

MEDICINA BASADA EN EVIDENCIAS EN HEPATOLOGÍA I

*Evaluación de estudios de Tratamiento:  
Utilidad de Pentoxifilina en el tratamiento  
de la hepatitis por alcohol grave*

ENRIQUE VEAS R.<sup>(1)</sup>, FRANCISCO BARRERA M.<sup>(1)</sup>, IGNACIO SALAS S.<sup>(1)</sup>,  
MANUEL ESPINOZA S.<sup>(1)</sup>, JUAN IGNACIO QUINTERO B.<sup>(1)</sup>, ALEJANDRO SOZA R.<sup>(2)</sup>,  
LUZ MARÍA LETELIER S.<sup>(3)</sup> y ARNOLDO RIQUELME P.<sup>(1,2,3)</sup>

EVIDENCE BASED MEDICINE IN HEPATOLOGY I: CRITICAL ANALYSIS OF TREATMENT STUDIES. PENTOXIFYLLINE IMPROVES SHORT-TERM SURVIVAL IN SEVERE ACUTE ALCOHOLIC HEPATITIS

*The aim is to present basic concepts on clinical decision making about therapeutics in Hepatology, using Evidence-Based Medicine (EBM) criteria. A clinical scenario of severe alcoholic hepatitis is presented, from which tools for finding the best available evidence are reviewed; a critical analysis of a treatment trial is performed, followed by a discussion of the applicability of trial findings to the clinical situation. Using a model to search for evidence based on a clinical question, a randomized controlled trial on Pentoxifylline for the treatment of severe alcoholic hepatitis was identified. Internal validity criteria for a treatment trial are discussed; these include randomization, follow-up and analysis by intention-to-treat. Secondary criteria include blinding, baseline characteristics and similar therapy for both groups. The reviewed trial satisfied these criteria, which backs its methodological quality. The trial results were analyzed, and treatment effect indicators were calculated for mortality as main outcome: relative risk reduction (RRR) 46,9%, absolute risk reduction (ARR) 21,6%, number needed to treat (NNT) 5 (95% CI 3-30) and number needed to harm (NNH) 6 (95% CI 4-46). The clinical applicability analysis, using criteria which include relevance of the findings, adverse effects and statistical accuracy, was favorable to the use of Pentoxifylline in the presented clinical situation. We conclude that the use of EBM criteria allows to find good quality evidence, providing a valuable tool in clinical decision making.*

**Key words:** Evidence-based medicine, pentoxifylline, alcoholic hepatitis, cirrhosis.

---

<sup>(1)</sup> Unidad Docente Asociada de Medicina Hospital de Urgencia Asistencia Pública.

<sup>(2)</sup> Departamento de Gastroenterología.

<sup>(3)</sup> Unidad de Medicina Basada en Evidencia, Escuela de Medicina, Pontificia Universidad Católica de Chile.

## INTRODUCCIÓN

La definición más popular de la Medicina Basada en Evidencias es “la aplicación juiciosa de la mejor evidencia disponible a una decisión clínica particular”<sup>1</sup>. Esto implica que practicar la Medicina Basada en Evidencias no consiste en tratar a los pacientes sólo con las terapias que han demostrado reducir la mortalidad en grandes meta-análisis y estudios multicéntricos, pues es posible que en una determinada área de la medicina no exista evidencia de gran calidad. Ante la necesidad de ofrecer una terapia efectiva y segura a nuestros pacientes, debemos ser capaces no sólo de encontrar la mejor evidencia disponible, sino también de aplicarla en forma criteriosa a nuestra situación clínica.

## ESCENARIO CLÍNICO

Ingresa al Servicio de Medicina del Hospital de Urgencia Asistencia Pública de Santiago de Chile, un hombre de 48 años que consultó por compromiso del estado general y fiebre. Tiene antecedentes de alcoholismo crónico y relata una historia de 7 días de progresivo compromiso del estado general con marcada anorexia, vómitos, fiebre e ictericia progresiva. Niega estar consumiendo alcohol actualmente. Sin embargo, al interrogarlo dirigidamente sobre su historia de consumo, se obtiene que dejó de beber hace sólo 10 días, previo a lo cual estuvo 3 semanas consumiendo medio litro de pisco al día (175 gramos de alcohol al día).

En el examen físico de ingreso destacan: ictericia moderada de piel y mucosas; frecuencia cardíaca 94 por minuto, presión arterial 110/60 mmHg, temperatura axilar 37,3° C, saturación de oxígeno 93%; hepatomegalia sensible 2 cm bajo el reborde costal, ascitis leve a moderada y asterixis.

Los exámenes de laboratorio muestran 12.000/mm<sup>3</sup> leucocitos, bilirrubina total 8,2 mg/dL (conjugada 5,4 mg/dL), fosfatasa alcalina 150 UI/L, SGOT 688 UI/L, SGPT 170 UI/L, tiempo de protrombina (TP) 18,4 segundos (control normal 9,7 segundos), actividad de

protrombina 35%; creatinina 0,7 mg/dL. Se realiza una ecotomografía abdominal que muestra ascitis de moderada cuantía y ausencia de lesiones focales hepáticas.

El diagnóstico clínico es hepatitis alcohólica aguda y se calcula el score de Maddrey (score discriminativo de función hepática)<sup>2</sup>:

Score = 4,6 x [TP-TP control(segundos)] + Bilirrubina total (mg/dL)

Score = 4,6 x (18,4 - 9,7) + 8,2 = **48,2**

Con dicha clínica y laboratorio, se clasifica el cuadro como una hepatitis grave por alcohol. En la visita de ingreso se discute la terapia, y se plantea iniciar corticoides dado su potencial de reducir la mortalidad en pacientes con score de Maddrey mayor de 32<sup>2,3</sup>. Sin embargo, también se discutió la posibilidad de tratarlo con Pentoxifilina. El equipo decide buscar la evidencia respecto de este nuevo tratamiento, ya que los resultados con corticoides son controversiales<sup>4-8</sup>.

## BÚSQUEDA DE EVIDENCIA: PRINCIPIOS

El primer paso para tomar una decisión basada en evidencia, aunque parezca obvio, es encontrar la evidencia. Actualmente, la investigación biomédica produce una gran cantidad de información científica, disponible en grandes bases de datos electrónicas; la búsqueda “a tientas” en dichas bases sólo nos llevará a perdernos en un mar de información. Presentaremos un modelo con el cual podemos orientar nuestra búsqueda de la evidencia y herramientas con las cuales podemos realizarla.

El primer paso es explicitar qué es lo que queremos encontrar y el modelo más utilizado es sintetizar la situación clínica en una pregunta clínica.

## PREGUNTA CLÍNICA

Debemos formalizar nuestra inquietud en una pregunta que será la que oriente nuestra búsqueda. Mientras más acotada sea la pregunta, se obtendrán resultados más escasos, pero más cercanos a nuestra situación. En temas donde

existe abundante investigación es vital. Por ejemplo, si usted desea tomar una decisión respecto del tratamiento de un paciente con un infarto agudo de miocardio, y busca en Medline “acute myocardial infarction treatment”, obtendrá más de 26.000 artículos; al cabo de algunos minutos es muy posible que decida suspender la búsqueda.

En temas de menor “popularidad”, como es probable que suceda en el área de la Hepatología, una búsqueda amplia proporcionará una cantidad menor de resultados, pero igualmente difícil de manejar. Abocándonos a nuestra situación, si usted coloca en Medline “alcoholic hepatitis treatment” obtendrá algo más de 600 resultados. En ambos casos es necesario acotar nuestra pregunta clínica, pero en el tema “menos popular” corremos el riesgo de excluir demasiados artículos, o incluso de quedarnos sin evidencia, pues tal vez la única disponible en esa área, no coincide exactamente con los términos utilizados para la búsqueda.

Sugerimos el modelo P.I.C.O. desarrollado entre otros por Guyatt, para formular la pregunta clínica<sup>9</sup>:

- **P**atient (Paciente)
- **I**ntervention (Intervención)
- **C**omparison (Comparación)
- **O**utcome (Evento observado)

Esto permite formular una pregunta susceptible de ser contestada que contiene los elementos de el quién, el qué y el para qué.

### **P: Paciente**

Las características del paciente deben definirse lo mejor posible para obtener evidencia aplicable a nuestro caso real. De acuerdo a lo antes expuesto, la precisión con la que ha de definirse al paciente depende de la cantidad de evidencia que esperamos encontrar. Por ejemplo, si se trata de un adulto, pero presenta una enfermedad más propia de la infancia, formular la pregunta estrictamente en términos de pacientes adultos podría tener escasos resultados. El principal punto a definir es el problema del paciente; otras variables que pueden precisarse son edad, sexo, raza, etiología, gravedad, etapa de enfermedad, etc.

En este ejemplo se trata de un paciente con hepatitis alcohólica grave.

### **I: Intervención**

Es necesario precisar la intervención o tratamiento cuyo efecto se busca conocer. Existen múltiples tipos de intervenciones: fármacos, vacunas, cirugía, intervenciones ambientales, procedimientos invasivos, programas de manejo, entre otras. En este ejemplo la intervención es el tratamiento farmacológico con Pentoxifilina.

### **C: Comparación**

Es la intervención contra cuyo efecto se quiere comparar el resultado de la intervención a estudiar. Si no se desea explícitamente comparar una contra otra, no es imprescindible incluir este punto al momento de hacer la búsqueda electrónica. Pero, en términos estrictos, debería formularse la pregunta usando como comparación el placebo o el tratamiento estándar de la condición en estudio. En este ejemplo la comparación podría ser placebo o corticoides.

### **O: Evento Observado**

Es un término sin traducción directa al español que se refiere al evento cuya ocurrencia se espera comparar entre el grupo control y el experimental. Pueden definirse outcomes favorables o adversos. En general, al plantear la pregunta interesa buscar outcomes relevantes: mortalidad (o sobrevida), morbilidad, recidiva, calidad de vida, etc. En nuestro ejemplo el outcome es idealmente mortalidad pues ya ha sido estudiado con los corticoides y ocurre con alta frecuencia en este cuadro (entre 35 y 46%).

Finalmente, deben sintetizarse estos elementos en la pregunta final. En este ejemplo una formulación podría ser: “en pacientes con hepatitis alcohólica grave, ¿el tratamiento con pentoxifilina reduce la mortalidad comparada con corticoides o placebo?”

## **BÚSQUEDA EN INTERNET**

Existen muchos motores de búsqueda de información, y el que utilizaremos es Pubmed

(<http://www.pubmed.com>), del mismo organismo estatal estadounidense que mantiene la principal base de datos biomédica MEDLINE.

Pubmed posee una ventana de búsqueda en la cual se puede ingresar cualquier tipo de palabra, sin embargo, esto generará una búsqueda textual en toda la gran base de datos, lo cual entregará muchos resultados probablemente erráticos. Es preferible usar los términos estandarizados MeSH (Medical Subject Heading).

**MeSH:** son términos bajo los cuales se encuentran indexados los artículos. Con esta herramienta los artículos están relacionados con uno o más conceptos, y no necesariamente con una palabra. Usando un MeSH como término de búsqueda, se encontrarán por ejemplo todos los artículos sobre una patología, aunque ésta aparezca con distintos nombres en cada uno. En resumen, representa un “lenguaje” del Pubmed al cual debemos traducir las palabras comunes que componen nuestra pregunta clínica. Los MeSH pueden encontrarse con la herramienta “MeSH Database” de Pubmed. Varios MeSH pueden combinarse con el operador AND (Y) en la ventana de búsqueda de Pubmed.

Los MeSH más adecuados para nuestra búsqueda son “Hepatitis, Alcoholic” [MeSH], “Pentoxifylline” [MeSH], “Glucocorticoids” [MeSH], “Placebos” [MeSH] y “Mortality” [MeSH]. Combinándolos con el operador AND no se obtienen resultados. Acortando a “Hepatitis, Alcoholic” [MeSH] AND “Pentoxifylline” [MeSH] se obtienen 8 resultados. Esto ejemplifica lo antes expuesto respecto de las ventajas y desventajas de efectuar una búsqueda demasiado acotada.

Una vez que hemos encontrado evidencia pertinente, ¿cómo podemos determinar cuáles artículos representan la mejor evidencia disponible? Los estudios se efectúan con distintos diseños metodológicos; diferentes tipos de estudio implican evidencia de distinta calidad. Existen varias clasificaciones para la calidad de la evidencia, la de Oxford es una de las más utilizadas (Ver Apéndice 1). Para seleccionar sólo trabajos de buen nivel de evidencia, podemos usar la herramienta Límites de Pubmed.

**Límites:** la función “Limits” permite restringir los resultados de una búsqueda al filtrar según uno o más criterios, entre los cuales están: idioma, fecha de publicación, sexo y grupo etario estudiado, área de investigación, etc. El segundo límite, “Publication Types”, nos permite filtrar según el tipo de estudio. Si en los límites de Pubmed se selecciona “Randomized Controlled Trial (RCT)”, que sería un buen nivel de evidencia, obtenemos para nuestro caso 2 títulos, pero en sólo uno de ellos se puede leer el abstract.

Pentoxifylline improves short-term survival in severe acute alcoholic hepatitis: a double-blind, placebo-controlled trial. Akriviadis E, Botla R, Briggs W, et al. *Gastroenterology* 2000; 119: 1637-48<sup>10</sup>.

Al leer el resumen se observa que el artículo aparentemente puede aplicarse al caso, pues según los autores la pentoxifilina (PTX) redujo la mortalidad en pacientes con hepatitis por alcohol grave; sin embargo, antes de instaurar la medida debemos revisar el trabajo.

## ANÁLISIS

En el análisis crítico de un estudio de tratamiento hay tres pasos fundamentales:

- Evaluar la validez interna del trabajo.
- Análisis de resultados.
- Valorar la aplicabilidad del estudio a la situación clínica (validez externa).

## VALIDEZ INTERNA

Se refiere a la calidad metodológica del estudio; un buen diseño del estudio reducirá el sesgo, sin llegar a eliminarlo, y sólo dejará al azar la distribución aleatoria de los pacientes asignados a grupo control o experimental. Los estudios que cumplan con los criterios de validez interna, podrán ser considerados como “válidos” o “ciertos”. Su evaluación incluye criterios primarios y secundarios.

### 1) Criterios primarios:

**Randomización o aleatorización:** La asignación de un paciente al grupo de intervención

o al de control es al azar para evitar sesgo, logrando grupos comparables; por lo tanto los grupos son en todo iguales, y en teoría cualquier diferencia en los *outcomes* es atribuible a la intervención. Sin randomización, nada garantiza que los investigadores, deseosos de demostrar la efectividad de un tratamiento, no asignen preferencialmente a los pacientes más sanos, que lograrán mejores resultados. En este estudio un coordinador, que no era uno de los investigadores, seleccionaba al azar sobres sellados con la decisión de tratamiento. Este método es muy utilizado, pero se ha criticado que en ocasiones, se puede ver al trasluz el tratamiento indicado, introduciendo así un sesgo.

**Inclusión de pacientes en el análisis:** Puede desglosarse en dos puntos; el seguimiento de los pacientes y el análisis por intención de tratar.

A) Seguimiento: Si el número de eventos en ambos grupos fuera similar, el “perder de vista” incluso a pocos pacientes falsearía las conclusiones. Esto queda claro en el análisis “*worst case scenario*” (la peor situación posible): si 2/20 pacientes mueren en cada grupo pero en uno se perdieron 2, la peor situación posible sería que ambos murieron, por lo cual el resultado sería 2/20 vs 4/20, el doble de mortalidad en vez de la misma mortalidad.

En este trabajo, varios pacientes no acudieron a controles pero finalmente se logró obtener datos de sobrevivida a 6 meses para todos los que vivían al momento del alta; un paciente que se perdió precozmente en el grupo PTX fue excluido del análisis.

B) Intención de tratar (*intention to treat*): Los pacientes se consideran para el análisis siempre en el grupo al cual se randomizaron, aunque abandonen o cambien el tratamiento; se preserva el beneficio de la randomización. Si un paciente se agrava y cambia de tratamiento, probablemente empeore el pronóstico global del nuevo grupo al que fue asignado. En este estudio los autores declaran haber realizado todos los análisis por intención de tratar; sin embargo, el paciente que se perdió precozmente en el grupo PTX fue excluido del análisis en forma

incorrecta, pues alcanzó a recibir tratamiento y debió ser analizado en el grupo experimental. Incluyéndolo en el análisis, las tasas de eventos habría sido 13/50 (26%) en lugar de 12/49 (24,5%). En este caso, el peor escenario no altera significativamente el resultado final de la investigación.

## 2) Criterios secundarios:

**“Ciego”:** La expresión “doble ciego” (médico y paciente) no garantiza calidad. Muchos participantes deberían ser ciegos: el paciente, el tratante (que idealmente no debe ser investigador), el que analiza los *outcomes* de todo tipo (clínicos, patólogos, laboratoristas, entrevistadores). El término “doble ciego” debería ser reemplazado por la enumeración de los diferentes individuos que fueron ciegos dentro del estudio. Además, se debe considerar que algunos trabajos nunca podrán ser ciegos por factores inherentes a la intervención, como sabor, efectos adversos, equimosis, entre otros que pudieran ser reconocidos por el paciente o el tratante. En este estudio los investigadores describen al trabajo como “doble ciego”; se entregaron cápsulas similares de PTX y placebo (no era un placebo habitual, se usó Vitamina B12 por la similitud de las cápsulas). Los *outcomes* fueron analizados por los investigadores (no describen métodos para asegurar que fueran ciegos). Pudiéramos inferir que dada la alta frecuencia de intolerancia gástrica, el tratante podría sospechar que su paciente está recibiendo PTX y no placebo, pero éste es un dato difícil de extraer de cualquier estudio de intervención.

**Características basales:** Si ambos grupos son diferentes, un tratamiento puede verse “beneficiado” o “castigado” al haberse dado a pacientes con distinta gravedad, comorbilidades, predisposición a complicaciones, etc. Una randomización bien hecha debería evitarlo pero, por ejemplo, si el *n* es pequeño existe la posibilidad de que la randomización falle y los grupos sean distintos en algún aspecto. Por lo tanto todo trabajo debe informar las características basales de ambos grupos. Habitualmente se dispone en la “Tabla 1” de la publicación, y en este caso no hubo diferencias significativas.

**Tratamiento similar a ambos grupos:** Si al inicio o durante el estudio los grupos reciben distinto tratamiento aparte de la intervención (co-intervenciones), los outcomes podrían no depender de la intervención sino de estas diferencias. No debería suceder si hay buena randomización y ciego, pero debe informarse. En este estudio no se entregaron estos datos. El protocolo no incluía corticoides pero no se indica si alguien llegó a recibirlos (no se informan posibles violaciones de protocolo).

## ANÁLISIS DE LOS RESULTADOS

Comienza con entender la forma en que éstos son dados a conocer, y si es necesario hay que traducirlos a términos más claros, relevantes o “decidores” (los investigadores pueden preferir expresarlos de manera de magnificar los efectos).

Existen tres posibles resultados para una intervención, cada uno de los cuales se puede describir matemáticamente de varias formas (fórmulas en el Apéndice 2):

### A) El tratamiento experimental reduce el riesgo de un evento desfavorable.

**RR: riesgo relativo.** Describe qué fracción del riesgo del grupo control representa el riesgo de eventos en el grupo experimental. Se calcula como el riesgo de eventos en el grupo experimental dividido por el riesgo de eventos en el grupo control.

**RRR: reducción de riesgo relativo.** Es la reducción porcentual del riesgo de eventos desfavorables entre el grupo experimental y control. Es un indicador muy usado por los investigadores, pues entrega valores numéricos altos.

**RRA: reducción de riesgo absoluto.** Es la diferencia aritmética entre las tasas de eventos del grupo experimental versus el control. Entrega valores numéricos menores, es poco usada para presentar resultados pero es muy útil para evaluar la magnitud real del efecto del tratamiento.

**NNT: Número necesario para tratar.** Es el número de pacientes que necesitan ser tratados para evitar un evento adverso adicional.

Se calcula a partir de la RRA. Indicador más usado en medicina basada en evidencias para cuantificar el impacto real del efecto del tratamiento.

### B) El tratamiento experimental aumenta el riesgo de un evento favorable.

**IBR: incremento del beneficio relativo.** Aumento en porcentaje de eventos favorables al comparar el grupo experimental con el control. La diferencia con RRR es sólo conceptual, sin embargo, matemáticamente no son equivalentes, pues si de 100 sujetos mueren 20 con intervención y de 100 sujetos mueren 40 en el grupo control, tenemos: RRR de muerte de 50%, en cambio si lo analizamos como sobrevida tenemos: 80% y 60 % donde el IBR no es 50 % sino 33%. La conversión de evento positivo a negativo es estable con los Odds ratio (OR), pero no con los RR. Esa es una de las razones por la que algunos estadísticos insisten en el uso de OR.

**IBA: incremento del beneficio absoluto.** Es la diferencia aritmética entre el porcentaje de eventos en el grupo experimental comparado con el grupo control.

**NNT: número necesario para tratar.** Número de pacientes que necesitan ser tratados para crear un evento favorable adicional.

### C) El tratamiento experimental aumenta el riesgo de un evento desfavorable

**IRR: incremento del riesgo relativo.** Aumento en porcentaje del riesgo de eventos desfavorables entre el grupo experimental y el control.

**IRA: incremento del riesgo absoluto.** Diferencia absoluta entre la tasa de eventos desfavorables cuando en el grupo experimental la intervención produce más daño que lo observado en el grupo control.

**NNH: número necesario para hacer daño.** Es el número de pacientes, entre los cuales, al recibir todos la intervención experimental, existirá un daño adicional al comparar con el grupo control. Casi nunca es reportado como tal en los trabajos. Es muy útil, por ejemplo, al comparar dos tratamientos que tienen

un NNT similar. Puede calcularse para un outcome desfavorable (mortalidad) o para un efecto adverso del tratamiento.

Analicemos los resultados aplicando estos conceptos al trabajo.

¿Cómo se informaron los resultados?

“Doce (24,5%) de 49 pacientes que recibieron PTX y 24 (46,1%) de 52 pacientes que recibieron placebo, murieron durante la hospitalización índice (P = 0,037; riesgo relativo 0,59; 95% IC, 0,35-0,97)”.

Se informó el porcentaje de un evento desfavorable en el grupo experimental y el control; se informó el riesgo relativo del grupo experimental versus el control.

Apliquemos las fórmulas 1, 2, 3 y 4 (Apéndice 2).

Tasa de evento (mortalidad) en cada grupo:

Tasa de eventos en grupo experimental (TEGE):  $12/49 = 24,5\%$

Tasa de eventos en grupo control (TEGC):  $24/52 = 46,1\%$

**RR:**

$$\frac{\text{TEGE}}{\text{TEGC}} = \frac{24,5\%}{46,1\%} = 53,1\%$$

El riesgo de morir en el grupo experimental es el 53,1% del riesgo control (el cálculo manual difiere del reportado por los investigadores, de 59% baja a 53,1%).

**RRR:**

La reducción de riesgo relativo describe qué porcentaje del riesgo basal representa la diferencia entre el riesgo de ambos grupos.

$$\frac{\text{TEGE}-\text{TEGC}}{\text{TEGC}} = \frac{46,1 - 24,5}{46,1} = 46,9\%$$

“La reducción de riesgo es el 46,9% del riesgo del grupo control”. Los trabajos suelen enunciarlo de forma más llamativa: “el tratamiento reduce en un 46,9% el riesgo de morir en relación al grupo control”.

RR y RRR son complementarios:  $46,9 + 53,1 = 100$ , es decir RRR puede calcularse simplemente como 100 (riesgo basal) – riesgo relativo.

**RRA:**

Describe en cuánto se reduce el riesgo respecto al riesgo basal del grupo control en tér-

minos absolutos, no como una fracción del riesgo basal.

$$\text{TEGE} - \text{TEGC} = 46,1 - 24,5 = 21,6\%$$

“El riesgo de morir se reduce en un 21,6% en los pacientes del grupo de tratamiento”.

El valor es menos de la mitad del RRR pero puede ser aún menor cuando el evento es poco frecuente en la población estudiada.

Ej: si TEGE = 2% y TEGC = 4%

$$\text{RRR} = (\text{TEGE} - \text{TEGC}) / \text{TEGC} = (4 - 2) / 4 = 0,5 = 50\%$$

$$\text{RRA} = 4 - 2 = 2\%$$

**NNT:**

A partir de RRA, usando la fórmula 4:

$$1/\text{RRA} = 1/0,216 \text{ o } 100/21,6 = 4,63$$

“Es necesario tratar a 5 pacientes para evitar una muerte adicional”.

En el ejemplo anterior con pocos eventos:

$$\text{RRA} = 2\%; \text{NNT} = 1/0,02 \text{ o } 100/2 = 50$$

Con RRA y NNT podemos cuantificar el impacto real del tratamiento.

**Precisión de la estimación de efecto:** La precisión está determinada por la homogeneidad en el resultado obtenido por varios estudios randomizados que pretendan responder la misma pregunta, tanto en la dirección del efecto como en su magnitud. El valor p informado es 0,037, es decir “la probabilidad de que la conclusión sea debida al azar es 3,7%” respecto a hipótesis nula o en otras palabras, si se realizaran 100 estudios aleatorizados, se espera que 3 a 4 estudios muestran resultados opuestos al resto.

Actualmente se prefieren los intervalos de confianza (IC) para el valor entregado; este último se denomina “estimador puntual” y está al centro de una amplia gama de posibles valores que se obtendrían si se hiciera el estudio cien veces.

Calculemos el IC para el RRA (que sería el estimador puntual en este caso):

$$\pm 1,96 \times \sqrt{\frac{\text{TEGC} \times (1-\text{TEGC})}{\text{N}^\circ \text{ pacientes GC}} + \frac{\text{TEGE} \times (1-\text{TEGE})}{\text{N}^\circ \text{ pacientes GE}}}$$

$$\pm 1,96 \times \sqrt{\frac{0,461 \times (1-0,461)}{52} + \frac{0,245 \times (1-0,245)}{49}}$$

$$\pm 0,18127 = \pm 18,1\%$$

Aplicado este intervalo de confianza a un RRA de 21,6% obtenemos que en el 95% de los casos RRA estará entre 3,4% y 39,7% → **NNT = 5 (entre 3 y 30)**.

### Efectos adversos:

El trabajo reporta efectos adversos en la tabla 3 junto con datos de mortalidad. Sólo hubo diferencias significativas en dolor epigástrico o vómitos:

Dolor abdominal epigástrico y vómitos: 13/49 (26,5%) vs 5/52 (9,6%),  $P = 0,037$ .

Se calcula IRA (fórmula 8):  $26,5 - 9,6 = 16,9\%$  (95% IC entre 2,2% y 31,6%).

Y con ello NNH (fórmula 9):  $1/0,169$  o  $100/16,9$  → **NNH = 6 (entre 4 y 46)**.

### APLICACIÓN

Hemos visto que el trabajo es de calidad metodológica aceptable salvo algunos reparos y hemos analizado sus resultados, los que parecen ser favorables. Sin embargo, esto no garantiza que podamos aplicar sus resultados a nuestro caso.

La determinación de aplicabilidad requiere poner atención a múltiples aspectos. Algunas condiciones que podrían hacer no aplicable a la clínica un buen estudio:

- Existen diferencias importantes entre los pacientes estudiados y nuestros pacientes reales (no debería ocurrir si nuestra búsqueda fue acotada): ej. el trabajo se hizo con pacientes de otro sexo, raza, edad, etiología de la enfermedad, gravedad, etc.
- Resultados de baja relevancia: El efecto favorable neto es poco significativo (por tamaño estadístico o por importancia del *outcome*).
- Efectos adversos excesivos: El daño que estos causen podría sobrepasar al beneficio.
- Costo: Es imposible aplicar la intervención por costo (un análisis costo-beneficio puede ser útil, pues eventualmente el ahorro que se logra evitando complicaciones y mejorando los *outcomes* sea mayor).
- Precisión de la estimación del efecto: Si el número de pacientes es pequeño los resultados pueden ser poco confiables, con un intervalo de confianza amplio.

Revisemos la aplicabilidad en este caso.

- La población es atingente (hepatitis por alcohol grave, el score de Maddrey era criterio de inclusión; la mayoría eran hombres, edad  $\pm 40$  años).
- Los resultados son importantes: Hay reducción de gran magnitud de un *outcome* importante como es mortalidad con un NNT de 5.
- Efectos adversos: NNH de 6 (mayor al NNT) para dolor epigástrico o vómitos, lo cual parece un efecto adverso menos relevante que el *outcome* que queremos reducir (mortalidad).
- Costo: En nuestro medio el fármaco está menos disponible que los corticoides, pero probablemente su costo se compense al ser más seguro (no tiene los riesgos de la terapia esteroidea), aunque para definirlo se requeriría un análisis costo-beneficio. Esto, sumado a la controversia aún existente respecto a la utilidad de los corticoides en esta patología.
- Precisión de la estimación de efecto: El intervalo de confianza de  $\pm 18,1\%$  para la RRA implica una estimación bastante imprecisa pero siempre a favor del tratamiento. El NNT es de 6 con un rango de 3 a 30. Aún con el peor NNT posible (30) el beneficio sería adecuado en una patología que tiene una mortalidad de hasta 46%.

Hasta ahora el análisis de aplicabilidad es favorable al uso de PTX con nuestro paciente. Sin embargo, persiste una duda: ¿será superior a los corticoides? ¿Por qué se usó placebo?

Los investigadores refieren que han realizado 3 RCT previos con corticoides sin éxito, por lo que decidieron realizar la comparación contra placebo, destacan además que existen varios meta-análisis con resultados contradictorios<sup>3-7</sup>. Durante nuestra búsqueda en Internet no encontramos estudios que compararan directamente PTX y corticoides. Por lo tanto, intentemos una comparación indirecta, con los resultados de uno de los meta-análisis favorables a los corticoides: Corticosteroids improve short-term survival in patients with severe alcoholic hepatitis (AH): individual data analysis of the last three randomized placebo controlled



double blind trials of corticosteroids in severe AH. Mathurin P et al, J Hepatol 2002; 36: 480-77.

El *outcome* analizado fue la *sobrevida* a 28 días. En total son 3 estudios que acumulan 102 pacientes que recibieron placebo y 113 que recibieron corticoides.

La sobrevida fue 84,6% con corticoides vs 65,1% con placebo.

Traduzcamos: mortalidad 15,4% vs 34,9%.

RRA: 19,5% (es igual que el IBA) → 5,128 ≈ 6

No se evaluaron efectos adversos.

En suma: una comparación indirecta de PTX y corticoides muestra que el NNT de PTX es mejor, y ante la falta de datos con corticoides, PTX probablemente es más segura.

## CONCLUSIÓN

Si bien la evidencia que respalda el uso de PTX es escasa (hasta ahora sólo un estudio randomizado y controlado), es de buena calidad metodológica y de buena aplicabilidad. Por el contrario, la evidencia que respalda a los corticoides es controversial, el beneficio parece ser similar o inferior al de PTX, y conlleva los riesgos del tratamiento esteroideal.

## DISCUSIÓN

En el escenario clínico que motivó este ejercicio, se decidió iniciar tratamiento con Pentoxifilina, la evolución del paciente fue favorable y fue dado de alta al séptimo día. Durante su evolución, presentó inicialmente una discreta alteración de la función renal, con elevación al doble de la creatinina basal, hasta 1,4 mg/dL. Sin embargo, se aumentó el aporte de volumen parenteral y la creatinina disminuyó hasta 0,8 al momento del alta. Al respecto, se puede destacar que el trabajo, si bien no se diseñó para ello, muestra un beneficio adicional del tratamiento con Pentoxifilina, que consistió en una reducción de la incidencia de síndrome heparotorrenal y de la mortalidad por dicha causa, con un riesgo relativo de 29%. En la discusión, los autores destacan que es probable que este efecto obedezca al antagonismo del factor de necrosis tumoral alfa (TNF- $\alpha$ ) que la Pentoxifilina ha demostrado previamente. El

equipo tratante discutió cuál fue el real papel de la Pentoxifilina en impedir la aparición de un síndrome heparotorrenal; ¿se debió al tratamiento o a la hidratación? Si bien la respuesta probablemente requiera a su vez un estudio randomizado, planteamos al lector como ejercicio aplicar los conceptos revisados en este artículo y por ejemplo, usar los datos del trabajo para calcular la real magnitud de este nuevo beneficio o realizar su propia búsqueda electrónica tras nueva evidencia en esta área.

## RESUMEN

**El objetivo de este artículo es entregar conceptos básicos acerca de la toma de decisiones sobre tratamiento en Hepatología, basada en los criterios de la Medicina Basada en Evidencias (MBE). La metodología propuesta se basa en un caso clínico de hepatitis por alcohol grave, a partir del cual se revisan herramientas para encontrar la mejor evidencia disponible; se realiza el análisis crítico de un estudio sobre tratamiento y se discute su aplicabilidad clínica. Con un modelo de búsqueda de evidencia basado en una pregunta clínica, se identificó un estudio randomizado y controlado sobre la utilidad de Pentoxifilina esta patología. Se discuten los criterios de validez interna de un artículo de tratamiento, que incluyen criterios primarios (randomización, seguimiento adecuado y análisis por intención de tratar) y criterios secundarios (ciego, características basales similares y tratamiento similar a ambos grupos). El estudio revisado cumplió con estos criterios que respaldan su calidad metodológica. Se analizaron sus resultados, calculando, a partir de ellos, indicadores de magnitud del efecto considerando como evento principal la muerte: reducción de riesgo relativo (RRR) 46,9%, reducción de riesgo absoluto (RRA) 21,6%, número necesario para tratar (NNT) 5 (95% IC 3-30) y número necesario para hacer daño (NNH) 6 (95% IC 4-46). El análisis de aplicabilidad clínica, con criterios que incluyen la relevancia de los resultados, los efectos adversos y la precisión estadística del estudio, resultó favorable al uso de**

**Pentoxifilina en la situación presentada. Se concluye que el uso de los criterios de MBE en búsqueda y análisis crítico permite identificar evidencia de buena calidad, constituyendo una valiosa herramienta en la toma de decisiones clínicas.**

*Palabras claves: Hepatitis alcohólica, Pentoxifilina, Cirrosis y Medicina basada en evidencias.*

## APÉNDICE 1

Niveles de Evidencia (clasificación de Oxford para estudios de tratamiento) <sup>11</sup>

**Nivel 1a:** Revisión sistemática de estudios randomizados y controlados (con homogeneidad, es decir concordancia general entre los resultados de los estudios).

**Nivel 1b:** Estudio individual randomizado y controlado con intervalo de confianza estrecho.

**Nivel 1c:** Estudios “todos o ninguno”. Significa que antes de estar disponible el tratamiento todos los pacientes morían, pero con el tratamiento algunos sobreviven; o que sin el tratamiento algunos morían pero ahora ninguno muere.

**Nivel 2a:** Revisión sistemática de estudios de cohorte (con homogeneidad).

**Nivel 2b:** Estudio de cohorte individual o estudio randomizado controlado de baja calidad.

**Nivel 2c:** Investigación de *outcomes* o estudios epidemiológicos.

**Nivel 3a:** Revisión sistemática de estudios de casos y controles (con homogeneidad).

**Nivel 3b:** Estudio individual de casos y controles.

**Nivel 4:** Series de casos y estudios de cohorte o casos y controles de baja calidad.

**Nivel 5:** Opinión de expertos sin valoración crítica, o basada en fisiología, investigación básica o “primeros principios”.

## APÉNDICE 2

### Fórmulas

**Fórmula 1: Riesgo relativo (RR) =  $\frac{TEGE}{TEGC}$**

TEGE: Tasa de eventos en grupo experimental.  
TEGC: Tasa de eventos en grupo control.

**Fórmula 2: Reducción de riesgo relativo (RRR) =  $\frac{TEGE-TEGC}{TEGC}$**

**Fórmula 3: Reducción de riesgo absoluto (RRA) = (TEGE-TEGC)**

**Fórmula 4: Número necesario para tratar (NNT) = (1/RRA) o (100/RRA%)**

El resultado se aproxima hacia arriba (p. Ej. 5,3 = 6).

**Fórmula 5: Incremento del beneficio relativo (IBR) =  $\frac{TEGE-TEGC}{TEGC}$**

**Fórmula 6: Incremento del beneficio absoluto (IBA) = (TEGE-TEGC)**

**Fórmula 7: Número necesario para tratar (NNT) = (1/IBA) o (100/IBA%)**

**Fórmula 8: Incremento del riesgo relativo (IRR) =  $\frac{TEGE-TEGC}{TEGC}$**

**Fórmula 9: Incremento del riesgo absoluto (IRA) = (TEGE-TEGC)**

**Fórmula 10: Número necesario para hacer daño (NNH) = (1/IRA o 100/IRA%)**

## BIBLIOGRAFIA

- 1.- SACKETT D, ROSENBERG W, GRAY J et al. Evidence based medicine: what it is and what it isn't. *BMJ* 1996; 312: 71-2.
- 2.- MADDREY W, BOITNOTT J, BEDINE M et al. Corticosteroid therapy of alcoholic hepatitis. *Gastroenterology* 1978; 75: 193-9.
- 3.- RAMOND M, POYNARD T, RUEFF B et al. A randomized trial of prednisolone in patients with severe alcoholic hepatitis. *N Engl J Med* 1992; 326: 507-12.
- 4.- IMPERIALE T, O'CONNOR J, MCCULLOUGH A. Corticosteroids are effective in patients with severe alcoholic hepatitis. *Am J Gastroenterol* 1999; 94: 3066-8.
- 5.- IMPERIALE T, MCCULLOUGH A. Do corticosteroids reduce mortality from alcoholic hepatitis? A meta-analysis of the randomized trials. *Ann Intern Med* 1990; 113: 299-307.
- 6.- DAURES J, MERAY P, BORIES P et al. Place de la corticothérapie dans le traitement des hépatites alcooliques. Resultat d'une meta-analyse. *Gastroenterol Clin Biol* 1991; 15: 223-8.
- 7.- CHRISTENSEN E, GLUUD C. Glucocorticoids are ineffective in alcoholic hepatitis: A meta-analysis adjusting for confounding variables. *Gut* 1995; 37: 113-8.
- 8.- MATHURIN P, MENDENHALL C, CARITHERS R et al. Corticosteroids improve short-term survival in patients with severe alcoholic hepatitis (AH): individual data analysis of the last three randomized placebo controlled double blind trials of corticosteroids in severe AH. *J Hepatol* 2002; 36: 480-7.
- 9.- MCKIBBON A, HUNT D, RICHARDSON W et al. Finding the evidence. En: Guyatt G. *Users' Guides to the Medical Literature*. American Medical Association 2002: 14-24.
- 10.- AKRIVIADIS E, BOTLA R, BRIGGS W et al. Pentoxifylline improves short-term survival in severe acute alcoholic hepatitis: a double-blind, placebo-controlled trial. *Gastroenterology* 2000; 119: 1637-48.
- 11.- PHILLIPS B, BALL C, SACKETT D et al. Oxford-Centre for Evidence Based Medicine. [www.cebm.net/levels\\_of\\_evidence.asp#levels](http://www.cebm.net/levels_of_evidence.asp#levels) (conexión 29.8.04).