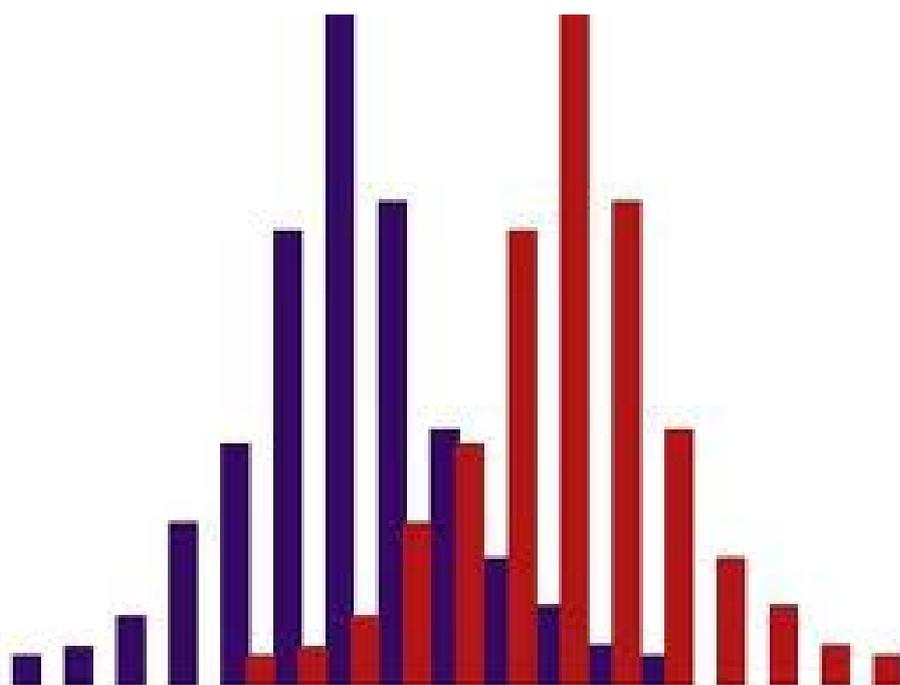


PRÁCTICA BASADA EN LA EVIDENCIA EN SALUD MENTAL DE LA INFANCIA Y LA ADOLESCENCIA

John Hamilton & Füsun Çuhadaroğlu-Çetin

Edición: Matías Irrázaval & Andres Martin

Traductores: Fernanda Prieto-Tagle & Jaume Morey



John Hamilton MD, MSc

Médico Senior, The
Permanente Medical Group,
Inc., Sacramento, California,
EEUU

Conflictos de interés: no se
declaran.

Füsun Çuhadaroğlu-Çetin
MD

Profesor de Psiquiatría del Niño
y el Adolescente, Facultad de
Medicina de la Universidad de
Hacettepe, Ankara, Turquía

Conflictos de interés: no se
declaran.

Agradecimientos: Profesor Eric
Youngstrom, Universidad de
Carolina del Norte, Profesor
Peter Szatmari, Universidad
McMaster, y Dr Eric Daleiden,
PracticeWise LLC, quienes
fueron revisores de gran

Esta publicación está dirigida a profesionales en formación o con práctica en salud mental y no para el público general. Las opiniones vertidas son de responsabilidad de sus autores y no representan necesariamente el punto de vista del Editor o de [IACAPAP](#). Esta publicación busca describir los mejores tratamientos y las prácticas basadas en la evidencia científica disponible en el tiempo en que se escribió, tal como fueron evaluadas por los autores, y éstas pueden cambiar como resultado de nueva investigación. Los lectores deberán aplicar este conocimiento a los pacientes de acuerdo con las directrices y leyes de cada país en el que ejercen profesionalmente. Algunos medicamentos pueden que no estén disponibles en algunos países, por lo que los lectores deberán consultar la información específica del fármaco debido a que ni se mencionan todas las dosis, ni todos los efectos no deseados. Las citas de organizaciones, publicaciones y enlaces de sitios de Internet tienen la finalidad de ilustrar situaciones, o se enlazan como una fuente adicional de información; lo que no significa que los autores, el Editor o IACAPAP avalen su contenido o recomendaciones, que deberán ser analizadas de manera crítica por el lector. Los sitios de Internet, a su vez, también pueden cambiar o dejar de existir.

©IACAPAP 2018. Esta es una publicación de acceso libre bajo criterios de [Licencia Creative Commons Atribución No Comercial](#). El uso, distribución y reproducción a través de cualquier medio están permitidos sin previa autorización siempre que la obra original esté debidamente citada y su uso no sea comercial. Envíe sus comentarios acerca de este libro digital o algún capítulo a jmrey@bigpond.net.au.

Cita sugerida: Hamilton J, Çuhadaroğlu-Çetin F. Práctica Basada en la Evidencia en Salud Mental de la Infancia y la Adolescencia. (Irrázaval M, Martin A, eds. Prieto-Tagle F, Morey J. trad.). En Rey JM (ed), *Manual de Salud Mental Infantil y Adolescente de la IACAPAP*. Ginebra: Asociación Internacional de Psiquiatría del Niño y el Adolescente y Profesiones Afines 2018.

Este capítulo explica cómo usar datos de estudios con menor sesgo y menor error aleatorio para tomar buenas decisiones clínicas para ayudar a niños y adolescentes que buscan tratamiento para sus problemas de salud mental. Específicamente, este capítulo mostrará cómo:

- Pensar que el sesgo y el error aleatorio no son deseables
- Entender los conceptos estadísticos más relevantes para alinear el trabajo clínico con la óptima evidencia científica
- Buscar en bases de datos electrónicas disponibles de forma gratuita para cualquier persona que tenga una conexión a Internet
- Utilizar técnicas adecuadas de entrevista diagnóstica para valorar si el niño o adolescente cumple los criterios diagnósticos para trastornos bien definidos
- Encontrar tratamientos con una mayor probabilidad de ser eficaces para esos trastornos.

En resumen, este capítulo muestra a los profesionales de salud cómo alinear lo que hacen con *lo que funciona*: esto facilitará que una mayor proporción de niños y adolescentes mejoren su funcionamiento más rápidamente y sustancialmente, comparado con decisiones alternativas. Pero antes de descubrir lo que realmente funciona, es necesario superar dos grandes enemigos de la verdad: *el sesgo y el error*.

SESGO Y ERROR ALEATORIO

El problema del sesgo

El sesgo interfiere con la validez — el grado en que un estudio mide lo que pretende medir — al crear desviaciones sistemáticas (no aleatorias) de la verdad subyacente (Guyatt & Rennie, 2002). El sesgo no es lo mismo que el *error aleatorio*. El error aleatorio disminuye a medida que aumenta el tamaño de la muestra, mientras que el sesgo existe en cualquier tamaño de la muestra. Considere las imágenes (a) a (d) en la Figura A.6.1, donde el círculo sombreado en el centro representa el efecto verdadero. En el estudio (a), existe poca o nula desviación sistemática del efecto verdadero (poco sesgo o ausencia de sesgo), así como poca

ayuda al corregir y editar los borradores anteriores. La responsabilidad final del manuscrito es de los autores, incluyendo los errores. Puede contactar al autor enviando un correo a john.hamilton@kp.org

Error

Todos los experimentos o decisiones clínicas están sujetos a un error aleatorio, un error sistemático o a ambos. Los *errores aleatorios* se deben a cambios desconocidos e impredecibles en los instrumentos de medida, o en las condiciones ambientales. Por el contrario, los *errores sistemáticos* (o sesgos), son imprecisiones que están consistentemente en la misma dirección. Pueden ocurrir porque hay algo erróneo en el instrumento de medición (p.ej., en una escala de evaluación, un criterio diagnóstico) o porque el instrumento es usado de manera errónea por el investigador (p.ej., un profesional que sistemáticamente sobrestima – o subestima – el riesgo de suicidio).

Figura A.6.1 Ejemplos de sesgo y error aleatorio en muestras de datos (Kabai, 2011a).

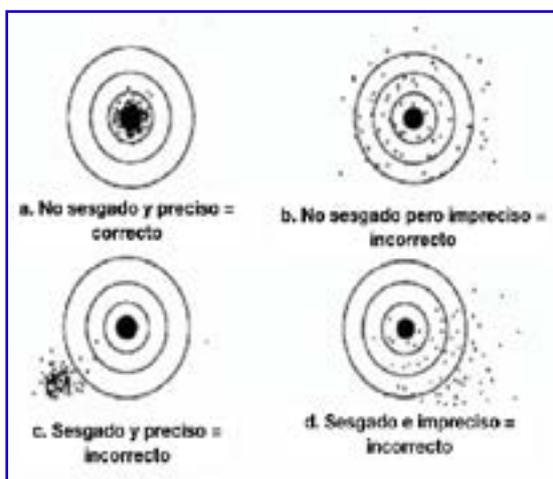
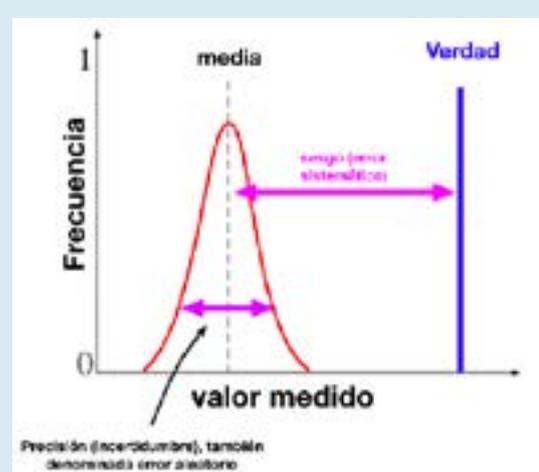


Figura A.6.2 Representación esquemática del sesgo y el error aleatorio (Kabai, 2011b).



dispersión (error aleatorio). En (b), también existe poco sesgo, pero gran error aleatorio, y en (c) existe más sesgo pero poco error aleatorio. En (d) existen gran sesgo y error aleatorio.

Encontrar evidencia con bajo sesgo y bajo error aleatorio

La evidencia más útil para guiar las decisiones clínicas es la evidencia con *un sesgo mínimo y un error aleatorio mínimo*. La Figura A.6.2 presenta una representación esquemática del sesgo y el error aleatorio. La Tabla A.6.1 muestra las fuentes de sesgo más frecuentes y las estrategias para reducir cada uno de ellas, mientras que la Tabla A.6.2 muestra las fuentes más frecuentes de error aleatorio, y las estrategias para reducirlo.

Sesgo
El sesgo es un riesgo siempre presente en la investigación, incluso en los ECAs bien diseñados. Por ejemplo, resultados positivos se publican mucho más frecuentemente en los ensayos financiados por la industria farmacéutica (78%), en comparación con aquellos sin patrocinio de la industria (48%) o financiados por un competidor (28%) (Kelly et al, 2006).

Tabla A.6.1 Fuentes frecuentes de sesgo y estrategias para reducir las.

Fuente de sesgo	Estrategias para reducir las fuentes de sesgo
<ul style="list-style-type: none"> Los pacientes del grupo de tratamiento y control tienen pronósticos distintos 	<ul style="list-style-type: none"> Asignación aleatoria al grupo control o intervención
<ul style="list-style-type: none"> Efecto placebo del tratamiento 	<ul style="list-style-type: none"> Pacientes deben ser ciegos a la condición control o tratamiento activo
<ul style="list-style-type: none"> Los cuidadores añaden otros tratamientos 	<ul style="list-style-type: none"> Cuidadores deben ser ciegos a la condición control o tratamiento
<ul style="list-style-type: none"> Los evaluadores de resultados están sesgados 	<ul style="list-style-type: none"> Evaluadores de los resultados deben ser ciegos a la condición de control o tratamiento de los casos que evalúan
<ul style="list-style-type: none"> Pérdida de seguimiento 	<ul style="list-style-type: none"> Seguimiento de todos o de al menos una gran proporción de los casos que participan en el estudio

Tabla A.6.2 Dos causas frecuentes de error aleatorio y estrategias para reducir las.

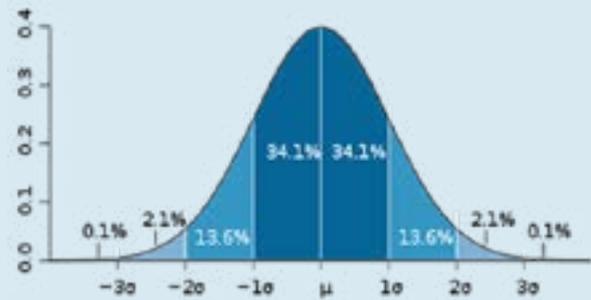
Fuentes de Error Aleatorio	Estrategia para Reducirlo
<ul style="list-style-type: none"> Instrumentos de medición imprecisos 	<ul style="list-style-type: none"> Mejorar la precisión de los instrumentos
<ul style="list-style-type: none"> Tamaño pequeño de la muestra 	<ul style="list-style-type: none"> Aumentar el tamaño de la muestra

CONCEPTOS ÚTILES EN LA PRÁCTICA BASADA EN LA EVIDENCIA

Media

Todo lo que se necesita para entender muchas de las ideas centrales usadas en los enfoques basados en la evidencia es una serie de conceptos básicos de estadística. En primer lugar, recuerde la curva *normal* o gaussiana, que describe la

Figura A.6.3 Gráfico de una distribución normal (gaussiana), con una media representada por la letra griega μ y una desviación estándar, σ , representada por cada banda coloreada. Observe que un poco más de un tercio (34.1%) de la población de esta distribución se encuentra a una desviación estándar sobre la media y un poco más de un tercio (34.1%) también se encuentra a una desviación estándar por debajo de la media. Sólo alrededor de un 2% de la población no está incluido dentro de 2 o más desviaciones estándar por encima de la media, y sólo alrededor de un 2% no está incluido dentro de 2 o más desviaciones estándar por debajo de la media (Wikipedia).



distribución de muchas de las características de la naturaleza y que puede derivarse matemáticamente utilizando la teoría de la probabilidad. El valor más frecuente es el mismo que la *media*, representada por la letra griega μ .

Desviación estándar (DE)

La DE, o desviación típica, es una medida de dispersión: refleja la variación o “cuánto tienden a alejarse” los valores concretos del promedio, “la media”. Una DE pequeña respecto a la media indica una distribución en la que los datos están agrupados estrechamente. Una DE grande respecto a la media indica una distribución en la que los datos tienen un rango amplio y están alejados respecto a la media. A veces la DE se describe como la *cantidad típica en que los casos difieren entre sí*. Se espera que dos de cada tres casos se sitúen a +/- una DE de la media.

Tamaño del efecto

En un estudio, el tamaño del efecto mide la magnitud del efecto de un tratamiento relativo a las diferencias entre los casos en general. Es la diferencia de las medias estandarizadas entre dos grupos en un estudio. Técnicamente, se mide como el coeficiente de la diferencia debida al tratamiento comparado a lo que se denomina la desviación estándar *combinada*. Esta última se refiere a la combinación de los valores del grupo control y experimental para obtener una desviación estándar común.

$$\text{Tamaño del efecto} = \frac{[\text{Media del grupo experimental}] - [\text{Media del grupo control}]}{[\text{Desviación estándar combinada}]}$$

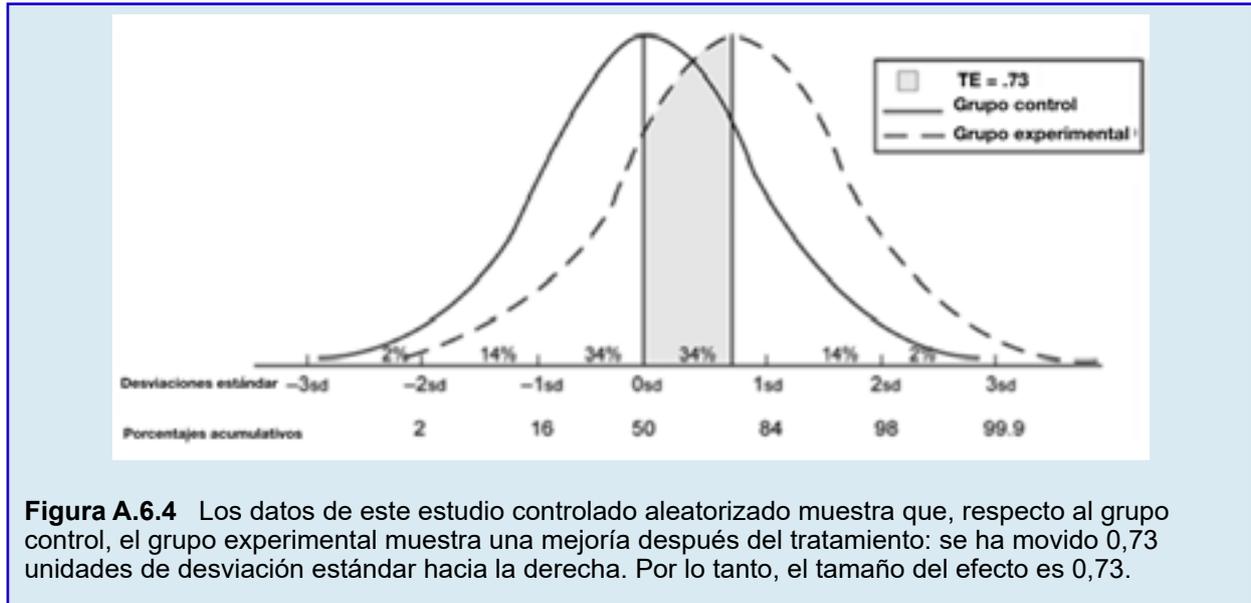
Ejemplo de tamaño del efecto

Un estudio (McCracken et al, 2002) intentó evaluar si un antipsicótico atípico (risperidona) era más eficaz que un placebo en reducir agresividad, rabietas y conductas autolesivas en niños de 5 a 17 años que presentaban un trastorno del espectro del autismo. Después de 8 semanas, las puntuaciones en una escala de irritabilidad fueron de 11,3 (DE 7.4) para el grupo experimental, y de 21,9 (DE 9,5) para el grupo control. ¿Cuál es el tamaño de efecto?

Tamaño de efecto = $(11,3-21,9) / 9,5$, es decir, 1,1. Este es un tamaño de efecto grande y muy significativo. Los tamaños de efecto mayores a 0,5 a menudo son considerados clínicamente significativos, pero los tamaños de efecto menores también pueden tener significancia clínica si el tratamiento puede aplicarse fácilmente a grandes poblaciones (p.ej., el uso de la aspirina para prevenir los infartos cardíacos, donde el tamaño de efecto es pequeño, pero cuando se da en poblaciones grandes, las consecuencias para la salud pública son significativas).

En las páginas web de estadística se puede encontrar cómo calcular la desviación estándar (p.ej., <http://www.leeds.ac.uk/educol/documents/00002182.htm>).

El tamaño de efecto pone el énfasis en el aspecto más importante de una intervención – el tamaño del efecto – más que en su significancia estadística, que se determina tanto por el tamaño del efecto como por el tamaño de la muestra.



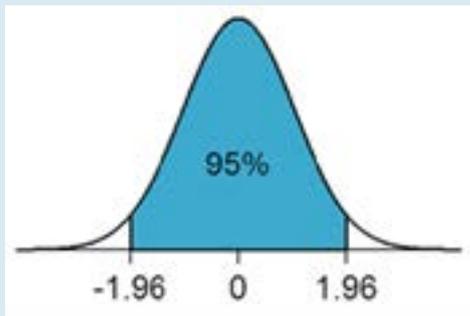
La Figura A.6.4 muestra una representación típica de los datos de un tratamiento experimental en psiquiatría infantil con un magnitud de 0,73. Observe que las dos curvas se superponen considerablemente.

Reducción absoluta del riesgo (RAR)

La reducción absoluta del riesgo es una medida utilizada para comparar dos alternativas diferentes: *¿En qué medida un tratamiento reduce el riesgo de un resultado “negativo” específico (como por ejemplo, que un niño deprimido continúe estando deprimido), en comparación con un tratamiento alternativo o con un placebo?* Por ejemplo, en el Estudio de Tratamiento de Adolescentes con Depresión (TADS, por sus siglas en inglés), a las 12 semanas, el 65,2% de los adolescentes que recibieron un placebo no experimentaron mejoría en comparación con 39,4% de los adolescentes que recibieron fluoxetina. Por lo tanto, el RAR de no mejorar para los adolescentes con un trastorno depresivo mayor que recibieron fluoxetina, comparado con placebo, fue de $65,2 - 39,4 = 25,8\%$. En otras palabras, la fluoxetina reducía considerablemente el riesgo de no mejorar: un 25,8%, o alrededor de 1 de cada 4 casos tratados.

La reducción absoluta del riesgo es útil por dos razones. En primer lugar, como se planteó anteriormente, la RAR puede ser utilizada para calcular el *número necesario a tratar* (NNT). La fórmula es simplemente $NNT = 1/RAR$, donde el valor obtenido se redondea al número entero más cercano. En segundo lugar, la RAR no exagera el efecto del tratamiento, como sí lo hace la *reducción relativa del riesgo* (RRR), que simplemente se calcula dividiendo el riesgo en el grupo de tratamiento por el riesgo en el grupo control. En el TADS, el riesgo relativo al comparar los resultados en los adolescentes que recibieron fluoxetina con los que recibieron placebo, es el riesgo de que los síntomas depresivos valorados como “sin mejoría significativa” en el grupo experimental, que recibe la fluoxetina, dividido por el riesgo de que los adolescentes no mejoren en el grupo que recibe placebo, o $39,4 / 65,2 = 0,60$, generalmente expresado como un 60%. Observar que esta cifra es bastante diferente a la reducción absoluta del riesgo calculada en el párrafo anterior (25,8%). En general, la RAR es un parámetro estadístico más útil que la RRR.

Figura A.6.5 Un aspecto importante asociado al "1,96" en la fórmula para calcular los intervalos de confianza del 95%: el 95% del área de una distribución normal está a 1,96 unidades de desviación estándar de la media.



Intervalo de confianza (IC)

En lugar de comenzar con una definición, puede ser más útil un ejemplo para entender este concepto fundamental. Supongamos que la ministra de salud le dice que está preocupada por los adolescentes deprimidos y le pide que determine el nivel *promedio* (es decir, medio) de síntomas depresivos en niños de 13 años en la ciudad donde usted vive. Usted acepta el desafío y comienza el proyecto haciendo que niños de 13 años de la ciudad completen un cuestionario de síntomas depresivos. La puntuación media es de 39,8 en los 49 cuestionarios completados. Usted calcula la desviación estándar, que es de 9,0. La ministra de salud quiere que usted estime, basándose en los datos de los 49 cuestionarios, lo cerca que está su media del valor *verdadero de depresión en esa población*, también llamado valor *poblacional*. Es decir, si fuera posible evaluar a *todos* los niños de 13 años de esa ciudad con este cuestionario de depresión, el valor que obtendríamos como "valor verdadero" estaría muy cercano al que hemos obtenido. Como puede ver, el valor *verdadero* o *poblacional* es un concepto muy útil, pero en realidad casi siempre es impráctico evaluar a todos los niños. Simplemente por casualidad, usted puede haber hallado una muestra de niños con un nivel de felicidad, ya sea bien por exceso o defecto, diferente del nivel de la población total, a pesar de tomar precauciones para no sesgar la muestra (por ejemplo, seleccionando solamente a estudiantes de una misma escuela, o solamente a estudiantes de una cierta clase).

Usted le informa a la ministra que la puntuación media de su muestra es 39,8, como estimación aproximada. Sin embargo, ella quiere saber más, y le pregunta, "¿Qué aproximada es esa estimación? ¿Cuál es el valor verdadero más alto? ¿Y el más bajo?" A lo que usted responde, "Le puedo dar un rango o intervalo, pero necesito saber con qué frecuencia usted quiere que yo esté en lo cierto. ¿Quiere que adivine correctamente un 80% de las veces en que haga la estimación? ¿Quiere que sea un 95%? ¿O incluso un 99%? Cuanto más alta quiera que sea la probabilidad de que la respuesta que yo le de contenga el valor verdadero para la población, más amplio será el intervalo que tenga que darle." La ministra dice, "Muy bien, quiero que sus respuestas en este tipo de estimaciones sean correctas en el 95% de las ocasiones." La ministra acaba de solicitar el intervalo de confianza del 95% de su media calculada.

Ahora es cuando un estadístico puede serle útil. El estadístico le dice que una buena estimación para calcular los intervalos de confianza en las distribuciones normales es la siguiente, donde N es el número de individuos que generan los datos (49 en este caso) y \sqrt{N} significa raíz cuadrada:

Los intervalos de confianza te pueden convertir en un mago de los datos





¿Cual debe ser la altura del listón requerido por los investigadores para que se considere que un niño o adolescente ha “mejorado” cuando se calcula el NNT?

Los listones más altos aumentarán el valor del NNT debido a que menos niños y adolescentes los superarán.

Intervalo de confianza (IC), límite superior = Media calculada + 1.96 x SD/ SQRT N

Intervalo de confianza, límite inferior= Media calculada - 1.96 x SD/ SQRT N

De esta forma, los intervalos de confianza son:

$$IC_{\text{superior}} = 39.8 + 1.96 \times 9.0 / \text{SQRT } 49 = 39.8 + 1.96 \times 9.0 / 7 = 39.8 + 2.52 = 42.32$$

$$IC_{\text{inferior}} = 39.8 - 1.96 \times 9.0 / \text{SQRT } 49 = 39.8 - 1.96 \times 9.0 / 7 = 39.8 - 2.52 = 38.28$$

Ahora usted tiene una respuesta para la Ministra.

En resumen, los IC se expresan en tres números, el primero es la media, nuestra mejor estimación del valor verdadero. Los otros dos números son evidentemente los números que determinan los límite inferior y superior del intervalo. Cuanto más seguros queramos estar, más amplio será el intervalo de confianza: los intervalos de confianza amplios *garantizan* que el número correcto está allí en alguna parte. *Los intervalos de confianza son importantes porque el análisis estadístico de estudios se basa en la idea central de que, a pesar de hacer observaciones sobre una muestra limitada de sujetos, la inferencia extraída del estudio se podrá aplicar a la toda población.* El propósito principal de un intervalo de confianza es indicar la precisión o imprecisión de la muestra estudiada como una estimación del valor real en la población. Por lo tanto, los intervalos de confianza son útiles, e incluso necesarios, cuando se hacen inferencias a partir de los resultados de un estudio a la población general. En el ejemplo, sólo 49 niños habían completado el cuestionario de depresión, pero la ministra quería generalizar los hallazgos a todos los niños de 13 años en la ciudad. Hay que tener también en cuenta que, incluso si usted y la ministra escogieran su muestra de manera adecuada — p.ej., de escuelas muy diferentes en diferentes partes de la ciudad, para evitar sesgos — aún existiría el problema de la variación aleatoria del nivel de síntomas depresivos. Simplemente, por casualidad, uno puede haber seleccionado un grupo de niños con niveles considerablemente mayores o menores de depresión de los existentes en toda la población de niños de esa edad, incluso si no hubo sesgos sistemáticos significativos en el método de muestreo.

Número necesario a tratar

El número necesario a tratar, a menudo abreviado como NNT, se utiliza para resumir con un único número la eficacia de un tratamiento comparado con el placebo. El NNT se define como *el número de personas que debemos tratar para*

Mann et al (2004) realizaron un metanálisis para estimar la eficacia del acamprosató en la consecución de la abstinencia entre las personas alcohólicas. Ellos observaron que al considerar todos los estudios, un 36,1% de los participantes que tomaban acamprosató habían logrado la abstinencia a los 6 meses, en comparación con el 23,4% de los que recibieron el placebo. ¿Cuál es el NNT?

En este caso, el NNT = $1 / (0,234 - 0,361) = 7,9$.

Es decir, se necesitaría tratar a ocho pacientes (es costumbre redondear el valor al número entero más cercano) con acamprosató para que un paciente adicional se abstenga del alcohol a los 6 meses, en comparación a aquellos que no reciben el tratamiento.

Número necesario para dañar (NND)

El NND equivale a uno dividido por la tasa de daños (efectos secundarios) en los individuos no expuestos menos la tasa en los individuos expuestos.

En el Estudio de Tratamiento de la Depresión en Adolescentes (TADS, por sus siglas en inglés), los autores observaron que un 11,9% de los participantes que recibieron fluoxetina refirieron algún daño, en comparación a un 5,4% de los que recibieron placebo. ¿Cuál es el NND?

En este caso el NND = $1 / (0,119 - 0,054) = 15,4$.

Es decir, es necesario tratar a quince pacientes deprimidos con fluoxetina para que un paciente adicional sufra un efecto secundario atribuible a los medicamentos.

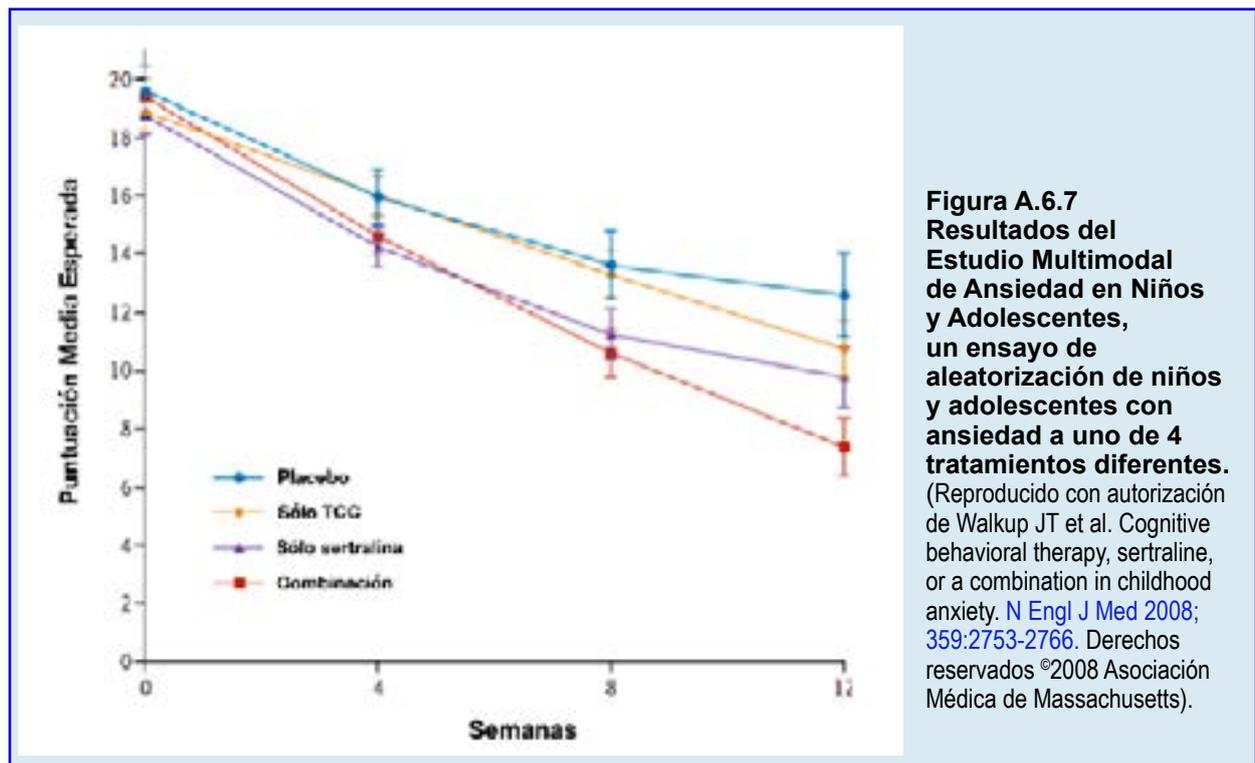
evitar un resultado negativo adicional (número de personas que se necesita tratar para que uno de ellos se beneficie comparado con el grupo control). Por lo tanto, un NNT bajo (por ejemplo 3 o 4) indica que un tratamiento es eficaz, dado que de cada 3 o 4 pacientes tratados uno responderá debido al tratamiento. Los resultados positivos se definen habitualmente por una mejoría en el estado de ánimo o en el funcionamiento, a menudo definido con un punto de corte en una escala.

Hay que tener en cuenta que el NNT depende en gran medida de la *diferencia* que debe producirse (la altura del listón) para considerar que un paciente está “mejor” o “curado.” *La magnitud de esa diferencia afectará en gran manera al NNT.* Por lo tanto, cuando se lee un artículo y se valora un NNT, siempre hay que tener en cuenta qué diferencia han establecido los investigadores (la altura del listón) para juzgar que un paciente mejora o está curado.

Los NNTs altos (por ejemplo, 25, 30 o incluso mayores) sugieren que el tratamiento es poco eficaz, o que el resultado es difícil de alcanzar. Por ejemplo, si nuestra meta es que un adolescente deprimido tenga una ausencia absoluta de síntomas depresivos, ésta será difícil de conseguir, y el NNT resultante será mayor que si hubiéramos elegido un resultado más modesto, por ejemplo, una reducción de síntomas depresivos del 50%.

Número necesario para dañar

El número necesario para dañar se abrevia NND, y representa la frecuencia con que ocurre un efecto secundario con un tratamiento específico. El NND se define como el número de personas que debemos tratar con un tratamiento X, para que una sola persona sufra un efecto secundario que no habría ocurrido si hubiera recibido el placebo. Por ejemplo, en el caso de medicamentos, cuando se presenta un efecto secundario que no se hubiera presentado si el paciente hubiera recibido un placebo. Por lo tanto, un NND alto indica que un tratamiento es seguro: se



necesita que muchas personas reciban el tratamiento para que una persona sea perjudicada por el tratamiento. En consecuencia, la decisión de recomendar un tratamiento necesita considerar los beneficios (NNT) y los riesgos (NND).

ANALIZAR UN ESTUDIO EVALUANDO LA ESTADÍSTICA BÁSICA

Consideremos los resultados de un estudio que evaluó la respuesta al tratamiento de los trastornos de ansiedad en niños y adolescentes utilizando los conceptos estadísticos revisados. Este estudio multicéntrico, denominado Estudio Multimodal de Ansiedad en Niños y Adolescentes (CAMS, por sus siglas en inglés), reclutó a niños y adolescentes de múltiples lugares que presentaban trastornos de ansiedad — trastorno de ansiedad generalizada, fobia social, o trastorno de ansiedad por separación — y los aleatorizó a una de las siguientes cuatro estrategias de tratamiento: TCC, sertralina, una combinación de TCC y sertralina, y placebo.

El gráfico de la Figura A.6.7 muestra las puntuaciones de los 4 grupos de tratamiento en la Escala de Evaluación de la Ansiedad Pediátrica (PARS, por sus siglas en inglés). Puntuaciones de más de 13 indican la presencia de un trastorno de ansiedad. Basándose en estos datos, la combinación de sertralina y TCC es el tratamiento más eficaz. Las *barras de error* en las puntuaciones de la PARS muestran la *dispersión* que existe en los datos. El extremo superior de una barra de error se entiende mejor como la media más el error estándar de la media (calculado como la desviación estándar dividida por la raíz cuadrada del número de participantes), y el extremo inferior de la barra de error es la media menos el error estándar de la media. Existen métodos más complicados para calcular el *error estándar* representado por las barras de error.

En este caso, la magnitud clínica del impacto del tratamiento en los resultados clínicos se evaluó calculando los tamaños de efecto para cada uno de los 3 grupos de intervención en comparación al grupo del placebo: ¿Qué diferentes son cada uno de los 3 grupos de intervención en la puntuación de la PARS en comparación con el grupo del placebo? Basándose en las mejorías observadas en la PARS, el tamaño de efecto a las 12 semanas fue 0,86 para la terapia combinada, 0,45 para la sertralina, y 0,31 para la TCC.

Los autores también calcularon el número necesario a tratar. Recuerde que para calcular el NNT debemos elegir un “límite (altura del listón)” que debe “superar” cada individuo. Cuanto más grande sea la diferencia, más difícil será de superar y, por lo tanto, generará un NNT más alto, mientras que las diferencias menores generarán NNT más bajos. Los autores escogieron que el paciente hubiera “mejorado” o “mejorado mucho” en el PARS como criterio de respuesta al tratamiento a las 12 semanas de estudio. Utilizando este criterio, los autores calcularon que el NNT para la sertralina era de 3,2 y de 2,8 para la terapia cognitivo conductual, y que el NNT para la combinación de ambos tratamientos era de 1,7. Este bajo NNT es alentador, porque demuestra que existen tratamientos eficaces para niños y adolescentes con trastornos de ansiedad.

Los autores también utilizaron la estadística para demostrar que sus resultados no fueron fruto de la casualidad o fortuna, o debido a eventos aleatorios. Para demostrarlo, los autores calcularon los intervalos de confianza para cada tamaño de efecto, y para cada número necesario para tratar. Los intervalos de confianza del

Fiabilidad

La fiabilidad o confiabilidad es la consistencia lograda por un instrumento de medición (p.ej., una prueba diagnóstica o un cuestionario) a través de individuos, el tiempo y los profesionales. La confiabilidad se mide utilizando el estadístico kappa (κ) para las variables categóricas (p.ej., diagnóstico) y el coeficiente de correlación intraclases (CCI) para las variables continuas (p.ej., puntuaciones de una escala). Los valores frecuentes para ambos van de 0 (no mejor que el azar) a 1 (concordancia perfecta). La confiabilidad sobre 0,5 es aceptable, y sobre 0,7 es excelente.

Validez

La validez es el grado en que un sistema de diagnóstico mide lo que pretende medir. La evaluación de la validez a menudo es indirecta, y se basa en la capacidad de una medida de coincidir con otras medidas del mismo concepto, o de predecir la evolución, los resultados, la respuesta al tratamiento, etc. Los coeficientes de correlación, y las cargas de factores son dos formas frecuentes de cuantificar al menos algunos tipos de validez.

Tabla A.6.3 Ejemplos de instrumentos de medición confiables* (ver también el Capítulo A.5)

Trastorno	Instrumento
Trastornos de ansiedad, incluyendo el trastorno de ansiedad generalizada y el trastorno de ansiedad por separación	<ul style="list-style-type: none"> • Cuestionario para Trastornos Emocionales Infantiles Relacionados con la Ansiedad (SCARED, por sus siglas en inglés) • Ver también los Capítulos F.1 y F.2
Trastornos depresivos, incluyendo el trastorno depresivo mayor y la distimia	<ul style="list-style-type: none"> • Escala de Evaluación de Depresión de Hamilton (HAMD-21, por sus siglas en inglés) • Inventario de Depresión para Niños de Johns Hopkins (HDCL-C, por sus siglas en inglés) • Ver también el Capítulo E.1
Trastorno por déficit de atención con hiperactividad	<ul style="list-style-type: none"> • Escala Swanson para el TDAH (SWAN ADHD, por sus siglas en inglés) • Escala de Evaluación SNAP-IV para Educadores y Padres • Escala de Evaluación Diagnóstica de TDAH para Padres de Vanderbilt • Ver también el Capítulo D.1
Trastorno obsesivo-compulsivo	<ul style="list-style-type: none"> • Escala de Síntomas Obsesivos-Compulsivos en Niños de Yale-Brown (CY-BOCS, por sus siglas en inglés) • Ver también el Capítulo F.3

*Estos instrumentos fueron desarrollados y evaluados en Europa y Estados Unidos. En países alejados de estas áreas geográficas, estos trastornos pueden ser sustancialmente diferentes. Además, recuerde que los resultados de un solo instrumento son sólo un aspecto para llegar a una conclusión, o incluso una hipótesis preliminar, sobre el diagnóstico; también son de gran utilidad otras fuentes de información, como la entrevista al niño o adolescente, y a sus padres, así como también los antecedentes y registros médicos y psiquiátricos.

95% para el tamaño de efecto calculado de 0,86 fue de 0.56 a 1.15 para los niños y adolescentes que recibieron la combinación de sertralina y TCC, en comparación con los que recibieron el placebo, que es un efecto robusto en psiquiatría infantil. Según la convención de Cohen (1988), los tamaños de efecto de 0,8 o superiores evidencian un efecto grande, y los de 0,3 y menores indican un tamaño de efecto pequeño.

FIABILIDAD Y VALIDEZ DE LOS MÉTODOS DIAGNÓSTICOS

LOS CIMIENTOS DE LA PRÁCTICA BASADA EN LA EVIDENCIA

Es importante alinear los métodos diagnósticos utilizados en la práctica clínica con los métodos de diagnóstico usados por los investigadores, ya que los niños y adolescentes que diagnosticados con un trastorno depresivo mayor en su práctica clínica, tenderán a mostrar la misma constelación de síntomas depresivos que los niños y adolescentes diagnosticados por los investigadores cuando evaluaban la eficacia de tratamientos para el trastorno depresivo mayor. De esta forma, usted podrá implementar en su práctica clínica con seguridad las terapias que han mostrado resultados positivos.

¿Qué fiables son los procedimientos diagnósticos que usted utiliza con los niños y adolescentes? Fiable significa que el uso repetido de una evaluación da resultados consistentes. Por ejemplo, si una entrevista diagnóstica estructurada y bien definida es readministrada después de un corto período de tiempo y produce resultados idénticos o muy similares, sugiere que la evaluación es fiable (fiabilidad test-retest). La fiabilidad significa que también se producen los mismos resultados o resultados muy similares cuando el instrumento es utilizado por diferentes observadores con el mismo niño o adolescente (fiabilidad interevaluador)

¿Son los sistemas diagnósticos que se utilizan para identificar trastornos en niños y adolescentes fiables? Un sistema diagnóstico válido es aquel que realmente mide lo que pretende medir. Por ejemplo, un sistema diagnóstico válido para evaluar la gravedad de la depresión es aquel que realmente mide la gravedad de la depresión. La evidencia que apoya la validez de un sistema diagnóstico para valorar la gravedad de la depresión en adolescentes incluye datos que muestran que los cuestionarios autoadministrados están altamente correlacionados con los informes de los entrevistadores formados que entrevistaron a los adolescentes utilizando el mismo sistema diagnóstico. La concordancia entre entrevistas diagnósticas estructuradas, que a menudo se utilizan en los estudios de investigación, y las evaluaciones clínicas varía según el trastorno, siendo entre baja y moderada para la mayoría de los trastornos (Rettew, 2008).

En resumen, en la práctica clínica, es fundamental tener un sistema diagnóstico fiable y válido. Desafortunadamente, las entrevistas clínicas no estructuradas no han demostrado ser fiables ni válidas para el diagnóstico de trastornos psiquiátricos en los niños y adolescentes. Las evaluaciones diagnósticas sin una clara estructura tienden a sacar conclusiones apresuradas, a menudo no detectan comorbilidades importantes y generan diagnósticos poco fiables. Este es un problema importante que, afortunadamente, puede resolverse.

Una buena forma de hacer un cribado para el trauma

A continuación se presentan algunos ejemplos respecto a cómo la entrevista semiestructurada K-SADS indaga sobre el trauma, pero sólo después de un período de “calentamiento” durante el que se recopila información básica y se facilita las respuestas. Tenga en cuenta que el uso de la K-SADS requiere entrenamiento, y la siguiente lista incluye sólo algunas de las preguntas sobre el trauma (Kaufman et al, 1997)

1. Eventos traumáticos

Guión: Voy a preguntarte una serie de cosas malas que a menudo les suceden a los niños de tu edad y quiero que me digas si alguna de estas cosas te han pasado alguna vez. Dime si alguna de estas cosas te ha sucedido, incluso si sólo ocurrió una vez.

- a. Accidente de coche: ¿Has estado alguna vez en un accidente de coche grave? ¿Qué pasó? ¿Te lastimaste? ¿Se lastimó alguien más?
- b. Otro accidente: ¿Has estado alguna vez en otro tipo de accidente grave? Por ejemplo, en un accidente en bicicleta, otros accidentes. ¿Qué pasó? ¿Te lastimaste?
- c. Incendio: ¿Estuviste alguna vez en un incendio? ¿Se ha incendiado tu casa o tu escuela alguna vez? ¿Alguna vez has encendido un fuego que se haya salido de control? ¿Qué pasó? ¿Se lastimó alguien gravemente? ¿Hubo mucho daño?

La K-SADS continúa en este punto con preguntas respecto a presenciar un evento traumático como una tormenta o un delito violento, violencia doméstica, o abuso físico o sexual.

Instrumentos diagnósticos fiables y válidos

Crear un sistema diagnóstico fiable y válido es una tarea difícil. Lo más útil es utilizar instrumentos diagnósticos específicos para distintas clases de trastornos. Existen instrumentos diagnósticos disponibles de forma gratuita en la Internet que ayudan a establecer el diagnóstico y su gravedad y, principalmente, sirven para establecer una línea de base que permitirá medir cambios en la sintomatología a lo largo del tratamiento. En la Tabla A.6.1 hay ejemplos de instrumentos diagnósticos para algunos de los trastornos más frecuentes.

Un ejemplo de como utilizar instrumentos diagnósticos para evaluar a un adolescente

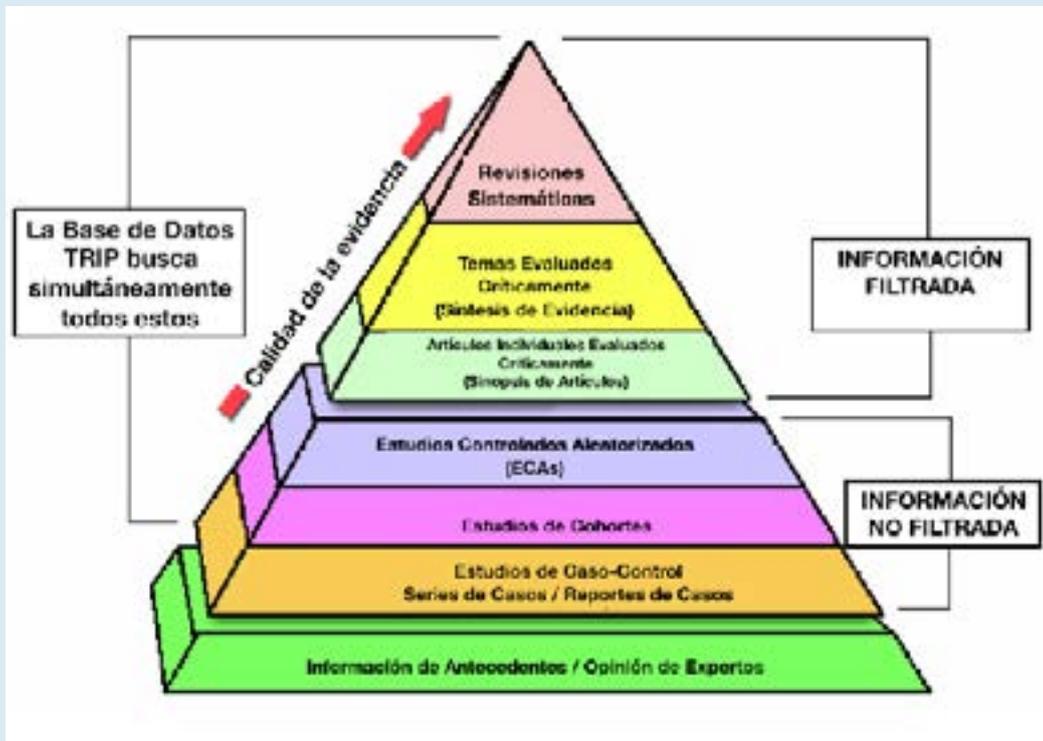
Es muy importante "traducir" correctamente los síntomas que describen los niños y adolescentes, sus padres y educadores, en un trastorno psiquiátrico, o al menos, en parte de un trastorno (*cuadro subclínico*). Para ello, hay que traducir las descripciones de lo que les está pasando en un trastorno conocido, evitando las diferentes fuentes de error en el proceso. Por ejemplo, consideremos a Sam, un niño de 14 años de edad, cuyos padres se quejan de que es muy irritable y tiene problemas para dormir. Su rendimiento académico ha disminuido y se ha vuelto más retraído. Esta constelación de síntomas puede encontrarse en más de un trastorno. Adolescentes pueden volverse irritables cuando están ansiosos o negativistas, pero irritabilidad también sugiere la posibilidad de un trastorno del estado de ánimo; de la misma forma, el retraimiento social puede tener múltiples causas. Afortunadamente, existen herramientas de evaluación para un amplio espectro de trastornos, fácilmente accesibles y de bajo coste, que son de gran ayuda para realizar una buena evaluación (ver Tabla A.6.1).

Si bien estos instrumentos nos ayudan a focalizar y definir con mayor precisión los síntomas en los dominios diagnósticos sugeridos por la historia clínica, también necesitamos una forma de valorar a Sam de una manera más general y cribar para otros trastornos menos obvios inicialmente, como haber sido

Las pirámides son una buena forma para representar los diferentes tipos de evidencia, ya que son amplias en la base, y estrechas en el extremo superior. Nuestra meta es llegar a la punta, donde la vista es óptima, minimizando los sesgos y el error aleatorio.



Figura A.6.8 Una pirámide de la evidencia, donde una mayor altura en la pirámide significa menos sesgo y menos error aleatorio.



Fuente: Universidad de California, San Diego, sistema de bibliotecas y su base de datos [TRIP database](#), la cual usa "filtros" (programas de software) para seleccionar la evidencia publicada más libre de sesgo y error aleatorio.

víctima de un trauma o abuso, consumo de sustancias, o presentar un trastorno del comportamiento que ha olvidado mencionar.

Para ello pueden ser útiles las preguntas de una evaluación integral de psicopatología ampliamente conocida. La Kiddie-SADS (K-SADS, por sus siglas en inglés) es una entrevista diagnóstica semiestructurada ampliamente utilizada, puesta generosamente a disposición de los usuarios que no la usen con finalidad de lucro, y para la investigación ética (Kaufman et al, 1997). La K-SADS no es una entrevista estructurada — en la que el entrevistador repite exactamente las preguntas escritas — sino semiestructurada, en la que el entrevistador puede elegir entre una serie de preguntas posibles para indagar en el mismo dominio sintomatológico. Los entrevistadores *no* tienen que hacer todas las preguntas sino que pueden parar cuando creen tener suficiente información. A continuación se presentan algunas preguntas de cribado de la K-SADS, diseñadas para indagar el posible consumo de alcohol:

- *¿Qué edad tenías cuando tomaste tu primera bebida alcohólica?*
- *¿Cuál es tu bebida alcohólica favorita?*
- *¿Tienes un grupo de amigos con los que habitualmente tomas alcohol, o generalmente tomas alcohol a solas?*
- *¿Dónde tomas alcohol generalmente? ¿En casa? ¿En fiestas? ¿En casa de un amigo? ¿En la calle? ¿En bares?*
- *¿Hay momentos especiales en los que es más probable que tomes alcohol que en otros? ¿En las fiestas escolares o en otras fiestas?*

- ¿Qué edad tenías cuando comenzaste a beber alcohol de forma regular, es decir, dos o más bebidas alcohólicas a la semana?
- ¿En los últimos seis meses, ha habido alguna semana en la que has consumido al menos dos bebidas alcohólicas?

Los autores de la K-SADS sugieren un umbral definido para seguir evaluando al niño o adolescente: si el adolescente o los padres refieren que el adolescente ha tomado 2 o más bebidas alcohólicas a la semana durante al menos 4 semanas. Los umbrales explícitos de la K-SADS para “contabilizar” un síntoma como presente o ausente son de gran utilidad, al igual que las preguntas que sugiere, que son claras y directas, pero ofrecen al entrevistador distintas posibilidades de obtener la información. La especificidad de las preguntas de la K-SADS, sus umbrales claros y la presencia de preguntas de cribado son útiles para los profesionales que tratan de mejorar la fiabilidad y validez de sus entrevistas clínicas. El uso de toda la entrevista requiere formación específica y puede ser poco práctica en muchos contextos clínicos, pero sus elementos, preguntas sugeridas y umbrales son útiles.

Una vez que se ha completado una buena — fiable y válida — evaluación diagnóstica del adolescente, la siguiente tarea es elegir el tratamiento.

ELECCIÓN DE TRATAMIENTO

El concepto de la pirámide de evidencia

La *pirámide de evidencia* es una forma práctica de resumir la calidad de la evidencia — es decir, la aproximación más cercana a la que es cierto. Las fuentes de evidencia con menor sesgo y menor error aleatorio se ubican hacia la cima de la pirámide. La pirámide representa la transición de los cientos o incluso miles de fuentes de “evidencia de nivel inferior”— los casos aislados, estudios no controlados e incluso las opiniones de expertos, que son ubicadas correctamente cerca de la base — a las relativamente raras revisiones sistemáticas de estudios aleatorizados controlados, y los metanálisis, que sintetizan toda la evidencia de mejor calidad y que se ubican correctamente hacia la cima de la pirámide. Debido a que hay una multitud de casos publicados y muchos estudios que probablemente contienen sesgos y errores aleatorios, y relativamente pocos que sintetizan cuidadosamente toda la evidencia disponible, la pirámide ofrece una buena representación de este concepto (Figura A.6.8).

Información básica

En la base de la pirámide esta la información básica sobre el tema a valorar (por ejemplo, la neurofisiología de la depresión o del espectro del autismo), y la opinión de expertos. Esto proporcionan evidencia de baja calidad. Esta información básica crea el escenario de lo que puede ocurrir: muchos dramas pueden representarse en el mismo escenario. Por ejemplo, si los estudios de imagen muestran cambios en el flujo sanguíneo en las áreas cortico-frontales en la depresión, esta información probablemente no nos ayudará a elegir el tratamiento más eficaz para la depresión.

Opinión de expertos

Del mismo modo, la *opinión de los expertos* puede ser errónea. El experto puede ser empleado por una compañía farmacéutica, puede sentirse presionado y, consciente o inconscientemente, puede favorecer un medicamento, minusvalorar

CONSORT significa Estándares Consolidados para el Informe de Ensayos (en inglés)

[Haga clic aquí](#) para acceder a la declaración CONSORT: información que debe incluirse cuando se describen los resultados de un ensayo aleatorizado.

resultados negativos, o sobrevalorar resultados positivos, sesgando los resultados, o simplemente el experto puede ser leal a un determinado tratamiento porque le es familiar. La idea central es que la opinión de un individuo, no respaldada por una recopilación de datos sistemática y con metodologías reconocidas, representa evidencia de baja calidad. Expresado en humor científico, "De Dios nos fiamos. Todos los demás, traigan datos".

Series de casos

Las series de casos y los casos individuales también se ubican cerca de la base de la pirámide. Un profesional que evalúa un solo caso, administra el tratamiento y luego evalúa los resultados, es muy vulnerable a los sesgos. Las series de casos producen datos débiles, porque no tienen un grupo de comparación, aunque esos datos pueden ser útiles para que los investigadores generen hipótesis para futuros estudios mejor diseñados.

Estudios de casos y controles

Los estudios de casos y controles son útiles para establecer una asociación entre la exposición a un factor de riesgo y un resultado. En los estudios de casos y controles, el resultado es conocido, por ejemplo, suicidio en adolescentes. El investigador crea entonces un grupo control de individuos sin el resultado (que no se suicidaron), pero con características similares a las de los individuos que se suicidaron, y determina la tasa de exposición a un factor de riesgo en los controles, en comparación con los que se suicidaron. Utilizando esta metodología, por ejemplo, se ha demostrado que la accesibilidad y disponibilidad de un arma de fuego en el hogar esta asociada a un mayor riesgo de suicidio en adolescentes (Brent et al, 1993). Esta es un dato útil cuando se aconseja a los padres de un adolescente potencialmente suicida.

Estudios de cohorte

Los estudios de cohortes son útiles para demostrar prospectivamente una asociación entre la exposición a un factor de riesgo específico y un resultado. Aquí el muestreo es por exposición a un factor de riesgo, no por resultado. Por ejemplo, se sabe que niños expuestos a ciertos tipos de abuso tienen un mayor riesgo de desarrollar depresión. Algunos estudios de cohortes utilizan bases de datos muy grandes, incluyendo incluso a todos los niños en un país y, por lo tanto, tienen gran poder estadístico (una capacidad de detectar incluso un efecto pequeño).

Estudios aleatorizados controlados

Los estudios controlados aleatorizados (ECAs) individuales son los principales bloques de construcción de los niveles más altos de la pirámide. Los ECAs tienen ventajas únicas como resultado de usar un grupo control elegido al azar y de un cegamiento deliberado de los sujetos, de los profesionales que administran el tratamiento y de los asesores que asignan los participantes a la intervención activa o al grupo control.

La cima de la pirámide están diseñada para eliminar las limitaciones de los ECA, como eliminar los datos de los ECAs llevados a cabo deficientemente, al mismo tiempo que combinan los datos de múltiples ECAs realizados correctamente, para generar un número mayor de sujetos estudiados, minimizando el riesgo de error aleatorio. Se tiene que tener cuidado de no ignorar los ECAs no publicados.



Este es el logo de PubMed, una fuente gratuita de evidencia. Posee filtros para encontrar revisiones sistemáticas, útiles como resúmenes de evidencia de alta calidad.

Evaluación crítica

La valoración crítica de artículos individuales elimina los datos de estudios con errores significativos. Por ejemplo, un estudio no cegó adecuadamente a los evaluadores, o perdió una gran proporción de la muestra en el seguimiento. Si es así, los resultados pueden contener un sesgo importante, y el estudio deben ser excluido. Los resultados de artículos sobre el mismo tema evaluados críticamente e identificados como libres de fuentes evidentes de sesgo, pueden ser agrupados para aumentar el tamaño de la muestra, con el fin de reducir el error aleatorio. Considere la posibilidad de utilizar listas disponibles de criterios para evaluar un estudio, como la lista de Normas Consolidadas para la Publicación de Ensayos Clínicos (CONSORT, por sus siglas en inglés).

El diagrama de la Figura A.6.9 (página 20) presenta algunos pasos críticos en la ejecución de un ECA, que se deben tener en cuenta al leer un artículo. Por ejemplo, en ocasiones un autor puede describir un estudio como un “ensayo aleatorizado”, a pesar de que queda claro en la sección de métodos que la asignación no fue realmente aleatoria. En la realización de un ECA existen muchas fuentes potenciales de sesgo, pero la mayoría de los profesionales simplemente no tienen la voluntad, el tiempo o la experiencia estadística y de investigación, para leer un ECA publicado. La evaluación crítica implica una observación crítica de cada etapa de un estudio, para garantizar que no se transmita ningún sesgo a los resultados finales. Asistir a un curso en medicina basada en evidencia y participar en un club de revistas para personas interesadas puede ayudar a mejorar la habilidad de realizar evaluaciones críticas.

Síntesis de evidencia y revisiones sistemáticas

Las síntesis de evidencia combinan los datos de EAC individuales que han sido evaluados críticamente por posibles fuentes de sesgo. La agrupación de datos tiende a disminuir el error aleatorio, generando resultados más precisos.

Por último, las revisiones sistemáticas constituyen un método especialmente organizado de síntesis de evidencia, basado en estándares acordados de síntesis. Con esta metodología, de la cual tal vez el mejor ejemplo es la Biblioteca Electrónica Cochrane de Revisiones Sistemáticas, el método utilizado para combinar los datos y establecer conclusiones está bien definido antes de comenzar la revisión. Esta metodología también es transparente, ya que el lector puede revisar cómo se obtuvieron los resultados. Un metanálisis es un tipo de revisión sistemática, que considera los datos de diversos estudios, para llegar a una estimación única del tamaño de efecto o del NNT.

UTILIZANDO LA INTERNET Y LA PIRÁMIDE PARA ENCONTRAR TRATAMIENTOS EFICACES

Contestando preguntas relevantes en su práctica clínica utilizando la Internet

Es importante ser curioso cuando uno realiza trabajo clínico: hacer preguntas relevantes al cuidado de los pacientes que uno ve ese día; es decir, hacer preguntas relevantes al caso específico y tratar de contestarlas utilizando la Internet para revisar la literatura publicada, centrándose en la evidencia libre de sesgo y error aleatorio.

Búsqueda en PsycINFO

PsycINFO es una excelente base de datos para la búsqueda de conceptos como "bullying" (acoso escolar) o "identity" (identidad), ya que tiene un tesoro de términos indexados.

Por ejemplo, al buscar el término "bullying" (acoso escolar), se despliegan los siguientes términos indexados en una columna a la izquierda de los resúmenes, además del número de éxitos para cada término indexado.

Términos Indexados:

- Bullying (acoso escolar) (2,808)
- Victimization (victimización)(1,175)
- Aggressive behavior (conducta agresiva) (845)
- Peer relations (relaciones con pares) (543)
- Schools (escuelas) (437)
- 5 más...

Al hacer clic en el hipertexto de la página, se accede a aquellos resúmenes que están disponibles sin coste. Use la pestaña "Advanced Search" (búsqueda avanzada) para buscar "bullying" (acoso escolar) y seleccione también "Only Show Content Where..." (sólo mostrar contenido donde...) para incluir la opción "Treatment Outcome/Randomized Controlled Trial" (resultado de tratamiento/Ensayo controlado aleatorizado) del menú desplegable.

Esto reduce los 2808 éxitos a sólo 10 (28 de Octubre, 2011).

El primer paso es hacerse una pregunta. Para esto, comience la pregunta con "¿Cuál es la evidencia de...?" Esto garantiza que la pregunta pueda ser contestada: Uno puede encontrar que no hay evidencia disponible, en cuyo caso la respuesta será "ninguna", o encontrar al menos algo de evidencia.

El objetivo de una búsqueda bibliográfica es obtener un resultado práctico, lo suficientemente bueno para informar sobre la atención clínica a un paciente sin pretender ser definitivo, dentro del tiempo y los recursos disponibles. La tecnología actual permite buscar resultados válidos y relevantes en las bases de datos rápidamente. Es fácil realizar una búsqueda sensible y específica en [PubMed](#), la página web de la Biblioteca Nacional de Medicina de los Estados Unidos. PubMed esta financiada por el gobierno de los Estados Unidos, no tiene costo para los usuarios, y no es necesario suscribirse. [PsycINFO](#), la página de Internet para estudios psicológicos, esta financiada por la Asociación Americana de Psicología, y ofrece *resúmenes* sin costo para los usuarios. CENTRAL, la base de datos de estudios aleatorizados de la Biblioteca Cochrane, tiene su base central en el Reino Unido. [CENTRAL](#) es la abreviatura utilizada para denominar el Registro Central de Cochrane de Ensayos Clínicos Controlados; y contiene los resúmenes de más de 700.000 ensayos controlados de PubMed y [EMBASE](#), una base de datos especialmente sólida en revistas europeas. EMBASE es disponible solamente para suscriptores y es financiada por Elsevier, una corporación con base en Ámsterdam. PubMed es un buen punto de partida para preguntas relacionadas con medicamentos, e incluye estudios que no son ECAs. PsycINFO es una buena alternativa para preguntas sobre temas psicológicos. CENTRAL es una buena base para buscar exclusivamente resultados provenientes de ECAs.

Terry: búsqueda de una solución para un niño con un TDAH de predominio inatento en 10 minutos

Un psiquiatra infantil está tratando de ayudar a un niño de 10 años con un TDAH de tipo inatento. En base a las puntuaciones en la SWAN (ver Tabla A.6.3) completada por su educadora, requirió 15 mg de metilfenidato tres veces al día para lograr una mejoría sustancial (una calificación de los educadores menor a 9 en las primeras nueve preguntas de la SWAN, los elementos inatentos). Sin embargo, perdió tanto peso en este régimen que tuvo que ser suspendido. Sus padres están cada vez más frustrados y buscan otra solución: "¿Hay otra alternativa?" El psiquiatra infantil quiere encontrar un fármaco distinto a los estimulantes o la atomoxetina, ya que ambos suprimen el apetito. Nuestro médico recuerda lo útil que ha sido la guanfacina en el tratamiento de niños que presentan un TDAH y un TND con síntomas hiperactivos/impulsivos en el pasado. ¿Pero, es eficaz la guanfacina en el TDAH de predominio inatento? ¿Qué evidencia existe respecto a la eficacia de la guanfacina en niños de 10 años con un TDAH de tipo inatento? En caso de que sea eficaz, ¿Cuál es el tamaño de su efecto?

1. Para comenzar en un lugar razonable, vaya a [PubMed](#). Haga clic en "Advanced"(avanzado), seleccione la opción "MeSH term" (Término MeSH) en el menú desplegable (MeSH significa Encabezados de Temas Médicos), escriba "attention deficit hyperactivity disorder" (trastorno por déficit de atención con hiperactividad) en el cuadro de búsqueda y haga clic en "Search" (buscar). El resultado debiera verse como la figura a continuación.

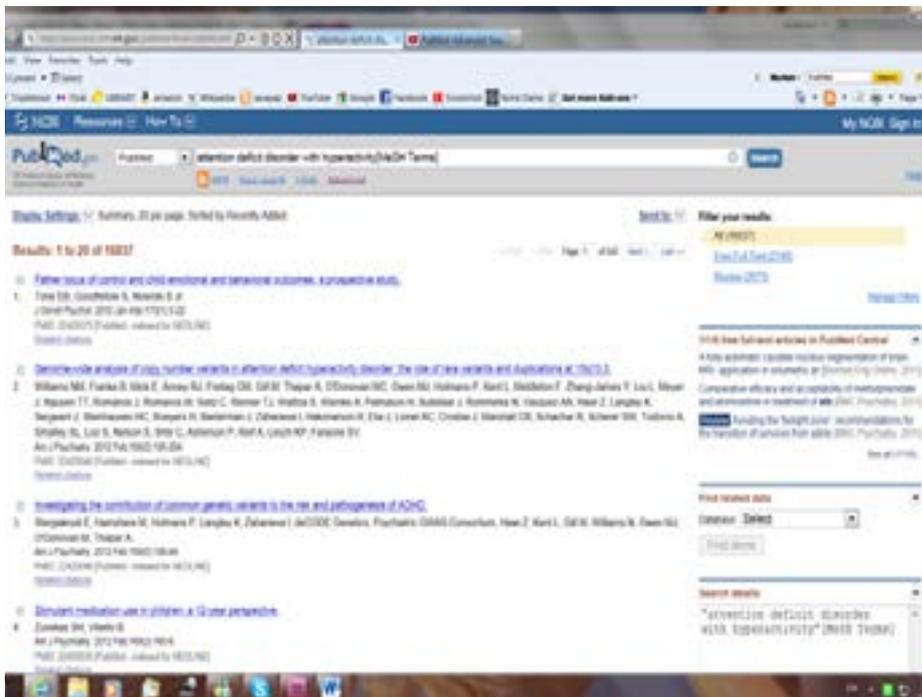
Como ejercicio, intente encontrar todos los ensayos aleatorizados que estudien el efecto del entrenamiento en relajación en adolescentes.



Haga clic en la imagen para ver un video tutorial sobre cómo hacer una búsqueda en PubMed (2:27)

Hay demasiados “éxitos” para poder leerlos todos — ¡Más de 16.000! Necesitamos limitar los resultados a un número más manejable, eligiendo sólo los resultados de la parte superior de la pirámide, para acceder a la *mejor* evidencia. A continuación explicamos cómo hacerlo.

2. Comience en la parte superior de la pirámide: busque todos los *metanálisis* publicados. Vuelva a la página de búsqueda y, después de repetir el paso anterior, escriba “guanfacine” (guanfacina) en el cuadro de búsqueda y seleccione “AND” (y) y “All Fields” (todos los campos), y luego haga clic en “Search” (buscar). Obtendrá 94 “éxitos”. Ahora haga clic en “Limits” (límites) y seleccione “Meta-Analysis” (metanálisis) en el nuevo menú, y haga clic en “Search” (buscar); obtendrá 2 resultados, pero este niño no pertenece a ninguna de las poblaciones estudiadas en estos artículos: uno



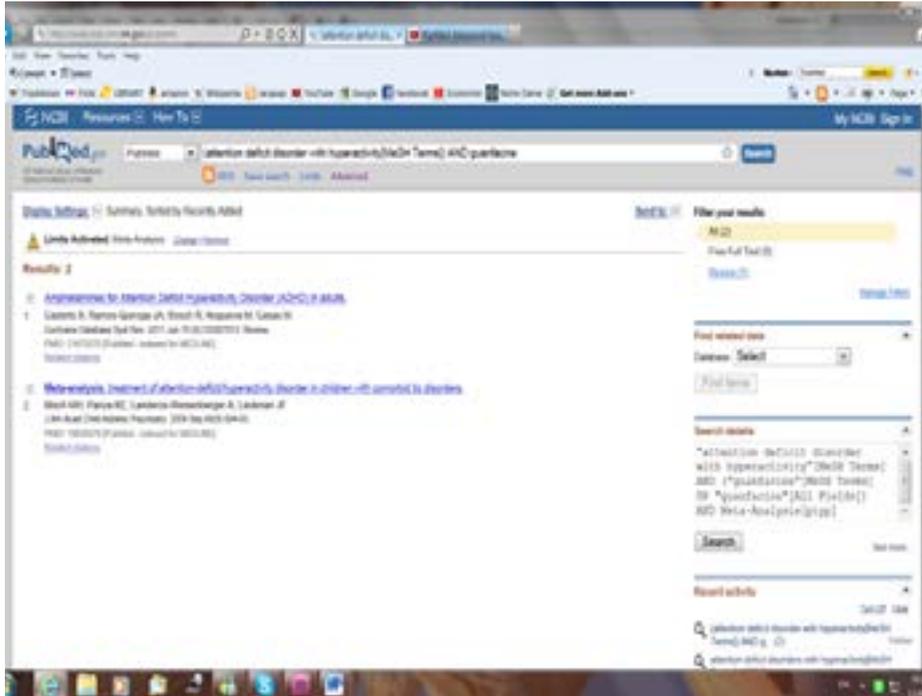
fue realizado en adultos, y el otro en niños con TDAH y tics, y nuestro paciente tiene un TDAH sin tics (ver la Figura). Hemos sido demasiado selectivos y ahora necesitamos ampliar nuestra búsqueda.

3. Baje un nivel más abajo de la pirámide, a los ECAs, cambiando en el cuadro de "Limits" (límites), "Meta-analysis"(metanálisis) por "Randomized controlled trials" (ensayos controlados aleatorizados), y aparecerán 9 artículos (ver Figura).

Los artículos 4 y 5 son los únicos que tienen una relación directa con un niño de 10 años, y sólo el artículo 5 está disponible a texto completo sin costo (Biederman et al, 2008). Este artículo fue financiado por las compañías farmacéuticas, una fuente potencial de sesgo, y se refiere a la guanfacina de liberación prolongada, que tiene un alto coste y no está disponible para este niño. Sin embargo, mediante una búsqueda rápida en Internet, encuentra que la vida media de la guanfacina en niños y adolescentes es de 13 a 14 horas, lo que sugiere que es posible imitar la preparación de liberación prolongada con una preparación asequible — la

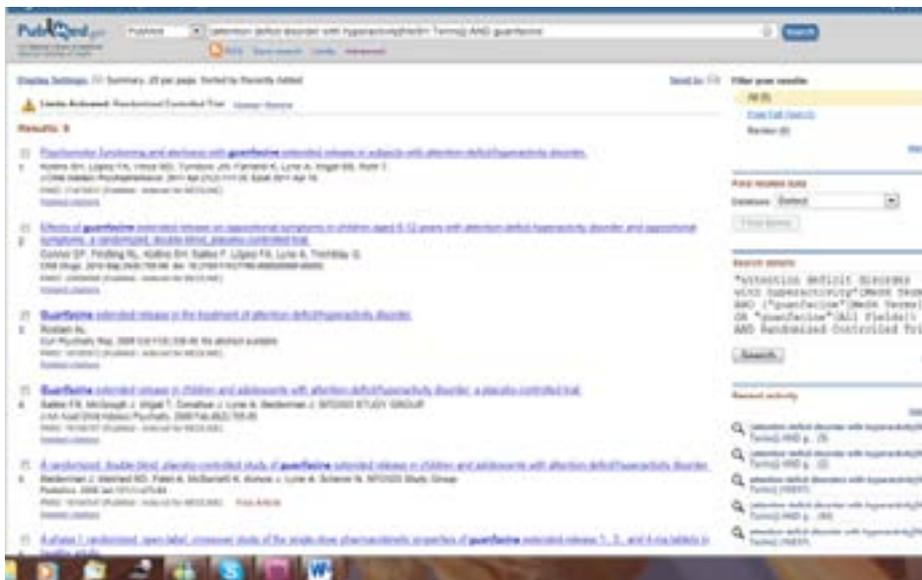
Búsqueda en la Base de Datos CENTRAL de la Biblioteca Cochrane

La búsqueda en CENTRAL de "bullying" (acoso escolar) resulta en 36 éxitos, todos ellos ECAs que investigan el acoso escolar. Por ejemplo, uno de ellos realizado en Australia en 2009 muestra resultados prometedores en adolescentes varones, que no apareció en la búsqueda en PsycINFO de ECAs de "bullying." (Berry K and Hunt CJ (2009), "Evaluation of an intervention program for anxious adolescent boys bullied at school". Journal of Adolescent Health 45:376-382). CENTRAL es útil para revisar ECAs de todo el mundo, y los resúmenes son disponibles de forma gratuita para todos.



guanfacina tradicional — tomándola al menos dos veces al día (Strange, 2008).

¿Cuál fue el tamaño del efecto de la guanfacina sobre los síntomas de inatención en niños con un subtipo combinado de TDAH? Los autores afirman "Los tamaños de efecto en los grupos de tratamiento GXR calculados *post hoc*



fueron de 0,58, 1,19, y 1,34 para los grupos de 0,05 a 0,08 mg/kg, 0,09 a 0,12 mg/kg, y 0,13 a 0 a 17 mg/kg, respectivamente." Sin embargo, estos tamaños de efecto fueron calculados para cambios en las puntuaciones totales de TDAH (síntomas inatentos e hiperactivos/impulsivos). Afortunadamente, los autores clarifican "Todos los grupos de niños que tomaron GXR mostraron mejoras significativas en las escalas de hiperactividad/impulsividad e inatención de la ADHD-RS-IV [una escala de evaluación], en comparación a los que tomaron el placebo. Los cambios



Como mantenerse al día con la medicina basada en la evidencia

Más de 5000 artículos biomédicos y los resultados de alrededor de 100 ensayos de tratamiento se publican en todo el mundo cada día, y el número sigue creciendo. Mantenerse al día con esta marea de información es desalentador. Para facilitar esta tarea, uno puede suscribirse sin costo a la [página web de Evidence Alerts \(Actualizaciones de la Evidencia\)](#) desarrollada por el Grupo BMJ y la Unidad de Investigación de Información en Salud de la Universidad de McMaster. Estas instituciones colaboran entre sí para proporcionar a los profesionales acceso a la mejor evidencia actual de la investigación, adaptada a los intereses del propio profesional, para apoyar decisiones clínicas basadas en la evidencia. Este servicio analiza más de 120 revistas médicas de gran calidad y más de 50.000 artículos cada año, calificándolos según su validez y relevancia, y envía regularmente a los suscriptores un correo electrónico con los artículos clave de su especialidad.

medios respecto a la línea de base en la hiperactividad/impulsividad en los grupos placebo y GXR de 2 mg, 3 mg, y 4 mg fueron -4,06, -6,94, -7,09, y -9,46, respectivamente. *Los cambios medios respecto a la línea de base en la inatención fueron -4,78, -8,46, -8,71, y -9,51, respectivamente* [se añadió la cursiva]." Si bien no se presenta ningún tamaño de efecto en cuanto a los cambios en la inatención por sí sola, estos datos sugieren que el tamaño de efecto será al menos igual que el tamaño de efecto para los síntomas combinados de inatención e hiperactividad/impulsividad.

Es razonable concluir que el niño de 10 años con un TDAH de tipo inatento puede responder con una reducción significativa de los síntomas a la guanfacina a una dosis de 1,0 a 1,5 o 2,0 mg dos veces al día.

Estrategias para encontrar información sobre un tema más amplio

En lugar de buscar evidencia sobre una pregunta clínica específica relativa a un niño o adolescente, imagine que está escribiendo una guía para su país sobre cómo tratar a niños y adolescentes con esquizofrenia. Una pregunta relevante podría ser: "¿Cuál es la evidencia de que hay tratamientos eficaces para niños y adolescentes con esquizofrenia?" Comience buscando los datos más actuales y relevantes, provenientes de ensayos aleatorios. Inicie su búsqueda en PubMed, respecto a estudios de tratamiento en niños y adolescentes con esquizofrenia.



Haga clic en la imagen para acceder a la página web del Centro de Medicina Basada en la Evidencia, que proporciona información y entrenamiento en la MBE.

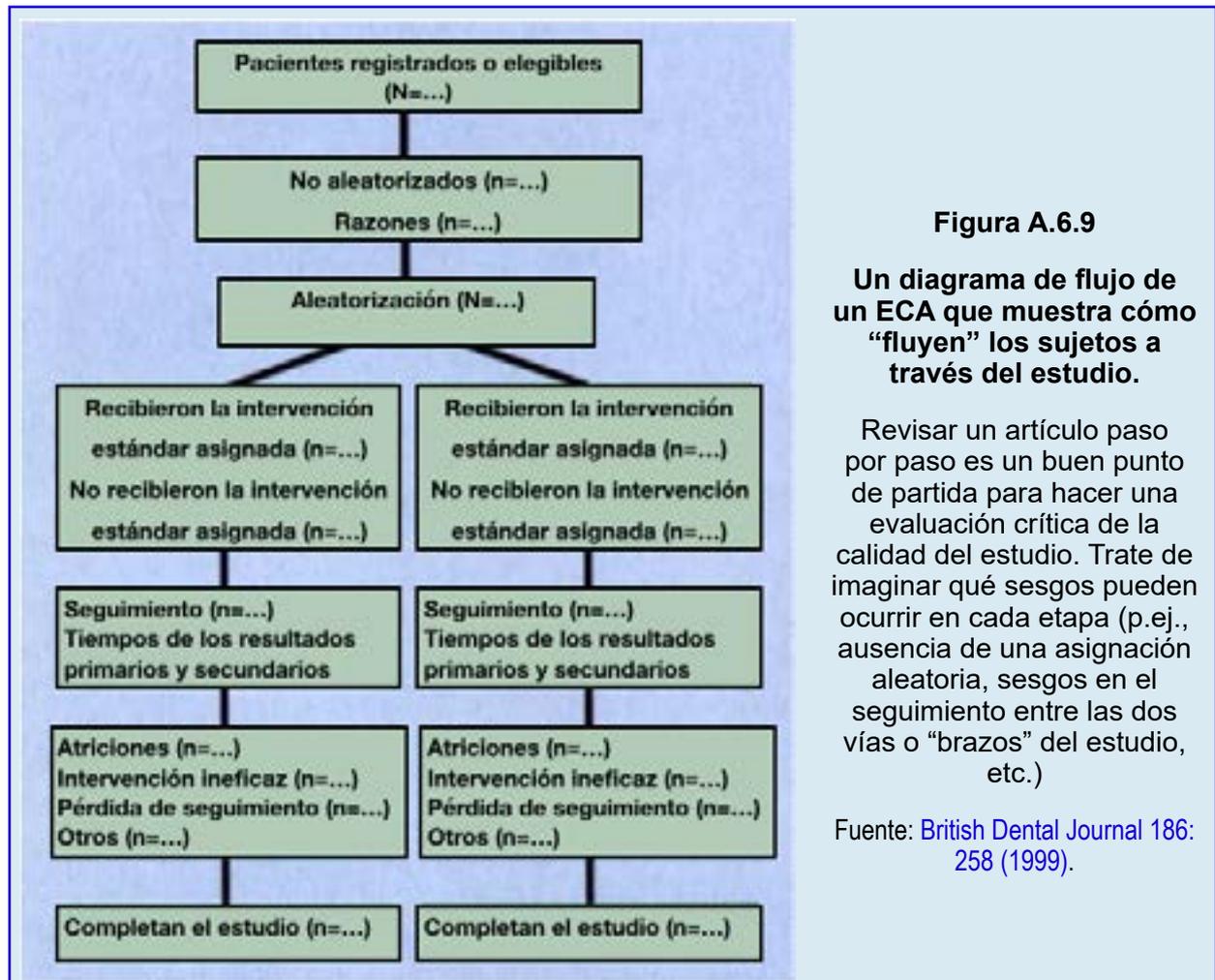


Figura A.6.9
Un diagrama de flujo de un ECA que muestra cómo "fluyen" los sujetos a través del estudio.

Revisar un artículo paso por paso es un buen punto de partida para hacer una evaluación crítica de la calidad del estudio. Trate de imaginar qué sesgos pueden ocurrir en cada etapa (p.ej., ausencia de una asignación aleatoria, sesgos en el seguimiento entre las dos vías o "brazos" del estudio, etc.)

Fuente: [British Dental Journal 186: 258 \(1999\).](#)

A continuación se indica cómo hacerlo (el número de estudios cambiará con el tiempo a medida que se agreguen nuevos estudios).

1. *Encuentre el encabezado de temas médicos para su tema de búsqueda.* Escriba "Schizophrenia" (esquizofrenia) en el cuadro de búsqueda de PubMed, una vez que haya utilizado el menú desplegable para cambiar la opción de búsqueda a "MeSH". Esto significa que ahora usted está buscando los Encabezados de Temas Médicos utilizados por la Biblioteca Nacional de Medicina de los Estados Unidos. Estos Encabezados de Temas Médicos son utilizados para clasificar cada uno de los artículos publicados en la literatura médica, haciendo posible encontrar todos los artículos sobre un mismo tema. En este caso, se recuperaron múltiples términos MeSH (12), descritos frecuentemente como 12 "éxitos". Cada uno de ellos contiene la palabra "schizophrenia" (esquizofrenia) en el título o en la descripción.
2. *Enviar los términos relevantes al cuadro de construcción de búsqueda ("Search Builder box"), combinándolos con el término "OR."* Sólo 6 de ellos parecen ser relevantes, ya que los últimos 6 se refieren a una proteína inusual. Después de seleccionar estos 6, haga clic en "Add to Search Builder" (añadir al constructor de búsqueda).
3. *Añadir a la búsqueda una versión truncada (tw) del concepto central, como una palabra seguida de "OR".* En este caso, añada "Schizophren*" (tw) (esquizofren*) al cuadro de búsqueda, seguido del operador "OR". Este paso recuperará todos los artículos que contenga en el título o en el resumen cualquier palabra que comience con las letras "Schizophren" (esquizofren). Este término adicional puede ayudar a encontrar artículos que no han sido clasificados correctamente bajo uno de los términos MeSH y que contenga información sobre la esquizofrenia, así como también artículos que sólo utilizan términos como "esquizofrénico" o "esquizofrénicos"
4. *Haga una búsqueda en PubMed con esta cadena de términos "OR".*
Se obtienen 125889 éxitos. Ahora es el momento de ser más específico y comenzar a filtrar esta montaña de referencias para encontrar las más relevantes.
5. *Haga clic en "Limits" (límites) y seleccione la opción "Children 0 to 18" (niños de 0 a 18 años)*
Aún hay 22545 éxitos.
6. *Haga clic en "Limits" (límites) y seleccione la opción "Meta-analysis" (metanálisis) y "Randomized controlled trial" (ensayo controlado aleatorizado).*
Aún hay 945 éxitos. Este es un número razonable para la tarea que se quiere hacer. Dos o tres personas pueden leer los títulos y resúmenes y elegir los más relevantes. Alternativamente, se puede volver atrás a "Limits" (límites) y seleccionar el cuadro "Published in last 5 years" (publicado en los últimos 5 años), reduciendo aún más el número de referencias. Observe también que muchas de las referencias corresponden a adultos, no a niños o adolescentes, y pueden ser descartados inmediatamente.

Si no hay tiempo para revisar tantos resultados, esta es otra alternativa:

1. Vuelva al paso 1 y observe que uno de los términos MeSH recuperados es "Schizophrenia, childhood" (esquizofrenia, infancia). Envíe sólo el término "Schizophrenia, childhood"[MeSH] al cuadro de búsqueda, solicitando al motor de búsqueda que solamente considere ese único término MeSH. Aparecen 1434 éxitos.
2. Ahora aplique los límites "Meta-analysis" (metanálisis) o "Randomized controlled trial" (ensayo controlado aleatorizado). Hay 8 éxitos. Estos resultados son de lectura obligatoria dado que los límites parecen haber eliminado demasiados éxitos.
3. Vuelva al paso 2 con los 1434 éxitos y aplique el límite "Last 10 years" (últimos 10 años): quedan 59 éxitos. La revisión de estos resultados revela una serie de artículos relevantes con temas como "Efecto de la clozapina y la olanzapina en el engrosamiento cortical en la esquizofrenia de inicio en la infancia", y la "Pérdida de asimetría es local más que global en la esquizofrenia de inicio en la infancia". Observe la dramática caída del número de éxitos, lo que sugiere que — como la página web de PubMed nos indica — "Schizophrenia, childhood" (esquizofrenia, infancia) es principalmente un término MeSH histórico; si bien los 59 éxitos recuperados de los últimos 10 años parecen ser altamente relevantes e interesantes para un informe de posición.

El mensaje para llevar a casa es *mantener las estrategias de búsqueda transparentes y claras, y siempre monitorizar cuántos éxitos resultan de la búsqueda; estos datos le permitirán ajustar continuamente la búsqueda a su propósito, tiempo e interés.*

Registre cuántos éxitos produce una estrategia en el contexto de los objetivos de su búsqueda: ¿Son los resultados manejables y adecuados para sus propósitos? Si no son manejables (demasiados éxitos y poco tiempo), continúe afinando nuevos resultados con más límites. Si le parece que no son suficientes, cambie la estrategia de búsqueda inicial para que sea más sensible. En este ejemplo, añadiendo el término "Schizophren*" (tw) añadió más de 20.000 éxitos. Otra alternativa para aumentar la sensibilidad es hacer clic en el hipertexto "Related citations" (citas relacionadas) debajo de las citas que parecen especialmente útiles; el software proporciona una serie de citas similares al tema de la cita que usted encuentra útil.

MBE CLÁSICA Y LA "PREGUNTA CONTESTABLE"

La "MBE tradicional" es nuestra frase para el enfoque desarrollado originalmente por David Sackett y sus colegas en la Universidad McMaster en Canadá, y más tarde en muchos otros lugares. Su trabajo fue fundamental para el enfoque que se analiza aquí. Sin embargo, distinguimos nuestro enfoque, que se adapta a la realidad de trabajar en salud mental infantil, de la "MBE tradicional", que fue desarrollada originalmente en el contexto de la medicina interna. Es importante conocer la "MBE tradicional", como el precursor fundamental de lo que se analiza aquí.

La "MBE tradicional" se puede resumir como un modelo de cinco pasos:

1. Definir una pregunta "contestable" sobre una situación clínica, que generalmente comienza con "¿Cuál es la evidencia de que..." un *paciente o problema* específico, tratado con una *intervención* específica resulta en qué *resultados*, en *comparación* con otra intervención o un placebo? Estas preguntas se denominan preguntas PICO (Paciente/ Intervención/Comparación/Resultados ("Outcomes")).
2. Buscar la evidencia para responder las preguntas PICO.
3. Evaluar críticamente la evidencia, haciendo preguntas sobre la calidad de los estudios.
4. Integrar la experiencia clínica y los valores del paciente con la evidencia, para decidir que tratamiento recomendar.
5. Evaluar el resultado.

La "MBE Tradicional" también aborda temas tan importantes como la sensibilidad y especificidad de las pruebas diagnósticas y cómo aplicar la evidencia (Sackett et al, 2000).

PRÁCTICA BASADA EN LA EVIDENCIA Y LA CULTURA LOCAL

La cultura — que no será definida para nuestros propósitos — es un factor importante que colorea cómo las ideas analizadas en este capítulo son recibidas tanto por los profesionales como por los pacientes. Los componentes culturales que afectan nuestra práctica clínica incluyen las creencias y valores, las prácticas de crianza, el papel de la familia extensa, y el impacto de la Internet.

Creencias y valores

Las creencias respecto al tratamiento de las enfermedades mentales son importantes, y es crucial que los profesionales entiendan la cultura en la que realizan su trabajo para conseguir que se acepten las intervenciones eficaces. En las sociedades donde hay una estigmatización más severa de los enfermos mentales, puede ser muy difícil convencer a un niño y a su familia de asistir a una entrevista psiquiátrica. Incluso si asisten, puede ser muy difícil para ellos aceptar el diagnóstico, así como el tratamiento sugerido. En algunas culturas se cree que las enfermedades mentales son espirituales, y por lo tanto los tratamientos espirituales son la primera opción. Las familias pueden aceptar el tratamiento sólo si el sanador espiritual lo aprueba. En uno de estos casos, un sanador espiritual turco había preparado un *muska* — un collar mágico de papel y tela envueltos para librarse de la enfermedad — para un adolescente psicótico que el sanador había enviado para que fuera hospitalizado. Durante el tratamiento, el adolescente abrió el *muska*. Encontró sólo una palabra garabateada: ¡*Haldo!*

Aún así, los clientes a menudo buscan curas mágicas e inmediatas. A veces es muy difícil explicar, y hacerles aceptar, que se requiere tiempo para un buen tratamiento, o la ventaja de usar un medicamento a pesar de sus posibles efectos secundarios. En culturas en las que la gente primero pregunta a sus vecinos o familiares sobre el problema, y se basan en sus experiencias personales o de terceros, el tratamiento basado en la evidencia sugerido por un profesional puede no encajar con las sugerencias de los vecinos, y se califica como "no debe ser usado" (el hijo del vecino se recuperó bien con otro medicamento, por lo tanto puede que éste no sea el adecuado).

Creencias sobre la crianza de los niños

Creencias como “los niños normalmente son más activos” puede ser una razón para que una familia rechace los medicamentos para tratar el TDAH, ya que la familia piensa que el medicamento ralentiza al niño y lo hace “anormal”. Otro creencia frecuente es que “los niños no deberían tomar pastillas durante mucho tiempo”. Esto a menudo interrumpe el tratamiento con medicamentos. O también, “la abuela dice que él es como su padre cuando era niño, y ahora su padre está bien. ¿Seguro que necesita tratamiento?”

Papel de la familia extensa

En las culturas tradicionales, los abuelos o los hermanos o hermanas mayores de los padres pueden tener la autoridad de decidir el tratamiento, incluso pueden negarse a permitir que los padres administren el tratamiento sugerido. En algunas ocasiones, padres que creen en los tratamientos basados en la evidencia tienen que administrar el tratamiento en secreto. Los miembros de la familia también pueden interferir con la dosis prescrita, basándose en sus opiniones personales.

La influencia de la Internet

Los padres o adolescentes que revisan constantemente la Internet pueden rechazar el tratamiento. Dado que la Internet puede tanto desinformar como informar, los pacientes y sus familias suelen encontrar más información errónea que orientación. Por ejemplo, buscan una palabra en Google y leen lo que se despliega, donde a menudo la información que se encuentra exagera los efectos secundarios, ignora los beneficios, o se opone a los medicamentos para tratar las enfermedades mentales. Proporcionar información sobre páginas de Internet con menos errores e información más realista y completa a los padres y adolescentes puede mejorar la adherencia al tratamiento.

En resumen, un papel clave del profesional es comprender tanto el sistema cultural de creencias y valores de sus pacientes como las ideas fundamentales de la epidemiología e investigación clínica, que constituyen la primera parte de este capítulo. La combinación de ambos, aplicada al paciente y a la familia en la consulta, es la esencia de la práctica basada en la evidencia.

REFERENCIAS BIBLIOGRÁFICAS

- Biederman J, Melmed R, Patel A et al (2008). A randomized, double-blind, placebo-controlled study of guanfacine extended-release in children and adolescents with attention-deficit/hyperactivity disorder. *Pediatrics* 121: e73-84
- Brent D, Perper J, Moritz G et al (1993). Firearms and adolescent suicide. A community case-control study. *American Journal of Diseases of Children*, 147:1066-1071.
- Guyatt GH, Rennie D (2002). *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice* Chicago: AMA Press
- Kabai P (2011b). *Bias and Random Error*.
- Kabai P (2011a). *More on Bias (systematic) and Random Error*.
- Kaufman J, Birmaher B, Brent D et al (1997). *Kiddie-SADS-Present and Lifetime Version (K-SADS-PL)*. Pittsburgh: Department of Psychiatry, University of Pittsburgh School of Medicine.
- Kelly R, Cohen L, Semple R (2006). Relationship between drug company funding and outcomes of clinical psychiatric research. *Psychological Medicine*, 36:1647-1656.
- McCracken JT, McGough J, Shah B (2002). Risperidone in children with autism and serious behavioral problems. *New England Journal of Medicine*, 347:314-321.
- Sackett DL, Straus SE, Richardson WS et al (2000). *Evidence-Based Medicine: How to Practice and Teach EBM*, 2nd edition. New York: Churchill Livingstone.
- Strange BC (2008). Once-daily treatment of ADHD with guanfacine: patient implications. *Journal of Neuropsychiatric Disease and Treatment*, 4:499-506.
- Walkup JT, Albano AM, Piacentini J et al (2008), Cognitive behavioral therapy, sertraline, or a combination in childhood anxiety. *New England Journal of Medicine*, 359:2753-2766
- Wikipedia (2011). *Standard Deviation*.
-