo incluyó a aquellos muy probablemente enfermos o no enfermos, ya que la prevalencia de la enfermedad en la población utilizada modifica los valores predictivos del *test* (ver “Introducción a la epidemiología clínica”). Si realizamos la PEG en un grupo de hombres fumadores mayores de 60 años y en 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, sería 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 repetir el estudio?

b) Aplicabilidad

- ¿Se discute cómo se modifican las probabilidades *post-test* 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?

3) Evaluación de artículos sobre causalidad

En realidad, la mayoría de los artículos que intentan evaluar la causalidad lo hacen buscando asociación entre una exposición (factor de riesgo o intervención) y un evento.

a) 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? Se refiere a 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.

4) Evaluación de artículos de revisión

a) Validez

- ¿La revisión es “sistemática” (metaanálisis 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 a preguntas concretas mientras que las no sistemáticas actualizan un tema en general.
- ¿Se utilizan criterios explícitos para la selección de los trabajos originales? Por ejemplo, solamente ensayos clínicos controlados o solo 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?

b) Aplicabilidad

- ¿Los resultados son aplicables en mi práctica?

5) Evaluación de artículos de pronóstico

a) Validez

- **¿Se incluyó una muestra representativa en un momento común (usualmente precoz) de la enfermedad?** Cuanto mayor sea el tiempo de observación de los participantes, más información tendremos sobre la historia natural del evento y su relación con la exposición de interés.
- **¿Fue el seguimiento completo y por un período suficiente?** Si el seguimiento no es razonablemente suficiente, o es incompleto para un porcentaje importante de los individuos (digamos, el 20% o más), esto puede limitar el número de eventos observados y, por ende, disminuir el poder del estudio.
- **¿Se aplicaron criterios objetivos para medir resultados de una manera “ciega” a la exposición?** En algunos estudios el seguimiento de la población expuesta difiere sutilmente del de la no expuesta. Esto puede constituir un sesgo, si la búsqueda del evento de interés depende de este seguimiento “más exhaustivo”. En este caso, se cae en el sesgo de vigilancia.
- **¿Se identificaron distintos subgrupos con diferente pronóstico? ¿Se ajustó por otros factores pronósticos? ¿Se validó el resultado en otro grupo de pacientes?** Es importante que los autores describan claramente qué otros factores tomaron en cuenta en el análisis estadístico, qué información proveen otros estudios similares y si existe algún indicio que sugiera subgrupos de diferente pronóstico, cuando discuten y sustentan su hipótesis.
- **¿Cuán probables fueron los resultados del seguimiento?** ¿Cuál fue el riesgo del evento en los grupos? ¿Hubo una cantidad suficiente de eventos? De no ser así, esto nuevamente limita el poder del estudio.
- **¿Cuán precisas son las estimaciones de pronóstico?** Esto se refiere, fundamentalmente, a la longitud del intervalo de confianza, ya que cuanto mayor es este, menos precisa es la estimación (y viceversa).

b) 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?



REFLEXIONES SOBRE LA PRÁCTICA CLÍNICA BASADA EN LA EVIDENCIA

Luego de revisar las acotaciones que los alumnos de PROFAM realizaron a este texto y al de epidemiología clínica, se pone en evidencia la dificultad que existe en explicar temas que muchos médicos sienten fuera de su práctica diaria. Sin embargo, la aplicación de la Medicina Basada en la Evidencia (MBE) y de la epidemiología

clínica no está reservada solo para los médicos que se dedican a investigar, sino que forma parte de cada una de nuestras decisiones en el consultorio. Es decir, cuando se decide darle, por ejemplo, aspirina y betabloqueantes a un paciente con enfermedad coronaria (EC) es porque existe literatura médica de buena calidad que asegura que los pacientes con EC se benefician con este tratamiento. Asimismo, cuando se decide pedirle un análisis a un paciente (por ejemplo, un hemograma), esta decisión depende de la probabilidad que se le asigna a dicho paciente de padecer determinada enfermedad (por ejemplo, anemia), tomando en cuenta la signosintomatología y la prevalencia de la enfermedad.

■ La MBE y la epidemiología clínica están estrechamente vinculadas con las decisiones que toma un médico de familia en el consultorio y forman parte de la práctica médica cotidiana.

1) La necesidad de estar actualizado y la gran cantidad de información científica

Cualquiera sea la especialidad, los médicos que trabajan en el área de la atención médica primaria tienen la responsabilidad de mantenerse actualizados. Esta tarea se ha vuelto más complicada a causa de los avances científicos y los aportes que la literatura médica brinda diariamente. Hasta hace unos años, se realizaban pocos estudios de investigación clínica, con lo que era más fácil “estar al día”. 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. Sin embargo, la mayoría de las veces no 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 de un paciente.

■ Para cualquier médico de nuestro tiempo es absolutamente imposible mantenerse al día con la información que aparece periódicamente, ya que sería necesario leer alrededor de 20 artículos diarios, los 365 días del año (Sackett).

2) La importancia de priorizar la información de buena calidad

La permanente presión que se ejerce sobre los investigadores de los centros académicos para publicar más lleva frecuentemente a la publicación de un sinnúmero de estudios de calidad objetable. Por este motivo, es necesario establecer prioridades y adquirir nociones básicas de revisión de la literatura para priorizar qué es lo que va a leer el médico en el escaso tiempo que tiene.

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

■ En un gran número ocasiones, las conductas y decisiones médicas están basadas en conocimientos incorporados durante la formación médica (que ya no están actualizados) o en la sobreinterpretación de experiencias sobre pacientes individuales, que pueden carecer de la necesaria objetividad y representatividad.

3) ¿Cómo aplicar la evidencia disponible en la práctica clínica cotidiana?

La mejor evidencia disponible es 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 resolución de problemas (aprendizaje basado en el problema). En nuestro medio, 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. Sus etapas son:

a) **La formulación de la pregunta:** del contacto con los pacientes surgen preguntas. Sin embargo, es fundamental organizar estas preguntas de tal manera que puedan ser respondidas usando la estrategia de resolución de problemas. Así, en lugar de preguntarse si valdrá la pena agregar anticoagulantes a un paciente con una miocardiopatía dilatada, se debería preguntar cuál es el riesgo de eventos embólicos en un grupo de pacientes similares a este (pronóstico), cuánto se modifica con el tratamiento y si el beneficio de esa intervención supera el riesgo de complicaciones (terapéutica).

La mayor dificultad para formular una buena pregunta radica en la habilidad o capacidad para ser lo más específico posible (formular preguntas precisas y puntuales). Para ello, hay que definir claramente el tipo de paciente, el problema, la intervención y el resultado por evaluar.

b) **La búsqueda de la información:** una vez esbozada la o las preguntas, se deben buscar las respuestas. Más allá de las fuentes habituales de información a las que se recurre diariamente (otros colegas, textos, publicaciones médicas, etc.) existen grandes bases de datos electrónicas, que se van haciendo cada vez más accesibles a través de bibliotecas, colegios médicos, instituciones o la Internet. Una de las bases de datos más conocida es el MEDLINE. Esta permite 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 sensiblemente el número de trabajos a revisar). Estas estrategias de búsqueda pueden ser fácilmente aprendidas e incorporadas para ser utilizadas en forma rutinaria.

c) **La revisión crítica de la literatura:** cuanto más clara y específica sea la pregunta, menor será la cantidad de estudios que se debe evaluar para responderla. 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.

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 la práctica clínica diaria del médico de familia.

No es necesario ser epidemiólogo o tener experiencia en investigación para poder realizar una excelente revisión. La estrategia consiste en responder una serie de preguntas estructuradas (guías) sobre los distintos tipos de estudios (de diagnóstico, pronóstico, tratamiento, calidad de atención médica, análisis económico, revisiones sistemáticas, guías de práctica clínica) 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 de las guías 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 son objeto de continua revisión y refinamiento (ver más atrás). Obviamente, la experiencia se va adquiriendo a medida que se realiza este proceso una y otra vez. Una vez aprendidos los pasos, la mayoría de los trabajos se podrán evaluar rápidamente.

d) **La evidencia obtenida:** una vez identificada la información como válida, relevante y aplicable, esta podrá ser utilizada directamente en la toma de decisiones sobre un paciente. Además, podrá ser tomada en cuenta para desarrollar guías de práctica, para enriquecer la discusión sobre pacientes individuales, para fundamentar las opiniones a la hora de decidir políticas de salud en una organización y para implementar programas de control de calidad.

En resumen, para poder resolver los problemas que presentan los pacientes todos los días en el consultorio, los médicos disponen de fuentes de información. Sin embargo, esta información puede llegar a desorientar al médico (ya sea por no encontrar una respuesta adecuada al paciente que tiene en el consultorio o por llevarlo a obtener conclusiones incorrectas). Es por esto que el médico de familia debe apuntar a buscar evidencia confiable realizando una pregunta puntual (o específica) y evaluando la validez y aplicabilidad de la información obtenida. Solo así el médico podrá utilizar la evidencia como punto de partida para su aprendizaje y resolución de problemas cotidianos.

Críticas a la Medicina Basada en la Evidencia

Muchas críticas se le realizan a la MBE. Se la tilda de ser una innovación peligrosa y se discute sobre los aspectos controvertidos de su aplicación. Alvan Feinstein, uno de los más destacados y polémicos epidemiólogos del mundo, resalta los aspectos controvertidos de su aplicación y discute la autoridad de las fuentes de información consideradas en la toma de decisiones, especialmente la referida a la calidad de los datos utilizados en los estudios, el tipo de análisis, el tipo de pacientes que participa y las distintas conclusiones que se extraen a partir de sus resultados. Expresa cierta preocupación al referirse a las fuentes de información que alimentan al “nuevo texto electrónico” de la práctica clínica y que algunos defensores de la MBE consideran el *gold standard*, como los ensayos clínicos randomizados y controlados y los metaanálisis, y alerta sobre el peligro de una nueva forma de autoritarismo dogmático en el área del conocimiento científico, originado esta vez no en la Roma galénica sino en Oxford.

Las distintas limitaciones derivadas de la metodología y del diseño de los estudios de investigación podría afectar la práctica clínica cotidiana basada en la experiencia, en el paciente individual y en los mecanismos fisiopatológicos.

Dificultades para aplicar la evidencia disponible en la práctica cotidiana

El diseño metodológico de los estudios muchas veces impone estrictas reglas que terminan limitando, en pos 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. Estas dificultades y la necesidad de adaptar los diseños para que incluyan (y no ignoren) la realidad clínica cotidiana, y así obtener “Evidencia Basada en la Medicina”, aparecen como pre-requisitos para la MBE.

Indudablemente, diseñar estudios que incluyan todas las variables por considerar en el cuidado de pacientes individuales sería ideal desde el punto de vista epidemiológico, pero es imposible desde el punto de vista práctico, y el balance entre la validez interna y externa puede inclinarse peligrosamente hacia la generalización de resultados obtenidos a través de métodos no sólidamente fundados.

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 randomizado y controlado (ECRC) es el paradigma dentro de los distintos diseños, ya que es el único que garantiza la distribución balanceada de los factores confundidores, tanto de los conocidos como de los desconocidos, entre los distintos grupos a través de la asignación aleatoria de la intervención.

El ECRC es el mejor diseño para establecer una relación causal inequívoca entre la intervención y el evento (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 (ver ventajas y desventajas de los estudios experimentales).

No todas las decisiones clínicas pueden estar basadas en la evidencia

La práctica clínica plantea innumerables situaciones en las que se toman decisiones, siendo la evidencia disponible para sustentarlas controvertida, o simplemente inexistente. Como sostiene Naylor, “estas costumbres se van transformando y adquirimos tanta confianza en nuestras conductas que llegamos al punto en que es fácil confundir la opinión personal con la evidencia y nuestra ignorancia con genuina incertidumbre científica”. 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 ECRC para la mayoría de nuestras decisiones cotidianas.

Los estudios de investigación observacionales o los ECRC no están diseñados para responder preguntas sobre el tratamiento de pacientes individuales. 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 porque 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 le permite al médico actualizar su conocimiento y trasladarlo a la práctica. También permite conocer la indicación o contraindicación de distintos tratamientos o intervenciones diagnósticas ante determinadas condiciones. 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. Estos principios deben integrarse con la propia experiencia, el conocimiento de la realidad sanitaria y la relación médico-paciente.

Es absolutamente cierto que la evidencia disponible puede no adaptarse adecuadamente a las necesidades con los pacientes individuales, ya que muchas de las condiciones presentes en los estudios de investigación no se pueden repetir 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.

Evidencia Orientada a los Pacientes

(adaptado de la revista *Evidencia en Atención Primaria* por la Dra. Karin Kopitowski)

En diferentes estudios se observó que en cada encuentro entre un médico y un paciente se generan preguntas o dudas. Generalmente, alguna de ellas puede ser respondida luego de una adecuada búsqueda y una evaluación crítica de la literatura. Como hemos dicho anteriormente, a la mayoría de los médicos nos falta tiempo para encontrar información adecuada, necesaria para resolver los problemas de los pacientes y, muchas veces, nos falta el conocimiento acerca de cómo encontrar las respuestas a las dudas que generan los pacientes.

La Evidencia Orientada a los Pacientes (EOPs) es una herramienta práctica para aprender a evaluar críticamente la literatura a partir de las dudas que generan los pacientes. Estas dudas son generadas por los problemas “reales” que presentan los pacientes.

Comenzaremos con un ejemplo y luego discutiremos su utilidad educativa y sus limitaciones clínicas.

Creando una EOPs: el escenario educacional

Durante un pase de médicos en el servicio de Medicina Familiar, un residente de segundo año comenta el caso de una paciente de 42 años. La paciente refiere episodios compatibles con migraña en los días previos a la menstruación y durante ella. Los síntomas se le han agravado en el último año, llegando a tener migraña en todas sus menstruaciones. La respuesta a los antiinflamatorios no esteroides (AINE) es variable. El residente se pregunta si existen estrategias de profilaxis especiales en este tipo de migraña con una aparición tan predecible como es la migraña asociada a la menstruación o catamenial.

El jefe de residentes le solicitó al residente que había atendido a la paciente que hiciera una búsqueda para contestar a la pregunta o duda que había surgido: en mujeres con migraña asociada a la menstruación, ¿existen intervenciones farmacológicas profilácticas que reduzcan la duración/intensidad de la cefalea? Para contestar dicha duda, el residente realizó una búsqueda en PubMed.

■ La situación que se acaba de describir es cada vez más frecuente en los diversos ámbitos educativos.

Veamos a continuación qué pasos se deben seguir para construir una EOPs:

- a) Escribir el caso clínico.
- b) Formular las preguntas que generó el paciente.
- c) Describir los términos empleados para realizar la búsqueda bibliográfica.
- d) Destacar los aspectos de relevancia incluyendo la respuesta a la pregunta.
- e) Resumir los hallazgos de los estudios que respondieron a la pregunta.
- f) Citar las referencias de los estudios.

EOPs de profilaxis para la migraña asociada a la menstruación

a) Escribir el caso clínico

Una mujer de 42 años consulta por episodios compatibles con migraña en los días previos a la menstruación y durante ella. Los síntomas han “ido y venido” por años pero, en el último año, ha tenido migraña en todas sus menstruaciones. La respuesta a los antiinflamatorios no esteroides (AINE) que tomó es variable (naproxeno, ibuprofeno, diclofenac). El síntoma llega a incapacitarla.

b) Formular las preguntas que generó la paciente

En mujeres con migraña asociada a la menstruación (población), ¿existen intervenciones farmacológicas profilácticas (intervención) que reduzcan la duración/intensidad de la cefalea (resultado)?

c) Describir los términos empleados para realizar la búsqueda bibliográfica

Se realizó una búsqueda en PubMed empleando como palabras clave “migraine AND menstrual AND drug therapy AND prophylaxis”. Se limitó la búsqueda a estudios de intervención empleando los filtros de terapéutica (*clinical queries*). La búsqueda arrojó 22 citas, algunas de las cuales lucían interesantes para responder a la pregunta.

**d) Destacar los aspectos de relevancia incluyendo la respuesta a la pregunta
(en este caso acerca de la migraña asociada a la menstruación)**

La migraña asociada a la menstruación o catamenial se refiere a migrañas que ocurren alrededor de la fecha de menstruación, clásicamente dos días antes y hasta tres días después del comienzo del sangrado. A pesar de que las mujeres con este cuadro pueden presentar migrañas con la ovulación o en otros momentos del ciclo, estas cefaleas suelen ser menos severas, menos duraderas y menos resistentes a los tratamientos que los ataques perimenstruales.

El dolor, en general, es percibido más intensamente alrededor de la menstruación. Esto refleja, probablemente, la disminución del efecto opioide endógeno de los estrógenos. Antes de la pubertad, la migraña ocurre en el 2.5 al 4% de los niños con mayor prevalencia en varones que en mujeres. Luego de la menarca, la migraña es mucho más prevalente en el sexo femenino. El pico de prevalencia es en la cuarta década de la vida. Para esta época, se estima que el 41% de las mujeres ha experimentado un episodio de migraña.

Luego de la menopausia, la prevalencia cae probablemente por la estabilidad hormonal. La menstruación debe ser uno de los gatillos más potentes en la migraña. Hasta un 70% de las mujeres migrañosas reconocen la asociación entre menstruación y sus crisis. Por lo tanto este tipo de migraña puede considerarse una de las condiciones incapacitantes más frecuentes en la edad reproductiva. Las manifestaciones clínicas son las que se observan en todas las migrañas. En general no se presenta con aura, aun en mujeres con migraña con aura en otro momento del ciclo. Fisiopatológicamente, algunos tratamientos propuestos se basan en las siguientes premisas:

- 1) Este tipo de migraña es precipitado por la caída del nivel de estrógenos.
- 2) La progesterona no juega ningún rol en este cuadro clínico.
- 3) Si se atenúa la disminución en la concentración de estrógenos, se podría prevenir la migraña.
- 4) Si se aumenta el déficit estrogénico, la migraña empeoraría.

El tratamiento abortivo en la crisis aguda es el mismo que para cualquier ataque de migraña. Si los resultados son poco satisfactorios, se pueden indicar estrategias preventivas con el objetivo de disminuir la duración, la frecuencia, los síntomas acompañantes y la discapacidad. La búsqueda realizada pretende responder la pregunta acerca de la eficacia/modalidad de intervenciones farmacológicas preventivas.

e) Resumir los hallazgos de los estudios que respondieron a la pregunta

Newman, LC, Lipton, RB, Lay, CL, Solomon, S. *A pilot study of oral sumatriptan as intermittent prophylaxis of menstruation-related migraine*. Neurology 1998; 51:307.

Este fue un estudio piloto abierto y realizado en 20 mujeres que recibieron sumatriptan y se comparó la severidad de la migraña con los síntomas basales. El diseño se considera no válido para responder a la pregunta.

Newman, LC, Mannix, LK, Sandy, S, et al. *Naratriptan as short-term prophylaxis of menstrually-associated migraine: A randomized, double-blind, placebo-controlled study* Headache 2001; 41:248.

El objetivo de este estudio fue evaluar la eficacia del naratriptan 1mg y 2.5 mg dos veces por día comparados con placebo en la profilaxis de la migraña catamenial. Se realizó un ensayo aleatorizado doble ciego de tres ramas. Se enrolaron mujeres de más de 18 años de edad con historia de migraña con o sin aura como ha sido definida por la Sociedad Internacional de Cefalea, de al menos 6 meses de duración. Se emplearon dos dosis diferentes de naratriptan o placebo dos veces por día durante 5 días, empezando dos días antes de la menstruación, durante 4 ciclos. Los puntos finales fueron migrañas asociadas con menstruación, total de días con migraña, severidad de la cefalea, tiempo perdido en el trabajo y calidad de vida. En el análisis por intención de tratar,* se evaluaron 206 mujeres (70 recibieron 1mg de naratriptan, 70 recibieron 2.5mg y 66 recibieron placebo); 171 mujeres fueron tratadas por los cuatro períodos. Hubo una diferencia significativa en el número de períodos libres de migraña en el grupo 1mg comparado con placebo (50% vs. 25%). El grupo de 1mg de naratriptan redujo significativamente el número de migrañas asociadas con la menstruación (2.0 vs. 4.0, P < 0.05) y días con dolor de cabeza asociado a la menstruación (4.2 vs. 7.0, P < 0.01). Más pacientes tratadas con naratriptan 1mg estuvieron libres de síntomas en todos los ciclos que recibieron tratamiento (23% vs. 8%). El naratriptan en dosis mayores no fue superior. Los autores concluyen que el naratriptan en dosis de 1mg dos veces por día sirve para profilaxis de migraña catamenial dado “alrededor” de la menstruación.

Goldstein, J, Ninan, M, Rapoport, AM. *The safety and tolerability of frovatriptan, a selective 5HT1B/1D agonist, used for prophylaxis of menstrually associated migraine* Neurology 2003; 60 (Suppl 1): A494.

El frovatriptan es un agonista del receptor serotoninérgico 5HT1B/1D. Tiene una larga vida media, lo que lo ha hecho candidato para profilaxis intermitente en la migraña catamenial. Se llevó a cabo un estudio aleatorizado doble ciego con un diseño “cruzado” (del inglés *crossover*). Las pacientes fueron tratadas en cada uno de 3 ciclos con frovatriptan 2.5 mg o placebo dos veces por día. El tratamiento duraba 6 días comenzando dos días antes de la menstruación. El resultado primario fue la incidencia de migraña catamenial durante los 6 días de tratamiento. 546 mujeres (edad promedio 37 años) participaron. La incidencia de migraña durante los 6 días fue de 67% para el grupo placebo, 52% para la dosis más baja y 42% para la más alta. Ambos regímenes fueron superiores al placebo. El régimen con mayor dosis fue superior al de menor dosis. Ambos regímenes disminuyeron la duración e intensidad de los síntomas y el uso de medicación de rescate. La conclusión de los autores fue que el frovatriptan dado preventivamente durante 6 días fue efectivo para disminuir la incidencia de la migraña catamenial. Más de la mitad de las pacientes que emplearon frovatriptan 2.5 dos veces por día no tuvieron migraña en esos días.

Comentario

Pareciera que una estrategia preventiva en el período perimenstrual sería útil para las mujeres con migraña catamenial que aparece con un ritmo predecible. Esta estrategia se ha llamado “mini profilaxis” y consiste en la administración de medicación desde 2 días antes y hasta 3-4 días después de la menstruación. Los trabajos encontrados tienen algunas limitaciones, especialmente en el reporte de los resultados que no incluyen los intervalos de confianza, lo cual impide conocer la precisión de la estimación de beneficio. En el trabajo de naratriptan los autores concluyen que no hay dosis respuesta y, dado el número de participantes, puede haber ocurrido que no hayan tenido poder para encontrar diferencias de eficacia entre dosis.

Si bien no se ha resumido, un pequeño ensayo ha mostrado la eficacia de naproxeno en dosis de 550mg dos veces por día administrado durante el período perimenstrual 6. Aunque el trabajo fue muy pequeño y pasible de numerosos sesgos, tal vez valdría la pena intentar mini profilaxis con antiinflamatorios no esteroides debido al relativo bajo costo y a su probada seguridad.

En pacientes en las cuales los antiinflamatorios fracasan o no son tolerados, puede emplearse naratriptan (o algún otro “triptan”) si se asume efecto de clase. El alto costo puede ser una limitación para ofrecer este tratamiento en nuestro medio.

Se debe aclarar que también se proponen tratamientos hormonales: ciclos largos de anticonceptivos (por ejemplo, 4 meses ininterrumpidos) o suplemento de estrógenos solos en el momento premenstrual. Estos tratamientos no han sido formalmente probados en ensayos clínicos y la recomendación surge de estudios observacionales que relacionan el nivel de estrógenos endógenos/síntomas y de extrapolaciones fisiopatológicas.

f) Citar las referencias bibliográficas

1. Bille, BS. Migraine in school children. *A study of the incidence and short-term prognosis, and a clinical, psychological and electroencephalographic comparison between children with migraine and matched controls* Acta Paediatr 1962; 51(Supp 136):1.
2. Launer, LJ, Terwindt, GM, Ferrari, MD. *The prevalence and characteristics of migraine in a population-based cohort: the GEM study*. Neurology 1999; 53:537.
3. Neri, I, Granella, F, Nappi, R, et al. *Characteristics of headache at menopause: a clinico-epidemiologic study*. Maturitas 1993; 17:31.
4. Johannes, CB, Linet, MS, Stewart, WF, et al. *Relationship of headache to phase of the menstrual cycle among young women: a daily diary study* Neurology 1995; 45:1076.
5. Calhoun, A. Four hypotheses for understanding menstrual migraine. The Female Patient 2004; 29:38.
6. Sances, G, Martignoni, E, Fioroni, L, et al. *Naproxen sodium in menstrual migraine prophylaxis: a double-blind placebo controlled study* Headache 1990; 30:705.
7. Sulak, PJ, Scow, RD, Preece, C, Riggs MW, Kuehl TJ. *Hormone withdrawal symptoms in oral contraceptive users* Obstet Gynecol 2000; 95:261.
8. Calhoun, AH. *A novel specific prophylaxis for menstrual-associated migraine* South Med J 2004; 97:819.

Valor educativo y clínico de las EOPs

La creación de la EOPs sirve para desarrollar la habilidad de revisar críticamente la literatura y mejorar las destrezas como médicos “al lado del paciente”. Dado que están basadas en el paciente, las EOPs interesan a los estudiantes y médicos en todos los niveles. Además, al estar basadas en la evidencia, promueven la adquisición y el desarrollo de habilidades para realizar búsquedas de información y analizarla críticamente, así como para aplicar la evidencia en un paciente individual.

Las EOPs no se limitan a tópicos de tratamiento. Pueden referirse a tests diagnósticos, utilidad del examen físico, etc. Este tipo de EOPs ponen en relieve conceptos tales como sensibilidad, especificidad, *gold standard*, acuerdo interobservador, etc. Pueden confeccionarse EOPs acerca del tratamiento de alguna enfermedad. En este caso, se ponen en relieve conceptos tales como estudios aleatorizados, eficacia, etc. Como herramienta educativa, las EOPs pueden enriquecer numerosas actividades (ateneos, pases, recorridas de sala). Pueden emplearse como material en diversos ámbitos de discusión clínica y son un buen punto de partida para ateneos, en los cuales puede discutirse la relevancia clínica y aspectos relacionados con el diseño, los resultados de los artículos, etc.

La difusión de las EOPs está motivada en el deseo de mejorar las habilidades para revisar críticamente la literatura y aprender Medicina Basada en la Evidencia. Como herramienta clínica, las EOPs tienen ciertas limitaciones. Dado que son concebidas en el ámbito asistencial (muchas veces sobrecargado), se basan en búsquedas rápidas para conseguir al menos un artículo útil y muchas veces se desconocen todos los artículos potencialmente útiles. Ciertos recursos pueden disminuir estas limitaciones: *American College of Physicians Journal Club*, revisiones sistemáticas como la Cochrane Colaboration, todos ellos disponibles solo en idioma inglés. La revista *Evidencia en Atención Primaria*, publicada en la Argentina, se está convirtiendo en otro recurso para la confección de EOPs. Citar el artículo o los artículos en los cuales se basó la EOPs alerta a los lectores subsiguientes, quienes pueden analizar críticamente la evidencia y corregir la EOPs o mejorarla.

Las EOPs son una aproximación para ayudar a los médicos a formular preguntas clínicas, buscar la evidencia relevante, analizarla críticamente, organizar la información, resumirla y practicar la Medicina Basada en la Evidencia. Las EOPs deben usarse como punto de partida y no como la “última palabra”, si se presentara un nuevo caso similar. La “nueva versión” de la EOP se genera más fácilmente y es más probable que sea útil “al lado del paciente”.

Autora del EOPs
Dra. Karin Kopitowski

D)

RESUMEN FINAL

La conducta de los médicos se basa en sus conocimientos y creencias sobre qué acciones son apropiadas para responder a un problema particular. 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. Estas fuentes 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** son utilizados para decidir *qué tipo* de información buscar y para evaluar la *validez y 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 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 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* de todo lo que llega a sus manos o le dicen tiene fundamento, vale la pena leer y va a ayudar a su práctica.

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

E)

BIBLIOGRAFÍA RECOMENDADA

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(6957): 761-4.

Covell DG, et al. Information needs in office practice: are they being met? *Ann Intern Med* 1985; 103(4): 596-9.

Davidoff F, et al. Evidence based medicine . *BMJ* 1995; 310(6987): 1085-6.

Evidence-based medicine. A new approach to teaching the practice of medicine . Evidence-Based Medicine Working Group. *JAMA* 1992; 268(17):2420-5.

Feinstein AR, et al. Problems in the “evidence” of “evidence-based medicine” . *Am J Med* 1997; 103(6): 529-35.

Gehlbach S. Interpreting the medical literature: practical epidemiology for clinicians . 2nd ed. New York: McMillan, 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(21): 2598-601.

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(1): 59-63.

Hennekens C, et al. Epidemiology in medicine . Boston, Toronto: Little, Brown, 1987.

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(5): 389-91.

Jaeschke R, et al. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA* 1994; 271(9): 703-7.

Jovell AJ, et al. Evaluación de la evidencia científica . *Med Clin (Barc)* 1995; 105(19): 740-3.

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(20): 1615-9.

Osheroff JA, et al. Physicians' information needs: analysis of questions posed during clinical teaching . *Ann Intern Med* 1991; 114(7): 576-81.

Oxman AD, et al. Users' guides to the medical literature. VI. How to use an overview . Evidence-Based Medicine Working Group. *JAMA* 1994; 272(17): 1367-71.

Oxman AD, et al. Users' guides to the medical literature. I. How to get started . The Evidence-Based Medicine Working Group. *JAMA* 1993; 270(17): 2093-5.

Rosenberg W, et al. Evidence based medicine: an approach to clinical problem-solving . *BMJ* 1995; 310(6987): 1122-6.

Rubinstein F. Evidencia en atención primaria: ¿es necesario? *Evidencia aten. primaria* 1997; 0(0): 1-2.

Rubinstein F. Medicina basada en la evidencia: su justa medida . *Evidencia aten. primaria* 1998; 1(2):2-3.

Rubinstein F. Práctica clínica y evidencia disponible: ¿cuál es la estrategia? *Evidencia aten. primaria* 1998; 1(1):1-2.

Sackett DL, et al. Clinical epidemiology: a basic science for clinical medicine . 2nd ed. Boston, Toronto: Little, Brown, 1991.

Sackett, DL, et al. Evidence-based medicine: how to practice and teach EBM . 2nd ed. Edinburgh: Churchill Livingstone, 2000.

Sackett DL, et al. Evidence based medicine: what it is and what it isn't. *BMJ* 1996; 312(7023): 71-2.