

# MEDICINA BASADA EN EVIDENCIAS



26 agosto 2014

Nehim Alberto Navarro Cisneros  
Pre interno Hospital Fray Antonio Alcalde  
Universidad Guadalajara Lamar

## Historia de la Medicina Basada en la Evidencia

La Medicina Basada en la Evidencia, cuyos orígenes filosóficos se remontan a la mitad del siglo XIX en París o incluso antes, resulta un tema de máximo interés para clínicos, médicos de salud pública, gerentes, planificadores y para el público en general. Actualmente se organizan a menudo seminarios sobre su práctica y enseñanza (uno de ellos, patrocinado por esta revista, tendrá lugar en Londres el 24 de abril); las enseñanzas de pre 1 y postgrado 2 la incorporan 3 (o se plantean cómo hacerlo); se han establecido o se planifican en Gran Bretaña, centros para la práctica basada en la evidencia en medicina de adultos, salud infantil, cirugía, patología, farmacoterapia, enfermería, medicina general y odontología; la Colaboración Cochrane y el Centro para la Revisión y Difusión de York, proporcionan revisiones sistemáticas sobre los efectos de la atención de la salud; se lanzan nuevas revistas sobre la práctica basada en la evidencia, y ha llegado a ser un asunto común en los medios de comunicación. Pero se ha mezclado este entusiasmo con algunas reacciones negativas 4-6. Las críticas van desde que se trata de una "vieja-historia", a que es una innovación peligrosa, perpetrada para los arrogantes para servir a los recortes de costes y para suprimir la libertad clínica. Como la medicina basada en la evidencia continúa su desarrollo y adaptación, ahora es un momento adecuado para clarificar la discusión de qué es y qué no es.

Al tratar de establecer un orden cronológico con los antecedentes de lo que hoy se conoce como Medicina Basada en la Evidencia, debemos diferenciar entre sus bases filosóficas, y su reciente desarrollo. Poco es lo que se conoce de sus orígenes, al decir de sus entusiastas promotores, que fueron calificados como escépticos post- revolucionarios de París de mediados del siglo XIX como Bichat, Louis y Magendie 1 . Respecto de la historia más reciente debemos hacer referencia a la Escuela de Medicina de la Universidad McMaster en Hamilton, Ontario, Canadá. Esta referencia es de suma importancia pues lo novedoso de la propuesta educativa que esta Universidad ha comenzado a desarrollar a fines de la década del '60, es su orientación comunitaria centrada en las personas, interdisciplinaria y con aprendizaje basado en problemas. Este tipo de aprendizaje intenta que el profesional a partir de la problemática del caso rastree los elementos de conocimiento necesarios para comprenderlo en el contexto en que se presentan en la realidad. El aprendizaje basado en problemas se estructura en tres estadios: la identificación del problema, la búsqueda de información y la resolución del problema. Más adelante veremos la estructura de la Medicina Basada en la Evidencia y podremos compararla con la del aprendizaje basado en problemas.

Esta reforma en educación médica pone énfasis en perspectivas poco exploradas en los modelos clásicos como son la exposición temprana a problemas clínicos, el aprendizaje de las ciencias básicas y clínicas en un mismo tiempo, el desarrollo de habilidades clínicas, de comunicación, y preocupación comunitaria, el aprendizaje de conceptos de epidemiología clínica, bioestadística y

ética, el aprendizaje en pequeños grupos y la introducción de múltiples evaluaciones para planear correcciones en cualquier momento del proceso educativo. La creación de esta Escuela de Medicina nos acerca a uno de los pioneros en el aprendizaje y práctica de la Medicina Basada Evidencia, David Sackett. Este médico estadounidense procedente de un pequeño pueblo rural al oeste de Chicago, que recibió su primer entrenamiento en la Universidad de Illinois y comenzó su carrera como investigador clínico y básico en nefrología detectó la importancia de la epidemiología y las estadísticas mientras servía en las fuerzas armadas durante la crisis de los misiles en Cuba. Posteriormente graduado con honores en la Escuela de Salud Pública de Harvard fue convocado por John Evans en 1967 a McMaster para comenzar a desarrollar el Departamento de Epidemiología y Bioestadística. Entre las muchas iniciativas que ha promovido en la investigación médica canadiense se incluye la Task Force on Periodic Health Examinations de amplia repercusión mundial al resaltar el énfasis en la prevención basada en la evidencia<sup>2-3</sup>. A los 49 años, Sackett decidió entrenarse en Medicina General para poder aplicar sus predicas a la practica diaria y actualmente se desempeña como consultor general del Hospital John Radcliff y Director del Centro para la Medicina Basada en la Evidencia del Instituto Nacional de la Salud Ingles en Oxford.

En la búsqueda de la mejor evidencia se debe mencionar al entusiasta epidemiólogo Archie Cochrane, fallecido en 1988, quien preocupado desde la década del '70 al reconocer que los recursos de salud son siempre limitados sugirió que la efectividad de las prácticas relacionadas a la salud debe ser juzgada sobre la base de las pruebas procedentes de trabajos controlados. En 1978 escribió sobre su preocupación por la falta de resúmenes o revisiones que agrupen este tipo de trabajos. Sus ideas guiaron a un grupo de investigadores de Oxford, Iain Chalmers entre otros, para trabajar desde fines de la década del'70 en la construcción de una base de datos con revisiones sistemáticas de trabajos controlados y randomizados. Así en 1989 se publicó el libro "Effective Care in Pregnancy and Childbirth". Este esfuerzo se terminaría convirtiendo en la Colaboración Cochrane<sup>4</sup> fundada en 1992. La Colaboración Cochrane inicialmente con base en Oxford es una organización internacional que busca ayudar a quienes necesitan tomar decisiones bien informadas en atención de salud. Esta organización prepara, actualiza, promueve y facilita el acceso a las revisiones sistemáticas sobre intervenciones en salud. Sus acciones se basan en ciertos principios claves que incluyen la colaboración, basar su desarrollo en el entusiasmo individual, evitar la duplicación de esfuerzos, minimizar el sesgo, promover la actualización permanente, hacer un marc ado esfuerzo en producir material relevante, promover un amplio acceso y asegurar la calidad permaneciendo abiertos y sensibles a las críticas. Con Centros diseminados por el mundo ha dado a luz a The Cochrane Library<sup>5</sup>, obra de edición periódica en soporte electrónico que contiene el trabajo de grupos de revisores, así como a una intensa difusión vía Internet. Estos esfuerzos han llevado a una expansión en el desarrollo de meta-análisis en los últimos 10 años <sup>6</sup>.

Otras formas han sido utilizadas para presentar el concepto de Medicina Basada en la Evidencia. Por ejemplo, los españoles se refieren a una Medicina Basada en Pruebas haciendo referencia a que la traducción literal de la palabra evidence puede confundir. Nosotros hemos preferido mantener el nombre de Medicina Basada en la Evidencia en un intento de mantener el nombre con el que se ha contribuido a los cambios.

A continuación le brindamos una serie de términos utilizados cuando nos referimos a Medicina basada en evidencias.

**Mediana:** La mediana es el valor que divide a la muestra en dos una vez ordenados los valores de menor a mayor: por ejemplo, si tenemos cinco pacientes con 2,4,6,8, y 30 días de internación, la mediana de días de internación es 6 ya que quedan 2 valores por arriba y 2 por debajo del valor del medio o mediana. La media o promedio, muy influenciada por el valor extremo, es  $2+4+6+8+30/5= 10$ . Así como la media o promedio es una medida que se utiliza para expresar los resultados cuando la distribución de los valores sigue una distribución normal o simétrica, la mediana es una medida muy útil cuando la variable tiene una distribución no normal o asimétrica y es la que se debe considerar cuando hay valores extremos que influirían mucho sobre el promedio.

**Muestra:** Subgrupo de individuos (u observaciones) seleccionados de una población determinada (universo), con quienes se realiza el estudio. Es necesario que este grupo sea representativo de esa población para poder extrapolar y aplicar los hallazgos.

**Población:** o universo, es el número total de unidades, ya sean individuos, observaciones, etc. de la cual se obtiene la muestra para realizar un estudio.

**Prevalencia:** Se refiere a la proporción de personas dentro de una población que tienen un evento o una enfermedad en un momento determinado. Por ejemplo; en una encuesta se relevan los casos de tabaquismo de una población de 1000 personas. De ellas 320 contestan que actualmente fuman. La prevalencia de tabaquismo en esta población es de  $320/1000=32\%$ .

**Incidencia:** Son los nuevos casos de una enfermedad o evento que ocurren durante un período de tiempo en una población determinada. Por ejemplo; si durante el último año ocurrieron 200 nuevos casos de cólera en una ciudad de 10.000 habitantes, la incidencia de cólera en esa población es del 2% en el año. La incidencia representa el riesgo de sufrir el evento para esa población.

## Términos utilizados en estudios de tests diagnósticos

**Anclaje:** Se refiere a la probabilidad pre-test. Es la probabilidad de enfermedad que tiene el paciente antes de la realización del test. Este anclaje o probabilidad pre-test se puede homologar a la prevalencia de la enfermedad en esa población.

**Sensibilidad:** Proporción de pacientes enfermos que tienen el test positivo.

**Especificidad:** Proporción de pacientes sin enfermedad que tienen el test negativo

**Valor Predictivo Positivo:** Proporción de pacientes con test positivo que tienen la enfermedad. Depende fundamentalmente de la especificidad del test y de la prevalencia de la enfermedad en la población estudiada.

**Valor Predictivo Negativo:** Proporción de pacientes con test negativo que no tienen enfermedad. Depende fundamentalmente de la sensibilidad del test y de la prevalencia de la enfermedad en la población estudiada.

**Coefficientes de Probabilidad (likelihood ratio):** Expresan la chance de que determinado resultado del test provenga de un paciente enfermo comparado con un sano. A diferencia de los valores predictivos, estos coeficientes no se ven alterados al variar la prevalencia de la enfermedad en el grupo estudiado.

**Coefficiente de Probabilidad Positivo (positive likelihood ratio):**  $\text{Sensibilidad} / (1 - \text{Especificidad})$ . Es la chance de que un resultado positivo provenga de un enfermo frente a que provenga de un sano. Por ejemplo, un coeficiente de probabilidad positivo de 4 significa que es cuatro veces más probable que un test positivo provenga de un enfermo que de un sano. Cuanto más alto sea este valor, mayor será la probabilidad posterior de enfermedad.

**Coefficiente de Probabilidad Negativo (negative likelihood ratio):**  $(1 - \text{Sensibilidad}) / \text{Especificidad}$ . Es la chance de que un resultado negativo provenga de un enfermo frente a que provenga de un sano. Cuanto menor sea este valor, menor será la probabilidad posterior de enfermedad.

Test de referencia (Gold Standard): Es el procedimiento o estudio que se utiliza para definir el verdadero estado del paciente. Test universalmente aceptado para diagnosticar la enfermedad en cuestión y con el cual se comparan nuevos tests para diagnosticar la misma enfermedad (en general menos invasivos, menos costosos o más fáciles de realizar).

Curva ROC: Es el método más utilizado para evaluar nuevos test diagnósticos. Compara la sensibilidad y especificidad de distintos puntos de corte del test a evaluar con el gold standard o test de referencia. El poder de discriminación se expresa en área bajo la curva ROC; el test a evaluar tiene mayor poder cuanto más cercano al 1 sea el área bajo la curva. Un test con un área bajo la curva de 0.5 no aporta información .

Medidas de asociación: cuáles son y cómo expresan el impacto en los pacientes. Utilizaremos el siguiente ejemplo para explicar algunas medidas de asociación frecuentemente utilizadas en la literatura: Riesgo de accidente cerebrovascular (ACV) en hipertensos moderados a severos seguidos durante 60 meses. (Understanding clinical trials. BMJ, 1994, vol 309:755 - 756).

ACV NO ACV TOTAL Tratamiento 1800 13200 15000 Placebo 3000 12000 15000

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

Riesgo absoluto (RA): Es la incidencia (casos nuevos detectados) del evento de interés en cada grupo. En nuestro ejemplo : RA en pacientes tratados =  $1800 / 15000 = 0.12$  RA en pacientes no tratados =  $3000 / 15000 = 0.20$  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 pregunta: ¿cuántas veces más probable (o menos probable en estudios de tratamientos efectivos) es la ocurrencia del evento en los expuestos comparado con el grupo control? ( Riesgo absoluto del evento en los expuestos / Riesgo absoluto en los no expuestos). Si el riesgo de sufrir el evento es igual en ambos grupos, el RR = 1 (no diferencia), si es mayor en el grupo expuesto es  $> 1$  y si es menor es  $< 1$ .

$RR = RA \text{ en tratados} / RA \text{ en no tratados} = 0.12 / 0.20 = 0.6$ . 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).

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 de riesgo del evento en el grupo tratado.  $RRR = (1 - RR) \times 100$  o  $[(\text{riesgo en expuestos} - \text{riesgo en no expuestos}) / \text{riesgo en no expuestos}] \times 100$ . Se mide en porcentaje de reducción. En el ejemplo, la  $RRR = (1 - 0.6) \times 100 = 40\%$  o  $(12\% - 20\%) / 20\% \times 100 = 40\%$ . 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: Nos responde la pregunta: cuál es, en términos absolutos, la reducción del riesgo de sufrir el evento que se le atribuye al tratamiento? (o el aumento del riesgo, en caso de un factor nocivo). Se obtiene calculando la diferencia de los riesgos absolutos en una y otra población.  $RRA = \text{Riesgo en no tratados} - \text{Riesgo en tratados}$   $RRA = 0.20 - 0.12 = 0.08 = 8\%$ . La reducción del riesgo en este estudio significa que de cada 20 ACV que ocurren en esta población de hipertensos moderados a severos, 8 pueden ser prevenidos con el tratamiento antihipertensivo o, lo que es lo mismo, el beneficio atribuible al tratamiento es del 8 %.

Numero de Pacientes Necesario para Tratar (NNT): Es el número de pacientes que necesitamos tratar para evitar un evento: a cuántos pacientes hay que someter al tratamiento para obtener un beneficio? Cuando una intervención es costosa, invasiva o esta acompañada de muchos efectos colaterales que comprometen la calidad de vida, habrá que evaluarla muy cuidadosamente para justificar su implementación. Al igual que el riesgo atribuible, el NNT incorpora el concepto de cuan frecuente es el evento en la población que no recibe la intervención.  $NNT = 1 / RRA$ .  $NNT = 1 / 0.08 = 12.5$  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.

Numero de Pacientes Necesario para Dañar (NNH): Se utiliza cuando una intervención aumenta el riesgo de un evento adverso. Similar al concepto de NNT pero cuando el tratamiento experimental provoca efectos adversos. Es el número de pacientes que necesitamos tratar para 'provocar' un evento: a cuántos pacientes sometidos al tratamiento experimental hay que tratar (comparado con el tratamiento control) para provocar un evento adverso. Se calcula como la inversa del aumento de riesgo absoluto:  $1 / (\text{incidencia de evento adverso en el grupo experimental} - \text{incidencia del evento adverso en el grupo control})$

Odds Ratio (OR): Puede encontrarse en la literatura en español como Razón de Producto Cruzado. Se utiliza para reportar la chance que tiene un paciente del grupo expuesto de tener el evento comparado a un paciente del grupo control. Por ejemplo, 100 pacientes (a+b) están expuestos a un factor X y 20 de ellos (a) tienen el evento, mientras 80 no (b); otros 100 pacientes del grupo control no expuestos (c+d), 40 de ellos sufren el evento (c) y 60 no (d). El Odds (chance) en el grupo expuesto es  $a/b=0.25$ , el Odds en el grupo control es  $c/d=0.67$ , y el ODDS RATIO es  $(a/b)/(c/d)=0.25/0.67=0.37$ . El odds ratio se utiliza en muchos casos como una buena aproximación al RR (cuando los eventos son raros) y como medida para resumir una revisión sistemática u otros estudios comparativos de tratamiento, diagnóstico, pronóstico o etiología. También cuando se trata de estudios caso-control o a partir de modelos matemáticos multivariados como la regresión logística.

Intervalo de confianza: El intervalo de confianza (IC) cuantifica la incertidumbre. El IC 95% es el rango de valores en los cuales podemos estar 95% seguros que se encuentra el verdadero valor de una estimación (p.ej. un determinado riesgo) para esa población. A medida que el estudio tiene más pacientes, el IC es más angosto y preciso. Como ventaja frente al valor de p, que solo nos indica si hay o no diferencias desde lo cualitativo, el IC nos cuantifica la magnitud de la asociación.

Calidad de la evidencia

Grado I Evidencia obtenida a partir de al menos un ensayo clínico randomizado y controlado bien diseñado.

Grado II - 1 Evidencia obtenida a partir de ensayos controlados NO randomizados y bien diseñados. Grado II - 2 Evidencia obtenida a partir de estudios de cohortes ( 2a: prospectivo y 2b: retrospectivo o 2c: estudios caso control) bien diseñados, realizados preferentemente en más de un Centro o por más de un grupo de investigadores

Grado II -3 Evidencia obtenida a partir de múltiples series comparadas en el tiempo con o sin grupo control Incluye resultados "dramáticos" producidos por experimentos no controlados (ej. Penicilina en los años '40)

Grado III Opiniones basadas en experiencias clínicas, estudios descriptivos o informes de comités de expertos.



Relación entre los niveles de la calidad de la evidencia y el grado de recomendación de una práctica según la Fuerza de tareas preventivas de los EE.UU. (US Preventive Task Force) Grado de Recomendación Grado de Evidencia

A : Existe adecuada evidencia científica para adoptar la práctica. I II-1

B : Existe cierta evidencia científica para recomendar la práctica. II-1 II-2

C : Hay insuficiente evidencia para recomendar (o no) la práctica. III

D : Existe cierta evidencia para NO recomendar la práctica. II-1 II-2

E : Existe adecuada evidencia científica para NO adoptar la práctica. I II-1

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 C) y que muchas otras se siguen realizando aún cuando ya existe evidencia suficiente como para dejar de hacerlo (grado D y E).

Sesgo: Error sistemático de un estudio de investigación. Los tipos principales de sesgo son sesgo de selección (cualquier error que surge en el proceso de identificación de la población del estudio), sesgo de observación o información (error sistemático en la medición de información acerca de la exposición o resultados). El sesgo de selección no siempre se refiere a los pacientes incluidos en un estudio; también puede referirse a los artículos incluidos para realizar un meta-análisis: por ejemplo, elegir sólo artículos en idioma inglés deja fuera del análisis a cualquier artículo relacionado con el tema que esté publicado en otro idioma.

Sesgo de recuerdo : Es un tipo de sesgo de información. Se observa cuando los sujetos del estudio expuestos a un factor de riesgo o que tengan algún evento adverso recuerdan sus experiencias de una manera sistemáticamente diferente de los que no están expuestos o no padecen el evento adverso. Por ejemplo, en un estudio en personal hospitalario de exposición a gases anestésicos y riesgo de aborto espontáneo mostró que en el grupo no expuesto se reportaron sólo el 70% de todos los abortos mientras que en el grupo expuesto se reportaron el 100%.

Sesgo de publicación: Se refiere a la tendencia de las revistas científicas a publicar artículos con resultados positivos (los que encuentran diferencias significativas) mucho más fácilmente que aquellos con resultados negativos. Este tipo de sesgo puede ser un problema principalmente en los meta-análisis. Si los autores del meta-análisis no realizan una búsqueda exhaustiva, en la cual incluyan los estudios no publicados, las conclusiones del estudio pueden ser erróneas. Si no se buscan los artículos no publicados, que tienen más probabilidad de ser negativos, se puede basar la conclusión sólo en los artículos publicados y concluir falsamente que la intervención estudiada es positiva.

Factor de riesgo independiente: Un factor de riesgo es cualquier condición (una característica del comportamiento, del estilo de vida, la exposición a alguna sustancia, o alguna característica sociodemográfica) que se encuentra asociada a la ocurrencia de un evento o enfermedad determinada. Independiente significa que la asociación entre este factor y la enfermedad o el problema en cuestión no puede ser explicada por la influencia de otros factores. Por ejemplo, el tabaquismo es un factor de riesgo independiente para enfermedad coronaria. Esto significa que los tabaquistas tienen más riesgo de padecer enfermedad coronaria, aún tomando en cuenta los otros factores asociados a esta enfermedad como la hipercolesterolemia, la diabetes, la hipertensión, etc.

Confundidor: La presencia de un efecto confundidor se observa cuando la asociación entre una exposición y el evento en estudio se modifica en función de la presencia o ausencia de un tercer factor. En estas circunstancias la relación observada entre la exposición y el evento puede ser explicada parcial o totalmente por la variable confundidora. Para ser denominada como tal, éste tercer factor debe asociarse independientemente tanto al factor de exposición como al evento en estudio. P.ej: Existe asociación entre el uso de bastón y mortalidad. Es entonces un factor de riesgo?. El hecho es que la relación está confundida por la edad, que se asocia tanto al uso de bastón como a la mayor mortalidad y si incluyéramos la edad en el análisis, aquella asociación desaparecería.

Interacción: Es un fenómeno que ocurre cuando la relación entre dos variables difiere según la categoría o estrato de una tercer variable. P.ej: La medicación A disminuye la incidencia de arrugas en la población de un estudio. Cuando se evalúa la respuesta al tratamiento en los dos sexos, se observa que el efecto en las mujeres es distinto al de los hombres. En este caso se dice que existe una interacción entre el género del paciente y la respuesta al tratamiento y esto debe tomarse en cuenta en el análisis o reportarse la asociación en forma independiente para cada estrato.

Análisis multivariado o multivariable: Es un tipo de análisis estadístico que trata de evaluar la asociación independiente de una variable con un evento, teniendo también en cuenta la participación simultánea de otras variables. Por ejemplo si un análisis univariable analiza la relación entre el tabaquismo y la enfermedad coronaria, un análisis multivariable analizará la relación del tabaquismo y la enfermedad coronaria en presencia de otras variables como hipertensión, diabetes, hipercolesterolemia, etc.

Análisis de supervivencia: Es un tipo de análisis estadístico para estimar la ocurrencia de los eventos en función del tiempo (cuanto tarda en ocurrir en uno y otro grupo?). Así se pueden hacer inferencias acerca de los efectos de los distintos tratamientos, factores pronósticos, factores de exposición, etc. a lo largo del tiempo.

Intención de tratar: Es la manera más correcta de analizar un estudio randomizado. Cada paciente es analizado según el grupo al que fue asignado originalmente, independientemente de si abandonó el tratamiento o recibió algún otro. Si no se utiliza este tipo de análisis, se estará atentando contra la randomización y, por ende, contra la esencia misma del estudio experimental.

Eficacia: Mide la utilidad de la intervención a estudiar en condiciones ideales y con pacientes motivados y supervisados intensamente (estudio randomizado controlado)

Efectividad: Expresa la utilidad de la intervención en la práctica cotidiana, no en condiciones ideales.

Eficiencia: Se refiere a la utilidad de una intervención teniendo en cuenta también aspectos de costo-efectividad. A igual eficacia, la de menor costo será la más eficiente.

Kappa (Índice Kappa): Es una prueba estadística que mide el grado de acuerdo más allá del azar entre dos observaciones. Cuanto mayor sea este índice (máximo de 1) mayor será el grado de acuerdo para un signo determinado y por ende podrá utilizarse con más confianza en la práctica cotidiana.

Poder estadístico: Un estudio debe tener un diseño apropiado para que, en el caso de existir diferencias entre los grupos comparados, estas puedan ser demostradas. Cuando nos encontramos ante un resultado negativo (no diferencia) debemos preguntarnos: las poblaciones

realmente no son diferentes, o el estudio no tuvo la capacidad suficiente como para demostrarlo? Esta capacidad del estudio para detectar diferencias cuando éstas existen se denomina poder estadístico, y está principalmente relacionado con el número de participantes en el estudio, el número de eventos y el tiempo de seguimiento.

Validez externa (generalizabilidad) : Un estudio es válido externamente cuando sus resultados pueden ser extrapolables a otras poblaciones. Por ejemplo, un estudio sobre dolores musculares realizado en un centro de derivación reumatológico o de rastreo de clamidias en una clínica de ETS no es generalizable a una población de mujeres en atención primaria.

Meta-análisis: Es un tipo de revisión sistemática de la literatura. Utiliza un método estadístico de síntesis de los resultados de dos o más estudios primarios que evaluaron la misma hipótesis de una manera similar, utilizando criterios explícitos de inclusión y exclusión de los estudios analizados.

Revisión sistemática: Es un tipo de revisión que intenta responder una pregunta específica utilizando estudios ya publicados en la literatura. Tiene una metodología de búsqueda con criterios explícitos. La diferencia con el meta-análisis es que no combina matemáticamente los datos de los distintos estudios para expresar los resultados.

Estudios abiertos: Se refiere a un estudio en el que tanto los pacientes como los médicos tratantes conocen el tratamiento recibido. Por este motivo pueden ser más propensos a sesgos.

Estudios ciegos: Son aquellos en los que el paciente (simple ciego), o tanto el paciente como el médico (doble ciego) desconocen el tratamiento instaurado. Esto disminuye la probabilidad de sesgos, o errores sistemáticos de los estudios.

Estudio de equivalencia: Habitualmente los estudios que comparan intervenciones tratan de demostrar que éstas son diferentes entre sí o mejores que el placebo. En el caso de un estudio de equivalencia, lo que se trata de probar es que los tratamientos 'no son diferentes'. Estos estudios se plantean cuando se quiere evaluar un nuevo tratamiento versus el tratamiento estándar (que es más caro o más invasivo). Es fundamental que estos estudios tengan un poder suficiente (en general cercano al 90%) para detectar diferencias clínicamente relevantes si las hubiere.

Consentimiento Informado: Es una autorización voluntaria dada por un sujeto para participar de un estudio, programa, régimen terapéutico, etc; luego de que se le ha informado en detalle sobre el mismo.

Resultado intermedio: Son variables de menor relevancia clínica que un resultado final pero generalmente más fácilmente medible. Por ejemplo medir cambios en las cifras de tensión arterial sería un resultado intermedio, mientras que la mortalidad o la aparición de eventos cardiovasculares podría considerarse un resultado final.

Diseño factorial: Es un diseño especial de un estudio experimental que combina distintas intervenciones para poder evaluar más de una hipótesis al mismo tiempo. Por ejemplo, en el estudio de aspirina y warfarina (número 4 p.8), se randomizan los pacientes a cuatro grupos: 1) Aspirina+Warfarina, 2)Aspirina+Placebo, 3)Placebo+Warfarina, 4)Placebo+Placebo. Con este tipo de diseño se puede evaluar la eficacia tanto de la aspirina como de la warfarina, solas o combinadas, en la prevención de eventos vasculares.

Análisis de costo-efectividad: es un método para evaluar y comparar distintas estrategias de salud que tiene en consideración tanto la efectividad de una intervención como sus costos. Los resultados de estos estudios se expresan como valor en pesos (o dólares) \$ / año de vida salvado; \$/año de vida ajustado por calidad (\$ / QALY), etc.

QALY: Año de vida ajustado por calidad. Es una manera de comparar distintas estrategias que evalúan la expectativa de vida de los pacientes tomando en cuenta la calidad de vida. Diez años vividos en perfecto estado de salud aportan 10 QALYs, en cambio 10 años vividos con pobre calidad de vida aportan menos que 10 QALYs según como afecte la calidad de vida la patología estudiada.

Análisis de sensibilidad: Técnica utilizada principalmente en estudios de análisis de decisión y costo-efectividad. Sirve para evaluar cuan estables son las conclusiones del análisis al modificar los valores de algunas variables utilizadas.

Costo-efectividad incremental o marginal: Es la manera correcta de comparar las distintas alternativas en un análisis de costo-efectividad. La costo-efectividad incremental o marginal de una alternativa en relación con otra es el costo extra de ésta dividido por la diferencia de

efectividad. Por ejemplo, al comparar el tratamiento A con el tratamiento B, la costo-efectividad marginal del tratamiento A es:  $(\text{Costo de A} - \text{Costo de B}) / (\text{Efectividad de A} - \text{Efectividad de B})$ .

Modelo de Markov: modelo empleado cuando se requieren complejos cálculos de costos y resultados en grandes cohortes de pacientes. Los pacientes ingresan a diferentes "estados" que representan niveles de salud. Por ejemplo un modelo puede conceptualizar la salud de una población en tres estados: a) sanos; b) enfermos; c) muertos. La transición entre los distintos estados se realiza sólo una vez durante cada ciclo. La longitud de cada ciclo depende de las características de la enfermedad analizada (ciclos de menor duración en enfermedades rápidamente progresivas).

TABLA 1  
Niveles de evidencia de acuerdo a las recomendaciones del centro de Medicina basada en la evidencia de Oxford\*

| Grado de recomendación | Nivel de evidencia | Tratamiento/etiología  | Pronóstico   | Diagnóstico   |
|------------------------|--------------------|--|--|---|
| A                      | 1a                 | RS (con homogeneidad) de EAC   | RS (con homogeneidad) de estudios de cohortes prospectivos GPC validadas | RS (con homogeneidad) de estudios de nivel 1 o GPC validadas  |
|                        | 1b                 | EAC con IC estrechos   | Cohortes con seguimiento $\geq 80\%$                                     | Comparación independiente y ciega de un espectro apropiado de pacientes   |
| B                      | 2a                 | RS (con homogeneidad) de cohortes                                    | RS (con homogeneidad) de estudios de cohortes retrospectivos             | RS (con homogeneidad) de estudios diagnósticos de nivel $\geq 2$  |
|                        | 2b                 | Estudios de cohortes o EAC de baja calidad (tamaño muestral pequeño) | Estudios de cohortes retrospectivos o GPC no validadas                   | Ausencia de comparación objetiva<br>Espectro limitado de pacientes o pacientes no consecutivos<br>GPC no validadas                          |
|                        | 3a                 | RS (con homogeneidad) de estudios de casos y controles               |  |   |
|                        | 3b                 | Estudios de casos y controles  |  | No realización de la prueba de referencia a todos los sujetos, respetando la independencia, la ceguera y el apropiado espectro de pacientes |
| C                      | 4                  | Series de casos  | Series de casos  |   |
| D                      | 5                  | Opinión de expertos  | Opinión de expertos  | Opinión de expertos   |

RS: revisión sistemática; EAC: ensayo clínico con asignación aleatoria; GPC: guía de práctica clínica. \*Levels of Evidence and Grades of Recommendations (Revised on 18<sup>th</sup> November 1999) del Evidence-Based Medicine Centre (Oxford). Disponible en: <http://www.cebm.ox.uk>

**TABLA 1**  
**Niveles de evidencia y grados de recomendación**

| Grado recomendación | Nivel evidencia | Tratamiento, prevención, etiología y complicaciones                     | Estudios de historia natural y pronóstico  | Diagnóstico   |
|---------------------|-----------------|---|--|---|
| A                   | 1a              | RS de ECA   | RS de estudios de cohortes   | RS de estudios diagnósticos de calidad alta   |
|                     | 1b              | ECA individual  | Estudios de cohortes individuales  | Estudios de cohortes que validen la calidad de un test específico, con un estándar de referencia adecuado   |
|                     | 1c              | Eficacia demostrada por la práctica clínica y no por la experimentación | Eficacia demostrada por la práctica clínica y no por la experimentación (serie de casos) | Pruebas diagnósticas con especificidad tan alta que un resultado positivo confirma el diagnóstico y con sensibilidad tan alta que un resultado negativo descarta el diagnóstico |
| B                   | 2a              | RS de estudios de cohortes  | RS de estudios de cohorte retrospectiva o de grupos controles no tratados en un ECA      | RS de estudios diagnósticos de calidad mediana  |
|                     | 2b              | Estudio de cohortes individual y ECA de baja calidad                    | Estudio de cohorte retrospectiva o seguimiento de controles no tratados                  | Estudios exploratorios o reglas de decisión diagnóstica   |
|                     | 2c              | Investigación de resultados en salud, estudios ecológicos               | Investigación de resultados en salud   |   |
|                     | 3a              | RS de estudios de caso-control  |  | RS de estudios diagnósticos de calidad baja   |
|                     | 3b              | Estudios de caso-control individuales                                   |  | Estudio con pacientes no consecutivos, sin estándares de referencia aplicados de manera consistente   |
| C                   | 4               | Serie de casos y estudios de cohortes y caso-control de baja calidad    | Serie de casos y estudios de cohortes de pronóstico de poca calidad                      | Estudio de casos y controles de baja calidad o sin un estándar independiente  |
| D                   | 5               | Opinión de expertos sin valoración crítica explícita                    | Opinión de expertos sin valoración crítica explícita                                     | Opinión de expertos sin valoración crítica explícita  |

RS: revisión sistemática; ECA: ensayo clínico aleatorizado.

Fuente: Adaptación de la clasificación del Centro de Medicina Basada en la Evidencia de Oxford ([http://www.cebm.net/levels\\_of\\_evidence.asp](http://www.cebm.net/levels_of_evidence.asp)).

**Tabla 1: Jerarquía de los estudios por el tipo de diseño según US Preventive Task Force**

| Nivel de evidencia | Tipo de estudio  |
|--------------------|--|
| I                  | Al menos un ensayo clínico controlado y aleatorizado diseñado en forma apropiada   |
| II-1               | Ensayos clínicos controlados bien diseñados, pero no aleatorizados   |
| II-2               | Estudios de cohorte o caso control bien diseñados, preferentemente multicéntricos.   |
| II-3               | Múltiples series comparadas en el tiempo con o sin intervención, y resultados sorprendentes en experiencias no controladas.  |
| III                | Opiniones basadas en experiencias clínicas, estudios descriptivos, observaciones clínicas o informes de comités de expertos. |

Tabla 5. Clasificación de los niveles de evidencia según Sackett<sup>13</sup>

| Recomendación | Nivel | Terapia, prevención, etiología y daño                | Pronóstico  | Diagnóstico  | Estudios económicos  |
|---------------|-------|--|---|--|--|
| A             | 1a    | RS con homogeneidad y Meta-análisis de EC            | RS con homogeneidad y Meta-análisis de estudios de cohortes concurrente                 | RS de estudios de diagnóstico nivel 1  | RS de estudios económicos de nivel 1   |
|               | 1b    | EC individuales con intervalo de confianza estrecho  | Estudio individual de cohorte concurrente con seguimiento superior al 80% de la cohorte | Comparación independiente y enmascarada de un espectro de pacientes consecutivos sometidos a la prueba diagnóstica y al estándar de referencia | Análisis que compara los desenlaces posibles, contra una medida de costos. Incluye un análisis de sensibilidad |
| B             | 2a    | RS con homogeneidad de estudios de cohortes          | RS de cohortes históricas   | RS de estudios diagnósticos de nivel mayor a 1   | RS de estudios económicos de nivel mayor a 1   |
|               | 2b    | Estudio de cohortes individual. EC de baja calidad   | Estudio individual de cohortes históricas   | Comparación independiente enmascarada de pacientes no consecutivos, sometidos a la prueba diagnóstica y al estándar de referencia              | Comparación de un número limitado de desenlaces contra una medida de costo. Incluye análisis de sensibilidad   |
|               | 3a    | RS con homogeneidad de estudios de casos y controles |   |  |  |
|               | 3b    |  |   |  |  |

**Cuadro III**

**VENTAJAS Y DESVENTAJAS DE LOS ESTUDIOS DE COHORTE**

C 4

**Ventajas**

- Más cercanos a un experimento
- La relación temporal causa efecto es verificable
- Se pueden estimar medidas de incidencia
- Eficientes para evaluar exposiciones poco frecuentes
- Se pueden estudiar varios eventos
- Se pueden fijar criterios de calidad en la medición del evento
- Bajo riesgo de sesgo de selección (en especial en estudios prospectivos)

D 5

**Desventajas**

- Cuando se trata de eventos poco frecuentes la complejidad y el costo pueden aumentar considerablemente, ya que requiere estudiar y seguir un número grande de participantes
- Son estudios difíciles de realizar

in una medida de costo, pero análisis de sen-

in análisis de ad

de expertos acción crítica o basada en económica

is de resultados

Por homogeneidad se entre los estudios indi



# Pirámide de la Evidencia



## Bibliografía:

1. British Medical Association: Report of the working party on medical education. London: British Medical Association, 1995.
2. Standing Committee on Postgraduate Medical and Dental Education: Creating a better learning environment in hospitals: 1. Teaching hospital doctors and dentists to teach. London: SCOPME, 1994.
3. Sackett DL, Rosenberg WMC, Muir Gray JA, Haynes RB, Scott Richardson W. Evidence based medicine: What it is and what it isn't. *BMJ* 1996, 312: 71-2.
4. Sackett D, Haynes RB, Guyatt GH, Tugwell P. *Epidemiología clínica: ciencia básica para la medicina clínica*. 2a. Ed. Buenos Aires: Editorial Médica Panamericana 1991.
5. Cohen L, McMaster's pioneer in evidence-based medicine now spreading his message in England. *Can Med Assoc J* 1996, 154: 388-90.
6. Robinson A. Research, practice and the Cochrane Collaboration. *Can Med Assoc J* 1995; 152: 883- 9.
7. Rubinstein FA. En PROFAM: Programa de Educación a Distancia de Medicina Familiar y Ambulatoria- Fundación para el Desarrollo de la medicina Familiar y la Atención Primaria de la Salud .1ra. ed. Buenos Aires: Hospital Italiano de Buenos Aires. Entrega II capítulo 7 1997.
8. Hennekens CH, Buring JE *Epidemiology in Medicine*. Little Brown, 1987
9. Weinstein, MC, Fineberg. *Clinical Decision Analysis*. WB Saunders 1980.
10. Sackett DL, Richardson WS, Rosenberg W, Haynes RB (eds). *Evidence-Based Medicine*. London; Churchill Livingstone, 1997
11. Fletcher RH, Fletcher SW, Wagner EW. *Clinical Epidemiology*. Williams & Wilkins 1996.