

3. Análisis descriptivo

Como ya hemos comentado, el análisis descriptivo sirve para resumir los datos. Las dos formas de realizar un resumen de datos son mediante una **tabla de frecuencias** o mediante **representaciones gráficas**. Para realizar un resumen de los datos, disponemos de una serie de estadísticos (cantidad numérica calculada a partir de la muestra), que se dividen en:

- **Estadísticos de centralización:** estos estadísticos ofrecen información acerca de valores típicos o promedio de un conjunto de datos. Los estadísticos que estudiaremos en este módulo son:
 - **Media:** suma de los valores dividida por el tamaño muestral. Es sensible a valores extremos.
 - **Mediana:** es un valor que divide a las observaciones con el mismo número de valores. Es decir, deja a cada lado el 50% de las observaciones. No es sensible a valores extremos.
 - **Moda:** es el valor o valores más frecuente. A diferencia de la media y la mediana, la moda no es un valor único, podemos encontrar distribuciones unimodales, bimodales.
- **Estadísticos de dispersión:** indican en qué medida están juntos o separados del valor central. Cuanto más separados estén del valor del central mayor dispersión.
 - **Varianza:** mide el promedio de las desviaciones al cuadrado de las observaciones respecto a la media. Es sensible a valores muy alejados de la media. Además es de difícil interpretación, ya que al tener las unidades al cuadrado dificulta su interpretación directa. Por ello se aconseja utilizar la desviación típica o estándar.
 - **Desviación típica:** se calcula como la raíz cuadrada de la varianza. Tiene las mismas unidades de medida que los datos originales y suele presentarse como descripción de la variabilidad de los datos en un trabajo de investigación.
 - **Coficiente de variación de Pearson:** es una medida de dispersión relativa, permite la comparación de la dispersión existente entre distribuciones distintas. Con este valor podemos conocer si los valores de una variable están dispersos o concentrados. La comparación directa de las desviaciones típicas no es muy informativa ya que dependerá de las unidades de medida. Se calcula como la razón de la desviación típica y la media. Suele expresarse en porcentaje. Así por ejemplo, si la media de una variable vale 10 y la desviación típica 5, podemos afirmar que la desviación típica supone el 50% de la media.

• **Estadísticos de Forma:**

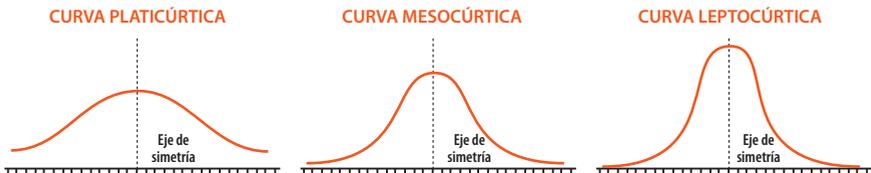
- **Asimetría:** las distribuciones pueden ser simétricas o asimétricas. Cuando hay simetría positiva la cola de la derecha es más prolongada y el coeficiente de asimetría será positivo. Cuando la asimetría es negativa, la cola de la izquierda es más larga y el coeficiente será negativo. Cuando la distribución es simétrica el valor es de cero. En distribuciones simétricas, media y mediana coinciden y la moda si es única también.

Figura 2. Simetría y asimetría de las distribuciones.



- **Apuntamiento o curtosis:** mide el grado de apuntamiento de la distribución respecto de la distribución normal, puede ser platicúrtica (curtosis<0), mesocúrtica (curtosis=0) y leptocúrtica (curtosis>0).

Figura 3. Grado de apuntamiento de la distribución.



• **Estadísticos de posición:**

- **Cuartiles:** son 3 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en cuatro tramos iguales, en los que cada uno de ellos concentra el 25% de los resultados.
- **Deciles:** son 9 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en diez tramos iguales, en los que cada uno de ellos concentra el 10% de los resultados.
- **Percentiles:** son 99 valores que distribuyen la serie de datos, ordenada de forma creciente o decreciente, en cien tramos iguales, en los que cada uno de ellos concentra el 1% de los resultados.

Bibliografía

Apuntes de bioestadística. Universidad de Málaga: <http://www.bioestadistica.uma.es/baron/apuntes/>

Borrás F, *Cuadernos de bioestadística* (1995). Cuadernos de bioestadística.

Cliffor R C, Taylor R.A. *Bioestadística* (2008).

Martínez González M.A, De Irala Estévez J, Faulín Fajardo F.J. *Bioestadística amigable* (2001).

Sokal R, Rohlf R. *Introducción a la bioestadística* (1980).

Unidad de bioestadística del Hospital Ramón y Cajal: http://www.hrc.es/investigacion/inves_unidadbio.htm

11

Probabilidad

MARI CARMEN VIVO MOLINA

Licenciada en Ciencias y Técnicas Estadísticas.
Profesora asociada de la Universidad Miguel Hernández.
Dirección General de Planificación, Ordenación Sanitaria
y Farmacéutica e Investigación. Consejería de Sanidad
y Política Social de la Región de Murcia.

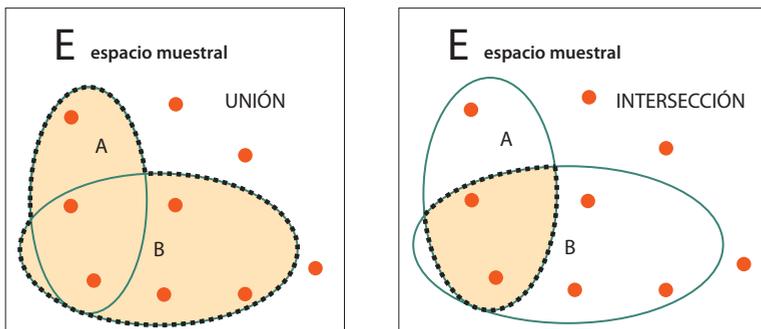
1. Introducción

Las pruebas diagnósticas suelen estar acompañadas de incertidumbre, la forma de cuantificar esta incertidumbre es mediante el cálculo de probabilidades. Una definición básica de probabilidad es la proporción de veces que un suceso ocurra si repetimos la experiencia un número muy grande de veces.

2. Conceptos de probabilidad

Quando realizamos un experimento aleatorio, obtenemos diversos posibles resultados. Al conjunto de todos los posibles resultados lo llamaremos **espacio muestral (E)**. Cuando estamos interesados en la ocurrencia de un resultado concreto, estamos hablando de un **suceso**. Llamaremos suceso complementario de A (\bar{A}), al suceso contrario ($P(A)=1-P(\bar{A})$), unión de los sucesos A y B ($A \cup B$), a los resultados experimentales que están en A, en B o en ambos sucesos. Sean dos sucesos A y B, las operaciones básicas de unión e intersección están representadas en la siguiente figura.

Figura 1. Unión e intersección de sucesos.



Llamamos probabilidad a cualquier función **P** que asigna a cada suceso **A** un valor numérico **P(A)**. Las reglas de la función de probabilidad son:

1. La probabilidad de un suceso seguro es 1.
2. La probabilidad toma valores entre 0 y 1.
3. La probabilidad de un suceso incierto es 0.

Regla de Laplace

Establece que siempre que los sucesos sean equiprobables, la probabilidad de obtener cada resultado será igual al cociente entre casos favorables y casos probables. Así la probabilidad de A se estima como:

$$P(A) = \frac{N^{\circ} \text{ de veces que ocurre } A}{\text{Número de veces que podría ocurrir } A}$$

Así, supongamos que de 1000 personas mayores de 50 años 100 pueden desarrollar cáncer de colon, la probabilidad de desarrollar cáncer de colon se estimaría como: $100/1000=0,1=10\%$.

Probabilidad condicionada

Se llama probabilidad de que ocurra un suceso A condicionada a B ($P(A/B)$), a la probabilidad de que ocurra el suceso A sabiendo que ocurre B. La fórmula de la probabilidad condicionada es:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

Diremos que dos sucesos son independientes cuando la ocurrencia de uno no influye en la ocurrencia del otro. Así, si A y B son dos sucesos independientes, cumplen:

1. $P(A \cap B) = P(A)P(B)$
2. $P(A/B) = P(A)$

Ejemplo: la siguiente tabla muestra los resultados de una muestra de 723 pacientes, donde cada uno de los 723 se caracteriza por ser fumador (F) o no (\bar{F}) y por presentar cierta enfermedad (Enf) o no (\overline{Enf}).

Tabla 1. Número de fumadores y no fumadores según presenten o no la enfermedad.

	Fumador	No fumador	Total
Presenta cierta enfermedad	428	82	510
No presenta cierta enfermedad	62	151	213
Total	490	233	723

La probabilidad de seleccionar un paciente que sea fumador $P(F)=490/723$, el suceso complementario o la probabilidad de seleccionar un paciente que no sea fumador es $P(\bar{F})=1-490/723$. La probabilidad de seleccionar una persona que fume y que presente la enfermedad, se calcula mediante la intersección de los dos sucesos, presentar la enfermedad y ser fumador. $P(Enf \cap F)=428/723$. La probabilidad de ser fumador o presentar la enfermedad $P(Enf \cup F)=490/723+510/723-428/723=572/723$. Y por último, si una persona es fumadora, la probabilidad de presentar la enfermedad es:

$$P(Enf/F) = \frac{P(Enf \cap F)}{P(F)} = \frac{\frac{428}{723}}{\frac{490}{723}} = \frac{428}{490}$$

Teorema de Bayes

Se llama partición a un conjunto de sucesos A_i tales que $A_1 \cup A_2 \cup \dots \cup A_n$, mutuamente excluyentes y que cubren todo el espacio muestral E, y sea B un suceso tal que $P(B) \neq 0$:

$$P(A_i/B) = \frac{P(B/A_i)P(A_i)}{\sum_{i=1}^n P(B/A_i)P(A_i)} \quad \forall i = 1, \dots, n$$

En general, la regla de Bayes permite usar $P(A/B)$ para calcular la $P(B/A)$.

La principal aplicación de esta prueba en ciencias de la salud es para el cálculo de índices predictivos. Este teorema tiene su aplicación en el cálculo de la probabilidad de un diagnóstico correcto.

Pruebas médicas

Las medidas de frecuencia más comúnmente utilizadas son: Prevalencia e Incidencia.

Prevalencia: probabilidad de que una persona tenga la enfermedad.

Incidencia: la incidencia refleja el número de nuevos “casos” en un periodo de tiempo.

Las pruebas médicas están diseñadas para establecer la presencia o ausencia de alguna enfermedad. Estas pruebas no son perfectas. El mejor o peor desempeño de una prueba puede evaluarse a través del cálculo de:

Sensibilidad (S): probabilidad de que una persona con la enfermedad tenga un resultado positivo.

Especificidad (E): probabilidad de que una persona que no tiene la enfermedad, dé un resultado negativo.

Estimación de probabilidades posterior: es la probabilidad de presentar la enfermedad después de conocer los resultados de la prueba diagnóstica, las principales medidas son:

- **Valor predictivo positivo (VPP):** es la probabilidad de que una persona que da positivo, en una prueba diagnóstica tenga la enfermedad. La probabilidad de identificar correctamente los casos positivos.
- **Valor predictivo negativo (VPN):** Es la probabilidad de una persona que da negativo en una prueba diagnóstica no tenga la enfermedad. La probabilidad de clasificar correctamente los casos negativos.

Tabla 2. Categorías de clasificación de la enfermedad según el estado real del paciente.

	Presencia de la enfermedad	No presencia de la enfermedad
Resultado positivo de la prueba	VPP	Falso Positivo (FP)
Resultado negativo de la prueba	Falso Negativo (FN)	VPN

- **Cociente de probabilidad positivo (CPP):** cuanto es más probable el resultado positivo entre los enfermos que entre los no enfermos. La relación entre la proporción de enfermos que son clasificados como tales por la prueba diagnóstica y la proporción de sanos que han sido clasificados como enfermos.
- **Cociente de probabilidad negativo (CPN):** cuanto más probable es el resultado negativo entre los enfermos, que entre los no enfermos. La relación entre la probabilidad de que la prueba sea negativa en presencia de la enfermedad y la probabilidad de que la prueba sea negativa en su ausencia. Así un resultado de 8,5 quiere decir que la probabilidad de que la prueba presente un resultado negativo en un individuo sano es 8,5 veces la probabilidad de que presente un resultado negativo en uno enfermo.

Riesgo relativo (RR): es el cociente entre el riesgo del grupo tratado y el riesgo del grupo control. Para evitar valores menores que 1, se define como el cociente entre el riesgo del grupo con mayor riesgo y el de menor riesgo.

Reducción absoluta del Riesgo (RAR): diferencia entre el riesgo del grupo control y el riesgo del grupo tratado.

Reducción relativa de riesgo (RRR): cociente entre la reducción absoluta del riesgo y el riesgo en el grupo control.

Número necesario a tratar (NNT): número de individuos que hay que tratar con el tratamiento experimental para producir, o evitar, un evento adicional respecto a los que se producirían con el tratamiento control. Se calcula como el inverso del riesgo relativo absoluto (RRA).

Razón de probabilidad (Odds Ratio): otra medida de la exactitud de un clasificador se basa en las odds (ventajas o preferencias) de las respuestas. Es el cociente entre la probabilidad de que ocurra un evento y la probabilidad de que no ocurra el evento (complementario). Indica cuánto más probable es la ocurrencia de un evento frente a la no ocurrencia. Puede tomar valores de 0 a infinito.

Existen dos tipos de odds que son de interés para analizar el rendimiento de un clasificador, que se pueden calcular en función de la sensibilidad y la especificidad (véase tabla).

Si toma el valor cero, esto quiere decir que no hay asociación entre las variables. Si el resultado es mayor que uno, la asociación es positiva, es decir, si la presencia del factor se asocia a mayor ocurrencia del evento. Si toma valor 1, no hay asociación entre la presencia del factor y el evento. Por ejemplo, si una droga con respecto al placebo produce una OR de 1,3, significa que la droga es 1,3 veces mejor que el placebo.

Este tipo de resultados se suele acompañar con intervalos de confianza, que desarrollaremos en el siguiente tema.

Ejemplo: el siguiente ejemplo muestra la capacidad predictiva de una prueba diagnóstica. Puedes comprobar los resultados usando el archivo CALCULADORA.xlsx, en el que además puedes obtener los intervalos de confianza al 95%.

Tabla 3. Resultados de una prueba para enfermos y no enfermos

		Enfermedad		Total
		No enfermo	Enfermo	
Resultado	Negativo	1	6	7
	Positivo	8	5	13
Total		9	11	20

Sensibilidad: $\frac{5}{1} = 0,45$. El 45,5% de las personas con la enfermedad están debidamente identificadas, menos de la mitad.

Especificidad: $\frac{1}{9} = 0,11$. El 11% de las personas que no tienen la enfermedad tienen un resultado negativo. Por lo tanto, un paciente que no tenga la enfermedad no es muy seguro que dé el resultado negativo.

Valor Predictivo Positivo: $\frac{5}{13} = 0,6923$. Si una persona da positivo en la prueba, la probabilidad de que tenga la enfermedad es de 69,23%.

Valor Predictivo Negativo: $\frac{0,11}{0,35} = 0,1428$. Si una persona da negativo la probabilidad de que no tenga la enfermedad es del 14,28%.

Prevalencia: 0,55.

Ejemplo: Supongamos, el siguiente resultado de cierto ensayo clínico.

Tabla 4. Número de pacientes fallecidos

		Fallecidos		Total
		Sí	No	
	Intervención	20	180	200
	Control	15	85	100
Total		35	265	300

$$RR = \frac{P(\text{Fallecer} | \text{Intervención})}{P(\text{Fallecer} | \text{Control})} = \frac{20/200}{15/100} = 0,667.$$

$$OR = \frac{\frac{P(\text{Fallecer} | \text{Intervención})}{P(\text{NoFallecer} | \text{Intervención})}}{\frac{P(\text{Fallecer} | \text{Control})}{P(\text{NoFallecer} | \text{Control})}} = \frac{20/200 / 180/200}{15/100 / 85/100} = 0,63.$$

$$RAR = \frac{20}{100} - \frac{15}{200} = 0,05. \text{ Por cada 100 pacientes que sigan el tratamiento se evitarán 5 muertes}$$

$$RRR = \frac{\frac{15}{100} - \frac{20}{200}}{15/100} = 0,33. \text{ La reducción del riesgo de 0,05 representa una reducción del 33,3\%.}$$

$$NNT = \frac{1}{0,05} = 20. \text{ Para evitar una muerte se necesitarían tratar a 20 personas.}$$

Tabla 5. Resumen de las medidas.

		CLASIFICACIÓN					
		POSITIVO	NEGATIVO				
ESTADO REAL	PRESENCIA	VERDADERO POSITIVO (VP)	FALSO NEGATIVO (FN)	TOTAL CON PRESENCIA (VP+VN)	SENSIBILIDAD (S) $\frac{VP}{VP + FN}$	Odds Presencia $\frac{VP}{FN}$	COCIENTE DE PROBABILIDAD POSITIVO $\frac{s}{1 - E}$
	AUSENCIA	FALSO POSITIVO (FP)	VERDADERO NEGATIVO (VN)	TOTAL CON AUSENCIA (FP+VN)	ESPECIFICIDAD (E) $\frac{VN}{FP + VN}$	Odds Ausencia $\frac{FP}{VN}$	COCIENTE DE PROBABILIDAD NEGATIVO $\frac{1 - s}{IE}$
		TOTAL RESPUESTA POSITIVA (FP+VP)	TOTAL RESPUESTA NEGATIVA (VN+FN)			Odds Ratio $\frac{VPVN}{FNFP}$	
		VALOR PREDICTIVO POSITIVO $\frac{VP}{VP + FN}$	VALOR PREDICTIVO NEGATIVO $\frac{VN}{VN + FN}$				

Todas estas medidas tienen asociado un intervalo de confianza (que explicaremos más adelante). En la web existen numerosas calculadoras que calculan los intervalos de confianza asociados a estas medidas.

3. Distribuciones de probabilidad

Antes de nada, hay que distinguir entre estimador y parámetro. En general, nunca tendremos acceso a la totalidad de la población, siempre trabajaremos con una muestra. Con la muestra solamente se pueden calcular estimadores. Un estimador es una función de los valores de la muestra.

La curva de la Normal

Muchas de las variables estudiadas en medicina y en ciencias de la salud suelen seguir una distribución normal. Esta distribución está indicada para datos que siguen una escala continua: peso, talla, edad, colesterol, tensión arterial, ácido úrico y además tiene la ventaja de que bajo ciertas condiciones, otras distribuciones acaban aproximándose a ella. La distribución normal teórica nunca se da exactamente en la realidad, solamente hay aproximaciones a ella. Si una variable x sigue una distribución normal de media μ y varianza σ^2 , se caracteriza por:

1. Tener forma de campana, mesocúrtica.
2. Ser simétrica.
3. Coincidir media, mediana y moda.
4. Tener el 95% de los individuos dentro del intervalo $\mu \pm 2\sigma$.

Un concepto importante a explicar es el procedimiento de tipificar o estandarizar la normal. Este procedimiento consiste en transformar una variable normal en otra que tenga una media igual a 0 y una desviación típica igual a 1. Sea x la variable antigua y z la variable tipificada, calculamos z de la siguiente forma:

$$z = \frac{x - \mu}{\sigma}$$

Supongamos que la media de tensión arterial de una población es de 120 y la desviación estándar es de 25, asumimos que sigue una distribución normal. Queremos calcular qué proporción de personas tienen una tensión arterial superior a 170.

$$z = \frac{170 - 120}{25} = 2$$

Buscando en las tablas, comprobamos que la probabilidad correspondiente al valor 2 es de 0,028, por lo tanto el 2,8% de los pacientes tendrá valores superiores a 170. A la hora de buscar en las tablas es necesario comprobar si nos proporciona la cola de la derecha o la cola de la izquierda.

Teorema Central del Límite

A grandes rasgos este Teorema afirma que las distribuciones muestrales se aproximan a la normal a medida que el tamaño de la muestra aumente. Esto quiere decir que para tamaños de muestra muy pequeños puede que no podamos aproximar a una normal.

Si x es una v.a de distribución cualquiera, con una muestra de tamaño grande $n > 30$, sigue aproximadamente una Normal. La importancia de esta propiedad radica en que es válida cualquiera que sea la variable aleatoria.

Distribución Binomial

Es la más simple de las distribuciones de probabilidad. Se refiere a sucesos en los que solamente ocurren dos posibilidades, que un paciente tenga un diagnóstico positivo o negativo, se trata de sucesos excluyentes (nunca se van a dar los dos sucesos a la vez). Esta distribución tiene dos características: el número de individuos observados y la probabilidad de alcanzar un determinado resultado, se denota como Bin (n, p).

Bibliografía

Apuntes de bioestadística. Universidad de Málaga: <http://www.bioestadistica.uma.es/baron/apuntes/>

Borrás F, *Cuadernos de bioestadística* (1995). Cuadernos de bioestadística.

Cliffor R C, Taylor R.A. *Bioestadística* (2008).

Martínez González M.A, De Irala Estévez J, Faulín Fajardo F.J. *Bioestadística amigable* (2001).

Sokal R, Rohlf R. *Introducción a la bioestadística* (1980).

Unidad de bioestadística del Hospital Ramón y Cajal: http://www.hrc.es/investigacion/inves_unidadbio.htm

12

Inferencia

MARI CARMEN VIVO MOLINA

Licenciada en Ciencias y Técnicas Estadísticas.
Profesora asociada de la Universidad Miguel Hernández.
Dirección General de Planificación, Ordenación Sanitaria
y Farmacéutica e Investigación. Consejería de Sanidad
y Política Social de la Región de Murcia.

1. Introducción

La inferencia estadística está relacionada con los métodos para obtener conclusiones o generalizaciones acerca de una muestra. De esta manera, estimamos lo que sucederá en la población de donde procede la muestra. Los dos procedimientos inferenciales son: estimación y contraste de hipótesis. Las dos formas de realizar estimación son:

- Estimación Puntual: nos proporciona un valor aproximado del parámetro. Una estimación por sí sola no nos dice lo buena que es la aproximación.
- Estimación por intervalo: nos proporciona un abanico de posibles valores, uno de ellos esperamos que sea el verdadero.

2. Intervalos de confianza

Un intervalo de confianza es un rango de valores en el que confiamos que se contenga el parámetro poblacional. El intervalo de confianza depende:

- Tamaño de la muestra. Cuanto mayor sea la muestra más estrecho será el intervalo de confianza.
- Nivel de confianza. Es la probabilidad con la que el verdadero valor estará incluido en el intervalo. A nivel general se usa el 95%.

La idea básica del intervalo de confianza al 95% es que si repitiésemos 100 veces el proceso de estimación, tendríamos la confianza de que en uno de los 95 intervalos se encontraría el verdadero valor de nuestra estimación.

El intervalo de confianza se puede realizar para cualquier parámetro de la población: media, mediana, proporción.

Para calcular el intervalo de confianza, necesitamos conocer el error estándar. El error estándar es una medida de la variabilidad del estimador. Depende del parámetro poblacional que queramos estimar y de la distribución de probabilidad.

3. Pruebas de hipótesis

Una hipótesis estadística es una asunción relativa a una o varias poblaciones, que puede ser cierta o no. Las hipótesis estadísticas se pueden contrastar con la información extraída de las muestras y, tanto si se aceptan como si se rechazan, se puede cometer un error. La hipótesis formulada con intención de rechazarla se llama hipótesis nula y se representa por H_0 . Rechazar H_0 implica aceptar una hipótesis alternativa (H_1).

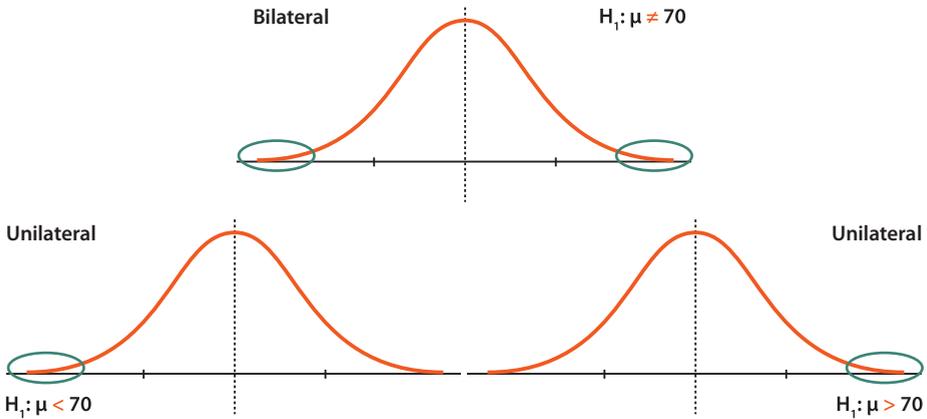
Así, por ejemplo, sabemos que la edad media de los individuos con la enfermedad X es 70 años. El contraste de hipótesis a formular sería:

$$H_0: \mu = 70$$

$$H_1: \mu \neq 70$$

Este tipo de contrastes se conoce como contraste bilateral. Si por ejemplo, en nuestra hipótesis alternativa quisiéramos contrastar si los individuos con la enfermedad x tienen una edad inferior a 70 años, nuestro contraste sería unilateral.

Figura 1. Tipos de contrastes.



La decisión de utilizar un contraste de hipótesis unilateral o bilateral dependerá de la información que tengamos sobre los datos que estamos analizando. En el caso de no tener ningún tipo de información, optaremos por un contraste bilateral.

Posibles resultados en un contraste de hipótesis

Por lo tanto, la hipótesis nula (H_0) es lo que contrastamos, los datos pueden refutarla y no debemos rechazarla sin una buena razón. La hipótesis alternativa (H_1) es siempre contraria a H_0 , los datos muestran evidencia a favor y no debe ser aceptada sin tener evidencia a favor.

La tabla 1 muestra las situaciones que se pueden dar en un contraste de hipótesis:

Tabla 1. Aciertos y errores al realizar un contraste de hipótesis.

	Ho Cierta	Ho Falsa
Rechazamos H0	Error tipo I (α)	Ok
No rechazamos H0	Ok	Error tipo II (β)

Error tipo I= P(Rechazar H0/Ho cierta)

Error tipo II= P(Aceptar H0/Ho falsa)

Potencia=1- β =P(Rechazar H0/Ho falsa)

Los errores tipo I y tipo II están inversamente relacionados, de manera que disminuir uno supondría aumentar el otro tipo de error.

El error de tipo I se conoce como nivel de significación. Típicamente se elige un valor pequeño, 5% o 1%. Todo experimento, en su definición y antes de elegir las muestras, debe llevar descrito cuál es el criterio con el que rechazaremos una hipótesis. Esto se traduce en fijar el nivel de significación del contraste.

El error de tipo II consiste en no rechazar H0 cuando es falsa. Este error se produce porque el tamaño muestral sea pequeño.

La probabilidad de que no ocurra el **error de tipo II**, cuando ésta puede calcularse, se denomina **potencia del contraste**. Es una medida de la habilidad de un contraste para detectar un efecto que está presente. Normalmente los programas estadísticos solamente hacen referencia a la significación de un contraste y no a la potencia.

¿Cómo decidimos si rechazamos o no H0? Significación estadística (P)

Para decidir si rechazamos H0 o no, necesitamos conocer la probabilidad (p) de que nuestros datos estén bajo el supuesto de H0. Si el valor de esta probabilidad es muy bajo, podríamos pensar que es muy raro encontrar tales datos si H0 fuera cierta. En cambio, si el valor de esta probabilidad fuera muy alto, no tendríamos ningún argumento en contra de H0 y no la podemos rechazar. Ahora la pregunta sería ¿qué valor de la probabilidad es alto y qué valor es bajo? El límite que se suele usar es 0,05, así con valores de $p < 0,05$ rechazamos H0 y decimos que el resultado es estadísticamente significativo, por lo tanto nos quedaríamos con H1. Por lo tanto la significación estadística es la probabilidad de tener una muestra que discrepe aún más que la muestra de H0. La significación es un indicador de la discrepancia entre una hipótesis nula y los datos muestrales. Cuanto más cercano sea a cero, más evidencia tenemos en contra de la hipótesis nula.

Supongamos que no encontramos diferencias significativas ($p > 0,05$) entre dos tratamientos que están siendo comparados. ¿Podríamos pensar que ambos tratamientos son iguales? La forma de expresar este resultado sería: *No se encuentran evidencias para rechazar H0*. Lo que queremos decir es que un test que no sea significativo nunca demuestra

que la hipótesis nula sea verdadera. Al obtener un valor de p no significativo, debemos fijarnos en el tamaño muestral, ya que un tamaño muestral pequeño reduce la potencia del test y no suele encontrar diferencias significativas.

Contrastes de hipótesis más utilizados

Los contrastes de hipótesis pueden ser paramétricos y no paramétricos. Los contrastes paramétricos son aquellos cuyos parámetros se basan en una distribución de probabilidad teórica conocida, que cumple los supuestos de normalidad y homocedasticidad (homogeneidad de varianzas). La forma de comprobar estos supuestos es:

- Normalidad: test de Kolmogorov-Smirnov, Shapiro Wilks y gráficos QQ normal. Si el tamaño muestral es mayor que 30, asumimos normalidad.
- Homocedasticidad: Test de Levene.

En cambio, los contrastes no paramétricos no asumen como conocida la distribución que sigue la población (no podemos suponer que siguen una distribución normal), por ello, no es necesario especificar la distribución de probabilidad de la población analizada. Estas técnicas se suelen usar cuando el tamaño muestral es pequeño ($n < 30$).

Según el tipo de variables que queramos comparar los test a realizar son:

Comparamos dos variables categóricas:

1. Contraste Ji-Cuadrado: en este contraste la hipótesis nula de que las variables a analizar sean independientes frente a que no lo sean. Para ser más exacto, este test trata de decidir si el número de resultados observados en cada clase corresponde aproximadamente con el número esperado. Para examinar el tipo de relación entre este tipo de variables usamos las tablas de contingencia.
2. Prueba exacta de Fisher: tiene la misma utilidad que la Ji-Cuadrado. Se utiliza cuando las frecuencias esperadas son inferiores a 5. En SPSS, como ya veremos en el tema siguiente, se calcula de la misma manera que el contraste de la Ji-Cuadrado.
3. Test de McNemar: lo usamos cuando a un único grupo de pacientes en el que se han realizado dos mediciones repetidas. Por lo tanto, cada paciente funciona como su propio control.

Variable dependiente: cuantitativa y variable independiente: categórica.

1. t-Student:
 - Para dos muestras independientes: permite comparar las medias para dos muestras distintas, tanto si cumplimos la homogeneidad de varianzas, como si no hay homogeneidad de varianzas (test t-Student, aproximación de Welch). Suponga-

mos que queremos comparar el nivel de colesterol de los hombres y las mujeres de una determinada población.

- Para dos muestras relacionadas: permite comparar las medias para dos muestras relacionadas. Supongamos que queremos analizar si el nivel de colesterol después de tomar un medicamento disminuye. En este caso diremos que los datos están relacionados.
2. Análisis de la Varianza (ANOVA): permite comparar las medias para más de dos grupos independientes. En este caso, la variable dependiente es cuantitativa y la variable independiente es el factor de agrupación. Supongamos que queremos medir la eficacia de cinco dosis de fármacos en el nivel de colesterol. Estamos comparando que las medias de los niveles de colesterol en las cinco mediciones sean iguales, frente a que al menos dos mediciones sean distintas, el contraste quedaría:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

$$H_1: \exists i, j / \mu_i \neq \mu_j$$

Como ya hemos comentado, para realizar los contrastes paramétricos debemos comprobar los supuestos de normalidad y homocedasticidad (test Kolmogorov y test de Levene, deben ser no significativos, $p > 0,05$). Para el caso de la ANOVA, lo que tiene que distribuirse como una Normal son los residuos (diferencia de cada valor y la media de su grupo).

3. U de Mann Whitney: es el análogo no paramétrico a la prueba t-Student para dos muestras independientes. Lo usamos cuando comparamos dos muestras y alguna de ellas tiene un tamaño menor a 30.
4. Wilcoxon: es el análogo no paramétrico a la prueba t- Student para dos muestras relacionadas.
5. Kruskal-Wallis: es el test análogo no paramétrico al análisis de la varianza. En este test, lo que realmente se compara no es la media sino las medianas. Lo usaremos cuando tengamos variables tipo ordinal, una muestra pequeña o no se cumpla el supuesto de normalidad de los residuos, ya que al ser un método no paramétrico no necesita los requisitos de los métodos paramétricos.

Variable dependiente: Cuantitativa y Variable independiente: Cuantitativa.

1. Correlación de Pearson (r): es un índice que mide la intensidad de dos variables cuantitativas, es un método paramétrico, por tanto necesita cumplir el supuesto de normalidad. Toma valores de -1 a 1. Los valores que puede tomar son:
 - $r=0$ No existe correlación lineal entre ambas variables.
 - $r > 0$ La correlación lineal es positiva. Cuando aumenta una variable la otra variable aumenta.

- $r < 0$ La correlación es negativa lineal. Cuando aumenta una variable la otra disminuye.
2. Correlación de Spearman: es un estimador no paramétrico, utilizado cuando las variables no cumplen el supuesto de normalidad, por ejemplo variables ordinales. A diferencia del coeficiente de Pearson, nos da una idea del grado de asociación de cualquier tipo, no solamente lineal.
 3. Regresión lineal Simple: este método lo usaremos cuando queramos estudiar la relación entre dos variables cuantitativas. Mediante los coeficientes de correlación podemos conocer el grado de asociación. Pero en ocasiones puede que nos interese encontrar una relación funcional entre las variables y saber con cuánta precisión podemos predecir el valor de la variable dependiente a partir de las variables independientes. De esta manera podremos llegar a predecir cuál va a ser el valor de una variable una vez que sabemos el valor de la otra. La recta de la regresión lineal es:

$$y = a + bx$$

Donde y es la variable dependiente y x la variable independiente. Los coeficientes a y b se calculan mediante el método de mínimo de cuadrados. El valor de a es la ordenada en el origen y el valor de b es la pendiente de la recta. El cálculo de b es el objetivo más importante del análisis de regresión lineal, mide el cambio de la variable y por cada unidad de aumento de la variable x . El signo de b puede ser positivo o negativo, para este último caso veríamos cómo decrece y a medida que aumenta x .

Al realizar un análisis de regresión simple obtendremos el valor del coeficiente de determinación que denotamos como R^2 , los valores de este coeficiente van desde 0 a 1. Cuanto más próximo a uno mayor es la fuerza de asociación de las variables.

Se interpreta como: *el porcentaje de variabilidad total de la variable dependiente y que es explicado por la variable x .*

La relación entre el coeficiente de determinación y el coeficiente de correlación de Pearson, es: $\sqrt{R^2} = r$. Así, mientras que R^2 es positivo, r puede tomar valores positivos o negativos, tendrá el mismo signo que la pendiente de la recta (b).

Bibliografía

- Apuntes de bioestadística. Universidad de Málaga: <http://www.bioestadistica.uma.es/baron/apuntes/>
- Borrás F. *Cuadernos de bioestadística* (1995). Cuadernos de bioestadística.
- Clifford R C, Taylor R.A. *Bioestadística* (2008).
- Martínez González M.A, De Irala Estévez J, Faulín Fajardo F.J. *Bioestadística amigable* (2001).
- Sokal R, Rohlf R. *Introducción a la bioestadística* (1980).
- Unidad de bioestadística del Hospital Ramón y Cajal: http://www.hrc.es/investigacion/inves_unidadbio.htm

13

Programa Estadístico SPSS

MARI CARMEN VIVO MOLINA

Licenciada en Ciencias y Técnicas Estadísticas.
Profesora asociada de la Universidad Miguel Hernández.
Dirección General de Planificación, Ordenación Sanitaria
y Farmacéutica e Investigación. Consejería de Sanidad
y Política Social de la Región de Murcia.

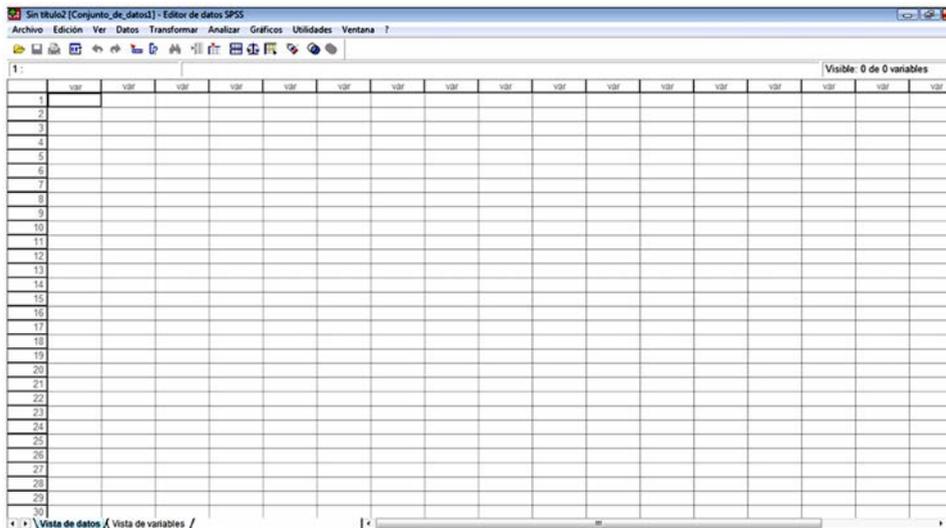
1. Introducción

En este capítulo realizaremos, por un lado, una breve introducción de los aspectos más básicos pero necesarios para la realización de un análisis estadístico, como son: creación de bases de datos, definición de variables aleatorias, transformación y recodificación de variables. Posteriormente, se darán las instrucciones necesarias para realizar los análisis vistos en el tema anterior.

2. Primeras nociones del programa (SPSS)

SPSS (*Statistical Package for Social Science*), es uno de los paquetes estadísticos más conocidos debido a su presentación amigable. La ventana que se nos abre al empezar a trabajar con SPSS está formada por celdas, donde cada fila hará referencia a un paciente y cada columna a cada una de las variables que hemos medido y que posteriormente vamos a analizar.

Figura 1. Ventana inicial SPSS.



Lectura de datos

Cuando queramos abrir un fichero de datos, desplegamos Archivo Abrir Datos, en el menú inicial de SPSS. Podemos abrir directamente un archivo en formato .sav que es el formato de datos de spss o en formato xls en excel.

Editor de datos

En la ventana principal de SPSS, abajo a la izquierda hay dos pestañas: Vista de datos y Vista de variables. En la vista de datos, visualizamos la base de datos con la que vamos a trabajar. En la vista de variables nos permite cambiar el nombre, tipo... de las variables con las que vamos a trabajar. En la casilla nombre, ponemos el nombre de las variables que queremos que salgan en los resultados, es aconsejable no dejar espacios y no poner caracteres raros. En la celda tipo, aparece un cuadro con tres puntos, al pinchar se abre una ventana en la que podemos seleccionar el tipo de variable con la que estamos trabajando, la anchura de la celda y los decimales que vamos a usar. Por defecto SPSS tiene una anchura de 8 y 2 decimales.

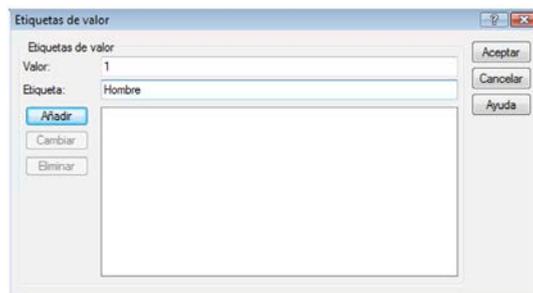
Figura 2. Descripción tipo de variable.



En la celda Etiqueta, podemos poner lo que a nosotros nos sirva para reconocer la variable, esto nos sirve a nosotros para reconocer las variables, porque lo que aparecerá en los resultados es el nombre que hemos puesto a la variable.

En valores, nos sale otra vez un cuadro con tres puntos que al pinchar se nos abre una ventana de etiquetas de valores. Las etiquetas son útiles cuando queremos identificar mediante etiquetas los valores de una variable. Por ejemplo, la variable Sexo, tenemos 1 y 2, sabemos que 1 es hombre y 2 mujer.

Figura 3. Etiquetas de valor.



Transformación de datos

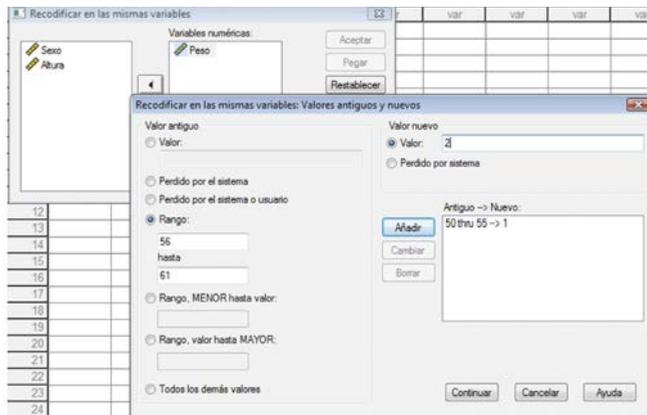
A veces necesitamos obtener una nueva variable, calculada a partir de variables que tenemos en la base de datos. Transformar \square Calcular variable. En la variable destino pondremos el nombre de la nueva variable que queremos calcular y en el cuadro de expresión numérica la fórmula que vamos a calcular.

Recodificación

Lo utilizamos cuando queremos agrupar los valores de una variable o renombrar los valores de una variable ya existente. Transformar \square Recodificar:

- En las mismas variables: nos permite recodificar los valores de una variable siendo la variable de destino la misma. Transformar \square Recodificar en las mismas variables. Seleccionamos la variable que queremos cambiar y pinchamos en el recuadro Valores antiguos y nuevos. Así, por ejemplo, supongamos que tenemos una variable peso y queremos recodificarla por intervalo, seleccionaríamos Rango y pondríamos el intervalo que queremos que nos recodifique. Como podemos ver, queremos que nos ponga en un intervalos aquellas personas que tienen un peso [50-55] y que no los recodifique con el valor 1. Luego le daríamos a añadir y seguiríamos el procedimiento.

Figura 4. Recodificación en las mismas variables.



- En distintas variables: nos permite recodificar los valores siendo la variable de destino una diferente a la que queremos recodificar. Para este caso, como vamos a crear una variable nueva, nos pide el nombre de la variable resultado. Es importante, después de poner la variable resultado, pinchar en el botón cambiar. Y a continuación, nos aparecerá la misma ventana de la recodificación anterior.

Figura 5. Recodificación en distintas variables.



Filtrar Datos

En ocasiones, puede que necesitemos trabajar solamente con una parte de nuestra base de datos. Datos \square Seleccionar Casos \square Si satisface la condición \square Si. Por ejemplo, queremos analizar las variables para los mayores de 15 años: $\text{Edad} > 15$. Si además de ser mayores de 15 años, queremos que sean mujeres, pondríamos $\text{Edad} > 15 \ \& \ \text{Sexo} = 2$, como muestra la figura 6.

Figura 6. Seleccionar casos.



Estadística Descriptiva con SPSS

Analizar \square Estadísticos Descriptivos \square Frecuencias. Aquí seleccionamos la variable con la que vamos a trabajar y a continuación pulsamos Estadísticos y nos aparecen todas las medidas de tendencia central, medidas de dispersión y medidas de posición. Para el cálculo de percentiles podemos elegir qué percentiles queremos que nos calcule. De manera que si queremos que nos calcule el percentil 50, seleccionamos Percentiles y ponemos en el cuadro de al lado, el valor que queremos que nos calcule (50).

Figura 7. Calculo Estadísticos.



Representaciones Gráficas: Gráficos □ Cuadros de diálogo antiguos. Se despliegan todas las posibilidades de gráficos que hay en SPSS.

Inferencia con SPSS

Intervalo de confianza para la media: Analizar □ Estadísticos Descriptivos □ Explorar. Seleccionamos en Dependientes la variable cuantitativa de la que queremos calcular el intervalo de confianza de la media. En estadísticos, podemos elegir el nivel de confianza, por defecto está 95%.

Test Ji-Cuadrado. Supongamos que queremos realizar un estudio sobre el cáncer de pulmón y su relación con el hábito de fumar. En este caso nuestra hipótesis nula sería que la incidencia de cáncer de pulmón es la misma en los fumadores que en los no fumadores. Tenemos dos variables, una será fumador, con dos categorías si o no, según fume o no, y otra variable que será cáncer con también dos categorías, según desarrolle o no el cáncer.

Analizar □ Estadísticos □ Tablas de Contingencia □ Estadísticos □ Chi-Cuadrado.

Figura 8. Tabla de contingencia.

Tabla de contingencia Cáncer * Fumador

Recuento		Fumador		Total
		Si	No	
Cáncer	Si	11	2	13
	No	1	11	12
Total		12	13	25

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	14,547 ^a	1	,000		
Corrección por continuidad ^b	11,652	1	,001		
Razón de verosimilitudes	16,571	1	,000		
Estadístico exacto de Fisher				,000	,000
Asociación lineal por lineal	13,965	1	,000		
N de casos válidos	25				

a. Calculado sólo para una tabla de 2x2.

b. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 5,76.

En primer lugar, SPSS nos ofrece una tabla de contingencia que relaciona ambas variables. Los valores observados son los que aparecen en cada una de las casillas. SPSS también ofrece la posibilidad de calcular los porcentajes de cada una de las celdas. La siguiente tabla nos muestra: el valor del estadístico de Ji-Cuadrado (14,54) y el valor de p para este contraste (0,00), también muestra el estadístico exacto de Fisher, en este caso, nos ofrece el valor de p para el contraste unilateral y el bilateral (ambos 0,00). En general, el estadístico de Fisher es más exacto que el test de la Ji-Cuadrado y se puede utilizar siempre. Mientras que el test de la Ji-Cuadrado solamente cuando las frecuencias esperadas sean superiores a 5. Así que al tener un valor de $p < 0,05$ (significación bilateral asociada al estadístico exacto de Fisher), rechazaríamos H_0 y podríamos afirmar que la incidencia de cáncer difiere entre fumadores y no fumadores. Recordar que, aunque nos den dos valores de significación para el test de Fisher (unilateral y bilateral), estamos realizando un contraste bilateral.

Test de McNemar: prueba no paramétrica para dos muestras relacionadas. De la misma forma que hemos calculado el test de la Ji-Cuadrado, en el apartado Estadísticos, marcamos la opción McNemar. Al aplicar este test, obtendremos una tabla de contingencia y otra tabla con el valor de p (Sig.exacta(bilateral)). Este test también se puede obtener Estadísticos Pruebas no paramétricas 2 muestras relacionadas. Seleccionamos las dos variables que queremos comparar.

Test t-Student. Antes de aplicar este test, debemos comprobar para la variable cuantitativa o dependiente:

- Normalidad: en el caso de que los tamaños muestrales no sean superiores a 30, aplicaremos el test de Kolmogorov Smirnov: Analizar Pruebas no paramétricas Ks de 1 muestra. Si el valor de p (Sig. asint.bilateral) es mayor que 0,05, no podemos rechazar la hipótesis de que nuestra variable se distribuye como una normal. En caso contrario, podemos optar por la transformación de la variable (no lo hemos visto en este módulo) o por aplicar su análogo no paramétrico.
- Homocedasticidad (solo en el caso de dos muestras independientes): en este caso usamos el test de Levene.

Test t-Student para muestras independientes: suponiendo que la variable dependiente cumple el supuesto de normalidad. Analizar Comparación de medias t muestras independientes.

En el cuadro contrastar variables, pondremos la variable dependiente, y en variable de agrupación, meteremos la variable independiente dicotómica. Así, supongamos que queremos comparar los niveles de colesterol entre hombres y mujeres. El nivel de colesterol irá en el cuadro contrastar variables y la variable de agrupación será la variable sexo. Los resultados que nos devuelve el programa serán en forma de tabla llamada prueba de muestras independientes, donde nos informa del resultado de la homogeneidad de

varianzas: test de Levene. Si el valor de significación estadística es menor que 0,05 consideramos que las varianzas no son iguales, en caso contrario ($p > 0,05$), asumimos homocedasticidad de varianzas. Según los resultados de este test, observaremos la primera fila (se han asumido varianzas iguales, $p > 0,05$) o, en caso contrario, la segunda fila (no se han asumido varianzas iguales $p < 0,05$). La siguiente tabla muestra la comparación del peso medio de dos poblaciones independientes: hombres y mujeres. Asumimos que las varianzas son iguales, ya que el test de Levene tiene una $p > 0,05$ (0,577). A continuación, la comparación de medias de pesos tiene una $p < 0,05$ (0,002), por lo tanto existen diferencias significativas entre el peso de hombres y mujeres. Un resultado muy importante es el intervalo de confianza asociado a la diferencia de medias, ya que nos informa en que sentido va la diferencia, la magnitud, y si el estudio tiene o no potencia, ya que cuanto más ancho sea el intervalo, menos potencia.

Figura 9. Prueba t-Student para muestras independientes.

		Prueba de muestras independientes									
		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias						95% Intervalo de confianza para la diferencia	
		F	Sig.	t	gl.	Sig. (bilateral)	Diferencia de medias	Error tip. de la diferencia	Inferior	Superior	
Peso	Se han asumido varianzas iguales	.338	.577	4,375	8	.002	17,58333	4,01873	8,31613	26,85054	
	No se han asumido varianzas iguales			4,654	7,762	.002	17,58333	3,77804	8,82444	26,34223	

Test t-Student para muestras relacionadas: supongamos que queremos medir el peso antes y después de seguir una dieta y queremos comprobar si existen diferencias. El procedimiento en SPSS será igual que en el caso de muestras independientes pero en este caso seleccionamos muestras relacionadas. Obtendríamos una tabla para los estadísticos de las muestras y los resultados de la prueba, con una media de la diferencia, una desviación típica, un intervalo de confianza que nos da información sobre la potencia del estudio y el valor de p. Si el valor $p < 0,05$, quiere decir que existen diferencias significativas en las medias del peso antes y después de seguir la dieta.

Prueba U de Mann-Whitney para dos muestras independientes: análogo de la prueba t-Student para dos muestras independientes. Lo usamos cuando el tamaño muestral es inferior a 30 o cuando tenemos una variable en escala ordinal en vez de cuantitativa. Por ejemplo comparar cómo dos tipos de dieta (1 y 2) ayudan a reducir peso. Tenemos que la variable reducción de peso está codificada en tres casos: 0 no hay reducción de peso, 1 reducción moderada y 2 alta reducción de peso. Analizar Pruebas no paramétricas dos muestras independientes. En este caso, obtendremos la significación unilateral y bilateral. Si solamente queremos analizar si existen diferencias entre ambos grupos, nos fijaremos en la significación bilateral. Al realizar este test no obtenemos un intervalo de confianza, como en el caso paramétrico.

Test de Wilcoxon para datos relacionados: análogo de la prueba t-Student para dos muestras relacionadas. Al igual que el test U de Mann-Whitney, lo usamos cuando el tamaño muestral es inferior a 30 o cuando tenemos una variable en escala ordinal en vez de cuantitativa. Analizar Pruebas no paramétricas Dos muestras relacionadas.

Análisis de la varianza (Anova): supongamos que queremos comparar si existen diferencias en las medias de la hipertensión de unos pacientes en función de cinco tratamientos recibidos. Estamos comparando más de dos muestras. El procedimiento en SPSS es el siguiente: Analizar Pruebas no paramétricas Anova de un factor. En opciones, pulsaríamos la opción Contraste de homogeneidad, ya que para aplicar este contraste los datos deben ser homocedásticos. Para el caso de que no podamos suponer varianzas iguales, seleccionaremos en opciones dentro de Anova de un factor, los test para la comparación de medias de Welch o el test de Brown-Forsythe.

Los resultados obtenidos, para el ejemplo de los niveles de hipertensión en función de cinco tratamientos aplicados, se encuentran en la siguiente figura:

Figura 10. Test Anova.

hipertension		ANOVA			
	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	1295,083	4	323,771	3,643	,011
Intra-grupos	4887,900	55	88,871		
Total	6182,983	59			

Con un valor de $p=0,011$, podemos afirmar que existen diferencias significativas en las cinco medias.

Una vez que sabemos que hay diferencias entre los tratamientos, nos interesa saber ¿en qué tratamientos hay diferencias? Para ello, en la opción Post-hoc hay una relación test para comparación múltiples, no se recomienda pedirlos todos porque cada uno de ellos tiene unas características, pero la explicación de cada uno de estos tests excede de los objetivos de este módulo.

Test de Kruskal-Wallis: el análogo no paramétrico a la tabla de Anova. Al igual que los demás test no paramétricos lo usamos cuando el tamaño muestral es inferior a 30 o cuando tenemos una variable en escala ordinal en vez de cuantitativa: Analizar Pruebas no paramétricas Comparación de k muestras independientes.

Correlación de Pearson y Spearman: Analizar Correlaciones Bivariadas. Seleccionamos las variables, para el caso del coeficiente de correlación de Pearson, el SPSS lo tiene marcado por defecto, mientras que si queremos calcular el de Spearman, lo tenemos que marcar. Las tablas de resultados son similares para ambos coeficientes, para cada una de las comparaciones que realiza nos da el valor del coeficiente y el valor de p. Para la compa-

ración de una variable consigo misma, está claro que el valor del coeficiente será 1 y para este caso no calcula el valor de significación (p).

Regresión Lineal Simple: Lo primero que debemos hacer es un gráfico de dispersión, para ver si tiene sentido plantearnos un modelo lineal entre dos variables. Gráficos □ Cuadro de diálogo antiguo □ Dispersión. Seleccionamos la variable dependiente en el eje Y, y la variable independiente en el eje X. Una vez que tenemos el gráfico en el visor de resultados pinchamos dos veces en el gráfico y se nos abre el editor de gráficos. Pinchamos en la opción señalada en rojo, para hacer el ajuste lineal.

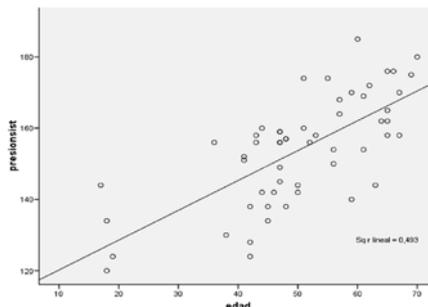
Figura 11. Regresión lineal simple.



Una vez que hemos visto el diagrama de dispersión vamos a realizar la regresión lineal simple. Como ya vimos en el tema anterior lo que buscamos es encontrar la recta que mejor se ajuste al diagrama de dispersión. Para ello tenemos que estimar los valores de a y b de la recta: $y=a+bx$. El procedimiento en SPSS, será: Analizar □ Regresión lineal □ Seleccionamos la variable dependiente y la independiente. En estadísticos seleccionamos: Intervalos de confianza, en Guardar seleccionamos Residuos no tipificados y en Gráficos: Histograma y Gráficos de probabilidad normal.

Supongamos que queremos analizar la relación entre la presión sistólica y la edad. En primer lugar hacemos el gráfico de dispersión para ver si percibimos algún tipo de relación lineal (el que no veamos relación lineal no quiere decir que no estén relacionadas, puede que la relación entre variables sea de otro tipo, cuadrática por ejemplo). Al observar el gráfico sí que percibimos una relación lineal positiva, conforme aumenta la edad aumenta la presión sistólica.

Figura 12. Diagrama de dispersión.



La primera tabla que obtenemos al realizar la regresión lineal simple es la siguiente (véase Figura 18):

Figura 13. Resumen del modelo de regresión lineal.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,702 ^a	,493	,484	10,937

a. Variables predictoras: (Constante), edad

b. Variable dependiente: presionsist

El valor de R se refiere al coeficiente de correlación de Pearson (0,702). SPSS lo presenta siempre sin signo, aunque puede ser negativo. El R cuadrado (0,493) es el coeficiente de determinación (recordar que el valor del coeficiente de correlación de Pearson al cuadrado es igual al coeficiente de determinación). El valor de R cuadrado se explicaría como que el 49,3% de la variabilidad de la variable presión sistólica queda explicado por la variable edad. Teniendo en cuenta que toma valores de 0 a 100, no es un valor muy bueno.

La siguiente tabla es la de la Anova de regresión:

Figura 14. Tabla Anova.

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	6175,385	1	6175,385	51,628	,000 ^a
	Residual	6339,452	53	119,612		
	Total	12514,836	54			

a. Variables predictoras: (Constante), edad

b. Variable dependiente: presionsist

Aquí destacamos el valor de la significación estadística $p < 0,05$, por lo tanto podemos afirmar que hay relación significativa entre las dos variables.

En la figura 15 podemos destacar el valor del coeficiente a de la recta de regresión que para nuestro ejemplo vale 111,834. Y el valor de debajo que corresponde con la pendiente de la recta o valor b de la recta (0,837). La siguiente columna son los errores de dichos valores. El valor de t y la significación hacen referencia al contraste de hipótesis, en el cual, la hipótesis nula indica que el coeficiente correspondiente (a o b) vale cero, frente a la hipótesis alternativa de que el coeficiente sea diferente de cero. En ambos casos, rechazamos H_0 , por lo tanto, los coeficientes toman valores distintos de cero. Y por último, tenemos información de los intervalos de confianza asociados a los dos coeficientes de la recta.

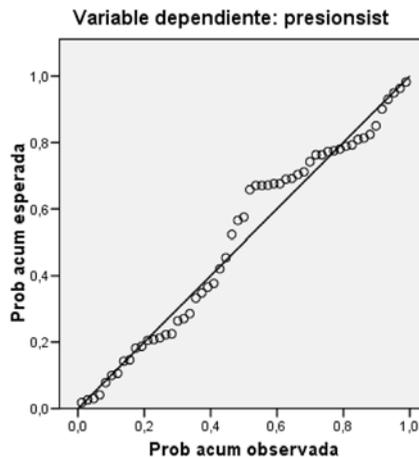
Figura 15. Coeficientes recta de regresión.

Modelo	Coeficientes ^a						Intervalo de confianza para B al 95%	
	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	Límite inferior	Límite superior	
	B	Error tip.						
1								
(Constante)	111,834	6,043		18,505	,000	99,712	123,956	
edad	,837	,117	,702	7,185	,000	,604	1,071	

a. Variable dependiente: presionsist

Como hemos pedido que nos guarde los errores tipificados, en la base de datos nos aparece una variable llamada RES_1, que hace referencia a los residuos tipificados. Con esta variable comprobamos el supuesto de normalidad de la regresión, mediante la prueba de Kolmogorov- Smirnov y el gráfico pp. Para comprobar la normalidad con el gráfico pp, todos los puntos (o la gran mayoría) deben de caer encima de la diagonal. Además, con el test de Kolmogorov-Smirnov obtenemos un valor de p no significativo (0,169). Por lo tanto, asumimos que los residuos se distribuyen de forma normal.

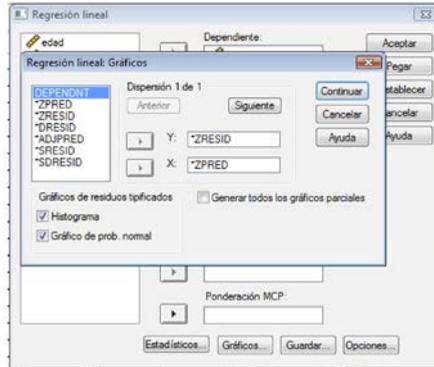
Figura 16. Gráfico pp de Normalidad.



La recta de regresión que conseguimos es: presión= 118,834+ edad*0,837.

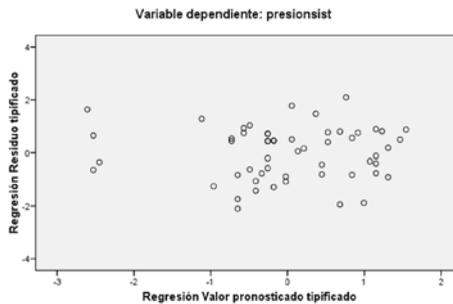
Otra forma de ver lo adecuado del modelo de regresión lineal es comprobar si los residuos normalizados siguen una distribución homogénea en un diagrama de dispersión. Para ello, volveríamos al análisis de regresión lineal y en Gráficos, pondríamos:

Figura 17. Ventana de SPSS para el cálculo de homogeneidad de varianzas residuales.



En el caso de que los residuos normalizados tuvieran varianzas homogéneas deberíamos obtener los puntos del diagrama de dispersión dispuestos de forma aleatoria, sin presentar ningún tipo de tendencia dentro de una banda horizontal.

Figura 18. Gráfico de dispersión de residuos.



Bibliografía

- Apuntes de bioestadística. Universidad de Málaga: <http://www.bioestadistica.uma.es/baron/apuntes/>
 Borrás F, *Cuadernos de bioestadística* (1995). Cuadernos de bioestadística.
 Clifford R C, Taylor R.A. *Bioestadística* (2008).
 Martínez González M.A, De Irala Estévez J, Faulín Fajardo F.J. *Bioestadística amigable* (2001).
 Sokal R, Rohlf R. *Introducción a la bioestadística* (1980).
 Unidad de bioestadística del Hospital Ramón y Cajal: http://www.hrc.es/investigacion/inves_unidadbio.htm

14

Orígenes, definición y fundamentos de la Práctica Clínica Basada en la Evidencia (PCBE)*

* La elaboración de este capítulo es una adaptación de: Sánchez Sánchez JA, Aguinaga Ontoso E, Navarro Mateu F. Introducción a la Atención Sanitaria Basada en la Evidencia. Grupo de Atención Sanitaria Basada en la Evidencia. Atención Sanitaria Basada en la Evidencia. Su aplicación a la Práctica Clínica. Consejería de Sanidad. Murcia. 2007. P. 16-46.

JUAN ANTONIO SÁNCHEZ SÁNCHEZ¹

ENRIQUE AGUINAGA ONTOSO²

1) Médico Especialista en Medicina de Familia y Comunitaria.
Centro Tecnológico de Información y Documentación
Sanitaria. Consejería de Sanidad y Política Social
de la Región de Murcia.

2) Médico Especialista en Medicina Preventiva y de Salud
Pública. Centro Tecnológico de Información y
Documentación Sanitaria. Consejería de Sanidad
y Política Social de la Región de Murcia.

1. ¿Qué es la práctica clínica basada en la evidencia/medicina basada en la evidencia?

La necesidad de sistematizar la revisión de la literatura científica para extraer conclusiones válidas, la falta de adecuación de las fuentes tradicionales de conocimiento, sumadas a la escasez de tiempo e imposibilidad del profesional de la salud a mantenerse actualizado en su campo de actuación están en el origen de un movimiento, que surge desde los profesionales clínicos y que trata de incorporar a la toma de decisiones los hallazgos más válidos derivados de los estudios clínicos y epidemiológicos. El término de Medicina Basada en la Evidencia (MBE) fue acuñado en 1992 en la Facultad de Medicina de la Universidad canadiense de Mc Master, por los mismos autores que habían difundido y aplicado en la docencia e investigación los conceptos de Epidemiología Clínica¹.

Más tarde, D. Sackett² definió la MBE como «La utilización consciente, explícita y juiciosa de la mejor evidencia científica disponible para tomar decisiones sobre el cuidado de los pacientes individuales». De acuerdo a los que proponen el término y le otorgan una importancia esencial, su práctica implica la integración de la experiencia clínica profesional, de las mejores evidencias científicas disponibles y con los valores y circunstancias únicas de nuestros pacientes³. Y hacerlo con un espíritu abierto de curiosidad y de escepticismo, extrayendo de la literatura científica las conclusiones más sólidas y contrastándolas con algunas prácticas tradicionales de la medicina, basadas en principios de autoridad de líderes académicos o científicos.

Posteriormente, este enfoque, este procedimiento de trabajo y de aproximarse al conocimiento científico para ser aplicado a los pacientes, se ha ampliado y aplicado a otras disciplinas, siendo el término de Práctica Clínica Basada en la Evidencia (PCBE) más global y abarcando a diferentes profesionales de la salud.

A la PCBE se le puede considerar como una aproximación coherente a la toma de decisiones clínicas. Hay autores que la definen como la aplicación del método científico para definir el tratamiento óptimo del paciente individual.

1.1. Lo que es y lo que no es la PCBE

Se puede afirmar sobre la PCBE que:

- No es algo imposible de practicar. Se trata de algo muy factible para ser realizada de forma cotidiana, hay experiencias de ello en diversas áreas de conocimiento (Atención Primaria, Psiquiatría, ...) y no está reservada a los centros de tercer nivel.
- No es un libro de cocina sobre medicina. No aporta recetas ni soluciones para todo.
- No es un caballo troyano que se introduce en el quehacer del profesional médico y le va a restar autonomía, sino un elemento más en la toma de decisiones con orientaciones explícitas. Es necesario relativizar lo que venga avalado por la evidencia, sin perder por ello la capacidad crítica.

- No trata de reducir los costes sino de buscar la mayor calidad en la atención individual. Los profesionales sanitarios que tratan de ejercer la PCBE no buscan una medicina barata, sino de calidad.
- No está restringida solo a los metaanálisis y a los ensayos clínicos para orientar decisiones terapéuticas. Permite también analizar la evidencia sobre otros aspectos como el diagnóstico o la etiología.

2. Cómo se practica la PCBE

2.1. Etapas en la práctica de la PCBE

Cuando se analizan las etapas de la práctica de la PCBE es útil distinguir las diferentes maneras que los clínicos incorporan la evidencia a su práctica clínica⁴: valorando críticamente la fuente original, utilizando las evidencias que previamente han evaluado fuentes fiables y basándose en fuentes actuales de ostensible calidad y autoridad.

La Práctica Clínica Basada en la Evidencia (PCBE), de forma clásica, se realiza en cinco etapas:

1. Formular preguntas clínicas, convertir la necesidad de información en una pregunta con respuesta (ver tema 16).
2. Realización de la búsqueda, seleccionar la fuente y buscar la mejor evidencia disponible con la mayor eficiencia (ver tema 17).
3. Lectura crítica de lo encontrado, valoración crítica de la evidencia para juzgar su validez, impacto y aplicabilidad.
4. Aplicación a la práctica clínica, integrar las evidencias con nuestra especialización clínica y los valores y circunstancias del paciente.
5. Autoevaluación, evaluar nuestra práctica clínica y de nuestra capacidad para ejecutar las 4 etapas y buscar maneras para mejorarlas.

Para incorporar la evidencia en su práctica, no es imprescindible que todos los profesionales desarrollen las cinco etapas y que cuenten con las competencias para ello; pero sí con las básicas que les permitan formular adecuadamente las preguntas clínicas, acceder a resúmenes de evidencia y a bases de datos específicas.

2.2. Nivel de evidencia y grado de recomendación

Una vez realizada la búsqueda recuperaremos diferentes tipos de estudios y dentro de cada tipo (por ejemplo, de todos los ensayos clínicos controlados) según la calidad del diseño (randomizado, doble ciego), variará la validez de las conclusiones.

El tipo de diseño de los estudios de investigación hace que éstos tengan diferentes probabilidades de incurrir en sesgos o errores sistemáticos que limiten su validez interna, permitiendo establecer según ello una jerarquía y una clasificación de los artículos científicos. La calidad o nivel de la evidencia indica hasta qué punto podemos confiar en que el estimador del efecto es correcto. La fuerza de una recomendación indica hasta qué punto podemos confiar en que poner en práctica la recomendación conllevará más beneficio que riesgo⁵.

En esta metodología se introduce la noción de que la evidencia científica se presenta de manera jerárquica y que algunos diseños de estudio están más sujetos a sesgos que otros y, por lo tanto, justifican en menor medida las decisiones clínicas.

2.2.1. Diferentes clasificaciones

La *"US Preventive Services Task Force"*, usando una metodología adaptada de la *"Canadian Task Force on the Periodic Health Examination"*, clasificó el nivel de validez de las evidencias de los distintos procedimientos utilizados para valorar una técnica diagnóstica o terapéutica y constituyeron las primeras clasificaciones utilizadas. Al graduar la calidad de la evidencia, se dio una importancia especial a aquellos diseños del estudio que por razones metodológicas están menos sujetos a error y sesgo. Así, en primer lugar, se situaron los metaanálisis y revisiones sistemáticas; luego, los Ensayos Clínicos Aleatorizados (ECA); y posteriormente estudios de cohortes, casos control y series temporales.

Con el paso del tiempo han ido apareciendo otras clasificaciones de la jerarquización de la evidencia, como las del *"Scottish Intercollegiate Guidelines Network"* (SIGN) y la de NICE: *Guideline Development Method* y hay recopiladas más de 100 hasta el momento actual. Aunque a nivel internacional no se ha consensuado una única clasificación y los distintos grupos que desarrollan o publican guías de práctica clínica basadas utilizan diferentes escalas, sólo difieren levemente unas de otras. Todas ellas se realizan en función del rigor científico del diseño del estudio y de su vulnerabilidad hacia los posibles sesgos inherentes a la investigación.

2.2.2. Elaboración de recomendaciones

Los grados de las recomendaciones que, además del nivel de evidencia, se encuentran en las Guías de Práctica Clínica tienen en cuenta otra serie de elementos, como el funcionamiento de los servicios sanitarios, el coste, las necesidades de formación, los recursos existentes y la valoración de los pacientes. Se suelen expresar en letras y suelen tener una correlación con los niveles de evidencia en que se basan.

Esta información ayuda a identificar las diferencias de la práctica clínica actual, la provisión de servicios y las experiencias de los pacientes, de tal manera que la orientación de las recomendaciones tenga el máximo impacto en los servicios y en los resultados. Expresa, en definitiva, la valoración de que los beneficios superan a los riesgos de su aplicación. Este proceso debe ser explícito, para que el juicio hecho a partir de las

evidencias sea lo más claro posible, pero expresa y los valores y creencias de los que recomiendan las recomendaciones. Aunque el nivel de evidencia será el mismo, el grado de recomendación puede variar según el contexto en que deba ser aplicada dicha recomendación.

2.2.3. La Clasificación GRADE

En el año 2000 y con el objetivo de abordar las deficiencias de los actuales sistemas de clasificación, personas de las diferentes instituciones que han desarrollado estas clasificaciones y/o elaboran Guías de Práctica Clínica crearon el grupo de trabajo «*Grading of Recommendations Assessment, Development and Evaluation*” (GRADE). Este grupo de trabajo ha desarrollado el sistema GRADE, el cual persigue hacer juicios más coherentes y que la comunicación de dichos juicios pueda respaldar opciones mejor fundamentadas en la atención sanitaria. Este sistema, actualmente en desarrollo, está ampliando su clasificación a estudios diagnósticos y está siendo adoptada por numerosas entidades e instituciones, empleándose en numerosas sociedades científicas al realizar sus recomendaciones o en sumarios de evidencia como *Uptodate* o *Clinical Evidence*.

El sistema de clasificación GRADE juzga la calidad de la evidencia para cada variable importante, teniendo en cuenta el diseño del estudio, la calidad, la consistencia y si la evidencia es directa. Para hacer juicios acerca de la fuerza de la recomendación considera el balance entre beneficios y riesgos, la calidad de la evidencia, la aplicabilidad y el riesgo basal de la población de interés.

3. Logros conseguidos por la PCBE

Las nuevas tecnologías y la aparición de Internet han revolucionado el acceso a la información y han transformado las bibliotecas de ciencias de la salud. El desarrollo de la medicina basada en la evidencia ha propiciado la aparición de nuevas fuentes de información que filtran y evalúan la investigación original en función de su calidad, relevancia y pertinencia clínica (ver tema 17).

El desarrollo de la PCBE ha supuesto importantes avances en la calidad metodológica de artículos originales que constituyen la base para tomar decisiones clínicas; en el desarrollo de metodología y listas de comprobación de calidad de los artículos (lectura crítica); en la agrupación y valoración de la calidad de los artículos: clasificación de la evidencia; y en los procedimientos para realizar revisiones sistemáticas y metaanálisis.

La colaboración Cochrane, con más de 4.750 revisiones sistemáticas realizadas, ha tenido una expansión considerable desde sus comienzos. Hay revistas que publican de forma periódica una evaluación crítica a la PCBE e incluso se han publicado libros de texto que analizan desde la perspectiva de la sociología, la historia, la repercusión y las críticas de la epidemiología clínica y de la MBE.

3.1. Ventajas para los Médicos y los pacientes de la PCBE

Son numerosas las ventajas que se han señalado para profesionales sanitarios y pacientes por la práctica de la PCBE y que resumimos:

Ventajas para los profesionales sanitarios:

- Facilita la integración de la educación médica con la práctica clínica y permite la actualización de conocimientos de forma rutinaria.
- Puede ser aprendida y practicada por los profesionales en cualquier nivel de su formación.
- El uso de la evidencia legitima a los clínicos. Mejora la comprensión de la metodología científica y el paso de meros lectores pasivos de artículos o libros médicos a realizar una lectura crítica que permita la elección de aquellos artículos metodológicamente correctos.
- Incrementa la confianza en las decisiones, que se toman tanto en el aspecto clínico como en el ámbito de la gestión sanitaria, reduciendo la incertidumbre en las decisiones del día a día.
- Aumenta la capacidad para utilizar fuentes de información bibliográfica y sus hábitos de lectura; así como las posibilidades actuales de búsqueda de información con las nuevas tecnologías de la información y comunicación.
- Facilita un marco de trabajo y discusión en equipo que permite reforzar la cohesión de los mismos.
- Facilita una mejor distribución de los recursos sanitarios.

Ventajas para los pacientes:

- Disminución de la variabilidad de la práctica clínica.
- Acceso a las prestaciones más eficaces.
- Ofrece una oportunidad de comunicación adicional y operativa con sus médicos al objetivar los pros y los contras de cada opción diagnóstica y/o terapéutica y vehicular su participación real en la toma de decisiones que le afecten.

4. Críticas, limitaciones y retos para el futuro

4.1. Críticas

En estos años de desarrollo de la PCBE, han sido numerosas las críticas recibidas. La principal, la más contundente y paradójica, ha sido el hecho de que no hay evidencia de que la práctica de la MBE sea más efectiva de conseguir mejores resultados en salud que la práctica habitual.

En los países en que la MBE se ha desarrollado más, ha tenido más repercusión en centros académicos que entre los clínicos. Se considera que se han realizado pocos esfuerzos por la integración en la práctica clínica.

4.2. Limitaciones en la práctica

Se necesita tiempo y entrenamiento para su aprendizaje y utilización de forma rutinaria e integrada en la práctica habitual. Se precisa una inversión importante en la implantación de Nuevas Tecnologías.

Se han realizado escasas evaluaciones de estudios económicos. Ninguna de las valoraciones de calidad ha incluido hasta ahora criterios de eficiencia ni requerimiento de estándares de estudios económicos.

A pesar del esfuerzo realizado en los últimos 20 años, el proceso de sintetizar la evidencia es lento y desalentador. Se ha estimado que sería preciso esperar hasta el año 2015 para producir las 10.000 revisiones que la Cochrane precisa para sintetizar la evidencia existente. Sin embargo, los clínicos precisan síntesis y revisiones de los miles de ensayos que son publicados cada año.

4.3. Retos para el futuro

Los ingredientes que recomienda David Sackett añadir a la lectura de su libro de Introducción a la MBE, para beneficio de los pacientes⁶, los consideramos imprescindibles y su aplicación anularía muchas de las críticas recibidas a este movimiento :

“Primero un dominio de las técnicas clínicas de la entrevista a pacientes, la recogida de la historia y la exploración física, sin las cuales no puede empezar el proceso de la MBE (mediante la generación de hipótesis diagnósticas) ni terminarse (integrando las evidencias válidas e importantes con los valores y expectativas de su paciente). Segundo, la práctica de un aprendizaje continuo, autodirigido y para toda la vida, sin el cual enseguida se quedaría peligrosamente desfasado. Tercero, la humildad necesaria para no mantenerse refractario tanto a la propia superación como a los avances de la medicina. Finalmente, deseamos que añada entusiasmo e irreverencia al empeño, ya que sin ellos se perdería toda la diversión que acompaña a la aplicación de estas ideas”.

Ninguna de las críticas encontradas aboga por el abandono de la práctica clínica basada en la evidencia y todas reconocen el avance que ha supuesto. La cuestión no resuelta es poder cuantificar lo que se ha modificado la práctica clínica y un aspecto pendiente sería cómo introducir la evidencia en la toma de decisiones clínicas. Es mucho lo que la MBE puede hacer por sustituir los datos por información, pero es necesario una visión más flexible, no dogmática, de lo que representa la evidencia

La tendencia, necesariamente, ha de ir dirigida a la integración con la práctica clínica. El mayor cambio esperado es la transferencia del conocimiento; asegurando que los clínicos cuenten con la evidencia actualizada para la práctica de cada día.

Los productos dirigidos a la práctica clínica basada en la evidencia son nuevos, en general bien contruidos y sometidos a un control de calidad internacional. Pero proporcionar evidencia no es suficiente para garantizar cambios en la práctica clínica: es algo necesario pero no suficiente. Los esfuerzos deben dirigirse a tratar de facilitar acceso inmediato a la mejor evidencia, en un formato amigable que sean hechos a medida de todos los implicados en las decisiones.

Como retos para el futuro se han señalado también la necesidad de incorporar la experiencia acumulada de los clínicos a los resultados aportados por las revisiones sistemáticas y a los de los Ensayos Clínicos, el desarrollo de la investigación cualitativa, y cómo medir lo que hacemos al intentar practicar la MBE y cómo utilizarlo mejor.

Por último queda mucho por trabajar en cómo incorporar los valores y expectativas de los pacientes, y cómo integrar toda la compleja información que se genera al tomar las decisiones clínicas, de una manera eficiente.

Bibliografía

1. Sackett D.L., Haynes R.B.; Guyatt G.H.; Tugwell P. *Epidemiología Clínica. Ciencia básica para la medicina clínica*. 2ª Edición. Editorial Médica Panamericana. Buenos Aires. 1994.
2. Sackett DI, Rosenberg WMC, Gary JAM, Haynes RB, Richardson WS. Evidence based medicine: what is it and what it isn't. *BMJ* 1996; 312:71-2.
3. Straus Sh., Richardson SW, Glasziou P., Haynes RB. *Medicina Basada en la evidencia. Cómo practicar y enseñar la MBE*. Elsevier. Tercera edición. Madrid. 2006.
4. Guyatt G, Rennie D, (editores). *Users' Guides to the Medical Literature. A Manual for Evidence-Based Clinical Practice*. Chicago: American Medical Association. 2002.
5. Marzo-Castillejo M y Alonso-Coello P. Clasificación de la calidad de la evidencia y fuerza de las recomendaciones. *Aten Primaria*. 2006; 37(1):40-50.
6. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Medicina Basada en la Evidencia. Cómo ejercer y enseñar la MBE*. Madrid: Churchill Livingstone España. 1997.

15

Identificación y focalización
de las preguntas que surgen
de la práctica diaria.

Diseño de preguntas susceptibles
de obtener respuestas aplicables
a la práctica diaria.

M^a CARMEN SÁNCHEZ LÓPEZ¹
ENRIQUE AGUINAGA ONTOSO²

1) Médico Especialista en Medicina de Familia y Comunitaria.
Centro Tecnológico de Información y Documentación
Sanitaria. Consejería de Sanidad y Política Social
de la Región de Murcia.

2) Médico Especialista en Medicina Preventiva y de Salud
Pública. Centro Tecnológico de Información y
Documentación Sanitaria. Consejería de Sanidad
y Política Social de la Región de Murcia.

1. Introducción

Durante nuestro periodo formativo y, fundamentalmente, durante nuestra práctica clínica, surgirán con frecuencia dudas o necesidades de información que respalden nuestra toma de decisiones y actos clínicos. Básicamente, las preguntas que pueden surgir al atender a un paciente van a ser de dos tipos (Ver ejemplos en tabla 1):

Tabla 1. Preguntas básicas y específicas.

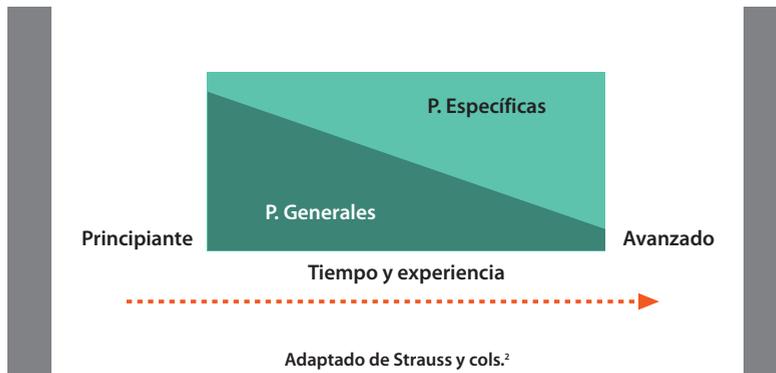
Preguntas generales	¿Cómo se diagnostica la diabetes? ¿Cuándo se considera que hay un buen control de la diabetes? ¿Cuál es el tratamiento de la diabetes?
Preguntas específicas	En pacientes con diabetes tipo 2 y mal control de su glucemia a pesar de tratamiento con dos fármacos orales, ¿la adición de insulina al tratamiento, frente a la asociación de un tercer antidiabético oral, mejoraría el perfil glucémico sin incrementar la posibilidad de efectos adversos?

- Preguntas básicas o generales, que serán aquellas que se refieren al conocimiento general que nos ayudaría a entender mejor una determinada patología, intervención o tratamiento. Estas preguntas, cuando están bien formuladas, suelen poseer dos componentes: un pronombre interrogativo (quién, cómo, cuándo...) con un verbo y la descripción del elemento de interés (patología, prueba diagnóstica, intervención terapéutica...).
- Preguntas específicas o de primera línea que, en este caso, son cuestiones que se refieren a un aspecto concreto del tema que nos interesa y cuyo conocimiento nos ayudaría en nuestra toma de decisiones. Como veremos más adelante, una pregunta específica bien formulada se compondrá de cuatro componentes esenciales que se resumen en el acrónimo PICO.

Las necesidades de información que irán apareciendo a lo largo de nuestra práctica asistencial serán básicas o específicas en base, fundamentalmente, a nuestra experiencia en un determinado tema. Así, y como norma general, cuando nuestra experiencia con una determinada afección es limitada, ante un paciente que la padece, las preguntas que nos planteemos serán de carácter general. Conforme avance nuestra experiencia clínica, nuestras posibles dudas ante una situación clínica serán cada vez más específicas. Sin embargo, puede ocurrir que, aún con limitada experiencia sobre un trastorno, necesitemos tener un conocimiento específico sobre éste y que, incluso tras amplia experiencia clínica, podamos requerir tener un conocimiento básico de ciertos aspectos (ver figura 1).

Centrándonos en las preguntas de primera línea, en este capítulo se pretende describir la importancia de transformar las dudas que se derivan de la práctica asistencial en

Figura 1. Tipos de dudas o preguntas en base a las necesidades de conocimiento.



preguntas clínicas susceptibles de ser contestadas con la evidencia disponible en la literatura. Este proceso de identificación de nuestras lagunas de conocimiento y su conversión en preguntas clínicas “contestables” constituye el primer paso para la aplicación de los principios que rigen la Práctica Clínica Basada en la Evidencia (ver tema 14).

2. Preguntas clínicas de primera línea

Como comentábamos, una pregunta específica, adecuadamente formulada, debería constar de cuatro componentes esenciales que se resumen en el acrónimo PICO, donde:

- **P** hace referencia al paciente, población o problema clínico de interés.
- **I** se refiere a la intervención principal que nos planteamos evaluar (una exposición, una prueba diagnóstica, un tratamiento, un factor pronóstico...).
- **C** indica la intervención con la que pretendemos comparar, si procede, la principal.
- **O**, del inglés *Outcome* (resultado), haría referencia al resultado o evolución clínica que tratamos de obtener al realizar una intervención frente a otra. (Ver ejemplo en tabla 2).

Tabla 2. Elementos esenciales de una pregunta clínica específica.

<p>En pacientes con diabetes tipo 2 y mal control de su glucemia a pesar de tratamiento con dos fármacos orales, ¿la adición de insulina al tratamiento, frente a la asociación de un tercer antidiabético oral, mejoraría el perfil glucémico sin incrementar la posibilidad de efectos adversos?</p>	<p>P: paciente con diabetes tipo 2, en tratamiento con dos antidiabéticos orales y mal control de la glucemia I: añadir insulina C: asociar un tercer antidiabético oral O: mejorar el perfil glucémico sin aumento de efectos adversos</p>
--	--

Cabe destacar que el elemento C puede estar ausente en preguntas bien formuladas en las que no sea cuestionada una comparación de la intervención a realizar. Su mayor relevancia se daría en aquellas preguntas centradas en intervenciones diagnósticas y terapéuticas en las que se plantea comparar una determinada prueba diagnóstica con otra o un tratamiento concreto con otra opción terapéutica.

3. ¿Por qué es conveniente aprender a transformar nuestras dudas en preguntas clínicas específicas bien formuladas?

El elaborar de forma adecuada una pregunta clínica puede ayudarnos de la siguiente forma:

- Permite concretar y definir adecuadamente aquello que queremos saber (características concretas del paciente, resultados específicos...).
- Permite conocer cuáles son los elementos clave (*keywords*) de nuestra necesidad de información.
- Ayuda a identificar el tipo de pregunta a la que nos enfrentamos (etiología, tratamiento, pronóstico, etc.).
- Ayuda a determinar el tipo de estudio a buscar (ensayo clínico, cohorte, casos-control, etc.) e identificar el recurso de evidencia más eficiente.

Figura 2. Proceso a seguir en la elaboración de una pregunta clínica.

a) Identificar los elementos básicos de la pregunta:

- **Paciente o problema:** ¿Cómo puedo describir un grupo de pacientes similar al mío?
- **Intervención:** ¿Qué procedimiento conviene ejecutar?
- **Con qué lo Comparo:** ¿Cuál es la alternativa?
- **Resultados (*Outcome*):** ¿Qué debo esperar que ocurra?

b) Identificar el tipo de pregunta:

- Tratamiento.
- Diagnóstico.
- Pronóstico.
- Etiología/Daño.

c) Determinar el tipo de estudio más adecuado:

- Tratamiento: Ensayo clínico.
- Revisión de un tratamiento: Revisión sistemática-Metaanálisis.
- Diagnóstico: Estudio transversal, validación pruebas diagnósticas.
- Pronóstico: Cohorte.
- Etiología: Casos-control/Cohorte.
- Creencias, expectativas: Investigación cualitativa.

d) Búsqueda de las evidencias:

De la Pirámide de las 5S a la de las 6S (Ver capítulo 16).

En definitiva, facilita la elaboración de la estrategia de búsqueda de mayor rendimiento para dar respuesta a nuestra duda.

4. ¿Cuál es el proceso que se recomienda seguir?

Cuando nos encontramos ante una situación clínica en la cual surge una duda es adecuado seguir una serie de pasos que van a facilitar y optimizar la obtención de la información necesaria (Ver figura 2).

4.1. Identificación de los elementos clave de la pregunta

Siguiendo el esquema PICO identificaremos los elementos clave de nuestra necesidad de información.

4.2. Identificación del tipo de pregunta clínica

Una vez concretada la duda, formulada la pregunta e identificados sus componentes básicos, el siguiente paso sería plantearnos el tipo de pregunta ante la cual nos encontramos. Los tipos de pregunta los podemos agrupar básicamente en cuatro grupos fundamentales: preguntas sobre tratamiento, diagnóstico, pronóstico o sobre etiología/daño.

4.3. Identificación del tipo de estudios más adecuados en base a las características de la pregunta

Conocer el tipo de pregunta nos ayudará a determinar cuál es el tipo de investigación (ensayo clínico, estudio de cohorte, casos-control, etc.) que nos va a resultar más útil para darle una respuesta y, por tanto, va a orientar y concretar nuestra búsqueda posterior de información. En la tabla 3 se describen de forma orientativa los estudios que de forma más apropiada aportarían información en función del tipo de pregunta.

Tabla 3. Ejemplos de estudios más apropiados en base al tipo de pregunta.

- Pregunta sobre tratamiento: ensayo clínico.
- Pregunta sobre diagnóstico: estudio transversal, de validación pruebas diagnósticas.
- Pregunta sobre pronóstico: estudio de cohorte.
- Pregunta sobre etiología: casos y controles/cohorte.
- Pregunta sobre creencias, expectativas: investigación cualitativa.

4.4. Diseño de la estrategia de búsqueda

En base a los elementos clave que constituyen nuestra pregunta, el tipo de pregunta ante el que nos hallamos y el tipo de estudio que consideramos es el más apropiado para obtener una respuesta, el siguiente paso sería definir la estrategia de búsqueda a realizar e identificar el recurso de evidencia más adecuado donde buscar, siguiendo el planteamiento de la Pirámide de las 6S (ver tema 16).

En la figura 3 puede verse un ejemplo del proceso de elaboración de una pregunta clínica.

Figura 3. Ejemplo del proceso de elaboración de una pregunta clínica.

A. Escenario clínico:

Andrés acude a nuestra consulta comentándonos que tras izar a su hija de 3 años comenzó de forma aguda con dolor lumbar intenso. Andrés tiene 36 años, trabaja como informático y reconoce tener hábitos sedentarios. No presenta antecedentes patológicos de interés aunque nos cuenta que hace unos 2 años presentó un cuadro similar desencadenado tras cargar peso a raíz de una mudanza. En aquel momento el proceso se resolvió en unos días tras tratamiento con analgésicos y reposo en cama. La exploración física muestra dolor con la palpación y flexo-extensión de tronco a nivel lumbar sin irradiación a miembros inferiores ni otros signos que denoten afectación radicular. El paciente ha iniciado por su cuenta tratamiento con analgesia oral y nos solicita la baja unos días para realizar reposo en cama. Ante su solicitud nos planteamos lo apropiado de dicha recomendación de reposo.

B. Formulación de la pregunta e identificación de sus elementos básicos:

En pacientes con dolor lumbar mecánico sin irradiación, ¿la indicación de reposo versus la indicación de mantenerse activo mejora la evolución clínica?

Los elementos clave de nuestra pregunta serían:

1. Paciente o problema, ¿cómo puedo describir un grupo de pacientes similar al mío?: Pacientes con dolor lumbar mecánico no irradiado.
2. Intervención, ¿qué procedimiento conviene ejecutar?: Reposo en cama.
3. Con qué lo Comparo, ¿cuál es la alternativa?: Mantenerse activo.
4. Resultados (*Outcome*), ¿qué debo esperar que ocurra?: Mejoría clínica en cuanto a alivio del dolor y reincorporación a actividad habitual.

C. Identificar el tipo de pregunta:

1. En este caso nos encontramos ante una pregunta sobre **tratamiento**.

D. Determinar el tipo de estudio más adecuado:

El tipo de investigación que mejor respondería a nuestra pregunta sería un **ensayo clínico aleatorizado** en el que se comparasen las opciones terapéuticas de reposo en cama versus mantener la actividad en pacientes con dolor lumbar mecánico sin radiculopatía.

E. Diseño de la estrategia de búsqueda:

En este caso las palabras clave que vamos a utilizar para definir nuestra estrategia de búsqueda serán "lumbalgia", "reposo en cama" y "mantenerse activo" (sus traducciones al inglés: "low back pain, bed rest, advice to stay active"). Y, utilizando estos términos y en base a la Pirámide de las 6S, comenzaríamos a buscar en Sistemas, posteriormente en Sumarios y Síntesis. El recurso ideal donde hallar información sería un **Guía de Práctica Clínica** sobre el manejo de la lumbalgia mecánica y, en su ausencia, una **revisión sistemática de ensayos clínicos aleatorizados** que comparan las dos estrategias diagnósticas planteadas.

Bibliografía

1. Navarro Mateu F, Martín JC. Formulación de preguntas clínicas e introducción a la estrategia de búsqueda de información. En: *Atención sanitaria basada en la evidencia: su aplicación a la práctica clínica*. Consejería de Sanidad de Murcia, 2007. pp. 47-71. [acceso 16 de enero de 2012]. Disponible en: http://www.murciasalud.es/recursos/ficheros/136606-capitulo_2.pdf
2. Straus SE y cols. *Medicina Basada en la Evidencia. Cómo practicar y enseñar la MBE*, 3ª ed. Churchill Livingstone: Elsevier, 2006.

16

Identificación de recursos de información relacionada con la PCBE y desarrollo de una búsqueda para localizar las mejores evidencias sobre un tópico clínico.

M^a DOLORES CASTILLO BUENO¹

ENRIQUE AGUINAGA ONTOSO²

1) Diplomada en Enfermería. Técnica de Gestión de Información Sanitaria. Centro Tecnológico de Información y Documentación Sanitaria. Consejería de Sanidad y Política Social de la Región de Murcia.

2) Médico Especialista en Medicina Preventiva y de Salud Pública. Centro Tecnológico de Información y Documentación Sanitaria. Consejería de Sanidad y Política Social de la Región de Murcia.

1. Introducción

El desarrollo de la Práctica Clínica Basada en la Evidencia (PCBE) se establece en cinco etapas (ver tema 14).

El objetivo del presente tema se centra en la segunda etapa (identificación y localización de la mejor evidencia), para ello propondremos una guía de recursos de información en la Web para localizar las mejores evidencias de utilidad y llevar a cabo la PCBE.

Por lo tanto, tras haber formulado la pregunta clínica, definiremos nuestra estrategia de búsqueda (ver tema 15), que debe ser lógica, sistemática y reproducible. Con ella seleccionaremos los descriptores (palabras clave) y operadores booleanos, delimitaremos el tipo de pregunta clínica que nos hemos formulado y el tipo de estudio más adecuado para responder ese tipo de pregunta.

A continuación, y antes de buscar directamente en las bases de datos, nos preguntaremos: ¿dónde podemos encontrar el tipo de estudio que nos interesa? (elección de la fuente, en qué base de datos). Con esta pregunta entramos de lleno en la segunda etapa de la PCBE.

2. ¿Dónde encontraremos el tipo de estudio que nos interesa?: Elección de la fuente

2.1. Existen muchas fuentes en Internet que permiten, con mayor o menor dificultad, localizar la mejor evidencia disponible para responder nuestra pregunta clínica: desde las tradicionales bases de datos bibliográficas (MEDLINE, Cochrane Library) a recursos de síntesis de la evidencia (Evidencia Clínica, UpToDate, etc.), pasando por guías de práctica clínica o revisiones sistemáticas (UpToDate, Guía Salud, etc., Cochrane Plus). Pero en muchas ocasiones no disponemos de mucho tiempo para poder buscar en todas ellas, por lo que la consulta en los recursos disponibles debe ser en función de la pregunta clínica, de los objetivos y del tiempo disponible, pues no todas las fuentes disponibles son adecuadas para poder responder a todo tipo de preguntas que podamos habernos hecho. A veces una sola consulta, en un solo recurso, podrá ser suficiente para localizar la evidencia que dé respuesta a una pregunta concreta, pero, en la mayoría de las ocasiones, será necesaria una consulta múltiple en diferentes recursos para poder completar la respuesta con mayor validez.

Un método útil para buscar información es preguntarnos si alguien antes ya se ha planteado nuestra misma pregunta y si ha realizado la recopilación de la información pertinente y la mantiene actualizada. Estas recopilaciones permiten delegar las necesidades de generar información basada en la evidencia en grupos o profesionales con mayor capacidad. Estas fuentes de información, elaboradas por estos grupos, es lo que se conoce como fuentes o recursos basados en la evidencia.

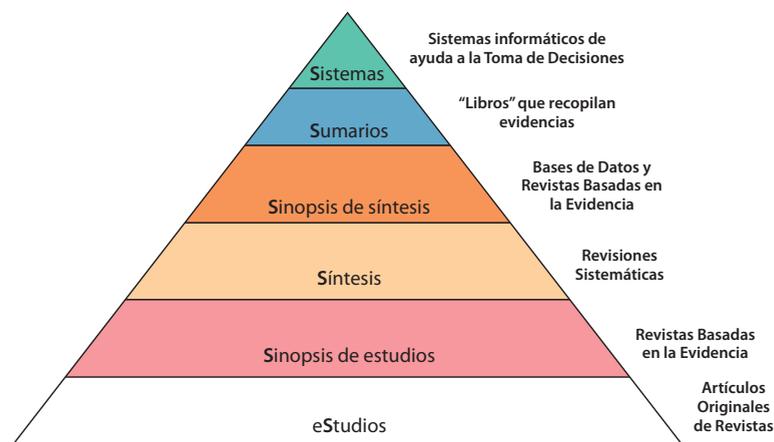
Además, en los últimos años se han creado muchos recursos prácticos que facilitan el acceso a la investigación de alta calidad. A estos recursos se les llama recursos “pre-evaluados” (*pre-appraised*) porque conllevan un proceso de filtrado previo. Todo esto contribuye a incluir en nuestra búsqueda sólo aquellos estudios que son de mayor calidad (ver tema 14). Por otra parte, se caracterizan por actualizarse periódicamente, por lo que la evidencia a la que accedamos estará siempre actualizada.

2.2. Modelo de las 6S (Tabla 1)

Para facilitar el uso de muchos de los recursos pre-evaluados, Haynes propuso el modelo jerárquico de las “4S”; que posteriormente refinó en el modelo de “5S” y actualmente ha quedado como el modelo de las “6S” o pirámide de las “6S”, que describe con mayor precisión y rigor los diferentes recursos. Aunque la mejor evidencia más actual no siempre puede estar disponible en un recurso pre-evaluado tan rápidamente como se quisiera.

En el modelo de las “5S” había las siguientes capas o escalones (comenzando por abajo): **estudios, síntesis, sinopsis, sumarios y sistemas**. El motivo de añadir una capa más y concretarlo en el de las “6S” es porque se concluyó que si la sinopsis de un único estudio y el resumen de una Revisión Sistemática (RS) eran equivalentes, según se indica en el modelo de las “5S”, en la jerarquía de la evidencia, una RS nos brinda mejor evidencia que un solo estudio, por lo que en el modelo de las “6S” se ha agregado una capa más para distinguir los dos tipos de sinopsis (de estudios o de Síntesis). Quedando construido el modelo de las “6S” de la siguiente forma, comenzando por la base (ver Figura 1):

Figura 1. Modelo de las “6S”. Basada en “DiCenso A, Bayley L, Haynes B. ACP Journal Club. 2009”⁽¹⁾.



2.2.1. Estudios originales individuales o primarios. Aunque hay un gran aumento de las fuentes de publicaciones secundarias, en muchas ocasiones necesitaremos buscar investigación original.

2.2.2. Sinopsis de estudios (capa añadida al modelo de las "5S"): proporcionan un breve, pero detallado, resumen de un único estudio original que puede dar suficiente e importante información para la práctica clínica. Van acompañados de comentarios orientados a resaltar la aplicabilidad clínica de sus resultados. Como ventajas, hay que señalar la garantía de que el estudio es de alta calidad y relevancia clínica para merecer ser resumido, además, son documentos breves y muy cómodos de leer.

2.2.3. Síntesis: Revisiones Sistemáticas (RS). Que son la recopilación completa de toda la investigación relacionada con una pregunta clínica determinada a través de un proceso de pasos sistematizados: formulación de una pregunta; identificación de los estudios pertinentes; valoración según su calidad; extracción de los resultados de los estudios pertinentes y resumen o síntesis de estos. El resumen de los artículos incluidos puede ser presentado en forma cualitativa o mediante el empleo de técnicas estadísticas para combinar cuantitativamente sus resultados, lo que se conoce como metaanálisis. Finalmente, se incluyen las conclusiones.

2.2.4. Sinopsis de síntesis: resumen de los resultados de una RS que proporciona información suficiente para apoyar una acción clínica. Están incluidas en revistas que rastrean, filtran y seleccionan artículos entre los publicados en las principales revistas.

2.2.4. Sumarios: documentos que integran las mejores evidencias disponibles de las capas inferiores de la pirámide (estudios, sinopsis de estudios, síntesis y sinopsis de síntesis) para desarrollar guías de práctica clínica (GPC) o documentos que abarcan una amplia gama de pruebas o evidencias sobre problemas clínicos específicos que se actualizan de forma periódica. Son considerados Sumarios los Servicios de Respuestas a Preguntas Clínicas, que consisten en un sistema de respuestas a interrogantes formulados por profesionales sanitarios. Las GPC son recomendaciones desarrolladas sistemáticamente para ayudar a los profesionales de la salud y a los propios pacientes a tomar decisiones sobre la atención sanitaria adecuada en circunstancias clínicas concretas. En las GPC cada recomendación debe ir acompañada de los Niveles de Evidencia (NE) y de los Grados de Recomendación (GR) (ver tema 14).

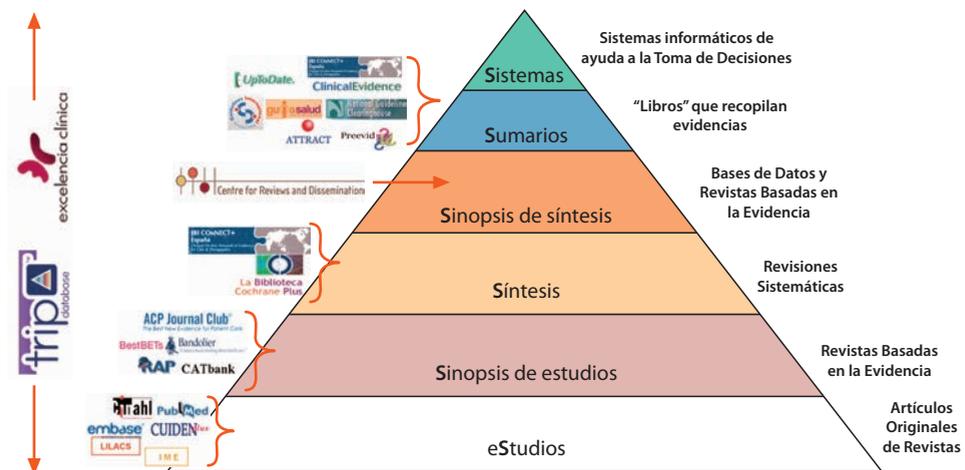
2.2.5. Sistemas: en los que las características individuales del paciente se ligan automáticamente a la mejor y más actualizada evidencia que coincida con las circunstancias específicas del paciente y del médico, y que le proporciona a este los aspectos clave de la gestión del caso. Por ejemplo, sistemas informatizados de apoyo a la toma de decisiones

(SIATDs). Estos integran y resumen todas las evidencias provenientes de la investigación, pertinentes e importantes sobre un determinado problema clínico. Se actualiza conforme aparecen nuevas evidencias. Estas se vinculan automáticamente a través de la historia clínica electrónica de cada uno de los pacientes. Los datos de cada paciente se introducen en un programa de ordenador y seorean con programas o algoritmos de una base informatizada de conocimiento médico, lo que resulta es la generación de recomendaciones para los médicos específicas para cada pacientes.

2.3. Elección de la fuente (Figura 2 y Tabla I)

A la hora de elegir la fuente y utilizando este modelo de toma de decisiones clínicas en la PCBE, la búsqueda se debe comenzar en el nivel más alto posible del modelo “6S”.

Figura 2. Clasificación de las fuentes según el modelo de las “6S”



2.3.1. Es decir, en una situación ideal, comenzaríamos en la capa de los **SISTEMAS**. Por ejemplo, existe un SIATDs para el manejo de anticoagulantes orales por enfermeras clínicas de atención primaria en el Reino Unido.

2.3.2. Si estos no existiesen, el siguiente paso sería buscar **SUMARIOS**, es decir, GPC o libros de sumarios con información basada en pruebas sobre problemas clínicos específicos. Ejemplo de Bases de Datos con Sumarios: UpToDate, Clinical Evidence, Guía Salud, JBICONNECT-Plus, Preevid, etc.

UpToDate: es una herramienta de información exhaustiva basada en la evidencia clínica y diseñada para proporcionar un acceso rápido a respuestas prácticas y fiables de las preguntas que surgen durante la práctica. Su contenido se renueva cada 4 meses.

Clinical Evidence: fuente internacional donde se encuentra la mejor evidencia disponible para una asistencia sanitaria eficaz. Sus revisiones se centran en resumir la evidencia en relación a los problemas clínicos comunes. Una comisión de expertos y de escritores especializados comprueba rigurosamente la evidencia más relevante y después la resume en formato rápido.

Entre las Guías de Práctica Clínica, destacamos: Guía Salud (repertorio de guías de práctica clínica del Sistema Nacional de Salud de España); *National Guideline Clearinghouse* elaborada por la *National Guideline Clearinghouse* (NGC) en cooperación con la AMA y la *American Association of Health Plans*, contiene más de 1.000 guías clínicas; y finalmente, *Guidelines International Network* (GIN).

Entre los Servicios de Respuestas a Preguntas Clínicas, respondidos según el esquema de la PCBE, el más conocido a nivel mundial es ATTRACT. Podemos destacar también el Servicio PREEVID, del Centro Tecnológico de Información y Documentación Sanitaria, donde pueden formular preguntas los profesionales sanitarios de la Consejería de Sanidad y del Servicio Murciano de Salud. El banco de respuestas está abierto a todo el mundo. Está clasificado por orden de respuestas y tiene un buscador por temas.

También encontraremos sumarios y *Best Practice* relacionados con los cuidados enfermeros en el JBICONNECT-Plus.

2.3.3. Cuando no se halla un sumario sobre el problema clínico que nos ocupa, entonces la siguiente opción sería buscar entre las **SINOPSIS DE SÍNTESIS**. Dado que a veces no se tiene tiempo de examinar detalladamente Revisiones Sistemáticas (RS), una sinopsis que resuma los resultados de una RS de alta calidad puede proporcionar información suficiente para apoyar la acción clínica. Estas las encontraremos en las Revistas Secundarias basadas en la evidencia (*ACP Journal Club*, *Evidence-Based Medicine*, *Evidence-Based Mental Health* y *Evidence-Based Nursing*), que cumplen estrictos criterios de calidad junto a los comentarios críticos sobre la calidad de estos. Ejemplo de donde localizar sinopsis de síntesis es el sitio del *Centre for Reviews and Dissemination* (CRD) donde existen 3 bases de datos con información preevaluada de gran calidad. Éstas son: la *Database of Abstracts of Reviews of Effects* (DARE), la *NHS Economic Evaluation Database* (NHS EED) y la *Health Technology Assessment* (HTA) *database*. DARE alberga gran cantidad de artículos primarios de alta calidad metodológica, analizados críticamente, según las pautas de la PCBE.

2.3.4. A continuación, si no hubiéramos encontrado una sinopsis de síntesis buscaríamos una **SÍNTESIS**: o Revisiones Sistemáticas (RS). La fuente de RS más conocida y fiable es la *Cochrane Library*, que es una organización internacional que prepara, actualiza, promueve y facilita el acceso a las RS, lo que ayuda en gran medida a quienes necesitan tomar decisiones en la práctica clínica. Se actualiza cuatro veces al año y es de acceso bajo suscripción pagada. Aunque, a través del centro Cochrane y coordinado por la Red Cochrane Iberoamericana, con centros colaboradores en distintos países de Ibe-

roamérica, el acceso es universal, traducido al castellano y gratuito en todo el territorio español. Esto es posible gracias a la suscripción realizada por el Ministerio de Sanidad y Consumo a la Biblioteca Cochrane Plus. El Ministerio también ha hecho una suscripción para la traducción al castellano y el acceso gratuito en todo el territorio español a RS del JBICONNECT Plus (Red clínica de Evidencia on-line sobre Cuidados y Procedimientos del Instituto Joanna Briggs (IJB), relacionadas fundamentalmente con temas de cuidados de enfermería, aunque también tiene contenidos sobre fisioterapia, nutrición, etc. El IJB es una organización internacional que promueve y apoya la síntesis, transferencia y utilización de la evidencia a través de la identificación factible, apropiada, significativa y eficaz de la práctica de los cuidados para su mejora. En el JBICONNECT Plus se pueden encontrar: *Best Practices Information Sheet*, Sumarios de Evidencias, herramientas para la valoración metodológica o para crear manuales de procedimientos o folletos de información para los usuarios, etc.

Otras dos bases de datos están disponibles gratuitamente en el sitio web del *Centre for Reviews and Dissemination* (CRD) de la Universidad de York, en colaboración con la *International Network of Agencies for Health Technology Assessment* (INAHTA). Una de ellas es la base de datos DARE, que alberga también una gran cantidad de artículos primarios de alta calidad metodológica, analizados críticamente según las pautas de la PCBE.

2.3.5. Si tampoco encontramos una síntesis relacionada con el problema clínico que se nos ha planteado, la siguiente capa a buscar es la de **SINOPSIS de estudios individuales**: que son revistas que rastrean, filtran y seleccionan artículos entre los publicados en las principales revistas médicas. Los estudios analizados y sus resúmenes se presentan de forma sistemática y suelen ser de mayor interés y calidad que el resumen del propio artículo original. Entre las más prestigiosas se encuentran: *ACP Journal Club* o *Evidence Based Medicine* del *BMJ* en su edición en español. También podemos considerar la publicación secundaria *Bandolier* o su versión española, *Bandolera*.

Asimismo se consideran sinopsis los Temas Valorados Críticamente (TVC) o CATs (*Critically Appraisal Topics*). Estos son resúmenes de artículos evaluados críticamente, que son elaborados a partir de una pregunta clínica específica. Como contrapartida, no necesariamente se valora un artículo basándose en una búsqueda amplia sobre un tema, por lo que no está carente de sesgos. Además, se abordan solo algunos aspectos específicos de una patología y su actualización no siempre es frecuente. Existen varios sitios donde localizar los TVC o CATs, por ejemplo: *CATbank* (Centro EBM Oxford), Best BETS, y la base de datos de *RAPs* del *JBICONNECT-Plus* (fundamentalmente a partir de preguntas clínicas surgidas durante los cuidados de enfermería).

2.3.6. A continuación, buscaríamos **ESTUDIOS** originales o primarios, ya que en ocasiones necesitamos encontrar investigación original. Los estudios originales o primarios se encuentran indexados en bases de datos como *Medline*, *CINHAL*, *EMBASE*, etc.

MEDLINE es la más conocida, a ella se puede acceder vía *PubMed* de forma gratuita. Hay gran cantidad de información contenida en *PubMed*, y el esquema de formulación de la pregunta clínica adquiere gran importancia (con sus 4 elementos PICO) (Ver tema 15) en la búsqueda de la información en esta base de datos (y en todas las que contienen los estudios originales). Dentro de las herramientas que ofrece Medline para una búsqueda más eficiente destacan los “*Mesh*” (*Medical Subject Headings*), que son términos estandarizados que describen el contenido del artículo y lo relacionan con otros que tratan temas similares. A través de los términos *Mesh* se indexan y catalogan todos los artículos que se incluyen en *Medline*, por lo que, para obtener los mejores resultados en la búsqueda, es necesario realizarla a través de los términos *MESH* adecuados. Otras herramientas de búsqueda en *PubMed* son: las “*Clinical Queries*” (permite realizar la exploración según el tipo de problema clínico); los *Babel MeSH* (ayuda a encontrar el término *MeSH* adecuado a las personas no muy familiarizadas con el inglés); o la búsqueda en *PubMed* vía PICO (formulario para la construcción de una consulta estructurada siguiendo el esquema PICO que realiza la misma búsqueda que en la página principal de *PubMed* y se obtienen los mismos resultados, pero el diseño nos ayuda a estructurar la consulta y hacerla más adecuada).

CINHAL, es una base de datos fundamentalmente sobre investigación enfermera y de otras ciencias, en inglés, que contiene más de 2,3 millones de registros que datan desde 1981.

EMBASE, contiene más de 11 millones de referencias (desde 1974). Incluye información de todos los ámbitos sanitarios: farmacología, toxicología, psiquiatría, medicina clínica y experimental, salud pública, ocupacional y medioambiental, etc.

En español son también de utilidad: la Base de Datos de Revistas indizadas en *SCIELO* (*IBECS*), el *Índice Médico Español* (*IME*), la base de datos de *Literatura Latinoamericana y del Caribe en Ciencias de la Salud*, *LILACS*; o *CUIDEN Plus* (Base de Datos Bibliográfica de la Fundación Index), que incluye la producción científica de la enfermería española e iberoamericana tanto de contenido clínico-asistencial en todas sus especialidades como con enfoques metodológicos, históricos, sociales o culturales.

2.3.7. Por otra parte, también habría que tener en cuenta los recursos de búsqueda simultánea de múltiples fuentes: los llamados **METABUSCADORES**. Destinados a aquellos que no conocen qué base de datos es la más adecuada para responder a una pregunta clínica, por ejemplo: *TripDatabase* (búsqueda en bases de datos en inglés) y *Excelencia Clínica*. Estos metabuscadores extraen y clasifican las evidencias a partir de una serie de bases de datos (por ejemplo, *National Guidelines Clearinghouse*, *Cochrane Library*, revistas de resúmenes, etc.). Sin embargo, la calidad de lo recuperado dependerá de la fuente, y muchas fuentes no ofrecen una valoración crítica de la evidencia.

TripDatabase, realiza en forma paralela una búsqueda de artículos primarios en *Medline* y en otros recursos de Medicina Basada en la Evidencia (síntesis de evidencias, servicios de preguntas clínicas, de GPC, etc.).

Excelencia Clínica realiza la búsqueda en las principales bases de datos en español y traduce la búsqueda automáticamente al inglés, a los que se puede acceder desde otra pestaña de la página. Incluye acceso a la Biblioteca Cochrane, alertas sanitarias, revistas secundarias, repositorios de GPC, etc. Lo que permite realizar las consultas desde un único punto.

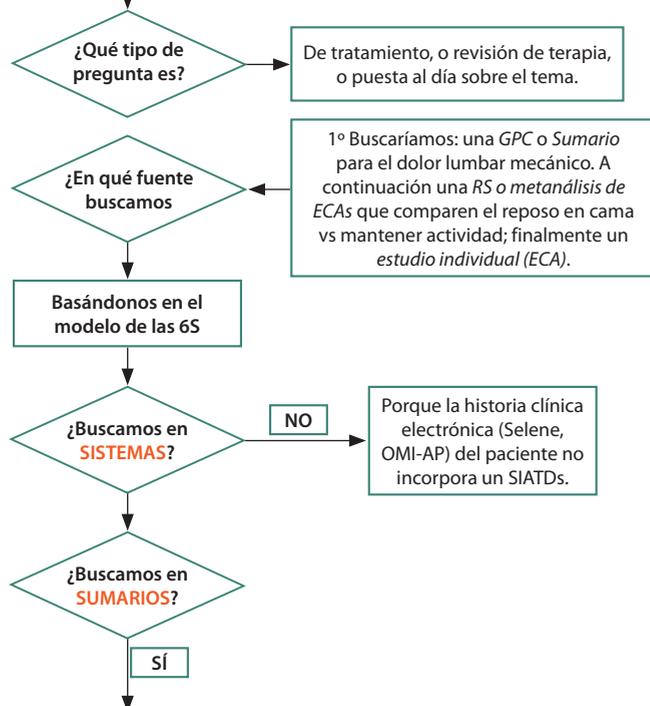
Tabla 1: Ejemplos de fuentes según recursos (según orden preferente de búsqueda).

RECURSOS	Ejemplos de FUENTES	
1. Sistemas	Programas de ordenador pareados con base informatizada de conocimientos médicos.	Triptatabase, Excelencia Clínica
2. Sumarios	Sumarios: UptoDate, Clinical Evidence, JBICONNECT-Plus. GPC: Guía Salud, National Guideline Clearinghouse, Guidelines International Network (GIN). Respuestas a Preguntas Clínicas: ATTRACT, Preevid.	
3. Sinopsis de Síntesis	Centre for Reviews and Dissemination (CDR): Database of Abstracts of Reviews of Effects (DARE), NHS Economic Evaluation Database (NHS EED) y Health Technology Assessment database (HTA).	
4. Síntesis	Cochrane Library, Librería Cochrane Plus (en castellano), JBICONNECT-Plus, Database of Abstracts of Reviews of Effects (DARE).	
5. Sinopsis de estudios	ACP Journal Club o Evidence Based Medicine del BMJ (en español). Bandolier (versión española, Bandolera). Temas valorados críticamente (TVC) o CATs: CATbank (Centro EBM Oxford), Best BETS y RAPs (JBICONNECT-Plus).	
6. Estudios	Medline (PubMed), CINHALL, EMBASE. En castellano; SCIELO, Índice Médico Español (IME), LILACS, Cuiden Plus.	

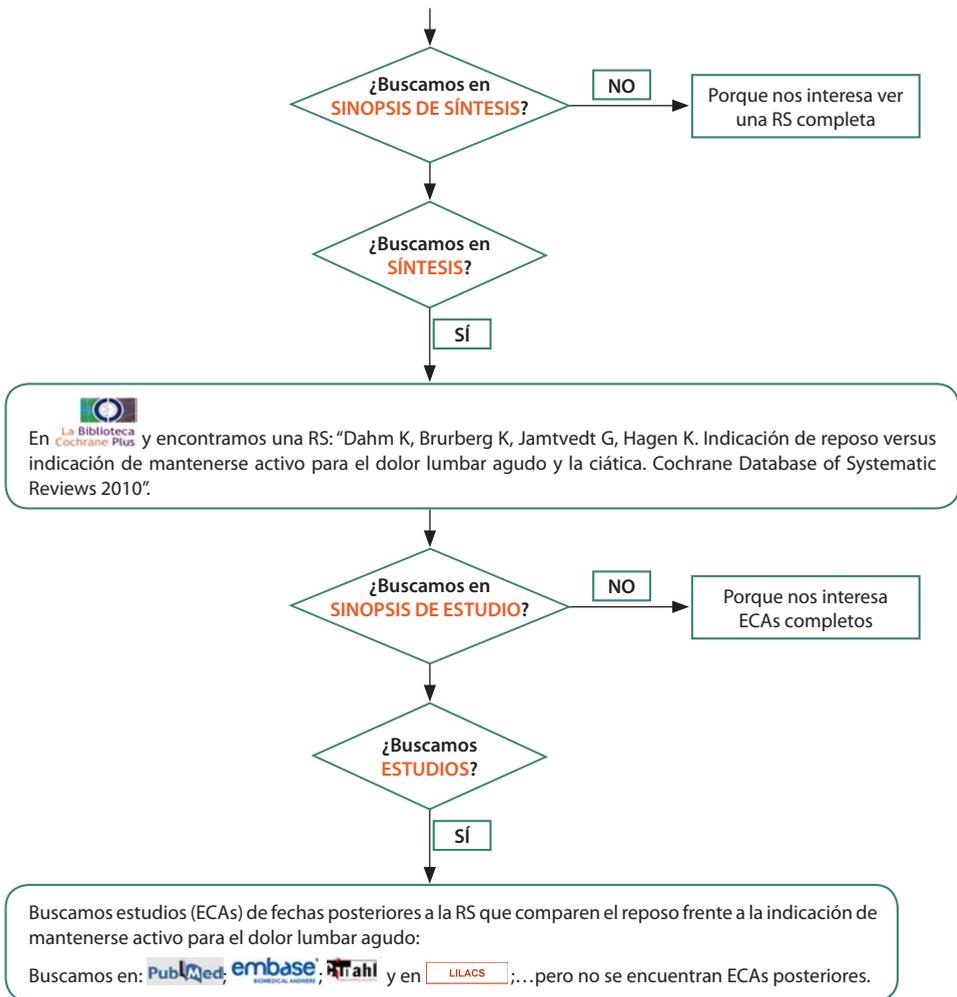
3. Desarrollo de una búsqueda para localizar las mejores evidencias sobre un tópico clínico.

ESCENARIO CLÍNICO: Andrés acude a nuestra consulta comentándonos que tras izar a su hija de 3 años en brazos comenzó de forma aguda con dolor lumbar intenso. Andrés tiene 36 años, trabaja como informático y reconoce que tener hábitos sedentarios. No presenta antecedentes patológicos de interés aunque nos cuenta que hace unos 2 años presentó un cuadro similar desencadenado tras cargar peso a raíz de una mudanza. En aquel momento el proceso se resolvió en unos días tras tratamiento con analgésicos y reposo en cama. La exploración física muestra dolor con la palpación y flexo-extensión de tronco a nivel lumbar sin irradiación a miembros inferiores ni otros signos que denoten afectación radicular. El paciente ha iniciado por su cuenta tratamiento con analgesia oral y nos solicita la baja unos días para realizar reposo en cama. Ante su solicitud nos planteamos lo apropiado de dicha recomendación de reposo.

PREGUNTA CLÍNICA: En pacientes con dolor lumbar mecánico sin irradiación, ¿la indicación de reposo versus la indicación de mantenerse activo mejora la evolución clínica?



Si buscamos en [guía salud](#), encontramos: "Guía de Práctica Clínica sobre Lumbalgia. Osakidetza. 2007".
 En [UpToDate](#) encontramos un resumen: "Lumbalgia aguda y crónica: Reposo en cama frente a actividad. 2009".
 En [UpToDate](#).....
 Si buscamos con el metabuscador [tripa](#), encontramos otra GPC posterior: "Low back pain. Early management of persistent non-specific low back pain. NICE clinical guideline 88. 2009'".



Bibliografía

1. DiCenso A, Bayley L, Haynes RB. ACP Journal Club. Editorial: Accessing preappraised evidence: fine-tuning the 5S model into a 6S model. *Ann Intern Med.* 2009 Sep 15; 151(6):JC3-2, JC3-3.
2. Fernández E, García AM. *Preparación de publicaciones biomédicas. UD 2: La búsqueda de información bibliográfica.* Bellaterra: Laboratori d'Estadística Aplicada i de Modelització (UAB); 2004.
3. Grupo de Atención Sanitaria Basada en la Evidencia. *Atención Sanitaria Basada en la Evidencia. Su aplicación a la Práctica Clínica.* Consejería de Sanidad de la Región de Murcia. 2006.
4. Cabrejos Perotti S, Cárdenas Bravo LE, Negro Álvarez JM, Aguinaga Ontoso E, Ivancevich JC, Bozzola CM. Buscando la mejor "evidencia" en Alergología. *Archivos de Alergia de Inmunología Clínica* 2008; 39(1):22-31.
5. Estrada Lorenzo JM, Barderas Manchado A. Recursos en Internet para el conocimiento y aplicación de la práctica clínica basada en la evidencia. *Matronas Prof.* 2006; 7(4): 17-24
6. García M, Puig M, Mateu MS, Bonet M. Fuentes de información de medicamentos. En: Carnicero J (Coordinador). *La gestión del medicamento en los Servicios de Salud. Informes SEIS* (8). Pamplona: Sociedad Española de Informática de la Salud, 2008. <http://www.seis.es>

17

Conocimiento de los servicios y funcionamiento de la Biblioteca Virtual Murciasalud (BVMS)

AGUSTÍN ROCA VEGA¹
ENRIQUE AGUINAGA ONTOSO²

1) Médico. Diplomado en Biblioteconomía y Documentación.
Técnico de Gestión de Documentación Sanitaria.
Centro Tecnológico de Información y Documentación
Sanitaria. Consejería de Sanidad y Política Social
de la Región de Murcia.

2) Médico Especialista en Medicina Preventiva y de Salud
Pública. Centro Tecnológico de Información y
Documentación Sanitaria. Consejería de Sanidad
y Política Social de la Región de Murcia.

1. Introducción. Concepto de Biblioteca Virtual

Una biblioteca virtual es aquella en que los recursos de información se encuentran disponibles en el formato digital, accesibles por medio de un ordenador a través de una red de comunicaciones. Algunos autores la denominan Biblioteca digital o electrónica.

La Biblioteca Virtual Murciasalud (BVMS) se encuentra incluida dentro del Portal sanitario de la Región de Murcia, Murciasalud (<http://www.murciasalud.es>), y está en funcionamiento desde el año 2003.

El Centro Tecnológico de Información y Documentación Sanitaria (CTIDS), perteneciente a la Consejería de Sanidad, se encarga de gestionar y coordinar la BVMS. Existe un órgano de asesoramiento de la BVMS, la **Comisión de responsables de Bibliotecas de la Consejería de Sanidad-Servicio Murciano de Salud**, creado en 2003 y en el que se encuentran representadas todas las Gerencias de salud, además de la Dirección General de Calidad Asistencial, Formación e Investigación Sanitaria y de la Subdirección General de Salud Mental y Asistencia Psiquiátrica.

2. La BVMS. Modos de acceso

El acceso a la BVMS está restringido a los profesionales sanitarios de la Región de Murcia y es gratuito para los mismos. Algunos servicios están limitados a los profesionales del sistema regional de salud. Es posible acceder de dos modos:

- a) A través de la **Intranet** de la Comunidad Autónoma de la Región de Murcia, incluyendo el SMS (Servicio Murciano de Salud), sin necesidad de identificarse.
- b) Desde cualquier ordenador ajeno a la Intranet, identificándose como usuario de Murciasalud. Para ello es necesario **estar registrado** como usuario de Murciasalud.

Se puede acceder desde Murciasalud o tecleando <http://www.murciasalud.es/bvms>.

3. Recursos y modos de acceso

Los recursos y modos de acceso a la BVMS en el año 2011 se pueden ver en la Tabla 1.

Tabla 1. Recursos disponibles y modos de acceso.

Recurso	Intranet	Extranet
CINHAL	Sí	SI
CUIDENPlus	Sí	SI
Embase	Sí	No
Enfermería al día (Ebsco)	Sí	Sí
Fisterrae	Sí	Sí
Guía terapéutica en A. Primaria basada en la selección razonada de medicamentos	Sí	Sí
Inf. Farmacoterap. del SMS	Sí	Sí
JBI COnect	Sí	Sí
Lexi-comp	Sí	No
Libros electrónicos	Sí	Parcial ¹
Martindale	Sí	Sí
MS-Med	Sí	Sí
Preevid	Sí	Sí
Revistas electrónicas	Sí	Parcial ²
Stockley	Sí	SI
UpToDate	Sí	No

4. Acceso al texto completo de las revistas electrónicas

Se puede acceder al texto completo de los artículos publicados en las revistas electrónicas contratadas de varios modos: desde el catálogo de revistas (por navegación o por localización del artículo –en este caso a través de un revolvedor de enlaces³–), desde el buscador MS-Med o directamente desde Pubmed.

4.1. El catálogo de revistas

El catálogo de revistas recoge todas las revistas contratadas por la BVMS así como un gran número de revistas que se encuentran de modo libre en Internet.

1 Algunos libros son accesibles desde la Extranet.

2 Más de 4.400 revistas, de las más de 4.600 disponibles en el catálogo, son accesibles desde la Extranet.

3 Revolvedor de enlaces: aplicación que a partir de los datos bibliográficos de un artículo muestra, si la biblioteca dispone de él, el enlace al texto completo de dicho artículo.

Modos de acceder a las revistas del catálogo:

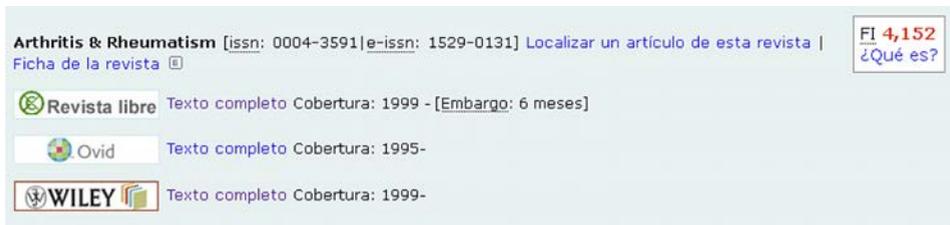
- **Por búsqueda** de ISSN (*Internacional Standard Serial Number*)⁴, título de la revista... El buscador **no es sensible a mayúsculas, minúsculas ni tildes**. Si introducimos más de una palabra el buscador recuperará las revistas que contengan en su título todas las palabras introducidas.
- **Por navegación**: alfabética (A-Z), materias, proveedores...

Figura 1. Catálogo de revistas, con el buscador y el resto de modos de acceso.



La búsqueda nos ofrece, para cada título recuperado, además del título, el ISSN y el período cronológico de cobertura de la revista, con el posible embargo⁵, y el enlace (o enlaces, ya que una revista podría accederse a través de dos o más proveedores) al texto completo de la revista.

Figura 2. Ejemplo de un resultado del catálogo de revistas.



Para acceder al texto completo de la revista pincharemos en alguno de los enlaces **“Texto completo”** del catálogo que, habitualmente, nos llevará al último número de la revista. Para acceder a un ejemplar anterior deberemos navegar por los ejemplares hasta localizar el deseado.

4 Número internacional que permite identificar de manera única una publicación en serie.

5 Algunas revistas tienen un período de embargo. Esto significa que no es posible acceder a los últimos n meses de la revista.

También podemos llegar al texto completo de un artículo a través del localizador de artículos pinchando en el enlace **“Localizar un artículo de esta revista”**. En el formulario que aparece introduciremos bien el PMID⁶ o el DOI⁷ del artículo, si disponemos de él, o bien sus datos bibliográficos, y el revolvingor de enlaces se encargará de indicarnos si existe un enlace al artículo a texto completo.

4.2. El buscador MS-Med

El buscador MS-Med nos permite, desde la BVMS, lanzar búsquedas en la base de datos Medline. En los resultados nos muestra, además de los datos bibliográficos de los artículos recuperados (título, autores, etc.), otros datos de interés como el factor de impacto (de la revista), enlace al texto completo, datos de cobertura cronológica de la revista, enlace al revolvingor de enlaces, datos del artículo en Pubmed y enlace al resumen, enlace para solicitar el artículo al SOD, posibilidad de enviarnos los resultados a nuestro correo electrónico.

Podemos buscar el artículo utilizando los datos bibliográficos del mismo (título, autor...) o bien el PMID.

Figura 3. Resultado de una búsqueda en Medline realizada con MS-Med.

Página 1: Resultados del 1 al 10 de 56 elementos encontrados.

anterior 1 2 3 4 5 6 siguiente

Bodies in mutation: understanding lipodystrophy among women living with HIV/AIDS.
Gagnon M, Holmes D
Research and theory for nursing practice. 2011 . Num:1 . Pag:23-38
PMID: 21469539. Resumen en Pubmed [Abre nueva ventana]

Coinfections associated with human immunodeficiency virus infection: workshop 1A.
Nair P.G, Owotade FJ, Leao JC, Hegarty AM, Hodgson TA
Advances in dental research. 2011 Apr . Num:1 . Pag:97-105
PMID: 21441489. Resumen en Pubmed [Abre nueva ventana]

Texto completo en BVMS [sólo Intranet]

Ver disponibilidad en la BVMS

Tramitar solicitud a la BVMS

Resolver enlace con LinkSolver

Ver disponibilidad en la BVMS

Tramitar solicitud a la BVMS

Resolver enlace con LinkSolver

Aplicar límites (refinar búsqueda)

Tramitar la solicitud de un artículo que no aparece en los resultados

Recibir las citas marcadas por email

Nueva búsqueda

4.3. Pubmed

Accediendo a Pubmed (buscador en Medline de la Nacional Library of Medicine de los Estados Unidos) desde el enlace que se proporciona en la página principal de la BVMS, o bien a través de la url <http://www.pubmed.com/?holding=iesctidslib> se nos mostrará un icono si el artículo está disponible en la BVMS.

⁶ Acrónimo de “PubMed Unique Identifier”, es un número único asignado a cada cita de un artículo de revistas biomédicas y de ciencias de la vida que recoge PubMed. Identifica de modo unívoco a un artículo.

⁷ Digital Object Identifier –Identificador de Objeto Digital– (o DOI) es un sistema parecido a los identificadores URI (direcciones de Internet).

Figura 4. En el lado derecho podemos ver el icono indicando que el texto completo del artículo está accesible desde la BVMS.



5. Obtención de artículos no disponibles a través de los recursos contratados

Para el acceso a los artículos no disponibles a través de los recursos contratados existe el Servicio de Obtención de Documentos (SOD), que permite a los usuarios realizar peticiones de artículos. Los artículos son obtenidos mediante convenios de colaboración con otras bibliotecas y centros de documentación, o bien mediante su adquisición a servicios especializados. El servicio es gratuito para el personal de la Consejería de Sanidad y del SMS. El tiempo entre que se solicita un artículo y se sirve es variable en función de donde se encuentre el artículo, y puede ir desde una hora hasta varios días, si bien la mayoría de los artículos se sirven en un período inferior a 24 horas.

Los artículos son servidos a los usuarios a través de una aplicación *on line* que permite a los usuarios volver a leer los artículos siempre que lo necesiten independientemente de donde se encuentren.

6. Libros on line

La BVMS recoge también algunos libros on line, entre los que destacan:

- Harrison, Principios de medicina interna en español.
 - Red book (inglés), una guía sobre las manifestaciones, etiología, epidemiología, diagnóstico y tratamiento de más de 200 enfermedades infantiles comunes.
- También se recogen libros publicados con licencia Open Access⁸.

⁸ Acceso inmediato, sin requerimientos de registro, suscripción o pago –es decir sin restricciones– a material digital educativo, académico, científico o de cualquier otro tipo, principalmente artículos de investigación científica de revistas especializadas.

7. Otros recursos electrónicos

- **UpToDate:** constituye un recurso de información clínica exhaustivo basado en la evidencia clínica y diseñado para proporcionar al médico un acceso rápido a respuestas prácticas y fiables de las preguntas que se suscitan durante el acto médico en la consulta.
- **Embase:** base de datos bibliográfica que contiene más de 11 millones de referencias, desde 1974 hasta la actualidad, de unas 5.000 revistas científicas de 70 países.
- **Fisterrae:** herramienta avanzada que integra recursos de apoyo en el punto de atención para profesionales sanitarios. Incluye una base de conocimientos clínicos, una guía farmacoterapéutica, calculadoras clínicas, bancos de imágenes...
- **Biblioteca Cochrane Plus:** incluye varias bases de datos, entre las que destacan las de revisiones sistemáticas, la de ensayos clínicos controlados y la de revisiones de efectividad y de evaluación económica.
- **Preevid:** preguntas realizadas al servicio Preevid por los profesionales sanitarios de la Consejería de Sanidad y el Servicio Murciano de Salud. Ofrece respuestas rápidas basadas en la evidencia a preguntas que surjan desde la actividad asistencial. La respuesta no es un consejo concreto o una recomendación específica, ni una revisión sistemática sobre un tema; sino que, tras una búsqueda protocolizada, resume las conclusiones de las referencias encontradas. Incluye un **banco de preguntas** con las preguntas de mayor interés realizadas por los profesionales desde marzo de 2004 (más de 1.200 preguntas).
- **Información farmacológica:**
 - **Guía terapéutica en Atención Primaria** basada en la selección razonada de medicamentos.
 - **Guía farmacoterapéutica del SMS.**
 - **Martindale:** guía completa de consulta farmacoterapéutica.
 - **Stockley:** alertas de interacciones e interacciones farmacológicas.
 - **Lexi-comp:** incluye evaluaciones de principios activos, interacciones, etc.
- **JBI CONNECT+:** acceso a distintos recursos que permiten la búsqueda de evidencias para la toma de decisiones.
- **CINAHL:** base de datos de revistas de enfermería. Registra las referencias y los resúmenes publicados en más de 900 revistas de enfermería.
- **Enfermería al día:** ofrece al personal de enfermería acceso inmediato a información importante de diagnóstico basado en la evidencia.
- **Legislación sanitaria:** base de datos que incluye la legislación sanitaria regional y nacional desde 1998. Se está realizando un trabajo de recopilación de legislación retrospectiva que pretende recoger la legislación desde 1982.

8. Opciones de personalización de la BVMS

8.1. Selección personal de revistas

Los usuarios pueden crear una o más selecciones personales de revistas de acuerdo a su propio criterio. Esto les permite disponer de un modo más sencillo de acceder a aquellas revistas que más utilizan.

8.2. Mis artículos

La aplicación mis artículos permite a los usuarios clasificar mediante el uso de etiquetas o palabras clave aquellos artículos que, a lo largo del tiempo, se les han ido sirviendo a través del SOD. Esto permite a los usuarios disponer de una gran flexibilidad para organizar los artículos. Esta aplicación permite además añadir comentarios sobre los artículos y realizar búsquedas, por palabras clave, autor, título, etc., en el repositorio de artículos del usuario.

9. Desarrollo de la BVMS

La BVMS se encuentra en continuo desarrollo, por lo que los recursos disponibles están sujetos a cambios, ya que son contratados o cancelados en función de las necesidades de los usuarios. También el diseño de la web puede sufrir modificaciones, ya que es posible que se realicen cambios para facilitar su utilización y mejorar la experiencia de los usuarios.

Bibliografía

Consejería de Sanidad. Ayudas y tutoriales sobre la Biblioteca Virtual. [citado 14 Sep 2011].
Disponible en: <http://www.murciasalud.es/bvms>

18

Introducción a la lectura crítica: concepto, utilidad y habilidades

ANTONIO RÍOS ZAMBUDIO

Médico Especialista en Cirugía General y Aparato Digestivo.
Profesor Asociado del Departamento de Cirugía de la
Universidad de Murcia. Hospital Universitario Virgen
de la Arrixaca de Murcia. Servicio Murciano de Salud.

1. Introducción

La lectura crítica de la literatura científica es una herramienta imprescindible para todos los profesionales sanitarios que deben tomar decisiones en su práctica diaria. La lectura crítica es parte del proceso de la Medicina Basada en la Evidencia (MBE) o, en términos generales, la Práctica Basada en la Evidencia.

El fundamento de la MBE es que todas las decisiones sobre diagnóstico, pronóstico, etiología y terapia estén basadas en evidencias sólidas procedentes de la mejor investigación clínico-epidemiológica posible. A este conocimiento derivado de la investigación debe añadirse la experiencia de la persona que la atiende, el entorno y los medios disponibles, y las preferencias de los pacientes.

Tradicionalmente se han intentado solucionar las dudas en el ejercicio profesional a través de consultas a libros de texto, consultas a revistas o preguntando a colegas. Pero este abordaje clásico presenta importantes limitaciones. Como alternativa, la MBE ofrece herramientas para poder resolver problemas clínicos mediante el análisis eficiente de la literatura científica. Herramientas metodológicas que, si se adquieren y ejercitan, ayudarán a valorar cualquier evidencia científica y a integrarla con los conocimientos y experiencia clínica para poder decidir sobre su aplicabilidad e idoneidad en un paciente concreto.

La Práctica Basada en la Evidencia requiere tener unos conocimientos previos sobre metodología de investigación, sobre estadística y sobre la patología o problema que estamos analizando. Es un proceso que consta de distintas etapas¹:

1. **Identificar el problema** y elaborar una pregunta clínica estructurada: se convierte la incertidumbre clínica en preguntas que pueden responderse con la investigación científica.
2. **Localizar la mejor evidencia científica**: hay que buscar con exhaustividad y eficiencia los trabajos científicos desarrollados sobre el tema de interés (Búsqueda bibliográfica).
3. **Lectura crítica de artículos**: evaluar críticamente la validez de un trabajo científico, saber interpretar la relevancia de los resultados y poder determinar su aplicabilidad a cada paciente.
4. **Aplicar los resultados a la práctica clínica**: intentar integrar la valoración realizada con nuestra experiencia clínica, considerando la aplicabilidad de los resultados a los pacientes.
5. **Evaluar los resultados** obtenidos en la propia práctica clínica y contrastarlos con los de referencia.

2. Concepto

La lectura crítica de un artículo científico tiene como objetivo verificar la validez (aproximación a la verdad), y la utilidad (aplicabilidad clínica) de los resultados del artículo analizado, con el fin de incorporarlos al cuidado de los pacientes y mejorar así la calidad de la asistencia. Para ello se requieren ciertos conocimientos y habilidades que es preciso adquirir y ejercitar.

3. Estructura de un artículo científico

Un artículo de investigación tiene los siguientes apartados:

1. **Identificación y resumen:** incluye título, autores, lugar de realización y resumen.
2. **Introducción:** incluye temas relacionados con el problema de estudio: a) estado actual del problema que se está investigando; b) investigaciones previas; c) propósito, pretensiones y argumentos que justifican la realización del estudio; y d) objetivo del mismo
3. **Materiales y métodos:** describe cómo se va a llevar a cabo el estudio. Debe contener suficiente información para que otro investigador pueda reproducir el estudio: a) tipo de diseño; b) descripción de los sujetos y cómo son seleccionados; c) intervención realizada; d) variables incluidas y cómo se miden los resultados; e) cómo se han controlado los sesgos; f) justificación del tamaño muestral; y g) plan de análisis estadístico.
4. **Resultados:** describe objetivamente lo que ha ocurrido con los pacientes: a) descriptivo de los resultados; b) diferencias entre grupos de estudio; c) significación estadística; y d) significación clínica.
5. **Discusión:** se resumen, interpretan y extrapolan los resultados; se discuten sus implicaciones y limitaciones; y se discuten y contrastan las hipótesis. Suele incluir: a) resumen escueto del estudio; b) propuestas de investigaciones futuras; c) comparación con otros estudios; d) limitaciones del estudio; e) generalización de los resultados; f) hallazgos no previstos; g) interpretación de los resultados por el investigador; y h) conclusión del estudio.
6. **Bibliografía:** Recoge las referencias exactas, puntuales y verificables necesarias para avalar el desarrollo del artículo.

4. Habilidades para la lectura crítica

4.1. Generalidades

La valoración de la validez interna del estudio implica analizar si la metodología empleada es la más adecuada para la consecución del objetivo del estudio y si es correcta. Cada tipo de estudio tiene una metodología distinta y la mejor evidencia es la que aporta el diseño más idóneo. Así, si lo que pretendemos valorar es una intervención (curativa o preventiva) el ensayo clínico aleatorizado será el diseño más correcto; si queremos analizar la evolución de una enfermedad, deberemos hacer un seguimiento de una cohorte de pacientes; si es una prueba diagnóstica lo que pretendemos evaluar deberemos realizar un estudio transversal comparándola con un patrón de oro; si queremos identificar la etiología de un problema debemos contar con un estudio casos-control.

4.2. Primera aproximación a la validez del artículo: localización y selección de artículos

La primera etapa consiste en la localización y selección de los artículos:

1. La **revista** donde se publica el artículo nos indica si ha sido evaluada por revisores o no.
2. El **título** dará una idea si el tema es interesante o útil para el problema a analizar.
3. El **resumen** incluye brevemente todas las partes del mismo: objetivo, diseño, número y principales características de los sujetos incluidos, tratamiento, principales resultados y conclusiones. Nos permite una primera valoración sobre si es válido, si los resultados son aplicables, y si sobre todo si son útiles para la resolución del problema planteado.

Esta nos selecciona una serie de artículos que abordan el problema de interés. En las siguientes fases se analizará la validez de la investigación.

4.3. Segunda aproximación a la validez del artículo: metodología

En la segunda aproximación, se realizan una serie de preguntas metodológicas claves para determinar la capacidad del artículo para poder contestar adecuadamente, y de forma fiable, a la pregunta planteada. Para ello se analiza en profundidad la metodología del estudio y se determina su calidad. Existen muchos tipos de investigación y a cada objetivo y tipo de pregunta clínica corresponde un diseño de investigación distinto.

Para la valoración metodológica existen diversos cuestionarios para valorar artículos, entre los que destacan los cuestionarios CASPe, a los que se puede acceder en su página web². Existen cuestionarios CASPe para valorar artículos primarios (artículos de proyectos

de investigación originales), cuestionarios para estudios cualitativos; y cuestionarios para valorar los artículos secundarios (artículos que resumen o sintetizan la información a partir de varios artículos primarios).

La segunda aproximación se centrará en dos o tres características metodológicas básicas para cada tipo de estudio que en los cuestionarios de valoración suelen aparecer como “Preguntas de eliminación”. Si estos requisitos básicos en el diseño de los estudios no se cumplen, los resultados obtenidos no serán fiables y la validez interna del estudio no se puede asegurar.

4.4. Tercera aproximación a la validez del artículo: análisis meticuloso

Una vez hecha la revisión preliminar, en la tercera aproximación se realiza un análisis meticuloso de los artículos, centrado fundamentalmente en la sección de material y métodos (Apartado A de las Guías CASPe).

Para todos los estudios hay que tener en cuenta que son correctos los siguientes puntos: a) Definición clara y precisa del objetivo; b) Adecuación del diseño a la pregunta de la investigación; c) Adecuación de los criterios de selección de la población de estudio; d) Tamaño de la muestra suficiente; e) Consideraciones sobre errores y sesgos, así como maneras en las que el estudio podría haberse mejorado; y f) Conclusión acorde con los resultados del análisis.

Una vez verificada la validez interna del estudio, se analizan los resultados y su precisión (Apartado B de las Guías CASPe). Los resultados obtenidos dependerán de la pregunta que responde el estudio y del tipo de diseño empleado. Además, hay que tener en cuenta que la relevancia de los resultados no viene determinada únicamente por la significación estadística obtenida, sino por la relevancia clínica observada.

Finalmente, se evalúa el grado en que los resultados pueden ser generalizados a poblaciones o pacientes específicos (Apartado C de las Guías CASPe). La aplicabilidad de los resultados de un estudio a nuestros pacientes, dependerá de si entre la población de estudio y la nuestra existen diferencias en las características de los sujetos y el entorno y, en el caso de que existan, si pueden condicionar la aparición de resultados distintos. También hay que tener en cuenta el posible impacto de los resultados sobre el manejo del paciente y tener en consideración el entorno y las preferencias de los pacientes.

5. Herramientas para la lectura crítica de la literatura científica (CASP)

CASP (*Critical Appraisal Skills Programme*) (Programa de habilidades en lectura crítica) es un programa creado por el *Institute of Health Sciences* de Oxford (Universidad de Oxford y NHS R&D) para ayudar a los “decisores” del Servicio de Salud a adquirir habilidades en la

búsqueda de información y en lectura crítica de la literatura científica en salud, de modo que pudieran obtener así la “evidencia científica” necesaria para tomar sus decisiones. CASP colabora con el centro para la medicina Basada en la Evidencia (Centre for evidence Based Medicine) de la Universidad de Oxford que enseña a los clínicos cómo tomar decisiones, basadas en la evidencia, sobre un paciente concreto.

En España existe una red CASP (CASP España-CASPe) con múltiples nodos distribuidos por el territorio y una sede coordinadora ubicada en Alicante. CASPe forma parte de una organización internacional llamada CASP Internacional con la que comparte filosofía y experiencias docentes y de organización, materiales desarrollados en conjunto, así como proyectos de investigación sobre la docencia.

CASP España tiene por misión proporcionar habilidades a los “decisiones” en salud para encontrar y evaluar críticamente las mejores pruebas científicas en las que fundar sus decisiones. Ello implica tanto la formación de los “decisiones” mediante métodos pedagógicos específicos, como la formación de personas que entrenen a los “decisiones” en la consecución de estas habilidades. Además de estas tareas de formación, desarrolla y elabora materiales docentes exclusivos y realiza labores de organización, difusión y gestión dentro de un marco de crecimiento sostenible.

CASPe es líder, en España y América, en la adaptación, diseño y actualización de los instrumentos necesarios para hacer lectura crítica de la literatura científica.

Bibliografía

1. Sackett DI, Richardson WS, Rosenberg W, Haynes RB. *Medicina basada en la Evidencia. Cómo ejercer y enseñar la MBE*. Madrid: Churchill Livingstone; 1997.
2. <http://www.redcaspe.org/caspe/lectura-critica/> (consulta el 29 de septiembre de 2011).

19

Características generales y conceptos metodológicos clave de la lectura crítica: control de sesgos

ANDRÉS CARRILLO ALCARAZ¹

ANTONIA LÓPEZ MARTÍNEZ²

1) Médico Especialista en Medicina Intensiva.
Hospital General Universitario Morales Meseguer de Murcia.
Servicio Murciano de Salud.

2) Médico Especialista en Medicina Intensiva.
Hospital Universitario Morales Meseguer de Murcia.
Servicio Murciano de Salud.

1. Introducción

La forma adecuada de realizar Medicina o, mejor, Práctica Clínica Basada en la Evidencia comprende cinco pasos que se abordan de forma secuencial:

- a. Convertir la necesidad de información en una pregunta con respuesta.
- b. Búsqueda de las mejores evidencias para contestar a esa pregunta.
- c. Evaluación crítica de los artículos obtenidos en la búsqueda.
- d. Integración de la valoración crítica con nuestra especialización clínica y con los datos biológicos/valores y las circunstancias exclusivas del paciente.
- e. Evaluar nuestra efectividad y eficacia para ejecutar los pasos anteriores y buscar maneras de mejorarlas en una próxima ocasión.

Debido a ello, el tercer paso, la evaluación crítica de los artículos científicos, es esencial para poder aplicar esta forma de ejercer la profesión. La Lectura Crítica de la Literatura Científica es el proceso por el cual verificamos la validez (aproximación a la verdad), el impacto (tamaño del efecto) y aplicabilidad (utilidad en nuestra práctica clínica) de los resultados de la investigación publicada, con el fin de incorporarlos al cuidado de los pacientes y mejorar así la calidad de la asistencia.

Una aproximación a la literatura crítica de artículos científicos debe contemplar, en primer lugar, la utilidad real de los diferentes tipos de artículos; y en segundo lugar, el análisis de las características generales de los mismos, sobre todo el grado de rigor con el que se ha realizado el trabajo, haciendo especial hincapié en los errores que fácilmente pueden producirse en el diseño de los mismos.

2. Identificación del tipo de investigación adecuado a cada problema

No todos los problemas que se plantean en la práctica clínica tienen respuesta mediante un mismo tipo de estudio. O con otras palabras, si nos interesa analizar los factores pronósticos de una enfermedad, raramente podremos encontrar la respuesta en un ensayo aleatorizado.

Así, si nos planteamos un estudio relativo a un tratamiento o una medida preventiva ¿cuál sería el diseño idóneo del estudio a realizar o analizar? La respuesta sería el ensayo clínico, controlado y aleatorizado (ECA). Los estudios que evalúan pruebas diagnósticas suelen ser estudios transversales donde, de forma controlada, se compara la técnica diagnóstica a estudio con la “verdad”, siendo ésta lo que se considera habitualmente como el diagnóstico de certeza de la enfermedad evaluada por la prueba diagnóstica. Pero, por otro lado, una prueba diagnóstica, al fin y al cabo una intervención, puede ser evaluada mediante un ensayo controlado y aleatorizado para valorar la utilidad de dicha prueba en

la consecución de un resultado o evolución de los pacientes. Cuando nos enfrentamos a la búsqueda de conocimientos sobre efectos nocivos, la evidencia puede aportarse con diferentes tipos de estudios. Por ejemplo, se podría plantear mediante un ECA, donde un grupo de pacientes formado por el azar recibiría el supuesto agente nocivo y otro grupo de pacientes no lo recibiría. Al final del estudio, la comparación de ambos grupos y la aplicación de una prueba estadística no determinaría si el compuesto estudiado es o no nocivo. Pero los problemas éticos derivados de un estudio con un diseño como el anterior, impediría su realización. Pero a veces estos efectos nocivos se han detectado en ECA que buscaban efectos beneficiosos de una intervención (por ejemplo, la demostración de la acción nociva de los antiarrítmicos orales para prevenir las taquiarritmias ventriculares post-infarto, cuando se buscaba un incremento de la supervivencia con estos fármacos), en este caso la evidencia contra el uso de estas intervenciones es máxima. Cuando no resulta factible o no es ético asignar aleatoriamente a los pacientes a una exposición o ausencia de exposición a un supuesto agente causal/nocivo, los investigadores deben de buscar una alternativa al ECA, realizándose entonces estudios de cohortes o mediante un estudio casos y controles. Aunque las ventajas e inconvenientes de estos estudios han sido analizados en otros capítulos de este texto, puede resumirse la indicación de uno u otro según la frecuencia de presentación del resultado de interés, y el tiempo transcurrido entre exposición al agente nocivo y la aparición de la enfermedad. Igualmente, cuando lo que necesitamos es analizar el pronóstico de una enfermedad o situación, y relacionarlo con factores o variables de cualquier tipo, los tipos de estudios que precisamos son cohortes, y casos y controles.

En los últimos años y cada vez con mayor frecuencia se realizan y publican estudios que sintetizan toda la información existente (publicada y no publicada) sobre problemas concretos de una determinada situación de salud. Las revisiones sistemáticas y los metaanálisis, que ofrecen la mejor evidencia posible y la mayor validez externa de sus resultados, pueden realizarse con cualquiera de los diseños anteriormente mencionados. Debido a ello, las revisiones sistemáticas/metaanálisis de ensayos aleatorizados aportan la mejor evidencia para una intervención, pero también se puede reanalizar metaanálisis de estudios sobre pruebas diagnósticas, estudio de cohortes y casos controles.

3. Conceptos metodológicos en Lectura Crítica

Cuando se realiza lectura crítica la aproximación del lector es similar independientemente del tipo de diseño que se ha utilizado en el artículo. Los pasos a seguir son la evaluación de la validez interna, el impacto o relevancia y la utilidad o aplicabilidad clínica (que entre otras cuestiones analiza la validez externa del estudio). Las cuestiones que se evalúan para medir la validez interna de los diferentes estudios difieren entre un tipo u otro, sin embargo el impacto y la utilidad no es muy diferente, quizá a excepción de los

artículos de pruebas diagnósticas que tienen una terminología diferente para mostrar sus resultados.

Todos los aspectos de un estudio y por tanto de un artículo pueden estar sometidos a errores, de los cuales los más importantes están relacionados con el diseño del mismo. Los llamados sesgos son frecuentes en la investigación clínica y comprometen la validez interna del estudio y por tanto la validez externa y la aplicación de los hallazgos a otros pacientes.

Se puede resumir las principales características de la VALIDEZ INTERNA de los estudios así como los sesgos más frecuentes que se pueden encontrar y las medidas para controlarlos de la siguiente forma:

a) Objetivos. Cualquier estudio científico debe adecuar sus objetivos mediante la clarificación de sus elementos esenciales. La población estudiada debe de ser la adecuada con una definición clara de los criterios de inclusión y exclusión; el objeto de estudio (ya sea una intervención, prueba diagnóstica, evaluación de un efecto nocivo o el análisis de un factor pronóstico) debe estar perfectamente definida así como la estrategia a comparar. Finalmente el resultado del estudio debe ser claro, sencillo de medir, reproducible e importante.

b) Momento de inicio del estudio. Según el momento de inicio del estudio, estos pueden clasificarse como retrospectivos (inicio del estudio en el presente con datos que han ocurrido en el pasado) y prospectivos (cuando se inicia el estudio y los datos van recojiéndose a medida que van apareciendo) e incluso podría ser que el diseño de un estudio fuera una combinación de ambos. El rigor metodológico de los estudios prospectivos siempre es mayor, debido a que en los retrospectivos la calidad de los datos necesarios suele estar siempre comprometida y la presentación de datos no disponibles (*missing*) puede llegar a ser muy elevada.

c) Muestra. La inmensa mayoría de estudios se realizan mediante muestras, al ser imposible abarcar toda la población con el problema de interés. Es un buen indicador metodológico que cuando se plantea un estudio se realice previamente el cálculo del tamaño muestral del mismo que puede indicarnos la factibilidad o no del estudio. La forma de selección de la muestra estudiada se puede conseguir de varias formas y depende del tipo de estudio analizado. Así, en el ensayo controlado aleatorizado, los pacientes que son estudiados deben cumplir unos criterios de selección que suelen ser muy restrictivos en ocasiones y que pueden comprometer la validez externa, asignándose a uno de los grupos de estudio de aleatorización disponibles, pudiendo utilizarse las diferentes formas que existen para asignar a los pacientes a un grupo u otro. Además, es muy importante la ocultación de la secuencia de aleatorización (el investigador no debe saber a qué grupo va asignarse el próximo paciente) que no tiene que ver con el enmascaramiento de la intervención. Sin embargo, en los estudios de pruebas diagnósticas se prefiere los pacientes que de forma consecutiva cumplen los criterios de selección y son lo más parecidos

posible al espectro clínico habitual de la enfermedad, siendo lo ideal que los pacientes estudiados representen todo el espectro de la enfermedad, desde los estadios iniciales de la enfermedad a los más avanzados. En los estudios de cohortes donde se analizan factores de riesgo o agentes nocivos o daño, la muestra utilizada es de sujetos que están libres del desenlace estudiado en el momento de inicio del estudio. Finalmente los estudios de casos y controles suelen ser estudios que se realizan de forma retrospectiva y donde los sujetos ya se sabe que tienen la enfermedad o circunstancia a estudio comparándose con sujetos que no tienen la enfermedad, para detectar causas o factores que explican la aparición de enfermedad.

Un tipo de estudio especial son las revisiones sistemáticas/meta-análisis donde el objeto de estudio no son pacientes sino trabajos de investigación. En este tipo de estudio, lo ideal es analizar todo lo que se encuentra publicado y no publicado en relación con el problema de estudio. A la hora de la búsqueda de artículos a incluir es recomendable que se realice por dos personas independientes, así como la determinación de la calidad de cada uno de los estudios incluidos. Un sesgo característico en este tipo de estudios es el sesgo de publicación, donde el error por los investigadores viene determinado por la asunción de que han evaluado todo lo que “existe” sobre un problema, y no es cierto. Para controlar este sesgo disponemos de varias herramientas entre las que se cuenta la representación gráfica (*funnel plot*) o más complicadas con pruebas estadísticas más objetivas, como son la prueba de *Begg* o la de *Egger*.

d) Seguimiento. Varios aspectos en el seguimiento de los pacientes en los estudios de cohortes y en el ensayo controlado y aleatorizado (que al fin y al cabo es un estudio de cohortes especial) son esenciales para asegurar un correcto rigor metodológico. En primer lugar, la duración del seguimiento debe ser lo suficientemente prolongada para garantizar la presencia del evento de interés. En segundo lugar, las mediciones durante el seguimiento, así como las intervenciones que se realicen, en todos los pacientes deben de ser similares. Finalmente, hay que tener en cuenta las pérdidas en el seguimiento. Los pacientes se siguen durante un tiempo determinado que en ocasiones va a producir que algunos sujetos no acaben el estudio. Las pérdidas excesivas en el seguimiento pueden alterar los resultados del estudio, y deben ser analizadas no solo en cuanto a número sino en cuanto a motivos. El número de pérdidas en el seguimiento que compromete la validez interna de una cohorte nunca debe de superar el 20% del total, pero sería más exacto indicar que depende del número de sujetos perdidos y de la frecuencia del evento a estudiar. A medida que el evento es menos frecuente las pérdidas de seguimiento “toleradas” deben de ser menores.

Un problema especial en cuanto al seguimiento del paciente lo presentan las violaciones de protocolo de los ensayos aleatorizados, donde un paciente puede incluso ser asignado a un grupo de tratamiento y recibir el otro. Para evitar los sesgos derivados de estos problemas, se puede hacer análisis de sensibilidad, donde se analizan los resultados asumiendo que todos los pacientes con pérdidas de seguimiento, en ambos grupos de

estudio, han tenido un efecto beneficioso, y después se analizan los resultados asumiendo que los pacientes perdidos han presentado un resultado perjudicial. Si los resultados difieren mucho entre el análisis de los pacientes que han terminado el estudio, del que se obtienen asumiendo el beneficio de los perdidos y el que se obtiene asumiendo que los perdidos han ido mal, existe un problema de validez derivado de la pérdidas en el seguimiento. Actualmente se aconseja que los pacientes con desviaciones en el protocolo, y que incluso pueden recibir una intervención diferente a la que se le asignó inicialmente, deben de ser analizados en el grupo de asignación inicial (análisis por intención de tratar). Aunque se pueden presentar los resultados de los grupos que hayan cumplido de forma correcta el protocolo y el seguimiento, este análisis debe de ser complementario del anterior.

e) Enmascaramiento. Los resultados de cualquier estudio pueden verse alterados por las expectativas, conscientes e inconscientes, de los investigadores e incluso de los pacientes. Los ensayos controlados y aleatorizados son el paradigma de este problema. Para evitar este sesgo se intenta que ni los pacientes, ni los médicos o investigadores que los atienden, ni los que miden o procesan los resultados conozcan a qué grupo de intervención se ha asignado el paciente. A veces es imposible mantener el cegamiento del paciente ni del médico que lo atiende, pero siempre que se pueda debe mantenerse por la persona que mida resultados y maneje los datos informáticos. Pero este problema puede aparecer en otros tipos de estudio. En los de diagnóstico donde se comparan dos pruebas, no se debe conocer el resultado de una de ellas por el que realiza la otra prueba. De igual forma, en los estudios de cohortes y de casos y controles, los que miden factores de riesgo, nocivos o pronóstico, no deben conocer si el paciente tiene la enfermedad a estudio y viceversa.

f) Medición de variables y resultados. Un problema en cualquier tipo de estudio es el de la medición de las variables, que siempre debe realizarse de forma uniforme en todos los pacientes. Debido a ello, cualquier medición que se realice en un estudio clínico debe hacerse por personal experto, independiente, que desconozca a qué categoría o grupo pertenece cada uno de los pacientes. De esta forma se minimizaría la posibilidad de que el investigador “investigara” con más intensidad la aparición del evento a estudio, o de un variable de interés, en un paciente que en otro.

Otro problema derivado de la medición de las variables aparece en los estudios de pruebas diagnósticas donde la realización de las mismas debe realizarse por diferentes investigadores, y que pueden presentar un grado diferente de experiencia en la realización de dichas pruebas. Para este tipo de estudios es importante que el personal investigador esté debidamente formado y tenga amplia experiencia en la prueba a realizar.

Por último, la presencia de varios investigadores realizando el estudio y realizando mediciones de las variables obliga la realización de estudios de fiabilidad, concordancia entre los resultados obtenidos entre diferentes observadores, así como entre diferentes resultados en el mismo paciente, pudiendo realizarse mediante diferentes técnicas estadísticas de análisis de concordancia intra e interobservador.

g) Tipo de variable de resultado. Cuando realizamos un estudio precisamos la medición de una o varias variables de resultado. En el apartado anterior se ha nombrado algunos de los problemas que podemos encontrar en la medición de los resultados y algunas de las posibilidades para minimizar el riesgo de que ocurran. Pero el rigor científico de un artículo puede estar relacionado con la variable de respuesta, y lo que intentamos es realizar estudios que midan variables de respuestas adecuadas y con una suficiente relevancia clínica.

En muchas ocasiones los estudios se basan en mediciones de variables intermedias (por ejemplo, podíamos plantear un estudio sobre un fármaco en el tratamiento de determinada infección y la variable de resultado es la normalización de un parámetro analítico más que de una erradicación microbiológica o curación clínica). Un resultado a alcanzar en un estudio debe ser fácil de medir, que se pueda realizar mediante mediciones objetivas y que sea lo más relevante posible.

Por último, los estudios que se realicen no solo deben hacer hincapié en los efectos beneficiosos del producto, sino que debe tenerse en cuenta los posibles efectos perjudiciales e incluso el coste derivado del mismo. Mientras que los datos referentes a efectos secundarios o perjudiciales de una intervención son cada vez más frecuentemente aportados por los investigadores, el problema económico, mucho más difícil de calcular, solo se presenta en una minoría de trabajos.

h) Consideraciones estadísticas. Un adecuado rigor metodológico de un estudio en relación con el análisis de datos, comprende no solo la elección de una variable de resultado adecuada y relevante sino además de una serie de variables independientes que puedan explicar la evolución de la variable resultado debido a alguna o algunas de ella. El análisis de resultados en los grupos creados, intentando encontrar una diferencia entre ellos, y la realización de una prueba estadística que indique si esta diferencia es poco probable que esté relacionada con el azar, sino que, probablemente, sea debida a un factor (valor p de inferencia estadística) es un concepto demasiado reduccionista. En la actualidad, se precisa que los estudios no solo muestren una medida puntual en la comparación de resultados, por ejemplo, entre los dos grupos sino, además, que nos muestren la precisión con la que estos resultados han sido obtenidos. Debido a ello cualquier presentación de resultados que se realice, reducción absoluta de riesgo, reducción relativa de riesgo, riesgo relativo, número necesario para tratar para prevenir un evento, número necesario para evitar un evento perjudicial, debe presentarse con sus intervalos de confianza al 95%, tal como se ha comentado en capítulos previos.

Por último, aunque en general el análisis estadístico de la mayoría de estudios realizados suelen ser sencillo, a veces se precisan técnicas estadísticas complejas, con control de factores de confusión, análisis de subgrupos, análisis de gradación de dosis respuesta, estratificación de resultados etc., que hacen difícil la comprensión de los mismos, pero no debemos olvidar que el fin último de la aplicaciones de estas pruebas estadísticas, que muchas veces sobrepasan nuestros conocimientos, es el medio para mostrarnos una conclusión y no una conclusión en sí misma.

i) **Validez externa.** Finalmente, cualquier estudio realizado en el ámbito sanitario tiene como fin el que sus resultados puedan ser aplicados a otros ámbitos sanitarios. La validez externa del estudio permite que un tratamiento farmacológico en una determinada patología realizada en un país, habiendo mostrado un resultado beneficioso, pueda utilizarse en otro ámbito geográfico diferente. Unos adecuados criterios de selección, inclusión y exclusión, deben estar perfectamente descritos en todos los estudios. Pero también debe tenerse en cuenta que, a veces, las características sociodemográficas, educacionales o culturales pueden diferir mucho de un ámbito a otro y dificultar o incluso invalidar la validez externa del estudio. Este problema es mayor cuando se evalúan pruebas diagnósticas o terapéuticas en centros sanitarios con muchísima experiencia o con unos medios tecnológicos diferentes al resto de centros. Por tanto, a la hora de aplicar los resultados de un estudio deben tenerse en cuenta estos factores.

j) **Conclusiones.** En muchas ocasiones las conclusiones que se presentan en un artículo tienen que ver más con las expectativas creadas por los investigadores que con los resultados mismos del estudio. Las conclusiones deben limitarse a aquellas derivadas de la comparación entre los grupos creados, y a la hora de enunciarlas deben tenerse en cuenta los posibles sesgos que pueden haber cometido los investigadores y que podrían poner, si no todas, sí al menos algunas conclusiones en entredicho.

4. Puntos clave

1. La realización de un estudio de investigación y su plasmación en una revista científica debe hacerse con el mayor rigor científico posible para asegurar la ausencia de errores que comprometan las conclusiones alcanzadas.
2. La Validez Interna de un artículo científico esta sometida a múltiples sesgos en los diferentes pasos del proceso de investigación y deben tenerse en cuenta para poner los medios para evitarlos o minimizarlos.
3. La Validez Externa de un artículo, que puede ser independiente de la validez interna, debe ser máxima para preservar la aplicabilidad y utilidad del factor estudiado.

Bibliografía

1. Glasziou P, Del Mar C, Salisbury J. Evidence-based Medicine Workbook. Finding and applying the best evidence to improve patient care. London. *BMJ Publishing Group* 2003.
2. Greenhalgh T. How to read a paper. *The basic of evidence-based medicine*. Fourth edition. London. Wiley-Blackwell 2010, pp. 78-89.
3. *Atención sanitaria basada en la evidencia. Su aplicación a la práctica clínica*. Murcia. Consejería de Sanidad de la Región de Murcia. 2007.

20

Lectura crítica de un metaanálisis

FERNANDO NAVARRO MATEU¹

VÍCTOR SORIA ALEDO²

1) Médico Especialista en Psiquiatría. Jefe de Estudios de la Unidad Docente Multiprofesional de Salud Mental de la Región de Murcia. Subdirección General de Salud Mental y Asistencia Psiquiátrica. Servicio Murciano de Salud.

2) Médico Especialista en Cirugía General y del Aparato Digestivo. Facultativo Especialista de Área Cirugía General. Coordinador de Investigación. Hospital Universitario Morales Meseguer de Murcia. Servicio Murciano de Salud.

1. Introducción

Escenario Clínico

A lo largo de su residencia se ha encontrado con diversos pacientes consumidores de alcohol en diversas rotaciones. Al margen de la posibilidad de derivación al especialista correspondiente, usted tiene dudas sobre si es posible incorporar algún tipo de intervención que pueda incorporarse en su práctica habitual para mejorar el pronóstico de estos pacientes.

Usted ha comentado este tema con su tutor y con otros compañeros. Tras revisar el tema, se da cuenta de que se han publicado numerosos estudios sobre las intervenciones breves (consideradas como consejos educativos o intervenciones terapéuticas de baja intensidad con el objetivo de informar y reducir los peligros asociados al alcohol). Estos estudios sugieren su eficacia en la reducción del consumo de alcohol. Sin embargo, la mayoría de estudios sobre este tipo de intervenciones están realizados en países anglosajones. Dadas las diferencias entre los factores socio-culturales entre este tipo de países y los del ámbito mediterráneo, se plantea dudas sobre si este tipo de intervenciones, probablemente muy influenciadas por este tipo de factores, serían eficaces en nuestro entorno.

En este capítulo vamos a describir la forma de solucionar el escenario planteándolo desde la perspectiva de la Práctica Clínica Basada en Evidencias (PCBE) con el siguiente esquema: elaboración de la pregunta clínica, elaboración de una estrategia de búsqueda, lectura crítica del artículo seleccionado, resolución del escenario combinando los diferentes elementos y elaboración del informe resultante (CAT o *Critically Appraised Topic*) resultante.

2. Elaboración de la pregunta clínica y estrategia de búsqueda de información

Ante cualquier duda que se nos pueda plantear en nuestra práctica diaria, el primer paso es clarificar y recopilar la mayor cantidad posible de información sobre el paciente y sus circunstancias para poder construir una pregunta clínica según el formato PICO (ver tema 15). El objetivo de una pregunta bien construida es facilitar la elaboración de la mejor estrategia de búsqueda de información, que permita optimizar las posibilidades de encontrar información relevante.

En el tema 15 se han desarrollado en profundidad los pasos para la elaboración de preguntas clínicas en el formato PICO (*Patient, Intervention, Comparison and Outcome*). En nuestro escenario, el resultado podría ser:

	1 Paciente o problema	2 Intervención	3 Intervención a comparar	4 Resultado
Orientaciones	¿Cómo podría describir un grupo de pacientes similar al mío?	¿Cuál es la intervención de interés?	¿Cuál es la intervención alternativa con la que comparar?	¿Cuál es el resultado que puedo esperar de la intervención?
Ejemplo	Bebedores excesivos de alcohol. Tratamiento comunitario (ambulatorio) no especializado. Población española (mediterránea).	Intervenciones breves basadas en el consejo terapéutico por profesionales sanitarios.	Tratamiento ambulatorio habitual.	Reducción del consumo de alcohol.
Pregunta clínica: “En pacientes con consumo excesivo de alcohol y en tratamiento comunitario (ambulatorio) no especializado ¿son eficaces las intervenciones breves para reducir el consumo de alcohol frente al tratamiento habitual en España?”				

Además de plantear la pregunta clínica, es importante contestar a otros dos apartados complementarios para facilitar la elaboración de una adecuada estrategia de búsqueda: a) definir el tipo de pregunta y b) valorar el tipo de estudio más adecuado para contestarla.

Básicamente, las preguntas pueden agruparse bajo 4 grupos principales según el tipo de información que se busque. Estos grandes grupos son: tratamiento, pronóstico, diagnóstico y etiología o daño. La importancia de este tipo de reflexión radica en su relación con los diferentes diseños de investigación (meta-análisis y/o revisiones sistemáticas, ensayos clínicos aleatorizados, estudios observacionales y los estudios transversales).

El escenario planteado sugiere una duda relacionada con la eficacia de una intervención terapéutica, por lo que los estudios más adecuados que tendríamos que buscar de forma preferente serían los meta-análisis y/o revisiones sistemáticas y los ensayos clínicos aleatorizados. Suele ser preferible seleccionar un meta-análisis sobre un tema de interés, siempre que sea posible, frente a otros diseños, puesto que sistematizan e integran la información científica disponible en una sola publicación (ver tema 7).

Inicialmente nos planteamos utilizar la base de datos bibliográficos PubMed (ver el capítulo correspondiente sobre la utilización del PubMed) empleando los siguientes descriptores del Thesaurus: “Psychotherapy”[Mesh] AND “Alcohol Drinking”[Mesh]”. En una búsqueda realizada el 1 de septiembre de 2011 encontramos 925 artículos. Este número es excesivo para una rápida evaluación de los mismos. Es conveniente la utilización de algunos limitadores. En concreto, nuestro interés se centra en los meta-análisis. Con este limitador, el número de artículos desciende a 11 referencias bibliográficas.

Tras una rápida lectura de los títulos, detectamos varias referencias que despiertan nuestro interés. En esta fase, es importante centrarse en nuestra pregunta previamente

elaborada y no distraer nuestra atención. En todo caso, podemos recopilar la información encontrada para su posterior análisis. La referencia número 11 (*Effectiveness of medical advice for reducing excessive alcohol consumption. Meta-analysis of Spanish studies in primary care*). Ballesteros J, Ariño J, González-Pinto A, Querejeta I. *Gac Sanit*. 2003 Mar-Apr;17(2):116-22. Spanish) parece ser la más adecuada. La lectura del resumen de la publicación confirma nuestra apreciación. Además, es de acceso libre (confirmado por el siguiente texto "free article"). En este caso, está escrito en español, por lo que, según nuestro nivel de lectura científica en inglés (una competencia esencial en estos días), puede facilitarnos su lectura y comprensión.

3. Lectura crítica de un Meta-análisis

La lectura crítica de un artículo científico consiste en una lectura atenta, valorando aquellos puntos críticos que pueden poner en riesgo la validez del estudio antes de fijarse en los resultados obtenidos por los autores del artículo. La lectura crítica de estos aspectos metodológicos suele ser incómoda para la mayoría de los clínicos. Los impulsores del movimiento de la PCBE, conscientes de estas dificultades en la lectura crítica, editaron numerosos artículos de ayuda agrupados bajo la denominación de "Guías de Ayuda para los Usuarios de Literatura Médica". Estas guías fueron publicadas originalmente en la revista *JAMA* y posteriormente se tradujeron al castellano en la edición española de la misma revista. Otras fuentes (por ejemplo, CASPE <http://www.redcaspe.org/>) han publicado diversas adaptaciones o variaciones de las anteriores. En la tabla 1 se incluyen las preguntas sugeridas por los expertos para la lectura crítica de un artículo meta-analítico. Estas preguntas están agrupadas en 3 grandes apartados que permiten analizar la validez del estudio, sus resultados y su aplicación en la resolución del escenario de partida. Para una mayor profundización en los conceptos metodológicos utilizados es imprescindible la lectura atenta del tema 7 de este programa.

Tabla 1. Guía del usuario para un artículo meta-analítico

A) ¿Son válidos los resultados del estudio?

1. ¿Se aborda un tema focalizado?
2. ¿Fueron apropiados los criterios utilizados para la inclusión de los artículos a seleccionar?
3. ¿Es poco probable que se pasaran por alto estudios relevantes importantes?
4. ¿Se evaluó la validez de los estudios incluidos?
5. ¿Fueron reproducibles las evaluaciones de los estudios?
6. ¿Fueron similares los resultados de estudio a estudio?

B) ¿Cuáles son los resultados?

1. ¿Cuáles son los resultados globales de la revisión de conjunto?
2. ¿Con qué precisión se estimaron los resultados?

C) ¿Serán de utilidad los resultados para la asistencia de mis pacientes?

1. ¿Se consideraron todos los resultados clínicamente importantes?
2. ¿Los beneficios compensan los inconvenientes o los costes?

Oxman AD, Cook DJ y Guyatt GH. Guías para usuarios de la literatura médica VI. Cómo utilizar una revisión de conjunto. *JAMA* (Ed. Esp.) 1997: 62-67.

A) ¿Son válidos los resultados del estudio?**1. ¿Se aborda un tema focalizado?**

Cuanto más concreto y focalizado sea el tema analizado en un meta-análisis, más probable será que pueda dar una respuesta veraz a la pregunta de investigación planteada. En este caso, el objetivo principal del estudio (la actualización de la evidencia sobre la eficacia del consejo médico, como intervención breve en bebedores de riesgo, en España) está claramente descrito y es suficientemente concreto y focalizado.

2. ¿Fueron apropiados los criterios utilizados para la inclusión de los artículos a seleccionar?

Los criterios de inclusión y exclusión de los artículos a seleccionar utilizados por los autores deben ser claros y coherentes con el objetivo planteado. Los meta-análisis pueden analizar diferentes áreas de interés relacionadas con un mismo tema, por tanto, los diseños de los estudios seleccionados deben ser apropiados para responder a la pregunta de investigación planteada. Para analizar la eficacia de un abordaje terapéutico, el mejor diseño de investigación son los Ensayos Clínicos Aleatorizados (ECA). En este estudio, los criterios de inclusión/exclusión son explícitos y adecuados (en el ámbito de la atención primaria y ECA realizados en población española).

3. ¿Es poco probable que se pasaran por alto estudios relevantes importantes?

Existen numerosas fuentes de información científica, tanto de artículos publicados, como de literatura gris. Es importante que los autores expliciten la estrategia de búsqueda utilizada, así como todas aquellas fuentes consultadas. Cuantas más fuentes se hayan consultado, mejor (lo ideal es un mínimo de 3 fuentes independientes). Entre ellas destacan: la utilización tanto de bases de datos bibliográficas como MEDLINE y EMBASE, la verificación de la bibliografía de los artículos encontrados y los intentos por realizar contactos con expertos en el tema. Aunque la exhaustividad en la estrategia de búsqueda de información permite minimizar el sesgo de publicación, es conveniente analizar su posible existencia. Este análisis puede realizarse mediante diversas técnicas, tanto de tipo gráfico (realización de gráficos de embudo o “*funnel plot*”) o estadísticas. Los autores deben explicar el método utilizado.

En el estudio que estamos analizando, los autores realizaron búsquedas en diversas bases de datos bibliográficas internacionales y nacionales, en páginas web relacionadas, resúmenes disponibles de congresos y realizaron una consulta con expertos españoles en el tema. Aunque esta estrategia es completa, los autores no analizan expresamente el posible sesgo de publicación. Entre las limitaciones del estudio que presentan en la discusión del artículo mencionan su posible existencia. Si los autores hubieran intentado analizar su presencia y/o cuantificar su efecto hubiera mejorado la calidad metodológica del estudio.

4. ¿Se evaluó la validez de los estudios incluidos?

Es importante que los autores analicen la calidad de los estudios individuales que encuentran, puesto que la revisión de los expertos y su posterior publicación no la garantiza completamente. Existen numerosas escalas para la medir la calidad de cada uno de los diseños de investigación (por ejemplo, para analizar la calidad de los ECA pueden utilizarse diversas escalas, entre las que destacan la Escala de Jadad de 5 puntos (la más utilizada en los meta-análisis realizados por la COCHRANE) o la Escala CONSORT (*Consolidated Standards for Reporting of Trials*). Una vez analizada la calidad de los diferentes estudios, ésta puede ser utilizada de diferentes formas (como un criterio de exclusión de aquellos estudios de baja calidad, para realizar un análisis estratificado en función del nivel de calidad o para un análisis de sensibilidad).

En el artículo que estamos analizando no se menciona explícitamente la utilización de alguna de las escalas más conocidas para analizar la calidad de los estudios. Sin embargo, los autores han incluido diversos criterios de calidad aceptados para los ECA entre los criterios de inclusión/exclusión, como la comparación en paralelo de la intervención de interés frente a otra intervención no especificada y la asignación aleatoria de los sujetos a las diferentes intervenciones.

5. ¿Fueron reproducibles las evaluaciones de los estudios?

Se considera de buena calidad que, al menos, dos revisores independientes participen en el análisis y extracción de los datos de los estudios seleccionados. Esta estrategia protege frente a posibles errores individuales. El análisis del acuerdo entre revisores y/o la solución de los potenciales desacuerdos mediante el debate y consenso entre ellos, permite al lector estar más seguro de los resultados presentados. El artículo que analizamos cumple este criterio de forma explícita.

6. ¿Fueron similares los resultados de estudio a estudio?

Es habitual que en un meta-análisis se encuentren diferencias importantes en diversos aspectos (características de los pacientes, exposiciones, medición de los resultados, en los métodos utilizados, etc.). Estas diferencias pueden modificar los resultados obtenidos. Para analizar si estos resultados son similares entre sí en los estudios analizados y valorar la posibilidad de combinarlos cuantitativamente para obtener un indicador resumen del efecto obtenido, los investigadores deben analizar si las diferencias observadas son mayores de lo que se podría esperar si las diferencias observadas fueran debidas sólo al azar. Esto se puede medir estadísticamente mediante las pruebas de homogeneidad/heterogeneidad. Sin embargo, este tipo de pruebas no son muy potentes, de forma que una prueba no significativa no necesariamente descarta una heterogeneidad importante. Esto supone que aquellas diferencias clínicamente importantes, independientemente de la significación obtenida en este tipo de pruebas, debieran ser tenidas en cuenta por los autores en los análisis posteriores.

Los autores realizan la prueba de heterogeneidad con un nivel de probabilidad superior a 0,20. Además, realizan un análisis de sensibilidad teniendo en cuenta un aspecto muy importante en los estudios analizados, las pérdidas durante el seguimiento. Analizan los resultados según el número de pacientes que cumplen el protocolo del ensayo (es decir, sin incluir a aquellos pacientes que se pierden en el seguimiento) y según el número de pacientes inicialmente incluidos en cada uno de los grupos (Análisis por Intención de Tratar). Este último tipo de análisis es más conservador y se le considera más adecuado al aproximarse más a lo que sucede en la realidad clínica.

B) ¿Cuáles han sido los resultados del estudio?

1. ¿Cuáles son los resultados globales de la revisión de conjunto?, y

2. ¿Con qué precisión se estimaron los resultados?

Tras la valoración crítica de la validez del estudio, podemos centrar nuestra atención en los resultados del meta-análisis y en su precisión. Es conveniente familiarizarse con la interpretación de las diferentes medidas de efecto que pueden ser utilizadas en investigación. En los últimos años se prefiere utilizar el Número Necesario para Tratar (NNT) por ser considerado como una medida de interpretación más intuitiva y que permite comparar

los resultados de diferentes intervenciones en función de su importancia práctica (ver las tablas de NNT en www.murciasalud.es/). La precisión del resultado hace referencia al cálculo del intervalo de confianza del 95% (IC95%) de la medida de resultado seleccionada, es decir, aquel intervalo alrededor de la medida puntual obtenida en el que, con una probabilidad preestablecida del 95%, se incluirá el efecto real de la intervención.

En nuestro estudio, la intervención breve presenta un efecto positivo significativo, tanto para la reducción de la cantidad de alcohol ingerido por semana ($d = -0,46$; IC95% = $-0,63$ a $-0,29$) como aumentando el riesgo (medido como *Odds Ratio* –OR–) de reducir el porcentaje de bebedores. Para esta última medida realizan un análisis de sensibilidad para valorar el efecto de la pérdida de pacientes en el seguimiento. Para ello realizan los mismos análisis teniendo en cuenta aquellos pacientes que finalizan el protocolo (análisis por cumplimiento de protocolo) (OR = 2,59; IC95% = 1,74 a 3,85) y aquellos pacientes que fueron asignados a cada una de las intervenciones (Análisis por Intención de Tratar, AIT) (*Intention To Treat, ITT*) (1,55; 1,06 a 2,26). En general, los AIT proporcionan resultados más conservadores y son preferidos por medir el efecto de la intervención en una situación más semejante a lo que sucede en la práctica clínica real, donde también se producen retiradas y abandonos en los tratamientos.

Normalmente, cuando el resultado que se espera es la reducción de una enfermedad o, en este caso, del porcentaje de bebedores, se espera obtener una OR < 1. En este caso, obtienen una OR > 1 porque han calculado el riesgo de ser no bebedor (definido por un consumo < 35 U/semana en el varón y < 21 U/semana en mujeres). Por otro lado, los autores no calculan el NNT de la intervención. Nosotros podemos hacer todos estos cálculos fácilmente aplicando las fórmulas correspondientes o introduciendo los datos facilitados por los autores en alguna de las calculadoras estadísticas disponibles en la web (<http://www.infodoctor.org/rafabravo/mberecu.htm>). Por ejemplo, utilizando la calculadora específica para la evaluación de tratamientos desarrollada por la Unidad de Bioestadística Clínica del Hospital Ramón y Cajal de Madrid (<ftp://ftp.hrc.es/pub/programas/calcu/evaltrat/evaltrat.htm>) los datos que se obtienen pueden verse en la siguiente tabla. Obsérvese que la OR obtenida es < 1, porque en este caso calcula el riesgo de ser considerado como bebedor de riesgo y, al ser efectivo el tratamiento aplicado, este riesgo disminuye tras la intervención. En resumen, sería necesario realizar una intervención breve a 9 pacientes bebedores en riesgo de alcoholismo para que uno de ellos pase a ser considerado por debajo de los niveles de riesgo considerados en el estudio en un período de seguimiento entre 6 y 18 meses.

Datos					Resultados		
Grupo	Bebedor	No bebedor	Total	Riesgo	Valor	IC 95%	
Experimental	138	98	236	0,58	Reducción absoluta del riesgo (RAR)	10,56% 1,96% a 19,15%	
Control	165	74	239	0,69	Riesgo Relativo (RR)	0,84 0,73 a 0,97	
Total	303	172	475		Reducción relativa del riesgo (RRR)	15,30% 2,84% a 27,75%	
					Odd Ratio (OR)	0,63 0,43 a 0,92	
					Número necesario a Tratar (NNT)	9 5 a 51	

C) ¿Me resultarán útiles los resultados para la asistencia a los pacientes?

1. ¿Se consideraron todos los resultados clínicamente importantes?, y
2. ¿Los beneficios compensan los inconvenientes o los costes?

En este apartado es necesario reflexionar sobre la generalización de los resultados que obtienen los autores y su aplicabilidad en nuestro entorno asistencial. El análisis de los posibles beneficios en relación con los costes y riesgos asociados. Además, es muy importante tener en cuenta los valores y las preferencias de los pacientes para poder incorporarlos en la toma de la mejor decisión clínica disponible.

4. Resolución del escenario clínico

Tras la lectura crítica del artículo seleccionado, podemos concluir que las intervenciones breves en bebedores excesivos en riesgo de alcoholismo en el ámbito de la Atención Primaria en España son eficaces para la reducción del consumo semanal de alcohol y del número de bebedores excesivos en un período entre 6 y 18 meses de seguimiento tras la intervención. Tras el trabajo realizado, decide presentar este estudio en la próxima sesión bibliográfica del servicio para debatirlo con sus compañeros. Además, aprovecha la ocasión para realizar un CAT (*Critically Appraised Topic*) o informe de su lectura crítica con un formato establecido utilizando el software Cat-Maker (www.murciasalud.es/) (ver anexo I).

En resumen, mediante la aplicación de las estrategias que se proponen desde la Práctica Clínica Basada en las Evidencias (PCBE) hemos sido capaces de centrar el problema de interés mediante la elaboración de la pregunta clínica en formato PICO, plantear una estrategia de búsqueda de información bibliográfica eficaz y realizar la lectura crítica del artículo encontrado. Además, hemos sido capaces de elaborar un registro documental del trabajo realizado y ya tenemos prácticamente preparada la próxima sesión que tengamos que presentar en nuestro servicio.

Anexo I. CATS (*Critically Appraised Topics*)

Los *Critically Appraised Topics* (CATs) o temas evaluados críticamente son unos breves informes muy estructurados elaborados por y para los clínicos a partir de la lectura crítica de algún artículo científico. En ellos queda reflejado todo el proceso de práctica de la MBE (elaboración de preguntas clínicas, estrategias de búsqueda de evidencias, lectura crítica del artículo de forma estandarizada...). Su longitud no es mayor de un folio, donde se resume todo el proceso que se lleva a cabo (pregunta clínica, estrategia de búsqueda, lectura crítica...), de forma que, de un vistazo, podemos recordar los puntos clave del artículo original sin tener que realizar el esfuerzo de leerlo de nuevo.

Para la elaboración de los CATs se diseñó un programa "ad hoc", el CAT-Maker. Este programa es una aplicación informática que facilita la realización de un CAT desarrollada por el NHS R&D *Centre for Evidence-Based Medicine* de Oxford, basándose en el trabajo previo del Servicio de Medicina Interna de la Universidad de McMaster. Se puede acceder a una versión gratuita en <http://www.cebm.net/catmaker.asp>, junto con el acceso a un tutorial de ayuda. Ya se han desarrollado varias versiones del CAT-Maker, incluso se ha traducido una de ellas al castellano (disponible en el área de descarga del directorio de recursos de la sección de Práctica Clínica Basada en las Evidencias de MurciaSalud (<http://www.murciasalud.es/pagina.php?id=1683&idsec=22>)).

El CAT-Maker permite resolver tres situaciones posibles en las que nos podemos encontrar al realizar una lectura crítica de un artículo científico: i) facilita la realización de nuevos cálculos al disponer de calculadoras especialmente diseñadas para su utilización en el contexto de la lectura crítica desde una perspectiva de la MBE; ii) facilita la elaboración de un informe breve y estructurado que nos permite recordar el trabajo realizado tiempo después o documentar las sesiones de lectura crítica de artículos científicos realizadas en el servicio; y iii) el mismo informe permite compartir la información con otros compañeros de nuestro entorno o a través de internet.

Una ventaja añadida es el extraordinario valor educativo que tiene la elaboración de CATs, ya que permite la combinación de los elementos esenciales de la MBE (evidencia, experiencia y a los pacientes). Por supuesto, este valor educativo es mayor si realizamos personalmente los CATs que si leemos los realizados por otros. Aunque se pueden elaborar CATs a título individual, suelen realizarse en el contexto de un trabajo en equipo como forma de compartir la información elaborada. Su éxito se demuestra con la profusión de bancos de CATs que han proliferado en los últimos años en internet.

Sin embargo, los CATs no están exentos de importantes limitaciones que pueden minimizarse si se conocen bien:

1. Los CATs son informes críticos elaborados que suelen estar centrados en la lectura crítica de un artículo, por lo tanto, están sometidos a las limitaciones inherentes a esta situación. Suelen estar basados en búsquedas bibliográficas realizadas de

forma rápida y no exhaustiva de la bibliografía disponible y abordan aspectos parciales del tema que nos interesa. En otras palabras, no son representativas de las evidencias disponibles y no son revisiones sistemáticas sobre un tema.

2. La vida media de los CATs puede ser corta y quedarse obsoletos en cuanto se hayan publicado nuevas investigaciones. Aunque se suele incorporar una fecha propuesta de caducidad en el CAT, no deja de ser conveniente tener la precaución de realizar una actualización de la búsqueda realizada. Para ello podemos utilizar la misma estrategia de búsqueda que realizamos en su momento y que reflejamos en el CAT.
3. Pueden estar sometidos a posibles errores relacionados con la lectura del artículo original, en la transcripción de datos o conclusiones o en el cálculo de algunos indicadores. Si en nuestro grupo de trabajo utilizamos los CATs como forma de compartir información, sería conveniente su "lectura crítica" en el equipo antes de ponerlo a la disposición general.

A continuación presentamos, a modo de ejemplo, el CAT elaborado como informe de la lectura crítica realizada para la elaboración de este capítulo.

EL CONSEJO MÉDICO ES EFICAZ EN LA REDUCCIÓN DEL CONSUMO DE ALCOHOL EN ATENCIÓN PRIMARIA EN ESPAÑA

Conclusiones clínicas: La intervención breve en bebedores excesivos en riesgo de alcoholismo en Atención Primaria en España es eficaz para la reducción del consumo semanal de alcohol y del número de bebedores excesivos en un período de seguimiento entre 6 y 18 meses.

Cita bibliográfica: Ballesteros J, Ariño J, González-Pinto A y Querejeta I: Eficacia del consejo médico para la reducción del consumo excesivo de alcohol. Meta-análisis de estudios españoles en atención primaria. *Gac Sanit* 2003; 17(2): 116-22.

Pregunta clínica (PICO): En pacientes con un consumo excesivo de alcohol y en tratamiento comunitario (ambulatorio) no especializado, ¿son eficaces las intervenciones breves para reducir el consumo de alcohol frente al tratamiento habitual en España?

Estrategia y términos de búsqueda bibliográfica: En Pubmed. Palabras clave: "Psychotherapy" [Mesh] AND "Alcohol Drinking" [Mesh]. *Límites:* Meta-análisis.

La revisión:

- **Fuentes de datos:** Medline, Embase, PsycINFO, CINAHL y Current Contents). Índice Médico Español (INE), Páginas y enlaces web de interés (Sociedad Española de Medicina Familiar y Comunitaria). Resúmenes de Congresos de Socidrogalcohol, Sociedad Española de Toxicomanías. Consulta a expertos españoles en el área de la atención a la problemática alcohólica en asistencia primaria.
- **Selección de estudios:** Criterios de inclusión: a) estudios realizados en población española; b) comparación, por grupos paralelos de tratamiento, del resultado de las intervención breve frente a intervenciones no específicas en consumidores excesivos de alcohol que no presentaran criterios de dependencia alcohólica; c) desarrollados en el ámbito de Atención Primaria; d) asignación aleatoria a las intervenciones; e) seguimiento mínimo de 6 meses.
- **Extracción de datos:** Dos investigadores independientes y resolución de discrepancias por discusión entre ambos
- **¿Revisión múltiple e independiente de los informes individuales? Sí**
¿Prueba de heterogeneidad? Sí

Las pruebas:

La intervención experimental disminuye el consumo de alcohol/semana y aumenta el riesgo de ser clasificado como bebedor por debajo del punto de corte seleccionado (< 35 U/semana en el varón y < 21 U/semana en la mujer) tras la intervención (OR= 1,55; IC95%= 1,06 a 2,26 según el análisis por intención de tratar (AIT) y 2,59; 1,74 a 3,85 según el análisis por cumplimiento del protocolo (APC)).

	Seguimiento	Incidencia en Expuestos (Ie)	Incidencia en No Expuestos (Io)	RR Ie/Io	RRR (1-RR)*100	RAR (Io-Ie)	NNT	IC95%
Reducción del número de bebedores excesivos (Análisis por intención de tratar, AIT)								
	Entre 6 y 18 meses	58%	69%	0,84	16%	0,69-0,58 = 0,11	9	5 a 51
Reducción del número de bebedores excesivos (Análisis por cumplimiento de protocolo, APC)								
	Entre 6 y 18 meses	31%	54%	0,57	43%	0,54-0,31 = 0,23	4	3 a 7

Comentarios:

Como limitaciones del estudio cabe destacar la ausencia de análisis del posible sesgo de publicación. Se han detectado pocos estudios y una alta tasa de abandonos en el seguimiento. Los autores analizan este sesgo realizando un análisis de sensibilidad realizando los cálculos según el análisis por intención de tratar comparándolo con el análisis por cumplimiento del protocolo.

Realizado por: Fernando Navarro Mateu y Víctor Soria Aledo;

Email de contacto: fernando.navarro@carm.es

Viernes, 02 Septiembre 2011

Fecha de actualización: Septiembre 2012

21

Lectura crítica de un ensayo clínico aleatorizado

ANDRÉS CARRILLO ALCARAZ¹

VÍCTOR SORIA ALEDO²

1) Médico Especialista en Medicina Intensiva.
Hospital General Universitario Morales Meseguer de Murcia.
Servicio Murciano de Salud.

2) Médico Especialista en Cirugía General y del Aparato
Digestivo. Facultativo Especialista de Área Cirugía General.
Coordinador de Investigación. Hospital Universitario
Morales Meseguer de Murcia. Servicio Murciano de Salud.

1. Introducción

Escenario Clínico

Un médico internista atiende a un paciente de 20 años, en una consulta externa hospitalaria porque presenta una prueba de Mantoux positiva a raíz de tener contacto con un familiar diagnosticado de tuberculosis. El paciente era estudiante de enfermería y dos años antes le habían realizado esta prueba, a raíz de un caso de tuberculosis que había atendido en urgencias, con resultado negativo. Tras un estudio adecuado se descarta afectación clínica. El internista indica tratamiento con isoniacida durante 6 meses pero el paciente refiere que no quiere recibir durante tanto tiempo el medicamento ya que piensa que no va a seguirlo, por lo que el facultativo se plantea un tratamiento más corto pero con 2 fármacos tuberculostáticos para favorecer la adherencia del paciente al tratamiento. Pero el médico tiene dudas si esta pauta más corta con 2 medicamentos puede producir mayores efectos secundarios, e incluso un mayor riesgo de desarrollar la enfermedad tuberculosa. Debido a ello consulta con un compañero y éste le remite al siguiente artículo: *María Paloma Geijo, Carmen Rosa Herranz, David Vaño, Ángel Jesús García, Montserrat García, José Francisco Dimas. Pauta corta de isoniazida y rifampicina comparada con isoniazida para la infección latente de tuberculosis. Ensayo clínico aleatorizado. Enferm Infecc Microbiol Clin 2007;25(5):300-4.*

Tras la lectura cuidadosa y crítica del artículo, el médico toma una decisión con su paciente.

En este capítulo vamos a describir la forma de solucionar el escenario planteándolo desde la perspectiva de la Práctica Clínica Basada en Evidencias (PCBE).

2. Elaboración de la pregunta clínica y estrategia de búsqueda de información

En el quehacer diario de nuestra profesión se nos plantean múltiples dudas y preguntas que afectan a cualquier campo de la práctica clínica, siendo las más frecuentes las derivadas de un problema diagnóstico y, sobre todo, de un tratamiento. La progresiva introducción de nuevos tratamientos, farmacológicos y no farmacológicos, en todos los campos de medicina, así como la rápida incorporación en la práctica clínica diaria, hace que en muchas ocasiones nos veamos sobrepasados en la actualización de conocimientos necesarios para una correcta atención clínica. Los estudios de investigación, y por tanto los artículos que deberíamos consultar, que abordan la acción de una intervención, tratamiento, medida preventiva o incluso de una prueba diagnóstica, para valorar la eficacia o seguridad de esta intervención, son los ensayos clínicos, y de ellos los más utilizados son

los ensayos controlados aleatorizados. Hay que conocer además que existen estudios, revisiones sistemáticas, que lo que hacen es sintetizar de forma sistemática toda la información sobre una pregunta científica específica (tratamiento, medida preventiva, factor de riesgo, prueba diagnóstica) con el objeto de obtener estimadores combinados del efecto, a veces aplicando las técnicas estadísticas muy específicas (metaanálisis). Las revisiones sistemáticas/metaanálisis de los ensayos controlados aleatorizados son considerados los estudios que más evidencia aportan actualmente a la hora de adoptar decisiones clínicas.

Como se ha comentado en anteriores capítulos, la lectura crítica de un artículo se inicia ante una pregunta que surge, en nuestro medio, la mayoría de las ocasiones en la práctica diaria y los primeros pasos que damos son en orden de convertir esta duda en una pregunta estructurada para lograr una minimización en el tiempo de búsqueda, junto a una optimización de documentos recuperados. En nuestro caso, la pregunta podría ser *“en los pacientes con infección latente tuberculosa la administración de 2 fármacos tuberculostáticos durante un tiempo corto en comparación con el tratamiento con un solo fármaco durante más tiempo, presenta la misma tolerancia y adherencia sin incrementar el riesgo de desarrollar enfermedad”*.

En el escenario propuesto al inicio de este capítulo, los componentes de la pregunta clínica planteada puede proponerse según el formato PICO (*Patient, Intervention, Comparison and Outcome*) de la siguiente forma:

	P Paciente o problema	I Intervención	C Intervención a comparar	O Resultado
Orientaciones	¿Cómo podría describir un grupo de pacientes similar al mío?	¿Cuál es la intervención de interés?	¿Cuál es la intervención alternativa con la que comparar?	¿Cuál es el resultado que puedo esperar de la intervención?
Ejemplo	Pacientes con positividad reciente a la prueba de la tuberculina sin enfermedad demostrable.	Pauta “corta” de 3 meses con 2 fármacos tuberculostáticos.	Pauta “habitual, larga” con un solo fármaco.	Adherencia al tratamiento, seguridad, desarrollo de enfermedad tuberculosa.

Los pasos necesarios para elegir la base de datos donde buscar los artículos relevantes, así como la utilización de estrategias de búsqueda, palabras claves, limitadores, etc., han sido descritos en capítulos anteriores de este libro.

3. Lectura crítica de un Ensayo Controlado Aleatorizado

Existe una gran diversidad de ensayos clínicos, pudiendo clasificarse según el objetivo, metodología, número de centros sanitarios implicados o según la presencia o no de algunos sesgos frecuentes en este tipo de estudio. Quizá lo más importante sea diferenciarlos según sus objetivos (ensayos clínicos fases I a IV) y por la metodología empleada (ensayo controlado cruzado, secuencial o en grupos paralelos). Por volumen e importancia, los ensayos clínicos más analizados y con mayor peso científico son los ensayos clínicos controlados, aleatorizados y de grupos paralelos, en fase III (ensayos clínicos destinados a evaluar la eficacia y seguridad del tratamiento experimental intentando reproducir las condiciones de uso habituales y considerando las alternativas terapéuticas disponibles en la indicación estudiada) o fase IV (ensayos clínicos que se realizan con un medicamento después de su comercialización). Estos dos tipos de ensayo suelen realizarse de forma controlada (el investigador “controla” todas las características, dosis, momento de inicio y fin, vía etc., de la intervención) y aleatorizado (la asignación de cada sujeto a la intervención o no intervención se realiza únicamente por la acción del azar). El paradigma del ensayo clínico es, por tanto, el Ensayo Controlado Aleatorizado (ECA) y la evidencia científica que utilizamos para nuestra práctica clínica se basa fundamentalmente en él.

Se conoce como lectura crítica el proceso de evaluar e interpretar la evidencia aportada por la literatura científica, considerando sistemáticamente los resultados que se presentan, su validez y su relevancia. En otras palabras, el lector de un artículo debe cuestionar de forma “crítica” un artículo más allá de las opiniones y conclusiones de los autores para detectar en qué pueden haberse equivocado o no haber realizado de forma óptima el trabajo de investigación y que determinaría unas conclusiones diferentes, y por tanto el valor o relevancia del artículo.

Existen diferentes escalas o cuestionarios para medir la calidad de un artículo científico, y dentro de ellos los ECA, pero, por su sencillez y facilidad en el aprendizaje, preferimos las recomendaciones de la iniciativa CASP (*Critical Appraisal Skills Programme*. Programa de Habilidades en Lectura Crítica) que ha desarrollado el grupo de trabajo de Medicina Basada en Evidencia de Oxford y han sido adaptadas a nuestro idioma por el grupo CASP España (CASPe).

La evaluación de un ECA viene determinada por los siguientes pasos:

A) ¿Son válidos los resultados del estudio?

1. ¿Se orienta el ensayo a una pregunta claramente definida?
2. ¿Fue aleatoria la asignación de los pacientes a los tratamientos?
2. ¿Fueron adecuadamente considerados hasta el final del estudio todos los pacientes que entraron en él?
3. ¿Se mantuvieron ciegos al tratamiento los pacientes, los clínicos y el personal del estudio?
4. ¿Fueron similares los grupos al comienzo del ensayo?
5. ¿Al margen de la intervención en estudio, los grupos fueron tratados de igual forma?

B) ¿Cuáles son los resultados?

7. ¿Es muy grande el efecto del tratamiento o la intervención?
8. ¿Cuál es la precisión de este efecto?

C) ¿Pueden ayudarnos estos resultados?

9. ¿Pueden aplicarse estos resultados en tu medio o población local?
10. ¿Se tuvieron en cuenta todos los resultados de importancia clínica?
11. ¿Los beneficios a obtener justifican los riesgos y los costes?

Plantilla desarrollada por CASP*Oxford y adaptado por CASP España (CASPe). <http://www.redcaspe.org/hacer-hacemos/herramientas/> (Consultada el 25 de septiembre de 2011).

A) ¿Son válidos los resultados del estudio?

La primera cuestión es si los resultados del estudio pueden considerarse válidos, es decir, si el resultado obtenido está producido verdaderamente por efecto del tratamiento. Se trata de comprobar si el diseño del estudio permite asegurar que el efecto observado es debido al tratamiento y no a otras variables o circunstancias que puedan haber conducido a una falsa conclusión. Si la respuesta es negativa, es decir, si los resultados carecen de validez interna y no representan el verdadero efecto del tratamiento en los sujetos estudiados, no vale la pena perder más tiempo en la lectura del artículo, ya que indica que el estudio presenta algún error grave en su diseño que invalida los resultados. En la plantilla utilizada, las tres primeras preguntas que evalúan la validez de un artículo son de eliminación, lo que determina que si algunas de ellas fuera NO, el artículo se consideraría no válido.

Las preguntas a hacerse serían:

1. ¿Se orienta el ensayo a una pregunta claramente definida?

Esta primera cuestión hace referencia a si la pregunta que aborda el artículo en cuestión está definida mediante los términos empleados en el acrónimo PICO. Por tanto, a la hora de evaluar el artículo deberemos reflexionar sobre si en él está claramente definida la población a estudio, la intervención realizada y con qué se compara, y cuáles son los resultados considerados. La identificación de los elementos de la pregunta clínica nos permitirá, además de evaluar su validez interna (ausencia de sesgo), conocer su aplicabi-

lidad, es decir, la generalización de sus resultados a nuestros pacientes, en el caso de que la muestra utilizada sea representativa.

En nuestro ejemplo de artículo, la comparación de las dos pautas de tratamiento (pauta “corta” con 2 tuberculostáticos frente a una “larga” con un solo fármaco) en pacientes con infección latente tuberculosa como población, el diseño metodológico (ensayo controlado y aleatorizado) y los resultados (tolerancia, adherencia, conversión de infección a enfermedad) están claramente descritos.

2. ¿Fue aleatoria la asignación de los pacientes a los tratamientos?

Para que los resultados de un ECA puedan considerarse válidos, es decir, para que las diferencias en la respuesta observada en los grupos puedan ser atribuidas al diferente tratamiento que ha recibido cada uno de ellos, es necesario descartar cualquier otro factor que pueda ser responsable de la misma. Al diseñar un ensayo clínico es imprescindible formar grupos comparables, es decir, similares en todas las variables pronósticas, ya que si no, podría interferirse con el efecto del tratamiento. La asignación aleatoria de los sujetos a los grupos de estudio es la técnica que tiende a asegurar que se produce una distribución equilibrada de las variables pronósticas, tanto de las conocidas como de las desconocidas, entre los grupos del estudio. Si el proceso es realmente aleatorio, es decir, si sólo interviene el azar, ni el sujeto ni el investigador influyen en la decisión de qué intervención recibirá cada uno de los participantes, de forma que todos ellos tengan la misma probabilidad de recibir cualquiera de los posibles tratamientos.

En nuestro ejemplo se dice que se ha utilizado una aleatorización asignada en los centros de atención donde se reclutan los pacientes, aunque no queda claro qué tipo de aleatorización se utiliza (simple, estratificada, bloques, etc.).

Otro punto importante a la hora de la aleatorización es la ocultación de la secuencia de aleatorización, que consiste en que el investigador desconoce a qué grupo será asignado el próximo paciente que incluya en el estudio. Este punto es muy importante, sobre todo en los estudios abiertos donde los investigadores saben qué tratamiento recibirá el paciente. La ausencia de la ocultación de la secuencia de aleatorización puede condicionar un sesgo a la hora de asignar a los pacientes a un grupo u otro. La ocultación se puede hacer de forma sencilla, como es el caso de nuestro artículo, mediante sobres cerrados y opacos, hasta la utilización un sistema informático on line centralizado.

3. ¿Fueron adecuadamente considerados hasta el final del estudio todos los pacientes que entraron en él?

En el análisis de un ECA deben considerarse todos los pacientes que inicialmente fueron incluidos en el estudio, documentándose el número y causas de los que abandonan, se retiran, se pierden o se desvían de los protocolos. Es conveniente la presentación de un diagrama de flujo donde se represente de forma gráfica los pacientes evaluados, incluidos/excluidos, asignados a cada grupo, y el seguimiento. La importancia de los pacientes

que finalmente no han sido seguidos viene determinada porque si estos pacientes hubieran presentando un resultado adverso y ésta hubiera sido la causa del no seguimiento, las conclusiones pudieran haber sido muy diferentes. No hay un número claro de pacientes no seguidos que hace que se afecte la validez interna del estudio, ya que depende no solamente de las pérdidas sino también de la frecuencia del resultado estudiado, pero parece claro que a mayor porcentaje, mayor compromiso en la validez. Cuando los no seguimientos son frecuentes, se aconseja realizar análisis de sensibilidad que consiste en reanalizar los resultados asumiendo que todas las pérdidas, en uno y otro brazo, se consideren con resultados positivos y después como resultado negativo. Si los resultados son muy dispares entre los diferentes cálculos, la validez del estudio puede ser cuestionada.

También es importante, a la hora de valorar la validez, el conocer si el estudio se interrumpió precozmente antes de incluir a todos los pacientes inicialmente previstos. Todo ECA debe realizarse mediante un cálculo de tamaño muestral previo que asegure la viabilidad del proyecto, y a veces con análisis de resultados antes de acabar el estudio para asegurar la seguridad del paciente o la no necesidad de incluir más pacientes por futilidad o lo contrario. Los ECA con finalización precoz pueden tener problemas de validez y deben tomarse sus conclusiones con más precauciones.

Finalmente, en los ECA se producen con cierta frecuencia desviaciones del protocolo, cuando un paciente que fue asignado aleatoriamente a recibir un tratamiento deja de tomarlo, lo sigue parcialmente o recibe en realidad, por distintos motivos, el tratamiento del otro grupo. Actualmente se acepta que el enfoque correcto es analizar al paciente en el grupo al que fue aleatoriamente asignado. Esto se conoce como análisis por intención de tratar, cuyas ventajas son: conservar las propiedades conseguidas con la asignación aleatoria: la homogeneidad de los grupos; y aproximarse a la realidad de la práctica clínica, donde los cambios o incumplimiento de los tratamientos son habituales.

Analizando el ejemplo de nuestro estudio, varios problemas metodológicos encontramos en relación con este punto del análisis de la validez interna que comprometen seriamente al trabajo.

4. ¿Se mantuvieron ciegos al tratamiento los pacientes, los clínicos y el personal del estudio?

Las expectativas, tanto de los pacientes como de los investigadores, pueden influir en la evaluación de la respuesta. Este problema puede prevenirse mediante el uso de las técnicas de enmascaramiento o cegamiento, que implican que los pacientes y/o los observadores desconozcan el tratamiento que recibe cada uno de los sujetos, de forma que no puedan interferir con los resultados. La mejor técnica es la del doble ciego, según la cual ni los pacientes ni los investigadores conocen el tratamiento asignado. Su utilización requiere, en muchos estudios, el uso de una sustancia placebo, similar en todo al tratamiento activo excepto en que carece de acción farmacológica directa, en el grupo control. Sin embargo, no siempre es posible la aplicación de estas técnicas. En estas situaciones, debe

evaluarse la posibilidad de que los pacientes o los observadores puedan haber influido sobre la respuesta, lo que es más probable si ésta incluye un componente de subjetividad. Como mínimo, es recomendable que las personas que midan la respuesta desconozcan el tratamiento asignado, de forma que se reduzca el posible sesgo.

En nuestro ejemplo no se asegura el enmascaramiento de los pacientes lo que condiciona un sesgo importante para la validez del mismo...

5. ¿Fueron similares los grupos al comienzo del ensayo?

Para que un estudio sea válido, los grupos deben ser similares en todo excepto en el tratamiento que reciben. Un artículo debe informar a los lectores de las características iniciales de los grupos. La asignación aleatoria no asegura que los grupos sean similares, especialmente cuando su tamaño es reducido. Por ello, debe evaluarse la comparabilidad de los grupos por los factores pronósticos. De todas formas, la ausencia de diferencias estadísticamente significativas no implica necesariamente que los grupos sean comparables, por lo que debe prestarse especial atención a la magnitud de las diferencias. Si los grupos presentan alguna diferencia importante por algún factor fuertemente asociado con el pronóstico, aunque no sea estadísticamente significativa, la validez del estudio puede verse comprometida.

Los estudios deben presentar una tabla, como en el artículo utilizado como ejemplo, donde se presenten las principales características basales de ambos grupos, sobre todo de aquellas variables que pueden relacionarse con el evento estudiado.

6. ¿Al margen de la intervención en estudio los grupos fueron tratados de igual forma?

Para poder atribuir el efecto observado a la intervención objeto de estudio, los grupos deben ser comparables, no sólo al inicio del ensayo, sino también durante todo el seguimiento. Si uno de los grupos ha sido seguido con mayor atención que el otro, existe una mayor probabilidad de detectar cualquier suceso que ocurra durante el seguimiento, e incluso de modificar la intervención o de administrar otros tratamientos, interfiriendo con la evaluación de la respuesta. Este problema es especialmente importante en los estudios abiertos (sin enmascaramiento) y cuando el protocolo del estudio permite a los investigadores utilizar otros tratamientos concomitantes durante el estudio.

Una proporción importante de artículos publicados, como el ejemplo expuesto, no hacen referencia a este punto, lo que compromete la validez.

B) ¿Cuáles son los resultados?

La decisión de si un tratamiento es o no más eficaz que el de referencia no debe basarse exclusivamente en el resultado de una prueba de significación estadística. Es fundamental que el artículo proporcione una estimación del efecto del tratamiento para que el lector pueda evaluar si es de la magnitud suficiente para poder ser considerado como

un efecto de relevancia clínica. Dos son las cuestiones que deben evaluarse para ello: en primer lugar, la magnitud del efecto observado; y, en segundo, la precisión con que ha sido estimado.

7. ¿Es muy grande el efecto del tratamiento o la intervención?

8. ¿Cuál es la precisión de este efecto?

El propósito de un tratamiento es mejorar un resultado de salud, y cuanto mayor sea esta mejoría el fármaco puede considerarse más eficaz. Una diferencia entre tratamientos de muy escasa magnitud, carente de cualquier importancia práctica, puede ser estadísticamente significativa si se ha estudiado un número elevado de sujetos. Por ello, si el lector sólo conoce el valor de p , sabrá si el estudio ha demostrado o no que existe un efecto, pero no si éste es importante. Existen varias formas de representar el efecto de un tratamiento o intervención. Los más utilizados son la reducción absoluta de riesgo en el resultado entre los dos grupos analizados, o el Número Necesario para Tratar (NNT) por ser considerado como una medida de interpretación más intuitiva y clínica. Sin embargo, a veces los resultados se presentan como reducción relativa de riesgo o como riesgo relativo, lo que puede dificultar la comprensión del verdadero impacto de la intervención. Por otro lado, no hay que olvidar que prácticamente todos los estudios son muestras, y el verdadero valor de la intervención es desconocida. Se debe determinar la precisión del resultado de la intervención mediante el cálculo del intervalo de confianza al 95%, es decir, aquel intervalo alrededor de la medida puntual obtenida en el que, con una probabilidad preestablecida del 95%, se incluirá el efecto real de la intervención.

En nuestro estudio, los resultados se presentan en una tabla con los números y porcentajes de los diferentes resultados en ambos grupos, lo que permite calcular la diferencia absoluta en el riesgo de complicaciones y en la adherencia al tratamiento. Sin embargo, no se calcula la precisión de los resultados.

C) ¿Pueden ayudarnos estos resultados?

El fin último de leer un artículo es comprobar si sus resultados son extrapolables a nuestros pacientes y poder aplicarles este tratamiento. Para valorar este punto debe contemplarse 3 aspectos:

9. ¿Pueden aplicarse estos resultados en tu medio o población local?

La contestación a esta pregunta se realiza mediante una cuestión *¿Crees que los pacientes incluidos en el ensayo son suficientemente parecidos a tus pacientes?* Obviamente, si mis pacientes tienen las mismas características que los pacientes incluidos en el estudio, los resultados serán perfectamente aplicables. Pero ¿qué ocurre con los pacientes que no las tienen? Supongamos que se desea aplicar los resultados a unos sujetos de mayor edad o que tienen problemas de salud concomitantes. En esta situación, debe utilizarse el

juicio clínico y preguntarse si existe alguna razón por la que los resultados no puedan ser extrapolados. Si no existe ninguna razón importante, tendremos una mayor confianza en la generalización de los resultados.

Por ello, los artículos deben indicar claramente los criterios de selección, la procedencia y las características de los sujetos estudiados. Otra circunstancia es la diferencia entre el número de pacientes evaluados y los que finalmente se incluyen en el estudio. A mayor diferencia, más se compromete la validez externa. Por eso los mejores ECA son los que presentan unos criterios de inclusión muy amplios y de exclusión mínimos.

En este punto también es importante conocer que el análisis de subgrupos es muy controvertido, ya que la realización de múltiples análisis aumenta la probabilidad de encontrar un resultado estadísticamente significativo, simplemente por azar. Cuando un estudio presenta análisis de subgrupos hay que ser muy cauteloso al interpretarlos, y evaluar detenidamente si el tamaño del subgrupo es grande, si responde a una hipótesis establecida previamente al inicio del estudio, si se han analizado pocos subgrupos y si se han observado los mismos resultados en otros estudios. Si no se cumplen estas circunstancias, habrá que ser escéptico sobre la posibilidad de aplicar los resultados.

10. ¿Se tuvieron en cuenta todos los resultados de importancia clínica?

Lo ideal es que la respuesta sea Sí, ¿pero qué pasa cuando la respuesta es NO? Si un estudio contempla el análisis de unas variables de eficacia pero no de seguridad podríamos tomar una decisión en base a un resultado positivo, pero dejaríamos fuera de lugar la posibilidad de resultados negativos. Por tanto los mejores trabajos son aquellos que contemplan no solo los posibles beneficios de una intervención sino, además, los perjudiciales.

En nuestro ejemplo, el artículo se centra en resultados beneficiosos, como es la adherencia al tratamiento y, al mismo tiempo, en la tolerabilidad, efectos secundarios de administrar dos fármacos, por lo que este punto está plenamente contemplado.

11. ¿Los beneficios a obtener justifican los riesgos y los costes?

Aunque cada vez con más frecuencia, aún son pocos los artículos clínicos donde se contempla un análisis de los costes de una intervención. Los riesgos, efectos secundarios, deben contemplarse como resultados de seguridad y cada vez se presentan con más frecuencia. Pero no así el coste, no solo de la intervención sino otros costes: derivados de nuevas medidas de tratamiento, hospitalizaciones, bajas laborales, años de vida, etc.

Pese a ello, el lector debe pensar y deducir, con sus conocimientos en el tema, si el coste final de la intervención, junto a los riesgos que produce, puede provocar que ésta no deba realizarse pese a demostrarse un beneficio en los resultados obtenidos.

4. Resolución del escenario clínico

Tras la lectura crítica del artículo seleccionado, podemos concluir que la utilización de dos fármacos durante un periodo de tiempo corto conlleva mayor adherencia al tratamiento comparándose con una pauta larga de un solo fármaco, a costa de aumentar las complicaciones derivadas del tratamiento. Sin embargo, los numerosos sesgos metodológicos del estudio comprometen a la validez del mismo y sus conclusiones deberían comprobarse mediante posteriores estudios con mayor rigor metodológico.

5. Puntos clave

1. El ECA es el diseño que mejor evidencia aporta para la evaluación de la eficacia de una intervención.
2. La lectura crítica de un ECA debe analizar aspectos relacionados con la validez del estudio, su relevancia y su utilidad.
3. La validez interna del ECA viene determinada fundamentalmente por la presencia de unos claros objetivos (definidos por una adecuada descripción de la población estudiada, la intervención y su comparación, y los resultados analizados), así como una adecuada aleatorización y seguimiento de los pacientes, junto a un análisis de resultados mediante el principio de intención de tratar.

Bibliografía

1. Glasziou P, Del Mar C, Salisbury J. *Evidence-based Medicine Workbook. Finding and applying the best evidence to improve patient care*. London. BMJ Publishing Group 2003.
2. Greenhalgh T. *How to read a paper. The basic of evidence-based medicine*. Fourth edition. London. Wiley-Blackwell 2010, pp. 78-89.
3. Tello Royloa C. *Lectura crítica de un ensayo clínico. En: Atención Sanitaria Basada en la Evidencia. Su aplicación a la práctica clínica*. Murcia. Consejería de Sanidad de la Región de Murcia. 2007; pp.: 275-296.
4. Ibáñez Pradas, Modesto Alapont V. MBE en cirugía pediátrica. Lectura crítica de artículos. Tratamiento (I). *Cir Pediatr* 2005; 18:159-164.
5. Ibáñez Pradas, Modesto Alapont V. MBE en cirugía pediátrica. Lectura crítica de artículos. Tratamiento (I). *Cir Pediatr* 2006; 19:2-6.

22

Lectura crítica de un estudio observacional

ALFONSO MIGUEL MUÑOZ UREÑA¹

JULIO FONTCUBERTA MARTÍNEZ²

1) Médico Especialista en Medicina Familiar y Comunitaria.
Técnico de Salud Unidad Docente Multiprofesional de
Atención Familiar y Comunitaria Cartagena-Mar Menor.
Servicio Murciano de Salud.

2) Médico Especialista en Medicina Familiar y Comunitaria.
Jefe de Estudios Unidad Docente Multiprofesional
de Atención Familiar y Comunitaria
Cartagena-Mar Menor. Servicio Murciano de Salud.

1. Introducción

Escenario Clínico

Un médico de familia realiza revisiones programadas a sus pacientes diabéticos. Dentro de estas revisiones, periódicamente realiza fondos de ojo en un intento de detección precoz de retinopatía. Además, en el centro de salud donde trabaja se ha recibido la oferta de implantar un programa de retinografías desde atención primaria. Sin embargo, tuvo hace años una mala experiencia con la dilatación de una pupila en un paciente, que sufrió un ataque agudo de glaucoma. Por ello, realiza una búsqueda para poder detectar desde atención primaria ojos con cámara anterior poco profunda, susceptibles de padecer un ataque agudo de glaucoma tras dilatación de la pupila. Debido a ello hace una búsqueda bibliográfica y encuentra el siguiente artículo: *Trueba Castillo A, Negredo Bravo LJ, Cárdenas Valencia C, Gil de Gómez Barragán MJ, Arribas García RA. ¿Es útil la prueba de la linterna en atención primaria para detectar ojos con cámara anterior poco profunda? Aten Primaria. 2010; 42(03):149-53.*

Tras la lectura cuidadosa y crítica del artículo, el médico toma una decisión sobre el uso de esta prueba previo a la dilatación pupilar.

A continuación desarrollamos este escenario, desde la perspectiva de la Práctica Clínica Basada en Evidencias (PCBE), con el objetivo de iniciarse en la lectura crítica de artículos de validez de pruebas diagnósticas.

2. Elaboración de la pregunta clínica y estrategia de búsqueda de información¹

Siguiendo las etapas recomendadas en lectura crítica basada en la evidencia, debemos iniciar esta ante una pregunta que surge, la mayoría de las ocasiones, durante nuestro ejercicio profesional. Como se ha comentado anteriormente, esta pregunta debe ser estructurada para optimizar el tiempo de búsqueda, dada la información disponible en la actualidad. Desde un punto de vista práctico, las preguntas pueden ser de diversa índole. En nuestro caso surgiría ante la necesidad de elegir una prueba diagnóstica adecuada en función de distintos factores.

Las preguntas pueden ser generales (¿cuál es el tratamiento de la HTA en población diabética?) o específicas (¿en los pacientes diabéticos qué pronóstico tiene el tratamiento con IECA frente a betabloqueantes?). Ante una duda debemos esforzarnos en hacer una pregunta específica con cuatro componentes, formato PICO (*Patient, Intervention, Comparison and Outcome*). En el caso de evaluación de pruebas diagnósticas el componente In-

tervención se sustituye por la prueba diagnóstica que se quiere evaluar y el componente comparación se sustituye por la prueba patrón oro de referencia.

En nuestro caso la pregunta podría ser “¿en los pacientes con ojos que poseen cámara anterior poco profunda es útil la prueba de la linterna en atención primaria para evitar ataques agudos de glaucoma tras dilatación pupilar para exploración de fondo de ojo?”

Formato PICO (*Patient, Intervention, Comparison and Outcome*):

	P Paciente o problema	I Intervención (Prueba)	C Prueba a comparar	O Resultado
Orientaciones	¿Cómo podría describir un grupo de pacientes similar al mío?	¿Cuál es la prueba diagnóstica?	¿Cuál es la prueba patrón oro con la que comparar?	¿Cuál es el resultado que puedo esperar de la intervención?
Ejemplo	Adultos que se van a someter a exploración de fondo de ojo en atención primaria con dilatación pupilar	Prueba de la linterna: iluminación completa del iris	Test de Van Herick	Evitar ataques agudos de glaucoma secundarios a dilatación pupilar farmacológica

Las estrategias de búsqueda, utilizando el lenguaje científico y minimizando en la medida de lo posible la falta de artículos publicados referentes a nuestra búsqueda, han sido descritos en capítulos anteriores de este libro.

3. Lectura crítica de un Estudio de Evaluación de Pruebas Diagnósticas¹⁻³

Ante la posibilidad de utilización de una nueva prueba diagnóstica en la práctica clínica, se debe evaluar su validez. Los estudios de validez de pruebas diagnósticas pueden diseñarse con dos objetivos. El objetivo más utilizado es el de evaluar la validez diagnóstica de la prueba en cuestión. Para alcanzar este objetivo es necesaria la comparación de los resultados obtenidos en dos procesos independientes de medición. El primero, una prueba de referencia (patrón oro o *gold-standard*) que se asume clasifica de forma válida a los sujetos respecto a la presencia o ausencia de una condición clínica o enfermedad. El segundo proceso es la prueba diagnóstica que se somete a evaluación. En este trabajo evaluaremos un estudio de evaluación de pruebas diagnósticas siguiendo este primer objetivo. En cuanto al segundo tipo de diseños en la evaluación de pruebas diagnósticas, estudian el impacto de la prueba en la práctica clínica, toma de decisiones, coste-beneficio, son menos frecuentes y no se tratan en este capítulo.

Existen varias dimensiones sobre las que se puede evaluar la calidad de un estudio. Primero hay que valorar la existencia o no de sesgos en la estimación de la prueba diagnóstica, por déficit en el diseño, en el desarrollo y/o en la ejecución. En segundo lugar, se debe evaluar la capacidad de generalización de la prueba, es decir, la posibilidad de reproducir la misma en otro medio que no sea en el que se desarrolló el estudio, dependiendo entre otros factores del espectro de enfermedad de los sujetos, del ámbito de estudio...

Existen numerosas herramientas de evaluación de estudios de pruebas diagnósticas, guías con escalas publicadas en literatura científica como *JAMA (Journal of the American Medical Association)*, *BMJ (The British Medical Journal)*. Siguiendo una línea de continuidad en este trabajo, la evaluación de pruebas diagnósticas se desarrolla según las recomendaciones de la iniciativa CASP (*Critical Appraisal Skills Programme*. Programa de Habilidades en Lectura Crítica), adaptadas a nuestro idioma por el grupo CASP España (CASPe).

La evaluación de un estudio de pruebas diagnósticas viene determinada por los siguientes pasos:

A) ¿Son válidos los resultados del estudio?

1. ¿Existió una comparación con una prueba de referencia adecuada?
2. ¿Incluyó la muestra un espectro adecuado de pacientes?
3. ¿Existe una adecuada descripción de la prueba?
4. ¿Hubo evaluación "ciega" de los resultados?
5. ¿La decisión de realizar el patrón de oro fue independiente del resultado de la prueba problema?

B) ¿Cuáles son los resultados?

6. ¿Se pueden calcular los Cocientes de Probabilidad?
7. ¿Cuál es la precisión de los resultados?

C) ¿Pueden ayudarnos estos resultados?

8. ¿Serán satisfactorias en el ámbito del escenario la reproducibilidad de la prueba y su interpretación?
9. ¿Es aceptable la prueba en este caso?
10. ¿Modificarán los resultados de la prueba la decisión sobre cómo actuar?

Plantilla desarrollada por CASP*Oxford y adaptado por CASP España (CASPe). <http://www.redcaspe.org/que-hacemos/herramientas/> (Consultada el 25 de septiembre 2011).

A) ¿Son válidos los resultados del estudio?

La primera cuestión es si los resultados del estudio pueden considerarse válidos, es decir, si el resultado obtenido en la prueba diagnóstica a estudio es realmente lo que dice ser. Se trata de comprobar si el diseño del estudio permite asegurar que el resultado positivo o negativo de la prueba se corresponde con la realidad.

Las preguntas a hacerse serían:

1. ¿Existió una comparación con una prueba de referencia adecuada?

En todo estudio de validez de pruebas diagnósticas a los sujetos seleccionados se les debe realizar la prueba que se quiere evaluar y el patrón oro diagnóstico. Por tanto, el patrón oro o de referencia debe clasificar correctamente en enfermedad o ausencia de la misma, ya que los errores diagnósticos se achacarán a la prueba evaluada. Además, debe evitarse la separación en el tiempo de la realización de ambas pruebas, ya que la enfermedad puede evolucionar.

En nuestro ejemplo de artículo, se describe una prueba patrón oro, el test de Van Herick, realizada con lámpara de hendidura por un oftalmólogo; aunque existe otra prueba, como es la gonioscopia, que se descarta por su complejidad. La prueba de Van Herick tiene elevado Valor Predictivo Negativo, por lo que los autores justifican su elección. Las dos pruebas se realizaron a todos los pacientes incluidos en el estudio y se excluyeron a los que no realizaron la prueba de patrón oro.

2. ¿Incluyó la muestra un espectro adecuado de pacientes?

Los pacientes deben ser descritos adecuadamente así como la forma de seleccionarlos. Los criterios de inclusión y exclusión deben ser detallados, lo que permitirá la reproducibilidad en nuestro medio. Otro dato a tener en cuenta es la variación en las características clínicas y demográficas de los sujetos. El estudio debe incluir un espectro de sujetos lo más parecido al que se pretende aplicar la prueba, evitando provocar sesgos como el que se produciría si intentamos comparar la prueba en pacientes que se sabe que tienen la enfermedad frente a otros que se sabe que no la tienen.

En nuestro estudio queda definida la población incluida en el mismo como la población diabética, con unos criterios de inclusión y de exclusión establecidos. También se calcula un tamaño muestral, que posteriormente se comprueba no se alcanza. Además, no existe una aleatorización en el muestreo, sino que se selecciona un muestreo no probabilístico como es el muestreo consecutivo. Sin embargo, en los pacientes seleccionados es desconocida la presencia o ausencia de enfermedad.

3. ¿Existe una adecuada descripción de la prueba?

Una adecuada descripción de la prueba es necesaria para evaluar la validez de la misma, así como su estimación. Debe describirse cómo se ejecuta y qué se considera un resultado positivo así como un resultado negativo. Del mismo modo, hay que explicar quiénes son los que ejecutan la prueba y cómo se controla la variabilidad entre observadores. De esta forma también se obtiene información para la reproducibilidad de la prueba, es decir, la validez externa.

Durante la realización de la prueba quedan descritos los grados de la misma que se consideran positivos, así como en la prueba de referencia. Del mismo modo se describe

quienes realizan la prueba estudiada y la prueba de referencia, excluyendo a todo sujeto que no se realice la prueba de referencia con el mismo oftalmólogo. Los autores evalúan la diferencia entre observadores mediante el índice Kappa (valor del 75%), existiendo esta diferencia dado que uno de los observadores tiene menos experiencia clínica.

4. ¿Hubo evaluación “ciega” de los resultados?

Las personas que interpretan la prueba diagnóstica que se evalúa deben desconocer los resultados de la prueba de referencia y, de igual forma, los profesionales que interpretan el patrón oro deben desconocer los resultados de la prueba a estudio. De no realizarse así, se obtendría una sobreestimación del rendimiento diagnóstico. En ocasiones esta pregunta no es aplicable, dependiendo del orden de ejecución de las pruebas y de la objetividad de la misma. Cuanto más subjetiva sea una prueba más debe tenerse en cuenta el enmascaramiento de los resultados.

En el trabajo analizado los pacientes seleccionados eran evaluados mediante la Prueba de la Linterna (prueba evaluada) y posteriormente eran remitidos a un oftalmólogo que realizaba la prueba de referencia sin conocer los resultados de la mencionada prueba de la linterna.

5. ¿La decisión de realizar el patrón de oro fue independiente del resultado de la prueba problema?

Esta pregunta, junto con la anterior, están relacionadas con el enmascaramiento. La selección de los pacientes para la realización de ambas pruebas debe estar representada en una tabla, así como explicar la forma de reclutamiento.

En el estudio se detalla que todas las personas que se realizaron la prueba de la linterna fueron remitidas a realización de la prueba de referencia, considerando pérdidas todos los que no se realizaron la misma, sin considerar el resultado previo de la prueba evaluada.

B) ¿Cuáles son los resultados?

La decisión de si una prueba diagnóstica es válida o no debe considerar una serie de parámetros a incluir en los resultados. El diseño debe proporcionar una estimación de los resultados de la prueba evaluada en comparación con el *gold standard* para que podamos evaluar si son de la magnitud y precisión suficientes.

6. Se pueden calcular los Cocientes de Probabilidad

El objetivo de una prueba diagnóstica es detectar los que están realmente enfermos y, de igual forma, clasificar como sanos a los que realmente no padecen la enfermedad. Para poder determinar si una prueba tiene esta capacidad se utilizan una serie de parámetros como son la sensibilidad y la especificidad. Estos índices son inherentes a la prueba y no

están influenciados por la prevalencia de la enfermedad estudiada. Se pueden definir de la forma siguiente, construyendo para su mayor comprensión una tabla de contingencia:

		Enfermedad		
		Enfermo E+	Sano E-	
Prueba Diagnóstica	Positivo T+	a	b	Positivos = a+b
	Negativo T-	c	d	Negativos= c+d
		Total enfermos= a+c	Total sanos= b+d	Total N

Sensibilidad: es la proporción de verdaderos positivos identificados por la prueba, del total de enfermos.

$$\text{Sensibilidad} = S = \frac{a}{a + c}$$

Especificidad: es la proporción de verdaderos negativos identificados por la prueba, del total de sanos.

$$\text{Especificidad} = E = \frac{d}{b + d}$$

Sin embargo, en la práctica clínica nos interesa responder a la pregunta: ¿Si un paciente ha resultado positivo en la prueba, cuál es la probabilidad de que esté realmente enfermo? Para responder a esta pregunta se utilizan los índices de valor predictivo positivo y valor predictivo negativo, que sí están influenciados por la prevalencia de la enfermedad y por el número de enfermos que pueda existir en la muestra, por tanto, no reflejan exclusivamente la calidad de la prueba. Para evitar esto se utilizan los **cocientes de probabilidad**, el cociente de probabilidad positivo y el cociente de probabilidad negativo. Se expresan con estas fórmulas en función de la sensibilidad y de la especificidad:

$$CP+ = \frac{S}{1 - E} \quad CP- = \frac{1 - S}{E}$$

Valores mayores de CP+ indican mejor capacidad para diagnosticar la presencia de enfermedad. Del mismo modo, valores de CP- menores indican una mejor capacidad diagnóstica de la prueba.

En el artículo analizado se pueden hallar todos estos índices, como se refleja en la tabla 1:

Tabla 1. Características de la prueba de la linterna en comparación con la prueba de Van Herick.

Prueba de la linterna	Prueba de Van Herick	
	Positivo	Negativo
Positivo	10	11
Negativo	1	88
Sensibilidad (IC del 95%)	90,9 (69,4 al 100,0%)	
Especificidad (IC del 95%)	88,9 (82,2 al 95,6%)	
Valor predictivo negativo (IC del 95%)	98,9 (96,1 al 100,0%)	
Valor predictivo positivo (IC del 95%)	47,6 (23,9 al 71,4%)	
Cociente de probabilidad positivo (IC del 95%)	8,2 (4,6 al 14,7%)	
Cociente de probabilidad negativo (IC del 95%)	0,1 (0,02 al 0,7%)	

IC: Intervalo de confianza.

Se puede observar una sensibilidad y una especificidad elevadas, así como el valor predictivo negativo. Sin embargo, el valor predictivo positivo y el cociente de probabilidad positivo no son demasiado elevados. Esto nos indica que la prueba puede ser útil para descartar la enfermedad en los que tienen un resultado negativo, pero no es una prueba útil de *screening* para detectar la enfermedad y realizar posibles tratamientos.

7. ¿Cuál es la precisión de los resultados?

En la mayoría de los trabajos de investigación se trabaja con muestras, por lo que posteriormente a la obtención de los resultados hay que hacer una estimación a la población total. Esto genera la posibilidad de un error aleatorio, por lo que es necesario que los resultados queden reflejados con un Intervalo de Confianza, que nos permite establecer el rango de valores en que puede estar el valor real del índice en la población.

En la tabla anterior también quedan reflejados los intervalos de confianza, siendo intervalos con un rango amplio en la mayoría de los índices, excepto en el coeficiente de probabilidad negativo y en el valor predictivo negativo. Esto puede ser debido al pequeño tamaño muestral alcanzado en el trabajo.

C) ¿Pueden ayudarnos estos resultados?

El fin último de leer un artículo es comprobar si sus resultados son extrapolables a nuestros pacientes y poder aplicarles esta prueba diagnóstica.

8. ¿Serán satisfactorios en el ámbito del escenario la reproducibilidad de la prueba y su interpretación?

En este apartado hay que considerar si el ámbito de realización de la prueba en el artículo analizado difiere mucho del escenario en el que se aplicaría la misma. Para ello hay

que considerar la población seleccionada, los profesionales participantes en el estudio y los medios utilizados, entre otros factores.

El artículo selecciona pacientes diabéticos, por lo que si deseáramos utilizar en pacientes hipertensos la prueba de la linterna, requeriría una nueva evaluación. Sin embargo, el ámbito es en atención primaria y con pacientes diabéticos se podría reproducir el estudio así como la interpretación en nuestro medio.

9. ¿Es aceptable la prueba en este caso?

Además de evaluar la validez de la prueba, hay que considerar la disponibilidad de la misma, la viabilidad, la efectividad, las molestias y los costes.

En nuestro ejemplo, la prueba de la linterna es una prueba sencilla, asequible para atención primaria, que se basa en el hecho de que al dirigir una fuente de luz hacia el ojo desde su canto externo, la superficie del iris se iluminaría completamente si se encuentra en el mismo plano que la pupila. Por tanto, no produce ningún riesgo añadido para el paciente. En cuanto a la viabilidad de la realización de la prueba depende de la presión asistencial y del tiempo en consulta para la realización de la misma, además del entrenamiento previo de los profesionales para la realización correcta de la misma.

10. ¿Modificarán los resultados de la prueba la decisión sobre cómo actuar?

Desde el punto de vista del escenario en que se aplicaría la prueba diagnóstica hay que considerar si la actitud clínica cambiaría ante un resultado positivo o negativo. Si no es así, la prueba es poco útil. En esta cuestión hay que considerar la probabilidad de enfermedad antes o después de la prueba.

En nuestro ejemplo, la probabilidad de padecer un ataque agudo de glaucoma tras una dilatación pupilar en la población general es bajo, pero aumenta el riesgo con la edad. Por otro lado, es cierto que un ataque agudo de glaucoma puede desencadenar un cuadro clínico con repercusiones y secuelas posteriores y una detección de los pacientes con mayor riesgo minimizaría la posibilidad de iatrogenia.

Resolución del escenario clínico

Tras la lectura crítica del artículo seleccionado, podemos concluir que la utilización de la prueba diagnóstica de la linterna podría ser útil en la detección de pacientes con ángulo amplio que no tienen riesgo de padecer un ataque agudo de glaucoma. Sin embargo, el pequeño tamaño muestral, y la población final estudiada tras las pérdidas, así como los índices encontrados en el estudio, no permiten utilizar la prueba para detectar pacientes con ángulo estrecho, además de comprometer la validez del estudio, por los que sus conclusiones deberían comprobarse mediante posteriores estudios.

4. Puntos clave

1. El estudio de pruebas diagnósticas es un tipo de estudio descriptivo que trata de conocer la capacidad diagnóstica de una prueba, es decir, la capacidad para clasificar a una persona según la presencia o ausencia de enfermedad.
2. El diseño se realiza mediante la comparación entre la prueba a estudio y el patrón de referencia (*gold standard*).
3. Los resultados se evalúan en términos de sensibilidad, especificidad, valores predictivos o coeficientes de probabilidad. Para este cálculo se tiene en cuenta los resultados verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos.
4. El estudio debe incluir el espectro de pacientes más parecido posible al que se pretende aplicar la prueba
5. Enmascaramiento en la interpretación de las pruebas. La prueba a estudio debe valorarse sin conocer los resultados de la prueba de referencia y viceversa.

Bibliografía

1. Trueba Castillo A, Negro Bravo LJ, Cárdenas Valencia C, Gil de Gómez Barragán MJ, Arribas García RA. ¿Es útil la prueba de la linterna en atención primaria para detectar ojos con cámara anterior poco profunda? *Aten Primaria*. 2010; 42(03):149-53.
2. Abad Corpa E, Monistrol Ruano O, Altarribas Bolsa E, Paredes Sidrach de Cardona A. Lectura crítica de literatura científica. *Enfermería Clínica* 2003; 13 (1): 32-40.
3. Zamora J, Abreira V. Análisis de la calidad de los estudios de evaluación de pruebas diagnósticas. *NEFROLOGÍA*. 2008. Vol. 28, Supl. 2: 42-45.



Región de Murcia
Consejería de Sanidad y Política Social

Dirección General de Planificación, Ordenación
Sanitaria y Farmacéutica e Investigación



**FORMACIÓN
SANITARIA
ESPECIALIZADA**