

DISEÑO Y EVALUACIÓN DE GUÍAS DE PRÁCTICA CLÍNICA

Dr. Ochoa.

Zamora, septiembre de 2007.

“La libertad clínica ha muerto y nadie debe lamentarse por ello”
(Hampton J. The end of clinical freedom. Br Med J 1983; 287: 1237-1238.)

Introducción

Podemos **definir** las guías de práctica clínica como “declaraciones desarrolladas de forma sistemática para ayudar a los médicos y a los pacientes en la toma de decisiones, para una atención sanitaria apropiada en circunstancias clínicas concretas”. Surgen ante la necesidad de reducir la variabilidad injustificada en la práctica clínica y mejorar la calidad del proceso asistencial. Constituyen una ayuda en el proceso de decisión clínica al presentar las alternativas diagnósticas y terapéuticas entre las que elegir.

Las guías no tienen porqué limitar la libertad del personal sanitario en su práctica clínica, ya que deben ofrecer recomendaciones suficientemente flexibles para su adaptación a las características de un enfermo concreto (“no hay enfermedades sino enfermos”).

El desarrollo de guías conduce a la necesidad de evaluar científica y objetivamente los resultados de los actos médicos, sobre todo cuando se cuenta con diferentes aproximaciones diagnósticas o terapéuticas para una misma enfermedad. Esta evaluación, necesariamente crítica, de la práctica clínica conlleva para algunos profesionales ciertas connotaciones negativas. Existe un cierto temor a encerrar las, a menudo complejas, decisiones clínicas de un problema de salud dentro del texto de una guía. Se teme quizás que lo plasmado en una guía se vuelva contra los que lo han elaborado, en forma de los variados fantasmas que amenazan la práctica médica. Se prefiere el término “guía de práctica clínica” al término “protocolo” ya que éste sugiere “rigidez de carácter normativo”, aunque las posibles diferencias son exclusivamente académicas.

Las razones de esta **desconfianza** pueden ser variadas aunque, en muchos casos, puede deberse a que las iniciativas de protocolización, a menudo no provienen directamente del personal clínico que, por otra parte, debería ser el principal encargado de desarrollarlas. No existe en estos casos la necesidad percibida de la guía como una herramienta clínica útil, por encima de otras consideraciones. Cualquier iniciativa de creación de una guía, desde criterios únicamente economicistas, o bajo una estrecha perspectiva de gestión, está probablemente condenada al fracaso desde su inicio más temprano. Parecería, por tanto más

viable, introducir este tipo de criterios sobre herramientas clínicas previamente consolidadas y consensuadas.

Aún considerando todas estas circunstancias es evidente que las guías conllevan múltiples ventajas:

- 1.- Permiten mejorar el estado del conocimiento sobre las alternativas terapéuticas y la historia natural de los problemas de salud, obligándonos a una continua actualización.
- 2.- Proporcionan un marco común de actuación, que permite igualar las condiciones en que se presta la atención hospitalaria, haciendo comparables los resultados de diferentes centros.
- 3.- Proporcionan a los profesionales seguridad legal ante demandas.
- 4.- Permiten identificar y asignar el papel de cada uno de los profesionales sanitarios implicados en la asistencia de un determinado problema de salud.
- 5.- Son una poderosa herramienta educativa para residentes y médicos en formación.
- 6.- Cuando se diseñan y aplican correctamente permiten mejorar los resultados de nuestros pacientes.

Diseño de Guías de práctica clínica.

Fases del diseño de una guía de práctica clínica.

Para poder diseñar una guía de práctica clínica tenemos que planificar una serie de pasos previos, que garanticen el éxito del proceso:

- Seleccionar el problema de salud a evaluar.
- Determinar los miembros del grupo de trabajo responsable de su elaboración.
- Elaborar un calendario de trabajo ajustado al tiempo y recursos disponibles.
- Repartir las tareas dentro del grupo de trabajo.
- Determinar la estructura de la guía.
- Planificar una estrategia de difusión y aplicación de la guía.
- Especificar los criterios, plazos y metodología de evaluación.

En lo que constituye propiamente dicho la ejecución de la guía, podemos distinguir también distintas etapas:

- Delimitación del problema de salud y su ámbito de aplicación.

- Identificación de las posibles intervenciones preventivas, diagnósticas o terapéuticas implicadas.
- Especificación de los parámetros que se considerarán en la evaluación de los procedimientos (eficacia, efectividad, eficiencia, utilidad, equidad, etc.).
- Revisión sistemática, síntesis y jerarquización de la evidencia científica.
- Elaboración de una versión preliminar con recomendaciones para la práctica clínica.
- Evaluación de la factibilidad práctica de la guía en nuestro entorno.
- Revisión externa de la guía.
- Cuando sea apropiado prueba piloto.
- Elaboración de la versión definitiva.
- Difusión y plan de implantación.
- Evaluación y actualización de la guía.

La fase más importante del desarrollo de una guía es el apartado de revisión, síntesis y jerarquización de la evidencia científica, ya que en ella se sustentan las recomendaciones que se incluirán en la guía. Podemos destacar en este proceso varios pasos:

- Búsqueda bibliográfica.
- Valoración metodológica (Validez). Jerarquización de la evidencia.
- Síntesis de la evidencia.
- Valoración de la relevancia y aplicabilidad (estrategias de consenso)
- Elaboración y graduación de las recomendaciones.

A continuación abordaremos con más detalle alguno de estos apartados.

Formulación del problema y ámbito de aplicación.

El camino para elaborar guías de práctica clínica se inicia en la identificación de un determinado problema de salud. Dicho problema debe ser descrito con precisión y situado en el entorno epidemiológico propio donde la guía va a ser aplicada. Interesa conocer las características de la población afectada (edad, sexo, factores de riesgo asociados, etc.), su frecuencia, su morbi-mortalidad y su importancia sanitaria y social.

Diversas circunstancias condicionan la selección de una condición clínica concreta:

- a) Su impacto en el sistema sanitario, según criterios epidemiológicos y/o de consumo de recursos.

- b) La posibilidad de identificar intervenciones sanitarias a utilizar en su abordaje clínico.
- c) La falta de consenso sobre los criterios idóneos de aplicación de dichas intervenciones.
- d) La existencia de variabilidad en la práctica clínica relacionada.
- e) La disponibilidad de evidencia científica.
- f) La aparición de nuevas tecnologías.

Resulta especialmente importante concretar los objetivos operativos que se pretenden conseguir con la guía diseñada y el ámbito de aplicación (criterios de inclusión-exclusión, nivel asistencial).

Revisión de la literatura.

A la identificación del problema de salud debe seguirle una revisión del conocimiento existente sobre las alternativas diagnósticas o terapéuticas disponibles. Este conocimiento debería basarse idealmente en los resultados de estudios de investigación básica o clínico-epidemiológica previos. Los resultados de éstos y otros tipos de estudios tendrán que ser integrados a través de un proceso intuitivo con los conocimientos sobre los mecanismos fisiopatogénicos subyacentes.

El éxito de una guía va a depender en gran manera de la exhaustividad de la información revisada y de su grado de actualización. Para ello resulta fundamental una adecuada estrategia de búsqueda bibliográfica. Aunque excede de los objetivos de este documento analizar los distintos recursos disponibles, en la tabla 1 se presentan algunos de ellos.

La literatura identificada debe ser recuperada, analizada, resumida y presentada, a ser posible en tablas en las que, junto a las características metodológicas más importantes, figuren los resultados principales y el grado de evidencia de cada uno de ellos. Asimismo tendremos que hacer una evaluación crítica de los resultados. En la tabla 2 se presentan algunos de los conceptos metodológicos y epidemiológicos que tendremos que manejar. En la tabla 3 se presentan criterios concretos a valorar según el tipo de estudio. Más adelante abordaremos estas cuestiones.

Tabla 1.- Recursos de búsqueda de evidencia científica.

Recursos tradicionales de información biomédica (no MBE)

Bases de Datos de referencias bibliográficas de revistas “primarias”

- Medline (PubMed): <http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>
- EMBASE: <http://www.embase.com/>
- Índice Médico Español:
http://bddoc.csic.es:8080/inicioBuscarSimple.do;jsessionid=25C2409CA0A904B52B847B18187689B9?now=1175013894458&tabla=docu&bd=IME&estado_formulario=show
- Otras (SciSearch, CINAHL, SciELO, IBECs, CUIDEN)

Revistas médicas

- Revistas médicas accesibles a través de internet: <http://www.infodoctor.org/revis.htm>
- Freemedicaljournals: <http://www.freemedicaljournals.com/>
- HighWire Press: <http://highwire.stanford.edu/>

Metabuscadores de fuentes de información secundarias:

- TRIPdatabase-Turning Research Into Practice: <http://www.tripdatabase.com/>
- SUMSearch: <http://sumsearch.uthscsa.edu/espanol.htm>

Colaboración Cochrane: <http://www.cochrane.org>

- Cochrane Reviews: <http://www.cochrane.org/reviews/index.htm>
- Cochrane Groups: <http://www.cochrane.org/cochrane/revabstr/crgindex.htm>
- Biblioteca Cochrane Plus en español: <http://www.update-software.com/publications/clibplus/>

Revistas con resúmenes estructurados:

- ACP Journal Club: <http://www.acpjc.org/>
- Evidence-Based Medicine: <http://ebm.bmjournals.com/> (original con suscripción) o <http://www.medynet.com/elmédico/publicaciones/inicio.htm> (gratuita en español)
- AAP GrandRounds: <http://aapgrandrounds.aapjournals.org/>
- Archimedes: adc.bmjournals.com/cgi/collection/archimedes
- PedsCCM Evidence-Based Journal Club: http://pedscm.wustl.edu/EBJournal_Club.html
- Evidencia. Actualización en la práctica ambulatoria: <http://www.foroaps.org/hitalba-home-evidencia.php>
- Evidencias en Pediatría. Toma de decisiones clínicas basadas en las mejores pruebas científicas: www.aepap.org/EvidPediatr/index.htm

Archivos de temas valorados críticamente:

- Universidad de Michigan: <http://www.med.umich.edu/pediatrics/ebm/>
- Centre for Clinical Effectiveness: <http://www.med.monash.edu.au/publichealth/cce/>
- ARIF-Aggressive Research Intelligence Facilities: <http://www.arif.bham.ac.uk/>
- BestBETS-Best Evidence Topics: <http://bestbets.org>
- Archivos de TVC de la AEPap: <http://www.aepap.org/evidencias/tvc.htm>

Guías de práctica clínica:

a) Centros elaboradores:

- GPC de la American Academy of Pediatrics: http://aappolicy.aappublications.org/practice_guidelines/index.dtl
- Scottish Intercollegiate Guidelines Network: <http://www.sign.ac.uk/>
- New Zealand Guidelines Group: <http://www.nzgg.org.nz/>

b) Centros de almacenamiento:

- GPC de la National Guidelines Clearinghouse: <http://www.guidelines.gov/>
- GPC de la Canadian Medical Association: <http://mdm.ca/cpgsnew/cpgs/index.asp>
- GuiaSalud: <http://www.guiasalud.es/>

Informes de Agencias de evaluación de tecnologías sanitarias:

a) Internacionales:

- International Network of Agencies for Health Technology Assessment (INAHTA): <http://www.inahta.org/>

b) Nacionales:

- AETS del Instituto de Salud Carlos III: <http://www.isciii.es/aets/>
 - Agència d'Avaluació de Tecnologia y Recerca Mèdiques: <http://www.aatrm.net>
 - AETS de Andalucía: <http://www.juntadeandalucia.es/salud/orgdep/AETSA/default.asp>
 - AETS del País Vasco: http://www.osasun.ejgv.euskadi.net/r52-478/es/contenidos/informacion/publicaciones_osteba/es_1215/pubost.html
-

Tabla 2.- Conceptos metodológicos y epidemiológicos de interés en la evaluación de la literatura

- Evaluación de pruebas diagnósticas:

- Sensibilidad y especificidad
- Valores predictivos
- Cociente de probabilidad (*likelihood ratio*)
- Probabilidades preprueba y postprueba
- Índice de concordancia (Kappa)

- Evaluación de la eficacia de una intervención

- Medidas de fuerza de asociación:
 - Odds ratio* (OR)
 - Riesgo relativo (RR)
- Medidas de impacto:
 - Reducción de riesgo relativo (RRR) y absoluto (RRA)
 - Número necesario de pacientes a tratar (NNT)

Tabla 3.- Criterios de calidad en la evaluación de la literatura, según el tipo de estudio.

Ensayos clínicos:

- Definición clara de la población de estudio, intervención y resultado de interés.
- Aleatorización adecuada.
- Control de sesgos (enmascaramiento).
- Seguimiento completo (menos del 20% de pérdidas)
- Análisis correcto (análisis por intención de tratar).

Estudios de valoración de pruebas diagnósticas:

- Comparación con un patrón de referencia válido.
- Muestra representativa.
- Control de sesgos (enmascaramiento; control de sesgos de incorporación, verificación diagnóstica y revisión).
- Análisis correcto (datos que permitan calcular indicadores de validez).

Estudios de cohortes:

- Cohortes representativas de la población con y sin exposición, libres del efecto o enfermedad de interés.
- Medición independiente y válida de exposición y efecto.
- Seguimiento suficiente y completo.
- Análisis correcto (control de factores de confusión y modificadores de efecto).

Estudios de casos y controles:

- Grupo control representativo del nivel de exposición en la población, sin la enfermedad o efecto de interés, pero en riesgo de presentarlo.
- Medición independiente de la exposición
- Análisis correcto (control de factores de confusión y modificadores de efecto).

Estudios de costes:

- Pregunta bien definida: que compare costes y efectos, evalúe una sola opción o compare varias alternativas.
- Asunciones del modelo válidas.
- Análisis correcto (análisis de sensibilidad).

Elección de procedimientos diagnósticos.

El abordaje diagnóstico de una guía puede abarcar desde la descripción de los síntomas o signos que definen el problema de salud hasta cualquier prueba complementaria empleada en el cribado, diagnóstico precoz, confirmación diagnóstica, graduación, pronóstico o evolución. Para seleccionar una determinada prueba diagnóstica debemos considerar su precisión y validez, evaluadas a ser posible en las circunstancias donde la prueba va a ser aplicada. Una prueba imprecisa, sin una reproducibilidad o consistencia mínimas, difícilmente nos ayudará en el proceso diagnóstico. De igual manera resultaría cuestionable una prueba inválida, que no mide lo que realmente se pretende medir. Sin establecer jerarquías, ambas características resultan importantes a la hora de seleccionar una determinada prueba (a veces nos interesará más acertar aproximadamente que equivocarnos con precisión).

Nos interesará estar familiarizados con distintos indicadores de precisión (índice kappa, coeficiente de correlación intraclase, alfa de Crombach, etc.) o de validez (sensibilidad, especificidad, valores predictivos, cocientes de probabilidades, probabilidad preprueba y postprueba, etc.). Resulta muy importante realizar una evaluación crítica de los indicadores que nos presentan la literatura científica o comercial, y especialmente adaptar dichos indicadores a nuestro entorno.

El punto de partida de un proceso diagnóstico es habitualmente un paciente, con unas características de gravedad y comorbilidad concretas, que le confieren una probabilidad determinada de tener la entidad a diagnosticar (probabilidad preprueba). El objetivo de la realización de una prueba diagnóstica es, una vez conocido el resultado, modificar esa probabilidad hasta obtener una probabilidad postprueba. La magnitud y dirección de ese cambio va a depender de las características operativas de la prueba diagnóstica, pero en todo caso debemos tener en cuenta que el punto de partida, la probabilidad preprueba, va a resultar muy importante en ese proceso.

Consideremos el escenario diagnóstico más simple, en el que tanto el patrón de referencia como la prueba diagnóstica clasifican a los pacientes en dos grupos, en función de la presencia o ausencia de un síntoma, signo o enfermedad. En la tabla adjunta presentamos, como ejemplo, los resultados de un estudio sobre la validez de la detección de leucocitos en orina como prueba diagnóstica de infección urinaria. En este caso, el urocultivo, cuya positividad confirma la existencia de infección (para orinas obtenidas por técnica estéril), es el patrón de referencia.

Generalmente los resultados se expresan a partir de la proporción de aciertos de la prueba diagnóstica entre las poblaciones enferma (sensibilidad) y sana (especificidad). La sensibilidad es la probabilidad de que la prueba sea positiva si la condición de estudio está presente (patrón de referencia positivo), mientras que la especificidad es la probabilidad de que la prueba sea negativa si la condición está ausente. Sin embargo, la sensibilidad o la especificidad no nos facilitan el cálculo de la probabilidad postprueba. Para ello resulta más útil el empleo de los cocientes de probabilidades.

El cociente de probabilidades (CP) para un determinado resultado de una prueba diagnóstica está definido como la probabilidad de dicho resultado en presencia de enfermedad dividida por la probabilidad de dicho resultado en ausencia de enfermedad. Los CP resumen información de la sensibilidad y de la especificidad e indican la capacidad de la prueba para incrementar o disminuir la verosimilitud de un determinado diagnóstico. En la tabla adjunta podemos observar los CP de los resultados positivo y negativo de la prueba de la estearasa leucocitaria en orina y su relación con la sensibilidad y especificidad. Utilizando los CP se pueden calcular las probabilidades postprueba (valores predictivos) considerando la probabilidad preprueba de cada paciente individual. En la tabla 4 se presenta un ejemplo de análisis de una prueba diagnóstica.

Tabla 4.- Tabla de contingencia de análisis de una prueba diagnóstica.

		Urocultivo		
		+	-	
Test de la estearasa Leucocitaria	+	81 a	160 b	241
	-	21 c	427 d	448
		102	587	689

Estudio de evaluación de la validez del test de la estearasa leucocitaria en tira reactiva, para el diagnóstico de la infección urinaria (urocultivo positivo). Características operativas de la prueba con sus intervalos de confianza al 95% (IC 95%). Datos tomados de Lohr et al. (Lohr JA, Portilla MG, Geuder TG, Dunn ML, Dudley SM. Making a presumptive diagnosis of urinary tract infection by using a urinalysis performed in an on-site laboratory. J Pediatr. 1993;122:22-5.)

Sensibilidad (Se) = $a/(a+c) = 81/102 = 0,79$ (IC 95%: 0,71-0,87)

Especificidad (Es) = $d/(b+d) = 427/587 = 0,72$ (IC 95%: 0,69-0,76)

Valor predictivo positivo (VPP) = $a/(a+b) = 81/241 = 0,33$ (IC 95%: 0,27-0,39)

Valor predictivo negativo (VPN) = $d/(c+d) = 427/448 = 0,95$ (IC 95%: 0,93-0,97)

Cociente de probabilidad positivo = $Se/(1-Es) = 2,91$ (IC 95%: 2,47-3,44)

Cociente de probabilidad negativo = $(1-Se)/Es = 0,28$ (IC 95%: 0,19-0,42)

Probabilidad preprueba (Ppre) = $(a+c)/(a+b+c+d) = 102/689 = 0,14$

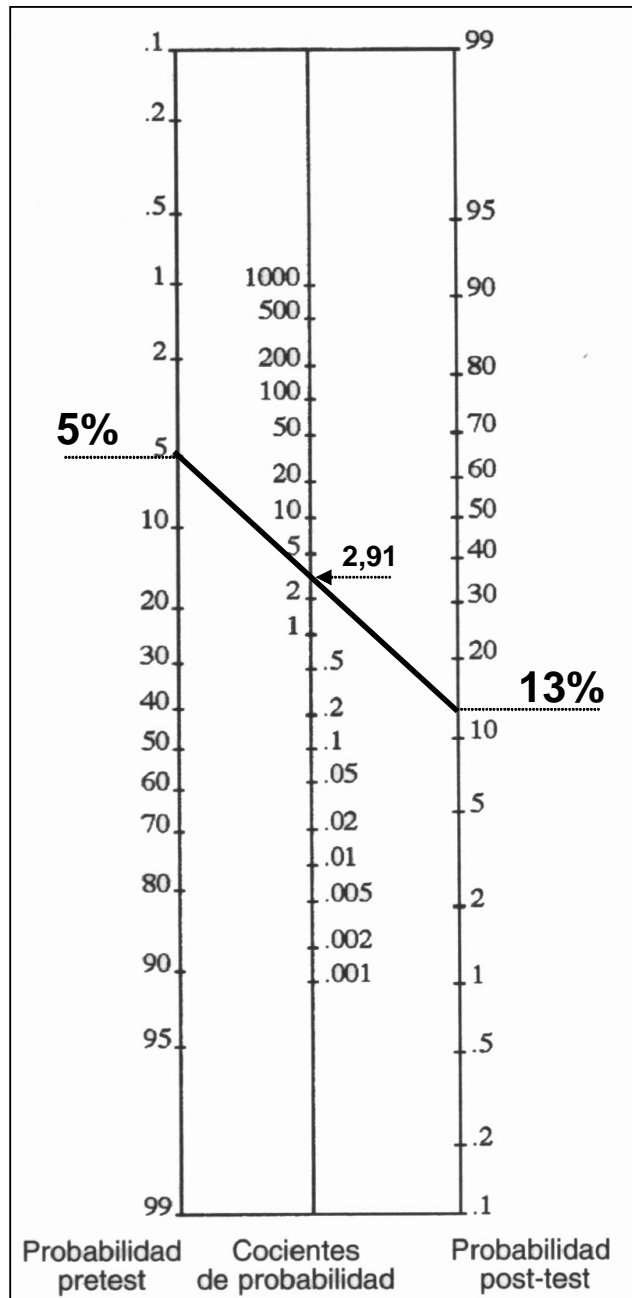
Para poder operar con los CP en el cálculo de probabilidades, éstas deben transformarse en ventajas (odds). Las ventajas u odds se calculan dividiendo las probabilidades por sus complementarios ($P/[1-P]$). Los pasos a seguir en el cálculo de la probabilidad postprueba son: 1) transformar la probabilidad preprueba en odds preprueba, 2) multiplicar la odds preprueba por el CP del resultado encontrado para obtener la odds postprueba, 3) transformar la odds postprueba en probabilidad (probabilidad = $\text{odds}/[1+\text{odds}]$). Este cálculo, aparentemente complejo, se simplifica mucho utilizando nomogramas desarrollados a tal efecto (ver figura adjunta). Si asumimos el CP positivo de nuestro ejemplo (2,91), podemos estimar la probabilidad postprueba de infección urinaria (ITU) en un escenario diferente al suyo. Para un lactante con fiebre sin foco (probabilidad preprueba estimada = 0,05; 5%) si la prueba de la estearasa leucocitaria es positiva la probabilidad postprueba de infección urinaria será 0,13 (13%). Esta estimación difiere considerablemente del valor predictivo positivo calculado con los datos del estudio (33%).

Cuando la prueba que se evalúa se mide en una escala cuantitativa, la sensibilidad y la especificidad varían en función del punto de corte que se utilice para separar los valores normales de los anormales. En estos casos, los resultados pueden representarse gráficamente como una curva ROC (iniciales del término inglés original Receiver Operating Characteristics), que permite conocer las características de la prueba según diferentes puntos de corte y que puede utilizarse para elegir el más adecuado.

En la figura 1 se representa un nomograma de Fagán que simplifica el cálculo de la probabilidad postprueba de ITU a partir de la estimación de probabilidad preprueba de ITU en lactante con fiebre sin foco (5%) y del cociente de probabilidades positivo de la estearasa leucocitaria en orina (Lohr et al.; ver tabla 4)

Al igual que en otros tipos de estudios, la valoración de la validez de las pruebas diagnósticas se hace sobre muestras, por lo que los resultados obtenidos son sólo estimaciones puntuales, sujetas a variabilidad aleatoria, y por lo tanto deben proporcionarse con sus intervalos de confianza. Estos intervalos de confianza tendrán que ser aplicados en el cálculo de la probabilidad postprueba para poder juzgar la utilidad de la prueba diagnóstica.

Figura 1.- Nomograma de Fagan para el cálculo de probabilidades postprueba a partir de probabilidades preprueba y cocientes de probabilidades.



En el proceso diagnóstico resulta particularmente útil establecer recomendaciones concretas de interpretación de los datos disponibles recurriendo a algoritmos diagnósticos, escalas de valoración o modelos probabilísticos. El diagnóstico debe expresarse de la forma más operativa y breve posible, aunque en algunos casos puede jerarquizarse como diagnóstico posible, probable y cierto.

Finalmente no podemos olvidar en la elección de pruebas diagnósticas las consideraciones económicas. En ocasiones un procedimiento diagnóstico muy válido, lo es a un coste personal o económico demasiado alto, especialmente si existen alternativas menos agresivas o más económicas con una validez aceptable.

Elección de procedimientos terapéuticos.

El tipo de estudio que mayor interés tiene a la hora de confeccionar recomendaciones terapéuticas es el ensayo clínico controlado y randomizado. La situación ideal es aquella en la que la eficacia o la efectividad de un determinado procedimiento ha sido evaluada en uno o más ensayos clínicos. Una revisión crítica de los resultados de esos estudios (incluyendo si es posible alguna revisión sistemática cualitativa o cuantitativa/metaanálisis) nos ayudará a decidir qué tratamientos son los indicados.

Los ensayos clínicos suelen utilizar como criterio de medición de los resultados de una intervención la presencia o ausencia de un evento adverso o favorable. Por ejemplo, si queremos comparar dos pautas de administración de antibióticos en ITU no complicadas, podemos contrastar la proporción de casos en cada grupo que presentan recurrencia de la infección tras el tratamiento. La forma más simple de estimar el efecto es calcular la diferencia absoluta entre ambas proporciones, lo que se conoce como reducción absoluta del riesgo (RAR). Sin embargo, en las publicaciones se emplea con frecuencia otra medida de efecto, la reducción relativa del riesgo (RRR), que ajusta la RAR a una escala relativa y que expresa la proporción de reducción de riesgo respecto al riesgo en el grupo control. En la tabla 5 se presenta un ejemplo en el que se comparan una pauta larga y una corta de cefradoxilo (10 días vs 1 día), con las estimaciones de RAR y RRR.

A través de estas medidas podemos estimar la diferencia entre las pautas comparadas; la pauta larga ofrece respecto a la corta una RAR del 24% y una RRR del 55%. Pero aunque estas estimaciones fueran las mejores disponibles, nunca podremos estar absolutamente seguros de que reflejan la verdadera reducción del riesgo, ya que son estimaciones puntuales obtenidas a partir de muestras. No obstante, es previsible que los verdaderos valores se encuentren en las proximidades de estas estimaciones, tanto más próximos cuanto más grande sea la muestra (menor error aleatorio, más precisión), pudiéndose concretar esa proximidad en lo que se conoce como intervalos de confianza.

Habitualmente se utiliza el intervalo de confianza al 95 % que quiere decir que dentro de ese intervalo se encontraría la verdadera reducción del riesgo en el 95% de los

casos (en 95 de cada 100 estudios similares). El intervalo de confianza contiene información de interpretación más intuitiva que el clásicamente referido nivel de significación estadística ($p < 0,05$), ya que el lector puede juzgar cuáles son las estimaciones más favorables y desfavorables de la reducción del riesgo (límites superior e inferior del intervalo). Un intervalo que comprende en su rango el valor nulo (para RAR y RRR: 0) equivale a ausencia de significación estadística. En todo caso la consideración de la significación estadística no debe distraer la atención sobre la valoración de la relevancia clínica del efecto.

Aunque la magnitud del efecto esperado en un paciente puede estimarse a partir de las medidas de reducción de riesgo, existe otra medida que ofrece una información más intuitiva de los posibles beneficios o perjuicios: el número necesario a tratar (NNT). El NNT es el número de sujetos que se necesitaría tratar con una intervención específica para producir, o evitar, un evento determinado. El NNT y su intervalo de confianza pueden calcularse fácilmente a partir de los inversos de la RAR y los límites de su intervalo de confianza. Con la intervención evaluada en nuestro ($RAR=0,24$) tendremos que tratar 4 sujetos ($NNT=4$) para beneficiar a uno.

No todos los ECA utilizan como criterio de medición variables dicotómicas. En ocasiones el efecto se refiere a una variable cuantitativa, por ejemplo, tiempo hasta la desaparición de la fiebre en niños tratados por ITU febriles. En estos casos, la medida del efecto de la intervención podrá estimarse calculando la diferencia entre los valores medios de dichas variables en los distintos grupos de estudio (diferencia de medias o medianas). Estas estimaciones puntuales y sus respectivos intervalos de confianza permitirán una interpretación directa y con sentido clínico de la dimensión de los resultados.

Tabla 5.- Tabla de contingencia de análisis de un ensayo clínico.

ITU no complicada	Medida de efecto		Proporción eventos
	Recurrencia 4-5º día		
Cefradoxilo	Si	No	
- 10 días (n=)	8	32	$P_i = 8/40 = 0,20$
- 1 día (n=)	16	20	$P_c = 16/36 = 0,44$

Ensayo clínico que compara dos pautas de tratamiento de ITU no complicadas en niños: cefradoxilo 10 días vs 1 día. Medidas de efecto con sus intervalos de confianza al 95% (IC 95%). Datos tomados de McCracken et al. (McCracken GH, Ginsburg CM, Namasonthi V, Petruska M. Evaluation of short-term antibiotic therapy in children with uncomplicated urinary tract infections. Pediatrics. 1981;67:796-801)

Reducción relativa del riesgo

$$RRR = \frac{P_c - P_i}{P_c} = \frac{0,44 - 0,20}{0,44} = 0,55 \text{ (55\%)} \quad (\text{IC } 95\% = 9-101)$$

Reducción absoluta del riesgo

$$RAR = P_c - P_i = 0,44 - 0,20 = 0,24 \text{ (24\%)} \quad (\text{IC } 95\% = 4-45)$$

Número necesario a tratar

$$NNT = \frac{1}{RAR} = \frac{1}{0,24} = 4 \quad (\text{IC } 95\% = 2 - 25)$$

En las tablas 6 y 7 se presentan dos ejemplos de revisión y presentación de evidencia. Mientras que en el primer ejemplo se presenta evidencia suficiente para poder emitir recomendaciones, en el segundo la información disponible es pobre y heterogénea.

Lamentablemente, existen muchos procedimientos que todavía no han sido evaluados en ensayos clínicos, por lo que nuestras decisiones tendrán que basarse en menores niveles de evidencia. Se han descrito diversos modelos de jerarquización de la evidencia científica en la que poder apoyar nuestras decisiones. En ellos se asigna el grado de evidencia en función del tipo de estudio que la ha generado. Podemos ordenar éstos de mayor a menor nivel de evidencia en: metaanálisis, ensayos clínicos controlados y aleatorizados, ensayos controlados no aleatorizados, estudios de cohortes, estudios de casos y controles y estudios descriptivos. En la tabla 8 se presenta un modelo simplificado de jerarquización y en la tabla 9 un ejemplo práctico de graduación de una serie de recomendaciones terapéuticas.

Tabla 6.- Presentación de evidencia sobre tratamiento de la otitis media aguda en la infancia

Grupos de comparaciones	Tratamientos Comparados	Años de los Estudios y N° niños	Diferencia de % curación (IC 95%)
No droga o placebo vs antibiótico	No antibiótico vs penicilina	1970 y 1981. N° niños: 242.	15,7 (4,7-26,7)
	No antibiótico vs aminopenicilinas	1968, 1970 y 1975. N° niños: 386.	12,9 (6,8-19,0)
	No antibióticos vs penicilina o aminopenicilinas	1968, 1970, 1981 y 1991. N° niños: 535.	13,7 (8,2-19,2)
Comparaciones con aminopenicilinas	Ampicilina vs penicilina	1969, 1970 y 1973. N° niños: 497.	-6,8 (-15,2-1,5)
	Ampicilina vs penicilina-sulfisoxazol	1967, 1968, 1975. N° niños: 462.	0,9 (-7,6-9,4)
	Aminopenicilinas vs eritromicina	1968, 1973 y 1987. N° niños: 525.	3,1 (-3,9-10,2)
	Aminopenicilina vs cotrimoxazol	1975 y 1983. N° niños: 275.	0,2 (-8,8-9,2)
	Amoxicilina vs cefaclor (*No existe homogeneidad entre los estudios)	1979, 1982, 1983 y 1984. N° niños: 453.	6,4 (-10,2-22,9)
	Amoxicilina vs cefixima	1989, 1991 y 1991. N° niños: 404.	-3,9 (-10,4-2,6)
Comparaciones con cefaclor	Cefaclor vs eritromicina sulfisoxazol	1987 y 1991. N° niños: 222.	7,0 (-6,5-20,4)
	Cefaclor vs amox-clavulanico	1985, 1986, 1987, 1984 y 1990. N° niños: 776.	2,8 (-1,3-6,8)
	Cefaclor vs cefixima	1987, 1987, 1991 y 1992. N° niños: 966.	1,2 (-2,4-4,79)

(Rosenfeld RM, Vertees JE, Carr J, Cipolle RJ, Uden DL, Giebink GS, Canafax DM. Clinical efficacy of antimicrobial drugs for acute otitis media: Metaanalysis of 5400 children from thirty-three randomized trials. J Pediatr 1994;124: 355-367).

Tabla 7.- Resumen de Ensayos Clínicos sobre tratamiento de Neumonías adquiridas en la comunidad en la Infancia.

Autor Año Lugar	Población de estudio Rangos de edad Características	Diseño	Pautas terapéuticas evaluadas (n° casos cada grupo) Dosis (d día) y reparto (x n° dosis)	Respuesta	Comentarios
Friis 1984 ⁽³⁰⁾ Dinamarca	136 niños 1 mes – 6 años Condensación RX Ingresados Casi todas víricas	Abierto	Cualquier antibiótico (72) Ampicilina en < 2 años Penicilina en ≥ 2 años Eritromicina en alérgicos No Antibiótico (64)	Reducción en 18 horas la duración de la fiebre en el grupos tratado No estadísticamente significativa.	Complicaciones: 1 mastoiditis en el grupo no tratado.
Paupé 1992 ⁽³³⁾ Francia	30 niños 1 – 12 años Incluye 8 bronquitis	Abierto	7-12 días: Cefetamet 10 mg/Kg/d x2 (10) Cefetamet 20 mg/Kg/d x2 (10) Cefaclor 10 mg/Kg/d x3 (10)	Éxitos (cura o mejoría) clínicos 100 % Cefetamet 10 100 % Cefetamet 20 90 % Cefaclor	Efectos adversos leves 30 % Cefetamet 10 40 % Cefetamet 20 50 % Cefaclor
Block 1995 ⁽⁷⁾ USA	260 niños 3 – 12 años <i>C.pneumoniae</i> 28 % <i>M.pneumoniae</i> 27 % Otros < 1,5 %	Ciego Simple (médico)	10 días: Claritromicina 15 mg/Kg/d x2 (133) Eritromicina 40 mg/Kg/d x2/x3 (127)	Exitos (cura o mejoría) clínicos 98 % Claritromicina 95 % Eritromicina Resolución RX 84 % Claritromicina 89 % Eritromicina	Efectos adversos leves 24 % Claritromicina 23 % Eritromicina No diferencias en eficacia ni efectos adversos
Klein 1995 ⁽³¹⁾ Francia Bélgica Alemania Suráfrica	348 niños 3 meses-11,5 años Sospecha Infección bacteriana (RX, T ^a >38°) moderada Incluye: 19 bronquio- litis y 37 bronquitis	Abierto	10 días: Cefpodoxima 8 mg/Kg/d x2 (234) Amoxicilina-Clavulánico 40/10 mg/Kg/d x3 (114)	Exitos (cura o mejoría) clínico- radiológicos 95,2 % Cefpodoxima 96,7 % Amox-Clav.	Efectos adversos leves 3,9 % Cefpodoxima 7 % Amox-Clav.

Tomada de: Ochoa Sangrador C, Solís Sánchez G, Vallano Ferraz A, Vilela Fernández M, Pérez Méndez C, Eiros Bouza JM, Inglada Galiana L, Guerra Romero L, por el Grupo Investigador del Proyecto FIS 97/0053 (Estudio de la variabilidad e idoneidad de los tratamientos antibióticos de las infecciones comunitarias respiratorias en la infancia). Tratamiento Antibiótico de las Infecciones Comunitarias Respiratorias en la Infancia. Valladolid 1999.

Tabla 8.- Escala simplificada de jerarquización de la evidencia.

NIVELES DE EVIDENCIA

Eficacia de una intervención:

- Nivel I: Basada en al menos un ensayo clínico aleatorizado y controlado, con diseño correcto, en la población objeto de la recomendación (incluye revisiones sistemáticas de estos ensayos clínicos).
- Nivel II: Sugerida por ensayos clínicos no aleatorizados o con otras limitaciones metodológicas (II-1), ensayos clínicos aleatorizados, realizados en población diferente, estudios de cohortes, estudios de casos y controles (II-2), y otros estudios no controlados con resultados espectaculares (II-3).
- Nivel III: Basada en estudios descriptivos.
- Nivel IV: Basada en opinión de expertos.

Validez de una prueba diagnóstica:

- Nivel I: Basada en al menos un estudio que realiza una comparación ciega de la prueba con un patrón de referencia válido, en una muestra de pacientes representativa de la que será objeto de aplicación (incluye revisiones sistemáticas de estos estudios).
- Nivel II: Basada en estudios que presentan sólo una de las siguientes limitaciones:
 - Comparación no ciega
 - Patrón de referencia imperfecto
 - Muestra de pacientes que no representa la que será objeto de aplicación
- Nivel III: Basada en estudios que presentan 2 o 3 de las limitaciones previas.
- Nivel IV: Basada en opinión de expertos.

GRADOS DE RECOMENDACIÓN:

1. Recomendación basada en una evidencia científica bien demostrada (Nivel de evidencia I; Recomendación grado A).
 2. Recomendación sugerida por una evidencia científica (Nivel de evidencia II; Recomendación grado B).
 3. Recomendación basada en experiencia clínica y opinión de expertos (Niveles de evidencia III y IV; Recomendación grado C).
-

Tabla 9. Criterios de idoneidad en la prescripción antibiótica de infecciones respiratorias agudas en la infancia. Grados de recomendación entre paréntesis. (Continúa en la página siguiente).

IRA	Primera Elección	Alternativas
Bronquiolitis	No antibióticos (A)	
Bronquitis aguda	No antibióticos (B)	Macrólidos (C) ^a
Faringo-Amigdalitis		
Infección por <i>S. pyogenes</i> confirmada o probable	Penicilina V oral (A) Penicilina Benzatina i.m. (A)	Amoxicilina (C) Ampicilina (C)
	Alérgicos a betalactámicos: Josamicina (C) Miocamicina (C) Eritromicina (C)	Otros macrólidos (C) Clindamicina (C)
Infección por <i>S. pyogenes</i> poco probable ^b	No antibióticos (B)	
Fracasos bacteriológicos demostrados	Amoxicilina-Clavulánico (C) Cefalosporina oral 2ª generación (C)	Clindamicina (C)
IRA no especificadas	No antibióticos (A)	
Laringotraqueitis		
Influenza		
Rinofaringitis		
Otitis media aguda		
Sin factores de riesgo	Amoxicilina (A) Amoxicilina-Clavulánico (B)	Cefuroxima axetil (C) Cefpodoxima (C)
	No tratar (en >18 meses) (C)	En alérgicos a beta-lactámicos Azitromicina (B) Claritromicina (C) Eritromicina (C)
Con Factores de riesgo ^c	Amoxicilina-Clavulánico (B) ^d	Amoxicilina (A) ^d Cefuroxima axetil (C) Cefpodoxima (C)
		En alérgicos a beta-lactámicos: Azitromicina (B) Claritromicina (C) Eritromicina (C)
		Pacientes hospitalizados: Cefotaxima (B) Ceftriaxona (B)
Sinusitis aguda	Amoxicilina (C) ^d Amoxicilina-Clavulánico (C) ^d	Cefuroxima (C) Cefpodoxima (C)
	No antibióticos (C) ^e	En alérgicos a beta-lactámicos (C): Azitromicina / Claritromicina
		Pacientes hospitalizados: Cefotaxima / Ceftriaxona (C)

Tabla 9 (Continuación).

IRA	Primera Elección	Alternativas
Neumonía		
< 3 años No hospitalizado	Amoxicilina-Clavulánico (B) ^d	No Antibióticos (B) Amoxicilina (B) ^{d,f} Cefuroxima axetil (B) Azitromicina (C) ^g
< 3 años Hospitalizado	Cefotaxima / Ceftriaxona (B) Amoxicilina-Clavulánico (B) ^d	No Antibióticos (B) Penicilina /Ampicilina (B) ^{d,f} Cefuroxima (B) Eritromicina / Claritromicina (C) ^h
≥ 3 años No hospitalizado	Macrólidos (B)	Amoxicilina (C) Amoxicilina-Clavulánico (C) Cefuroxima axetil (C)
≥ 3 años Hospitalizado	Cefotaxima / Ceftriaxona (B) Amoxicilina Clavulánico (B) ^d y/o Eritromicina / Claritromicina (C)	Penicilina /Ampicilina (B) ^d Cefuroxima (B)

Tomada de: Ochoa Sangrador C, Solís Sánchez G, Vallano Ferraz A, Vilela Fernández M, Pérez Méndez C, Eiros Bouza JM, Inglada Galiana L, Guerra Romero L, por el Grupo Investigador del Proyecto FIS 97/0053 (Estudio de la variabilidad e idoneidad de los tratamientos antibióticos de las infecciones comunitarias respiratorias en la infancia). Tratamiento Antibiótico de las Infecciones Comunitarias Respiratorias en la Infancia. Valladolid 1999.

^a Alternativa en casos persistentes (10 días).

^b Menores de 3 años o sin signos clínico-epidemiológicos de infección estreptocócica.

^c Factores de riesgo: Pacientes con otitis previas repetidas. Antecedente de antibioterapia (beta-lactámicos) previa reciente. Otitis en tratamiento con mala evolución. Niños menores de 18 meses que acuden a guardería.

^d Cuando existan factores de riesgo de infección por neumococo resistente a penicilina (infecciones respiratorias previas tratadas con betalactámicos, cuadro con mala evolución) usar dosis altas de amoxicilina (sola o en asociación).

^e En ausencia de signos inflamatorios locales (dolor facial, edema periorbitario) o afectación general.

^f Vacunados frente a *Haemophilus influenzae*.

^g En alérgicos a betalactámicos.

^h Lactantes menores de 6 meses con sospecha de infección por *C. trachomatis*.

Cuando los conocimientos disponibles no son sólidos o existe controversia podemos recurrir a métodos de consenso para tratar de obtener recomendaciones generales. Estos métodos están metodológicamente bien estructurados pero no exentos de limitaciones; deben ser considerados una alternativa válida hasta que aparezca evidencia científica suficiente. Los más extendidos son los métodos Delphi, la técnica de grupo nominal y las conferencias de consenso.

Además de la efectividad de un procedimiento terapéutico, en ocasiones convendrá considerar su eficiencia, utilidad o equidad. En estas situaciones tendremos que considerar los costes de los procedimientos alternativos utilizados (análisis de minimización de costes) y los beneficios obtenidos, referidos a unidades naturales (análisis de coste efectividad),

económicas (análisis de coste-beneficio), de satisfacción o de calidad (análisis de coste-utilidad). Asimismo podremos valorar el grado de accesibilidad real de cada opción seleccionada (análisis de equidad).

Aspectos éticos de la guía.

En la elaboración de toda guía deben considerarse los principios éticos que han de regular nuestro ejercicio asistencial: autonomía, beneficencia, no maleficencia y justicia. Cualquier procedimiento diagnóstico o terapéutico que conlleve riesgos o que no pueda ser considerado “rutinario” tendría que contar con el correspondiente documento de consentimiento informado.

Elección de la estructura de la guía de práctica clínica.

Las recomendaciones de una guía pueden estructurarse de distintas formas, dependiendo de las características particulares de problema de salud considerado, el grado de evidencia que apoya los procedimientos a utilizar y los objetivos concretos que se hayan planteado.

- Listado de recomendaciones. Relación de las opciones disponibles y sus indicaciones, más o menos jerarquizadas.
- Algoritmo diagnóstico-terapéutico. Diagrama de flujo que reproduce los pasos asistenciales que conducen a la solución de un problema, seleccionados a través de nudos de decisión encadenados. Este diagrama permitirá cuando sea posible realizar una asignación o reparto de tareas por niveles. En la figura 2 se presenta un ejemplo de algoritmo terapéutico.
- Vía clínica. Las vías clínicas han sido definidas como “La secuencia y cronología óptima de ejecución de las intervenciones de médicos, ATS y otros profesionales sanitarios en un diagnóstico o procedimiento particular, que permite minimizar retrasos en el uso de recursos y proporcionar al paciente los mejores cuidados posibles”. Serán útiles en aquellas situaciones en las que la secuencia y duración del proceso asistencial es previsible. En la figura 3 se presenta un modelo de vía clínica.
- Protocolos de aplicación sistemática. Se considerarán así las recomendaciones que recogen procedimientos terapéuticos o diagnósticos, de eficacia o validez bien establecidas, cuya indicación no se cuestiona o que requieren una aplicación

estrictamente guiada para ser útiles. Estos protocolos pueden ser trasladados directamente a órdenes médicas en determinadas situaciones clínicas consideradas estándar. Son aplicables aquí los procedimientos higiénicos, determinadas medidas preventivas y otros cuidados de enfermería.

Plan de implantación y consideraciones económicas.

Dado que la puesta en marcha de una guía supone la introducción de nuevos procedimientos asistenciales y la cancelación de otros previamente empleados, es necesario elaborar un cronograma detallado de las fases de implantación, información y consenso con los servicios y personas implicados, la preparación del material, la organización de la sistemática de trabajo, el reparto de tareas y a ser posible la estimación de gastos y beneficios.

Figura 2.- Algoritmo para el tratamiento antibiótico de la otitis media aguda

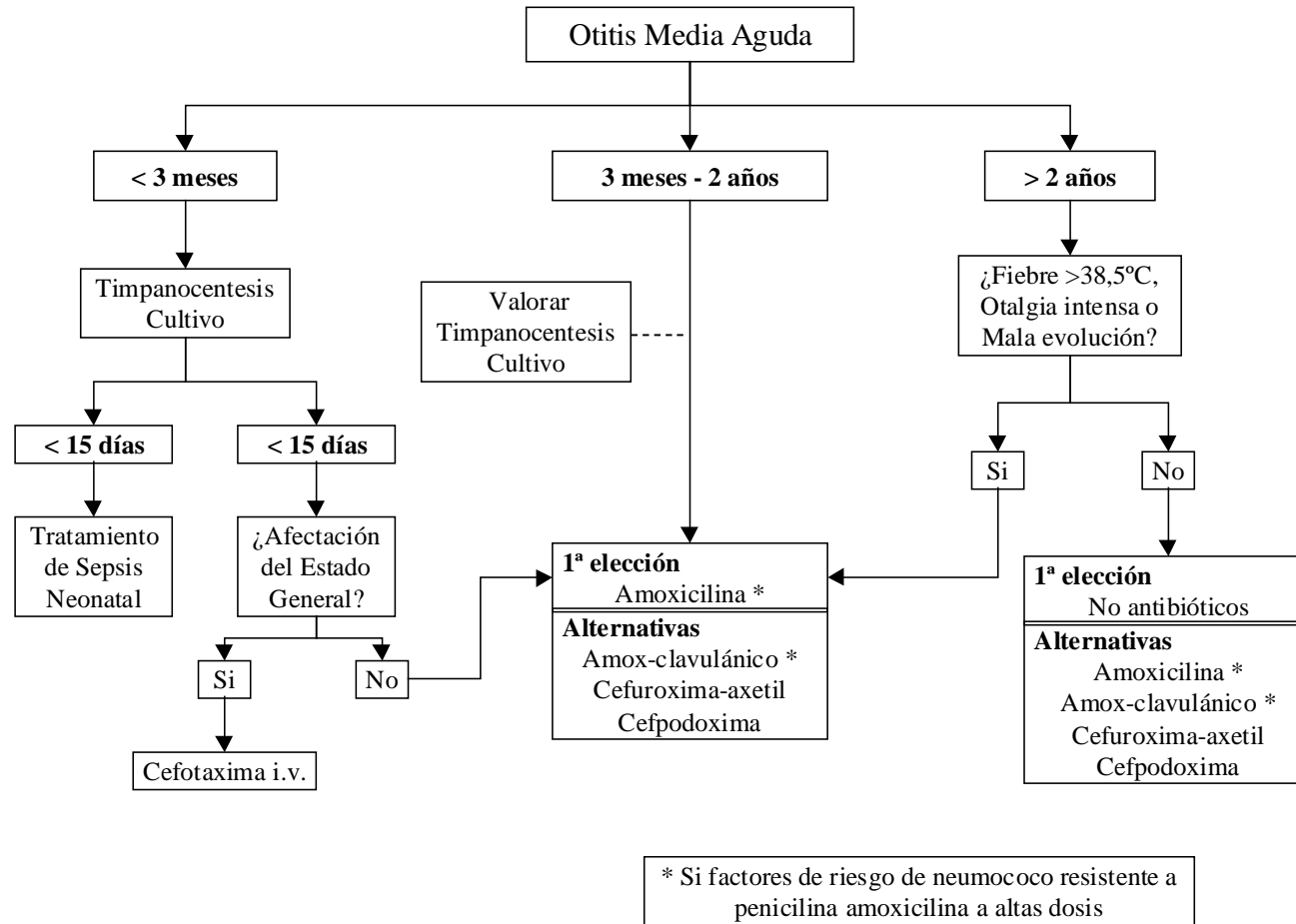


Figura 3.- Vía Clínica del paciente con crisis asmática hospitalizado

Servicio de Pediatría
Hospital Virgen de la Concha
Zamora

Nombre:	(etiqueta)
Nº Hª:	
Edad:	
Cama:	

Vía Clínica del paciente con crisis asmática hospitalizado

	Fase 1 – Status	Fase 2 – Transición	Fase 3 – Alta	Salidas de Protocolo (especificar) (indicar causa)
Monitorización:	Fecha: Hora: <input type="checkbox"/> Pulsioximetría continua <input type="checkbox"/> Valoración función respiratoria inicial y antes-después/ aerosol: - Frecuencia cardíaca - Escala de asma - Peak-Flow (% del basal)	Fecha: Hora: <input type="checkbox"/> Suspender pulsioximetría cuando no precise oxígeno Valorar medición esporádica <input type="checkbox"/> Continuar valoración función respiratoria antes-después/ aerosol:	Fecha: Hora: <input type="checkbox"/> Valoración final	<input type="checkbox"/>
Tratamiento:	<input type="checkbox"/> Oxígeno para Sat.O2 ≥94% <input type="checkbox"/> Salbutamol nebulizado: Ajustar según escala. Aprox.: - Cada 1- 2 h. 1 ^{as} 8 h. - Cada 2- 4 h. sigu. 16 h. Dosis: < 15 kg: 0,3 cc > 15 kg: 0,5 cc + 3 cc de SF <input type="checkbox"/> Si hay escasa respuesta valorar: Bromuro de ipratropio Dosis: 0,125 mg/ 4 h. ≤ 15 kg. 0,250 mg/ 4 h. > 15 kg. <input type="checkbox"/> Corticoides orales - Prednisona/ prednisolona Dosis carga: 2 mg/kg Mantenim.: 1-2 mg/kg/día (max.60 mg) en 2 dosis: ----- <input type="checkbox"/> Corticoides parenterales: - Metilprednisolona Dosis carga: 2 mg/kg Mantenim.: 0,5-1 mg/kg/6h. <input type="checkbox"/> Continuar o añadir medicación de mantenimiento (excepto broncodilatador de acción prolongada hasta el alta): -----	<input type="checkbox"/> Oxígeno para Sat.O2 ≥94% <input type="checkbox"/> Salbutamol cada 4-6 h. <input type="checkbox"/> Suspender Bromuro de ipratropio <input type="checkbox"/> Corticoides orales <input type="checkbox"/> Corticoides i.v. <input type="checkbox"/> Continuar o añadir medicación antiinflamatoria de mantenimiento:	<input type="checkbox"/> No precisa oxígeno <input type="checkbox"/> Salbutamol cada 6-8 h. <input type="checkbox"/> Corticoides orales <input type="checkbox"/> Continuar o añadir medicación antiinflamatoria de mantenimiento:	<input type="checkbox"/>
Actividad:	<input type="checkbox"/> Reposo en cama o actividad <i>ad libitum</i>	<input type="checkbox"/> Actividad <i>ad libitum</i>	<input type="checkbox"/> Actividad <i>ad libitum</i>	<input type="checkbox"/>
Nutrición:	<input type="checkbox"/> Dieta líquida según tolerancia o Normal. Si está indicado vía venosa para medicación	<input type="checkbox"/> Normal <input type="checkbox"/> Suspender vía venosa	<input type="checkbox"/> Normal	<input type="checkbox"/>
Educación	<input type="checkbox"/> Iniciar instrucción de padres y paciente sobre tratamiento y evaluación del asma	<input type="checkbox"/> Continuar instrucción	<input type="checkbox"/> Instrucción completada	<input type="checkbox"/>
Pruebas:	<input type="checkbox"/> Considerar gasometría capilar o arterial si: -Alteración estado mental -Aumento necesidad. Oxígeno -Pobre entrada de aire tras 4 dosis de salbutamol	<input type="checkbox"/> Radiografía de tórax: Sólo recomendada en: - Primera valoración - Sospecha complicaciones - Mala evolución	<input type="checkbox"/> Otras (indicar):	<input type="checkbox"/>
Interconsultas:	<input type="checkbox"/> Consulta a especialista en asma alergia o neumología	<input type="checkbox"/> Asistente social / A.Primaria	<input type="checkbox"/> Otras (indicar):	<input type="checkbox"/>
Objetivos:	<input type="checkbox"/> -Peak-flow > 70% -Bajar necesidades O2 para mantener Sat O2≥93%	<input type="checkbox"/> -Mínima o nula espasticidad -Sat. O2 ≥93% aire ambiental ≥94 despierto/≥92 dormido	<input type="checkbox"/> -Educación completada -Instrucciones / informe alta -Pauta inhalación en cámara	<input type="checkbox"/>
Otros:				

Evaluación de Guías de práctica clínica.

La planificación previa de rutinas de evaluación de toda guía es un elemento fundamental para poder controlar su calidad. Para ello deben definirse criterios, indicadores y sistemática de evaluación.

Serán criterios los elementos en que podemos desglosar la actividad que se pretende medir. Estos pueden ser de estructura (dotación necesaria), de proceso (procedimientos intermedios necesarios) o de resultado (circunstancia final que expresa el éxito o fracaso de la guía). Deben seleccionarse no más de 5 o 6 criterios para poder explorar la cobertura, cumplimiento y efectividad de la guía. Estos criterios pueden elegirse a partir de objetivos estándar (ideales) u objetivos empíricos (basados en la práctica real), y ser planteados a priori o bien a posteriori por consenso de expertos.

Los indicadores son la expresión o fórmula concreta que va a dar valor numérico medible a cada uno de los criterios una vez aplicados a la práctica asistencial. Debería existir al menos un indicador por criterio, pudiendo corresponder a los distintos tipos de criterios. Se conocen como indicadores “centinela” a aquellos que monitorizan la aparición de determinados acontecimientos adversos.

Los estándares representan el punto de corte del valor del indicador a partir del cual se considera cumplido el criterio. Es especialmente importante que se definan a priori los estándares de calidad y a ser posible que se correspondan con criterios e indicadores documentados bibliográficamente.

Para llevar a cabo la evaluación de la guía debe plantearse el tipo de estudio en que se van a cuantificar los indicadores de calidad. Se tratará habitualmente de estudios descriptivos sobre muestra o población entera, en un punto de corte concreto o estableciendo sistemas de registro mantenidos con evaluación periódica. Es conveniente que todo protocolo esté sujeto a modificaciones periódicas, guiadas a ser posible por los resultados de los estudios de evaluación.

Con respecto a la valoración de guías de práctica clínica ya publicadas, existen diferentes repertorios de evaluación. Estos repertorios exploran elementos formales y de contenido de las guías, y resultan útiles a la hora de considerar la validez y aplicabilidad de las mismas. El repertorio más extendido es el instrumento AGREE. Este consta de 23 criterios distribuidos en 6 áreas que analizan, respectivamente, los siguientes campos: 1) objetivo de la guía; 2) tipo de participantes implicados en su elaboración; 3) rigor en el

desarrollo; 4) claridad en la presentación; 5) aplicabilidad, y 6) independencia editorial de sus autores. En un estudio publicado en el que se evaluaron 61 guías de práctica clínica españolas se encontró que la calidad media de las mismas es muy baja. El análisis de cada uno de los elementos de este instrumento puede ser útil en el proceso de diseño de una guía de práctica clínica, aunque haya sido creado para guías ya elaboradas. En la tabla 10 se presenta una adaptación de los criterios del instrumento AGREE.

Tabla 10. Criterios de evaluación de guías de práctica clínica.

<p>A. ESTRUCTURA. Valora, mediante una escala de 4 puntos (de 4 “muy de acuerdo” a 1 “muy en desacuerdo”) los siguientes aspectos cualitativos del protocolo:</p> <ul style="list-style-type: none">◆ Alcance y objetivo. Preguntas clave:<ul style="list-style-type: none">– ¿Los objetivos del protocolo están específicamente descritos?– ¿Las preguntas clínicas están claramente descritas?– ¿Los pacientes en los que el protocolo se aplicará están perfectamente definidos?◆ Personas involucradas. Preguntas clave:<ul style="list-style-type: none">– ¿El grupo de trabajo incluye a individuos de todas las disciplinas requeridas?– ¿Los usuarios diana están perfectamente definidos?– ¿Se han considerado los puntos de vista, preferencias y posibles reacciones de la población diana (pacientes)?– ¿Se ha hecho alguna prueba piloto entre sus potenciales usuarios?◆ Rigor en el desarrollo. Preguntas clave:<ul style="list-style-type: none">– ¿Se han utilizado métodos sistemáticos de búsqueda de la evidencia?– ¿Los criterios para elegir la evidencia están descritos?– ¿Los métodos para formular las recomendaciones están claramente descritos?– ¿Al formular las recomendaciones, se han considerado los beneficios de salud, riesgos y efectos secundarios?– ¿Hay una relación explícita entre las recomendaciones y la evidencia que las soporta?– ¿El protocolo ha sido, antes de publicarse, revisado externamente?– ¿Hay establecidos fecha y procedimiento de actualización?◆ Claridad y presentación. Preguntas clave:<ul style="list-style-type: none">– ¿Las recomendaciones son suficientemente específicas?– ¿Las diferentes opciones para el diagnóstico y/o tratamiento están descritas?– ¿Las recomendaciones más importantes pueden ser fácilmente identificadas?– ¿El protocolo tiene herramientas (materiales adicionales) para su aplicación?◆ Aplicabilidad. Preguntas clave:<ul style="list-style-type: none">– ¿Se han tenido en cuenta las barreras potenciales en su aplicación?– ¿Se ha valorado la necesidad de recursos necesarios para cumplir las recomendaciones?– ¿Existen criterios de revisión que permitan su monitorización y evaluación posterior?◆ Independencia. Pregunta clave:<ul style="list-style-type: none">– ¿Los conflictos de intereses del grupo de trabajo se han descrito?
<p>B. EVALUACIÓN GLOBAL. Se realiza una valoración final de todo el protocolo evaluado, de forma subjetiva, clasificándolo como:</p> <ul style="list-style-type: none">– “muy recomendable”– “recomendable con condiciones o modificaciones”– “no recomendable”– “no se sabe”

Bibliografía.

- Rodríguez Artalejo F, Ortún Rubio V. Los protocolos clínicos. Med Clin (Barc). 1990; 95: 309-316.
- Díez Sebastián J, García Caballero J, Monjas Bonache A. Modelo para la elaboración de un protocolo o guía asistencial. Comisión de Tecnología. Unidad de Garantía de Calidad. Servicio de Medicina Preventiva. Hospital Universitario "La Paz". Madrid.
- Validity of the Agency for Healthcare Research and quality Clinical Practice Guidelines. How quickly do guidelines become outdated? JAMA. 2001; 286: 161-1467.
- Jovell AJ, Navarro-Rubio MD, Aymerich M, Serra-Prat M. Metodología de diseño y elaboración de guías de práctica clínica en atención primaria. Aten Primaria. 1997; 200: 259-266.
- Arquiga Thireau R. Protocolos: hasta dónde especificar las recomendaciones. Formación Médica Continuada en Atención Primaria. 1995; 2: 147-151.
- Jovell AJ, Navarro MD. Guías de práctica clínica. Formación Médica Continuada en Atención Primaria. 1995; 2: 152-156.
- Scottish Intercollegiate Guidelines Network. Sign50: a guideline developers' handbook. SIGN, Edinburgh. 2001.
- Miller J, Petrie J. Development of practical guidelines. BMJ. 2000; 335: 82-3.
- Rico Iturrioz R, Gutierrez-Ibarluzea I, Asua Batarrita J, Navarro Puerto MA, Reyes Dominguez A, Marin Leon I, et al. Valoración de escalas y criterios para la evaluación de guías de práctica clínica. Rev Esp Salud Publica. 2004; 78:457-67.
- Navarro Puerto MA, Ruiz Romero F, Reyes Domínguez A, Gutiérrez Ibarluzea I, Hermosilla Gago T, Alonso Ortiz del Río C, et al. ¿Las guías que nos guían son fiables? Evaluación de las guías de práctica clínica españolas. Rev Clin Esp. 2005; 205:533-40.
- The AGREE Collaboration. AGREE Instrument Spanish. version, www.agreecollaboration.org