

Artículo especial

Pediatría basada en la evidencia. Estudios de intervención

Evidence-based Pediatrics. Interventional studies

Dras. Valeria Berlín* y Graciela Demirdjian**

La serie de lectura crítica que iniciamos con los estudios de daño-etilogía,¹ continúa en este artículo en el que analizaremos un ensayo clínico aleatorizado. Este tipo de estudios (cuya fascinante historia revisamos hace poco²) es considerado el diseño óptimo para valorar los resultados (la eficacia) de cualquier intervención (terapéutica, preventiva, educativa, organizativa). Como modelo para el análisis se empleará un artículo publicado en la revista *Lancet* en 2006, cuyo objetivo es evaluar la eficacia de una vacuna para la prevención de la otitis media aguda.³

Con frecuencia, durante la práctica diaria surgen dudas concretas sobre los pacientes que estamos evaluando, los métodos diagnósticos que utilizamos, los mejores tratamientos disponibles, o el pronóstico esperable. Existen varias opciones para contestar dichos interrogantes: basarnos sólo en la experiencia (sin olvidar que la memoria es traicionera y que tendemos a vernos influidos por las experiencias recientes), basarnos en una autoridad experimentada en el tema en conflicto, o basarnos en la evidencia científica. El problema que surge con esta última opción es la falta de tiempo disponible para la lectura; y si a esto se le suma la dificultad en la interpretación de los resultados, parece una opción casi imposible de aplicar. El conflicto es aun más grave cuando se trata de la toma de decisión sobre un tratamiento, donde juegan numerosos factores que involucran la incertidumbre acerca de su eficacia, las preocupaciones respecto de su seguridad, los costos para la familia o el sistema de salud, y las presiones de la industria, colegas y familiares.

El objetivo de este artículo es facilitar la lectura crítica de trabajos científicos que evalúan la eficacia de intervenciones, y proporcionar herramientas para la interpretación y el análisis de los resultados,

dándole a éstos un sentido práctico y aplicable a nuestros pacientes. Para esta oportunidad se seleccionó un estudio complejo en su diseño y rico en terminología estadística, denominado ensayo clínico controlado y aleatorizado (ECA).

Diseño de estudios de intervención

Frecuentemente nos sentimos tentados de aplicar tratamientos cuyos buenos resultados han sido comunicados como exitosos en series de pacientes. Esta interpretación optimista de la bibliografía lleva a incorporar numerosas prácticas sin probada eficacia, potencialmente dañinas para los pacientes, y seguramente generadoras de costos injustificados para el sistema de salud. Como las enfermedades pueden evolucionar espontáneamente, y dado que muchas características de los pacientes influyen sobremanera la respuesta terapéutica, el único modo de probar la eficacia de una intervención es mediante un diseño comparativo que enfrente la intervención a evaluar contra una alternativa (como el tratamiento habitual, un placebo o no hacer nada).

Un ensayo clínico ("clinical trial") se define como toda investigación en la cual se intenta probar la utilidad de algún tipo de estrategia terapéutica. Como veremos más adelante, para que un estudio de estas características sea válido (es decir, esté libre de sesgos) la asignación de los pacientes al grupo experimental o al de control (testigo) debe ser realizada de manera aleatoria, lo que origina el llamado ensayo clínico controlado y aleatorizado (ECA), que se considera el diseño óptimo para la evaluación de intervenciones[#].

El término "intervención" se refiere a cualquier estrategia terapéutica, preventiva, educativa u organizativa que implique un "tratamiento" de sujetos o grupos en un sentido amplio.

* Residencia de Medicina Interna. Hospital Nacional de Pediatría "Prof. Dr. J. P. Garrahan". Ciudad de Buenos Aires.

** Investigación Tecnológica. Hospital Nacional de Pediatría "Prof. Dr. J.P. Garrahan". Grupo de Pediatría Basada en la Evidencia. Subcomisión de Epidemiología de la Sociedad Argentina de Pediatría.

Correspondencia:
Dra. Graciela Demirdjian
gdemir@intramed.net

El artículo analizado es un ensayo clínico controlado de grandes dimensiones (4.968 pacientes, compara el uso de la vacuna antineumococo con proteína D contra un grupo testigo [control] que recibe vacuna para hepatitis A a modo de placebo), y multicéntrico (incorporó pacientes en 50 centros pediátricos de la República Checa y Eslovaquia).

Cómo iniciar la lectura crítica

Como siempre, el primer paso es comprender de qué se trata el estudio. En este caso, ya desde el título (ver recuadro) puede reconstruirse la pre-

gunta que originó el estudio con sus 4 componentes (“PICO”): el Paciente (población), la Intervención (terapéutica o preventiva), el Comparador (la alternativa) y el “Outcome” o resultado (el evento o enfermedad que se quiere prevenir):

¿La vacuna antineumococo con polisacáridos de 11 serotipos diferentes conjugada con proteína D de H. influenzae previene la otitis media aguda (OMA) en niños?

¿Sigo leyendo?

¿Qué aspectos deben revisarse en un artículo

| | |
|--|--|
| <p>Pneumococcal capsular polysaccharides conjugated to protein D for prevention of acute otitis media caused by both <i>Streptococcus pneumoniae</i> and non-typable <i>Haemophilus influenzae</i>. A randomised double-blind efficacy study.</p> <p><i>Lancet</i> 2006; 367:740-748.</p> <p>Question: Can a novel vaccine, containing polysaccharides from 11 different <i>Streptococcus pneumoniae</i> serotypes each conjugate to <i>Haemophilus influenzae</i>-derived protein D, prevent acute otitis media ?</p> <p>Design: Randomised, controlled trial.</p> <p>Setting: 27 pediatric centers in the Czech republic and 23 in Slovakia.</p> <p>Participants: 4968 infants, ages 6 weeks to 5 months.</p> <p>Intervention: Infants were randomly assigned to receive either pneumococcal protein D conjugate or hepatitis A vaccine at the ages of 3, 4, 5, and 12-15 month and were followed up until the end of the second year of life. Middle ear fluid was obtain for bacteriologic culture and serotyping in children who presented with abnormal tympanic membrane or presence of middle-ear effusion, plus 2 predefined clinical symptoms.</p> <p>Outcomes: Protective efficacy against the first episode of acute otitis media caused by vaccine pneumococcal serotypes.</p> <p>Main results: From 2 weeks after the third dose to 24-27 month of age, 333 clinical episodes of acute otitis media were recorded in the protein D conjugate group (n= 2455) and 499 in the control group (n= 2452), giving a significant relative risk reduction of 33,6% (95% CI 20,8-44,3) in the overall incidence of acute otitis media (absolute risk reduction 6,8%, Number needed to treat= 14). Vaccine efficacy was also shown for episodes of acute otitis media caused by pneumococcal vaccine serotypes (52%, 95% CI 35.0–65.5 for the first episode and 57.6% (95% CI 41.4-69.3 for any episode). Efficacy was also shown against episodes of acute otitis media caused by non-typable <i>H. influenzae</i> (35.3%, 95% CI 1.8-57.4). The vaccine reduced frequency of infection from vaccine related cross-reactive pneumococcal serotypes by 65%, but did not significantly change the number of episodes caused by other non vaccine serotypes.</p> <p>Conclusions: Use of the <i>H. influenzae</i>-derived protein D as a carrier protein for pneumococcal polysaccharides not only allowed protection against pneumococcal otitis, but also against acute otitis media caused by non-typable <i>H. influenzae</i>. Whether this approach would also improved protection against lower respiratory tract infections warrants further investigation.</p> | <p>Guía del usuario para artículos sobre terapéutica o prevención</p> <p><i>JAMA</i> 1993; 270: 2598-2601.</p> <p>Validez interna</p> <p>1. <i>Criterios primarios:</i></p> <ul style="list-style-type: none"> (a) ¿La asignación de los tratamientos a los pacientes se hizo de manera aleatoria? (b) ¿Se han tenido en cuenta adecuadamente todos los pacientes incluidos en el ensayo y se los ha considerado a su finalización? <p>2. <i>Criterios secundarios:</i></p> <ul style="list-style-type: none"> (a) ¿Se ha mantenido el enmascaramiento respecto del tratamiento aplicado en lo que se refiere a los pacientes y el personal del estudio? (b) ¿Eran similares los grupos al inicio del ensayo? (c) ¿Aparte de la intervención experimental, se ha tratado a los grupos de la misma forma? <p>Análisis de resultados</p> <ul style="list-style-type: none"> (a) ¿Cuál ha sido la magnitud del efecto del tratamiento? (b) ¿Cuál ha sido la precisión de la estimación del efecto del tratamiento? <p>Utilidad de los resultados para la asistencia de mis pacientes</p> <ul style="list-style-type: none"> (a) ¿Se consideraron todas las variables de evaluación del resultado, clínicamente importantes? (b) ¿Compensan los probables beneficios del tratamiento los posibles efectos perjudiciales y sus costos? |
|--|--|

sobre tratamiento para saber si se lo descarta o si se lo lee? Los referidos en las 3 preguntas básicas que todas las guías de usuarios publicadas en el *JAMA* proponen: la validez del estudio, los resultados y si serán útiles en mis pacientes (ver recuadro).

GUÍA PARA EL ANÁLISIS CRÍTICO⁴

A. Validez interna: ¿Son válidos los resultados del estudio?

Para que las conclusiones sean verdaderas (válidas), el estudio debe haber sido realizado de manera metodológicamente correcta (validez interna). Para ello, deben estar presentes dos requisitos fundamentales:

1. Criterios primarios: si no se cumplen, dedique su tiempo a leer otra cosa, a menos que sea lo único que tenga...

a. *¿La asignación de los tratamientos a los pacientes se hizo de manera aleatoria?*

El principal criterio de validez de un estudio que evalúa la utilidad de una intervención es que la asignación del tratamiento haya sido al azar (aleatorizada o "randomizada", del término inglés: "random"). Este dato es tan importante que generalmente se menciona ya desde el título (como en el estudio que estamos analizando). Pero ¿por qué es tan trascendente este punto? Veamos algunos ejemplos para aclararlo.

Si se quisiera realizar un estudio sobre los beneficios de incorporar el desayuno en la escuela, y se dejara librado al criterio de las maestras la asignación de los niños, probablemente les darían el desayuno a los más pobres. De la misma manera, cuando se permite que los médicos tratantes asignen las intervenciones a los pacientes en un ensayo clínico, se tiende a elegir los casos más graves para la rama experimental. Así, se generan dos grupos que son diferentes desde el inicio, por lo que al comparar los resultados al final del estudio no se podrán atribuir las diferencias al tratamiento. Es más, es posible que la mayor gravedad de los pacientes en el grupo experimental impida evidenciar el efecto benéfico del nuevo tratamiento. En la mayoría de los casos, sin embargo, lo que ocurre es lo contrario: el investigador (consciente o inconscientemente) asigna el nuevo tratamiento a pacientes que espera respondan favorablemente, con lo que inclina la balanza a favor de la rama experimental. Se ha comprobado que los ensayos clínicos con asignación no aleatorizada tienden a dar más resultados falsamente positivos, sobrestimando la eficacia y promoviendo la incorporación de nuevas intervenciones que luego se muestran inefectivas.

Es por esta razón que la decisión de qué paciente se asigna a cada grupo debe quedar librada al azar. La asignación aleatorizada puede realizarse por medio de técnicas simples, como la utilización de sobres cerrados opacos, o más sofisticadas, como la selección informatizada, y debe estar oculta o enmascarada para los investigadores (no deben poder adivinar la secuencia). La gran ventaja de la aleatorización es que genera grupos equivalentes, es decir similares en todas aquellas características que pueden influir en la respuesta al tratamiento, tanto las conocidas (medibles) como las desconocidas (no medibles), lo que permite atribuir luego cualquier diferencia a la intervención.

En este estudio multicéntrico, la asignación al azar se efectuó por medio de un listado de aleatorización generado por una computadora, al que los investigadores podían acceder vía Internet.

b. *¿Se han tenido en cuenta adecuadamente todos los pacientes incluidos en el ensayo y se los ha considerado a su finalización?*

Este criterio evalúa dos componentes. El primero es si el seguimiento ha sido completo. Para valorarlo se debe observar bien que el número de pacientes incorporado al estudio y el que se incluye en el análisis sea el mismo. Frecuentemente ocurren pérdidas en el seguimiento de los pacientes, sobre todo en ensayos clínicos prolongados. Como regla general, se considera que un seguimiento es adecuado si es mayor al 80% (la pérdida superior al 20% invalida las conclusiones). Pero además de la magnitud de la pérdida debe evaluarse su distribución, ya que se sabe que quienes se pierden suelen ser los pacientes de peor evolución: una pérdida mayor en la rama experimental podría significar que hubo peores resultados que no se tendrán en cuenta en el análisis, lo cual sesga las conclusiones. Una forma de comprobar si la pérdida pudo haber generado un sesgo (una tendencia falsa en los resultados) es que los autores (o el lector) realicen un análisis del peor escenario, que consiste en asumir que todos los pacientes perdidos en la rama que se muestra más efectiva tuvieron mala evolución, y volver a analizar los resultados para ver si se mantienen las mismas conclusiones.

En el presente estudio se perdieron en el seguimiento sólo 19 niños en el grupo experimental y 12 en el grupo control (menos del 1%).

El segundo componente de este criterio evalúa si se han analizado los pacientes en el grupo al que fueron asignados aleatoriamente; este principio se denomina análisis por intención de tratar. Durante el desarrollo de un ensayo clínico, existen a menu-

do desviaciones del protocolo imposibles de evitar: los pacientes no cumplen (adherencia, cumplimiento u observancia terapéutica, o “compliance”) con la intervención asignada, o acuden a otro médico que cambia el tratamiento (cruces de grupo o “cross-over”) o agrega medidas coadyuvantes (co-intervenciones). La intuición nos dice que quienes no adhieren al tratamiento deberían ser eliminados del estudio por incumplimiento del protocolo, y que los que se cruzan de grupo deben ser analizados en el grupo contrario, es decir, debe realizarse un análisis por protocolo o “per protocol”. Pero si esto fuera así, estaríamos inclinando la balanza de nuevo a favor del grupo experimental, ya que frecuentemente los pacientes que no cumplen con el tratamiento o cambian de grupo suelen hacerlo por mala evolución o efectos adversos. Es por este motivo que, cuando se realiza el análisis de los resultados, los pacientes deben ser tenidos en cuenta en el grupo donde inicialmente fueron asignados, es decir según la intención de tratamiento, para respetar y no echar por tierra la aleatorización inicial.

En este trabajo los autores realizaron ambos análisis: por intención de tratar y por protocolo. Para el análisis de eficacia “per protocol” se excluyeron 15 pacientes en cada grupo por violaciones diversas del protocolo (la *Figura 1* del estudio muestra un flujograma[#] de los pacientes).

2. Criterios secundarios: si están presentes le agregan validez al estudio pero son menos esenciales o a veces no son aplicables.

a. *¿Se mantuvo un diseño enmascarado (“ciego”) para pacientes y médicos?*

En numerosas oportunidades saber si el paciente está recibiendo una medicación nueva activa o un placebo o nada influye psicológicamente en la respuesta del paciente, o en cómo el médico evalúa la respuesta al tratamiento o proporciona cuidados adicionales. Para controlar esta influencia de la subjetividad se aplica un diseño a ciegas (“ciego”) o enmascarado que implica un desconocimiento del tratamiento asignado. En los estudios con enmascaramiento simple, el único que no sabe es el paciente, mientras que en los de doble enmascaramiento desconocen la asignación tanto el médico como el paciente. Es importante resaltar que se trata de un criterio secundario dado que existen intervenciones

que no pueden enmascararse (como las quirúrgicas, las kinésicas, las psicológicas o las educativas); en estos casos, debe reforzarse especialmente la objetividad en la medición de la respuesta al tratamiento, con medidas de resultado estandarizadas exentas de subjetividad.

En el título de este estudio ya se anticipa que se utilizó un diseño con enmascaramiento doble (“double-blind”), que implica que ni los médicos ni los pacientes (en este caso las madres de los niños) sabían si habían recibido vacuna antineumococo (tratamiento activo) o antihepatitis (placebo).

b. *¿Fueron similares las co-intervenciones?*

Frecuentemente existen tratamientos adyuvantes o de sostén aplicados a los pacientes junto con la intervención principal. Estas co-intervenciones deben ser asignadas a ambos grupos de manera igualitaria para no desbalancear la equivalencia generada por la aleatorización y asegurar que, si existen diferencias al finalizar el estudio, éstas sólo puedan ser atribuidas a la intervención evaluada. El diseño a doble ciego protege contra la administración inequitativa de co-intervenciones dado que los médicos no pueden diferenciar a qué grupo pertenecen los pacientes.

En este estudio se administró una dosis de vacuna hexavalente (difteria, tétanos, pertussis, hepatitis B, polio, influenza) a todos los niños participantes en el estudio

c. *¿Eran comparables los grupos al inicio del estudio?*

Si bien la aleatorización tiende a generar grupos equivalentes, no garantiza la similitud entre los grupos; es decir: es improbable pero no imposible que aun por azar (especialmente si las muestras son pequeñas) los grupos sean diferentes en alguna característica relevante para la evolución (las potenciales “variables de confusión o “confounding”). Una manera de verificar esto es observar la composición final de ambos grupos. Esta información generalmente se muestra en la “Tabla 1” del artículo que despliega la distribución en ambas ramas de aquellos factores pronósticos o confusores medidos. Si alguna de esas variables estuviera distribuida de manera diferente entre los grupos, deberá evaluarse si esto pudo haber influido sobre los resultados, ya sea mediante el criterio clínico o con métodos de análisis multifactorial para ajuste estadístico.

La *Tabla 1* de este estudio muestra una distribución similar de ambos grupos en cuanto a los posibles confusores medidos (factores de riesgo de OMA, como edad, lactancia materna, número de niños convivientes y exposición fuera del hogar), lo que confirma la equivalencia de las muestras.

[#] Es una práctica cada vez más común que los autores incluyan un flujograma o diagrama de flujo (“flow diagram” o “trial profile”) que muestra la circulación de pacientes y los motivos de pérdida en las distintas etapas del estudio (elegibilidad, aleatorización y evaluación de resultados).

B. Análisis de resultados: si se cumple la validez interna, podemos pasar a analizar los resultados.

1. ¿Son los resultados importantes? (magnitud del efecto)

Antes de comenzar a analizar los números es importante tener en claro cuáles son los resultados o "outcomes" a los cuales se refieren los efectos del tratamiento. En general, los estudios terapéuticos tienen un resultado o criterio de valoración ("end-point") primario (el que se considera de mayor relevancia para el paciente) y otros resultados o criterios de valoración secundarios (otros efectos que se intenta medir). Para cada uno de ellos, los resultados podrán ser expresados con diferentes medidas estadísticas sobre magnitud del efecto. Veamos.

La eficacia de la vacuna fue medida en cuanto al primer episodio de otitis media aguda[#] por neumococo (resultado primario) o por *Haemophilus influenzae* (resultado secundario), durante el período de seguimiento desde la administración de la vacuna (3, 4, 5 y 12-15 meses) hasta cumplidos los 2 años de edad.

a. Riesgo relativo (RR, "relative risk")

Expresa el riesgo comparativo de sufrir un episodio (en este caso OMA) en el grupo experimental respecto del placebo. Es la misma medida empleada en los estudios de daño con diseño de cohorte, sólo que en caso de intervenciones terapéuticas se espera que su valor sea menor de 1. Recordemos que su cálculo se obtiene dividiendo la incidencia de eventos en ambos grupos:

$$RR = \frac{\text{incidencia en grupo experimental}}{\text{incidencia en grupo control}}$$

| | Eventos totales (episodios de OMA) | | |
|--------------------|---------------------------------------|-------------|-------|
| | Sí (fracasos) | No (éxitos) | |
| Grupo experimental | 333 | 2.122 | 2.455 |
| Grupo placebo | 499 | 1.953 | 2.452 |
| Total | 832 | 4.075 | 4.907 |

- Incidencia de OMA en el grupo tratado (GT)= $333 / 2.455 = 13,6\%$
- Incidencia de OMA en el grupo control (GC)= $499 / 2.452 = 20,4\%$
- $RR = 0,136 / 0,204 = 0,67$

Los autores proveen una definición operativa para el diagnóstico de OMA y aclaran que la timpanostomía es parte del manejo habitual en los países en los que se desarrolló el estudio.

La interpretación del RR es siempre en relación a la unidad, lo que significa que por cada 0,67 episodios de OMA en el grupo vacunado hubo 1 en el grupo placebo.

b. Reducción del riesgo relativo (RRR, "relative risk reduction")

Expresa qué porcentaje del riesgo basal se reduce debido a la intervención y se calcula restando el RR de 1:

$RRR = 1 - RR = 1 - 0,67 = 0,33 = 33\%$ (la vacuna reduce en un 33% el riesgo basal de OMA).

La RRR permite estimar cuánto redujo la nueva intervención el riesgo de base que existía sin tratamiento o con el tratamiento convencional. Es una medida fácil de entender para los pacientes, por lo que es muy usada en publicidad. Algunos autores consideran que una intervención tiene significación clínica cuando esta reducción del riesgo relativo está por encima del 25%.

Un problema con las medidas relativas que hemos visto hasta ahora es que no dan cuenta del impacto real que la intervención va a tener. Para esto se utilizan otras dos medidas estadísticas:

c. Reducción del riesgo absoluto (RRA, "absolute risk reduction" o "ARR")

Expresa cuál es el porcentaje de reducción del riesgo atribuible a los beneficios de la intervención y se calcula restando ambas incidencias:

$RRA = \text{incidencia en GC} - \text{incidencia en GT} = 0,204 - 0,136 = 0,068 = 6,8\%$ (comparada con placebo la vacuna reduce en un 6,8% el riesgo de OMA).

d. Número necesario a tratar (NNT, "number needed to treat")

Es el número de pacientes que debo tratar para que uno de ellos se beneficie o evite el criterio de valoración (end-point o punto final) estudiado. Se calcula como la inversa de la reducción absoluta del riesgo:

$NNT = 1 / RRA = 1 / 0,068 = 0,147 = 14,7 \approx 15$ (se necesita vacunar a 15 niños para prevenir un episodio de OMA).

Es lógico que lo deseable es un NNT lo más pequeño posible (lo ideal sería un NNT de 1, es decir beneficiar a cada paciente que se trata, aunque esto es casi utópico).

2. ¿Cuál es la precisión de la estimación del efecto?

En este punto es donde extrapolamos los resultados de la muestra a la población, o sea que consideramos el grado de error aleatorio que pueden tener las medidas estadísticas estimadas se-

gún el tamaño de la muestra estudiada. Cada vez son menos los estudios que expresan la significación estadística con valores de "p" (término con el cual la mayoría de los médicos se sienten cómodos), y crece el empleo de los intervalos de confianza (IC 95%). Estos expresan el rango de valores en los que podemos estar 95% seguros de que se encuentra el verdadero valor para la población. A medida que el estudio tiene más pacientes, el IC es más estrecho y preciso.

Recordemos que para que las diferencias sean consideradas estadísticamente significativas, el valor de p debe ser $< 0,05$, lo que equivale a que el IC 95% no debe contener la unidad neutra (o sea el valor de no diferencia: uno para RR y cero para RRR y RRA).

En nuestro ejemplo:

- El valor de p para el criterio de valoración (*outcome*) primario fue de 0,0104 en el análisis por intención de tratamiento y de 0,0156 para el análisis por protocolo, lo que implica que las diferencias fueron estadísticamente significativas.
- El IC 95% para el RR (no informado) es de 0,59 a 0,76 (no pasa por el uno, lo que significa que podemos estar 95% seguros de que la vacuna es protectora).
- El IC 95% para el RRR (no informado) es de 24-41% (no pasa por el cero, lo que significa que con una confianza del 95%, la vacuna reduce el riesgo basal de OMA en la población)[#].
- El IC 95% para el RRA (no informado) es de 4,7-8,9% (no pasa por el cero, por lo que podemos tener un 95% de confianza en que la población inmunizada tendrá menos riesgo comparada con los no vacunados).
- El IC 95% para el NNT (no informado) es de 12-22 (podemos prever que vamos a beneficiar con la vacuna a uno de cada 10 o 20 niños que inmunicemos).

Comentario: Hemos sintetizado aquí el análisis de los resultados de eficacia para el total de los episodios de otitis, y a los fines docentes hemos debido agregar los IC que no figuran en el artículo original. Para el análisis de las otitis discriminadas por tipo de germen los autores utilizan otros métodos estadísticos que valoran el tiempo al evento:

tasa o densidad de incidencia (número de episodios de OMA por 1.000 personas - año) y análisis multifactorial por regresión de Cox. Veremos esto próximamente al analizar estudios de pronóstico.

C. Validez externa: ¿Son útiles los resultados para la asistencia de mis pacientes?

a. ¿Son aplicables los resultados a mi práctica clínica?

Es importante considerar si los pacientes del estudio son comparables con los que están bajo nuestra atención. Es frecuente que los tratamientos (sobre todo las drogas) sean probados en adultos y que luego se asuma que la eficacia es análoga en los niños. Existen diferencias fisiológicas, metabólicas y farmacocinéticas para argumentar que esto puede no siempre ser así.

Los criterios de inclusión del estudio fueron niños sanos de 6 semanas a 5 meses reclutados por los pediatras y seguidos hasta completar los dos años de vida, por lo tanto, las conclusiones son aplicables a nuestra población (a menos que pensemos que las otitis en los países del estudio sean muy diferentes a las nuestras).

b. ¿Se consideraron todas las variables de evaluación del resultado clínicamente importantes?

Con frecuencia, los ensayos clínicos muestran que la intervención mejora parámetros intermedios o surrogantes (exámenes complementarios o pruebas fisiológicas), aunque esto no necesariamente implica un impacto clínico importante para el sujeto. Es necesario demostrar que los tratamientos proveen beneficios reales trascendentes para los pacientes (como reducir la morbilidad, la internación o la mortalidad).

En el estudio que nos ocupa, la respuesta a la vacuna se midió en términos de su prevención de otitis, que es lo que realmente importa.

c. ¿Compensan los probables beneficios del tratamiento los posibles efectos perjudiciales y sus costos?

En este punto, cada lector evaluará en su propio contexto si los resultados son lo suficientemente buenos como para aplicarlos a sus pacientes en función de efectos adversos, costos, y preferencias de sus pacientes en caso de que exista eficacia similar entre dos alternativas.

Confiamos en que esta guía facilite a los pediatras la lectura crítica de estudios de intervención, pero sobre todo, que los ayude a descartar tratamientos cuya eficacia no ha sido demostrada mediante estudios de calidad metodológica y a evitar posibles efectos adversos y gastos innecesarios a

Si la reducción del riesgo relativo de 33% tuviera un IC 95% (I entre -10 y +40), el resultado no sería significativo porque abarca el cero y la vacuna podría incluso aumentar el riesgo en la población.

los pacientes, y recursos a nuestro sistema de salud. La información sintetizada en este artículo sobre un tema tan complejo, se complementa con algunas fuentes de lectura recomendada que se incluyen en la bibliografía.⁴⁻⁹ ■

BIBLIOGRAFIA

1. González C, Palladino M, Kuperman S, Demirdjian G. Estudios de etiología-daño. Arch Argent Pediatr 2004; 102(1):53-59.
2. Demirdjian G. Historia de los ensayos clínicos aleatorizados. Arch Argent Pediatr 2006; 104(1):52-61.
3. Prymula R, Peeters P, Chrobok V, et al. Pneumococcal capsular polysaccharides conjugated to protein D for prevention of acute otitis media caused by both *Streptococcus pneumoniae* and non-typable *Haemophilus influenzae*: a randomised double-blind efficacy study. Lancet 2006; 367:740-748.
4. Guyatt GH, Sackett DL, Cook DJ. Guía para usuarios de la literatura médica. Cómo utilizar un artículo sobre tratamiento o prevención. JAMA (ed. esp.) 1993; 270:2598-2601.
5. González de Dios J, González Rodríguez P. Evaluación de artículos científicos sobre intervenciones terapéuticas. Evid Ped 2006; 2:90-97.
6. Kendall JM. Designing a research project: randomised controlled trials and their principles. Emerg Med J 2003; 20:164-168.
7. Guyatt GH, Sackett D, Cook DJ, for the Evidence Based Medicine Working Group: How to use an article about therapy or prevention. Disponible en: <http://www.cche.net/usersguides/therapy.asp>
8. López-Giménez F, Paniagua D, Lamas GA. La interpretación de los ensayos clínicos negativos. Rev Invest Clín 1998; 560(5):435-440.
9. Demirdjian G. Estudios de valoración de tratamientos. En: Programa de Educación a distancia en Metodología de la Investigación para Pediatría (PREMIP). Nivel 2. Módulo 2.7. [en vigencia].

“En la escuela de Proyectistas Políticos pasé un mal rato. Los profesores parecían, a mi juicio, por completo fuera de sus cabales, lo que es un espectáculo que nunca deja de deprimirme. Aquella infeliz gente andaba presentando planes para persuadir a los monarcas de que escogieran a sus ministros en razón de su sabiduría, capacidad y virtud; de enseñar a los funcionarios a considerar el bien común; de recompensar el mérito, las grandes aptitudes y los servicios eminentes; de instruir a los príncipes para que conozcan que su verdadero interés se asienta sobre los mismos cimientos que los de su pueblo; de seleccionar para los empleos públicos a las personas capacitadas para desempeñarlos; con otras muchas insensatas o imposibles quimeras que nunca hasta entonces habían sido concebidas por cabeza humana y que me confirmaron la vieja observación de que no hay nada tan extravagante e irracional que no lo hayan sostenido como verdad algunos filósofos. Para ellos, cualquiera que fuera capaz de hacer crecer dos espigas de trigo o dos briznas de hierba en el espacio en que antes nacía una sola merecería mejor trato y habría un servicio más esencial a su país que toda la casta de políticos”.

JONATHAN SWIFT (1667-1745), escritor irlandés, es el autor de *Los viajes de Gulliver*, una dura sátira contra la sociedad británica. Publicamos fragmentos de su visita a Lilibut, un país de sabios habitantes diminutos.