

# Del conocimiento científico de la investigación clínica a la cabecera del enfermo: las guías de práctica clínica y su implementación

Marta Aymerich / Emília Sánchez

Agència d'Avaluació de Tecnologia i Recerca Mèdiques. Barcelona. España.

*Correspondencia:* M. Aymerich. Agència d'Avaluació de Tecnologia i Recerca Mèdiques (AATRM). Parc Sanitari Pere Virgili. Esteve Terradas, 30. Edifici Mestral, planta 1. 08023 Barcelona. España. Correo electrónico: maymerich@aatrm.catsalut.net

*Recibido:* 16 de octubre de 2003.  
*Aceptado:* 10 de abril de 2004.

(Of scientific knowledge of clinical research at the bedside: clinical practice guidelines and their implementation)

## Resumen

Las guías de práctica clínica (GPC) están dirigidas a profesionales y pacientes, y proporcionan información para la toma de decisiones. Tienen como propósito disminuir la variabilidad de la práctica médica para garantizar un nivel óptimo de calidad y mejorar así la atención sanitaria.

Las GPC se basan en revisiones sistemáticas de la literatura biomédica y recomiendan líneas de actuación clínica de diferentes grados, según la calidad de la evidencia científica en la que se sustentan. Sin embargo, debe tenerse en cuenta que el conocimiento científico es sólo uno de los elementos que influyen en la práctica médica. Así pues, no es tarea fácil recomendar una línea de actuación en un contexto determinado (realidad) a partir de lo que ha sido investigado en condiciones bastante ideales (ensayo clínico aleatorizado); como tampoco lo es plasmar en una recomendación los resultados de la investigación, que depende del contexto (estudios de efectividad o de coste-efectividad). Estas dificultades siempre estarán presentes, porque cada paciente (y sus circunstancias) es diferente, pero pueden minimizarse si se llevan a cabo nuevos estudios que permitan reducir, un poco más, la incertidumbre en la toma de decisiones clínicas.

Finalmente, hay que comentar otra dificultad: conseguir la implementación de las GPC y en esto aún queda mucho camino por recorrer en el área de las ciencias de la salud. Aún así, seguro que si no hay un buen producto, de poco servirán las estrategias que permitan implementarlo. Por tanto, los esfuerzos deben concentrarse sobre todo en elaborar GPC de buena calidad.

**Palabras clave:** Guía de práctica clínica. Ensayos clínicos aleatorizados. Medicina basada en la evidencia.

## Abstract

Clinical practice guidelines (CPG) are targeted at health professionals and patients providing information for decision-making. They are aimed at reducing variations in medical practice in order to guarantee an optimum level of quality and improve health care.

CPG are based on systematic reviews of biomedical literature and they recommend different clinical intervention strategies depending on the quality of scientific evidence on which they are based. However, we must bear in mind that scientific knowledge is only one of the elements that influence medical practice. Thus, it is not easy neither to recommend an intervention strategy for a specific circumstance (reality) from what has been studied in some rather ideal conditions (randomized controlled trial), nor to translate the context-dependent research evidence -effectiveness or cost-effectiveness studies- into a recommendation. These difficulties will be always present, as each patient (and his/her circumstances) is different from the others; but they can be lessened if more research is carried out to reduce uncertainty in clinical decision making.

Finally, another difficulty that should be mentioned is that of CPG implementation, and in order to succeed in it, there is still much left to do in the health sciences field. Nevertheless, if there is not a good product to implement, the strategies used to do it will not be very helpful. Thus, we need to focus our efforts on the development of high quality CPG.

**Key words:** Clinical practice guideline. Randomized controlled trials. Evidence-based medicine.

La versión en catalán está disponible en la web de GACETA SANITARIA ([www.doyma.es/ga](http://www.doyma.es/ga))

*Editores de la serie:* Joan MC Pons y Xavier Castells, con la colaboración de la RED IRYSS (*Investigación en Resultados de Salud y Servicios Sanitarios*).

## Estandarizar la práctica clínica incorporando el conocimiento científico

Las guías de práctica clínica (GPC) son, en pocas palabras, la estandarización de procesos y procedimientos. El estándar representa la forma de obtener un resultado en un determinado momento y permite garantizar un nivel óptimo de calidad. A pesar de ello, no todos los procesos y procedimientos médicos son estandarizables en la misma medida, como tampoco hay un mismo grado de progreso en todas las especialidades de la medicina.

La variabilidad en la práctica clínica, presente al comparar zonas geográficas, se puso de manifiesto en publicaciones científicas a principios de los años ochenta en los Estados Unidos, estudiando, especialmente, procedimientos quirúrgicos (probablemente porque la práctica quirúrgica es más fácil de medir que la médica)<sup>1</sup>. Aquí y ahora la variabilidad continúa presente<sup>2,3</sup>. Como en toda varianza, en la variabilidad hay una parte inexplicable que se llama azar o se considera el resultado de una heterogeneidad entre individuos que aún no se puede medir y diferenciar. No obstante, se han mencionado algunos factores para explicarla, entre los que se pueden citar los siguientes: los factores propios de la población (de donde surge la demanda como una necesidad expresa), como los factores demográficos, la distribución de la morbilidad, las expectativas; los factores propios del sistema como los recursos existentes, su organización, su financiación y accesibilidad; y los factores propios del proveedor más directo: el grado de conocimiento (o, por el contrario, la ignorancia) y la incertidumbre (desconocimiento probabilístico)<sup>4</sup>. Son los factores del sistema y del proveedor los determinantes de la oferta, y este cociente entre demanda y oferta explica, en parte, la variabilidad, la sobreutilización o infrautilización y sus consecuencias (cirugía innecesaria, listas de espera). Las GPC, fundamentalmente, son útiles al proveedor directo en tanto que pretenden delimitar la ignorancia y la incertidumbre.

Como recopilaciones sistemáticas que ordenan y sintetizan el estado de conocimiento existente, las GPC van dirigidas a los principales actores del sistema, pacientes y profesionales, pero también a otros participantes, como gestores, planificadores, juristas, medios de comunicación y público en general. Así, permiten, tanto a los profesionales como a los pacientes, tomar decisiones informadas sobre las diferentes opciones existentes (preventivas, diagnósticas, terapéuticas, de rehabilitación y paliativas) y seleccionar, de acuerdo con el conocimiento, los valores y los recursos, la más adecuada para una situación específica o concreta. La utilidad para los gestores deriva de la posibilidad de realizar una previsión de los costes que supondría la aplicación de las recomendaciones de las GPC, puesto que con ellas se

pretende ofrecer una atención más coste-efectiva (lo que no es necesariamente sinónimo de ahorro para el sistema). Para los planificadores, las GPC ofrecen oportunidades para la compra de servicios y la evaluación de los resultados. De manera amplia, pues, podría considerarse que las GPC tienen 2 tipos de objetivos —clínicos y administrativos— y que proporcionan resultados —previsibles, documentados y reproducibles— extremadamente útiles en ambas áreas.

Las GPC son recopilaciones sistemáticas, porque no puede ser de otra manera. Hoy, el grado de conocimiento existente y la profundidad que éste ha adquirido en las más pequeñas materias resultan inalcanzables. Sólo hay que ver, para centrarse exclusivamente en el área biomédica, el número de revistas médicas existentes, de ensayos clínicos aleatorizados (ECA) o de publicaciones científicas que se llevan a cabo anualmente. Toda esta ingente cantidad de conocimiento publicado no se puede abordar sin las herramientas que permitan, accediendo a todas las fuentes posibles (bases de datos, «literatura gris», Internet), recuperar y seleccionar los trabajos científicos más significativos. Ya no es válido, a pesar de que continúa desempeñando un papel muy importante, confiar en la última edición del tratado de referencia que actualice la forma de abordar, diagnosticar o tratar una enfermedad concreta. Por mucho que se reduzca el período entre ediciones, el saber cambia y se modifica rápidamente, y hay que estar atento para actuar o procurar actuar en todo momento con el mejor y más efectivo conocimiento científico. Éste es un deber de carácter deontológico y milenar: actuar de la mejor manera para el paciente. Por ello, es preciso que todo este conocimiento, e incluso la experiencia propia o más cercana, se pase por el tamiz o filtro que permita «separar el grano de la paja»; separar, pues, lo que responde al más alto grado de rigor científico (experimentación y reproducibilidad) de lo que simplemente es anecdótico o espurio<sup>5</sup>.

### «Separar el grano de la paja», el elemento fundamental de la recopilación sistemática de la literatura biomédica

Las revisiones sistemáticas recopilan, de la manera más exhaustiva posible, toda la información bibliográfica al alcance sobre la investigación en un determinado tema y sintetizan los hallazgos, resultados, pruebas o evidencias de los estudios más importantes. La evidencia científica así obtenida es un conocimiento generalizable que procede de datos generados por grupos de pacientes y que, en el caso de las GPC, conduce a la formulación de recomendaciones basadas en dichos grupos. Dada la variedad en el tipo de diseños (experimentales, observacionales) y de estudios (bá-

sicos, clínicos, epidemiológicos) que se revisan, la calidad del conocimiento científico que proporcionan es, también, variable.

La relevancia de un estudio hay que determinarla en función de unos parámetros que hacen referencia a la metodología o el diseño del trabajo de investigación y sobre los que hay un acuerdo general: cuanto más riguroso es el método utilizado, más cercano a la «verdad» objetiva está el resultado y mejor es la evidencia científica. Con estos criterios se construye la denominada jerarquía de la evidencia y se definen diferentes escalas o niveles para medirla. A pesar de que las primeras propuestas para clasificar la evidencia científica se publicaron a finales de la década de los setenta<sup>6</sup>, después han aparecido otras procedentes de diferentes grupos<sup>7-9</sup>. Con independencia del número de niveles que presentan (por ejemplo, la utilizada por la Agència d'Avaluació de Tecnologia i Recerca Mèdiques (AATRM) tiene 9<sup>10</sup>), todas las escalas están basadas en el diseño clínico-epidemiológico del estudio y en determinadas características metodológicas cualitativas, como la existencia de comparabilidad entre los grupos que se estudian, el tamaño de muestra adecuado, el seguimiento suficientemente largo y completo, la importancia de la variable respuesta, el análisis adecuado de los datos y la pertinencia de las conclusiones. En relación con el diseño, el paradigma de corrección desde el punto de vista metodológico, en referencia a la validez de los resultados, es el estudio experimental y, en concreto, el ECA.

La jerarquía de la evidencia científica sitúa los estudios observacionales más sencillos en los niveles inferiores y los diseños que incorporan metodologías epidemiológicas y técnicas estadísticas más sofisticadas en los niveles altos. Debe aclararse, sin embargo, que esta clasificación es válida en condiciones óptimas, es decir, cuando el diseño del estudio es impecable y no hay defectos metodológicos. En esta situación —y a grandes rasgos— el nivel más alto lo ocupan los metaanálisis de ECA y los propios ensayos clínicos seguidos, en el nivel inmediatamente inferior, por los estudios prospectivos que incorporan un grupo control y permiten la comparación. En el siguiente nivel, se encuentran los estudios observacionales de cohortes y los de casos y controles, para acabar, en el nivel más bajo de evidencia, con las series de pacientes, los casos clínicos y las opiniones de expertos.

La posibilidad de establecer un orden o rango según la calidad del diseño permite clasificar los estudios, valora la validez de los resultados y confiere credibilidad a las recomendaciones que se puedan derivar. Como ejemplo, la tabla 1 recoge los niveles de evidencia científica definidos por la Agency for Healthcare Research and Quality de los Estados Unidos<sup>8</sup>.

No obstante, hay que recordar que la jerarquía o escala descrita no es inamovible, que el orden estable-

Tabla 1. Niveles de evidencia científica<sup>a</sup>

Nivel	Tipos de evidencia científica
Ia	La evidencia científica procede de metaanálisis de ensayos clínicos aleatorizados
Ib	La evidencia científica procede de al menos un ensayo clínico aleatorizado
Ila	La evidencia científica procede de al menos un estudio prospectivo controlado sin aleatorización y bien diseñado
Ilb	La evidencia científica procede de al menos un estudio cuasiexperimental bien diseñado
III	La evidencia científica procede de estudios observacionales bien diseñados
IV	La evidencia científica procede de documentos u opiniones de comités de expertos y/o de autoridades de prestigio

<sup>a</sup>Utilizados por la Agency for Healthcare Research and Quality (AHRQ) de los Estados Unidos.

cido es relativo y que, por tanto, puede cambiar. De hecho, se trata de un instrumento para medir la calidad de los trabajos científicos y, como tal, tiene algunas limitaciones. Una de ellas es que el nivel o grado de evidencia científica no está determinado sólo por el diseño de la investigación. Esto quiere decir que los diseños que aparecen en los niveles más bajos no siempre son superfluos o tienen poca utilidad. De la misma manera, la existencia de revisiones sistemáticas bien hechas o la disponibilidad de un ECA que incluyera un gran número de individuos frente a determinados metaanálisis conllevaría una discusión sobre la posición que tendrían que ocupar en la escala y, en consecuencia, sobre el nivel de evidencia científica.

Para valorar el rigor metodológico del diseño se dispone de listas de verificación o *check lists*, como las publicadas por el Evidence Based Medicine Working Group<sup>11-14</sup>, que ayudan a revisar la presencia o ausencia de las características propias del diseño en un estudio determinado. En el caso de los diseños experimentales, hay listados específicos que contienen desde 48 preguntas distribuidas en 15 apartados<sup>15</sup>, o 35 preguntas repartidas en 3 capítulos<sup>16</sup>, hasta cuestionarios con sólo 3 criterios de calidad esenciales<sup>17</sup>. Mediante la aplicación de estas listas se obtiene una puntuación cuantitativa que permite no sólo otorgar un valor al estudio revisado, sino también situarlo y ordenarlo en la escala de evidencia científica.

Otra limitación viene condicionada por la pregunta de investigación. Así, cuando se quiere estudiar el efecto de un factor de riesgo, no es éticamente aceptable el diseño del ECA (no se puede exponer a una exposición, que se supone perjudicial para la salud, a las personas incluidas en el estudio) y debe recurrirse a los estudios observacionales que, si están bien diseñados y se llevan a cabo correctamente, proporciona-

rán un buen nivel de evidencia. Obviamente, no será el nivel más alto, pero sí el mejor posible. Habrá que tener en cuenta este aspecto cuando en una GPC se quiera informar sobre prevención. Igualmente, cuando en una GPC se quiera recomendar sobre pruebas diagnósticas, habrá que basarse en estudios transversales que, a pesar de ser adecuados, no se encuentran en la parte alta de la jerarquía de la evidencia científica debido a su diseño.

Por tanto, no hay un diseño que sea el mejor, es decir, cada tipo de pregunta de investigación requiere la metodología más adecuada para poder disponer de una respuesta válida y útil<sup>18</sup>. La pregunta determina la metodología apropiada y la estrategia de diseño, el diseño condiciona la calidad de la evidencia científica y, en función de ésta, se consigue un nivel en la escala. Por eso, cuando deben tomarse decisiones clínicas o de salud pública o hay que hacer recomendaciones de práctica clínica basadas en la evidencia científica, conviene actuar con transparencia, explicitando los criterios que han determinado el grado o nivel de evidencia que apoya la decisión tomada.

### Del papel (revisión sistemática de la literatura biomédica) al papel (recomendaciones clínicas)

La ventaja de las revisiones sistemáticas de la literatura científica sobre las revisiones narrativas de algunas revistas o los capítulos de un libro es que la síntesis del conocimiento científico publicado es exhaustiva y rigurosa. Es decir, intenta incluir todos los estudios originales que se han llevado a cabo sobre la condición clínica de interés, pero, a la vez, no los incluye todos en el mismo nivel, sino que les da diferente peso según su diseño y calidad metodológica, como se acaba de comentar. Ahora bien, las GPC aún exigen un paso más: traducir la evidencia científica hallada (con la correspondiente gradación) a una recomendación clínica. En este punto hay 2 dificultades:

1. Cómo traducir lo que ha resultado eficaz en condiciones de práctica clínica ideal (ECA) a condiciones de práctica clínica real.

2. Cómo traducir lo que ha sido efectivo (en condiciones de práctica clínica real) en un contexto determinado —con unos recursos, profesionales y pacientes concretos— al contexto en el que se quieren aplicar las recomendaciones de una GPC (podríamos llamarlo dificultad de extrapolación contextual). Tradicionalmente, los profesionales traspasado el conocimiento científico a la práctica clínica implícitamente. Lo que se aprendía en los libros se trasladaba a la práctica de la manera que, subjetivamente, parecía más fidedigna. Actualmente, cuando esta traslación se hace explícitamente

en una GPC, a pesar de que se intenta realizar de forma objetiva y transparente, continúa habiendo cierto grado de subjetividad (aunque no sea individual, sino del grupo de trabajo multidisciplinar que elabora la GPC).

La primera de las dificultades mencionadas es difícil de salvar si no se dedican más esfuerzos a diseñar estudios que no sean tan restrictivos al definir las condiciones de práctica clínica ideal de un ECA. Como ya se ha reclamado, habría que dedicar más neuronas y recursos para incluir en los ECA, por ejemplo, a pacientes con comorbilidades o de edades extremas, o que éstos se llevarán a cabo en ámbitos diferentes del hospital universitario<sup>19,20</sup>. No hay que olvidar que los resultados de estas investigaciones u otras similares son los que se sintetizan, de manera genérica, en las GPC y el profesional sanitario debe seleccionar, entre todas las recomendaciones, la más correcta y útil para su paciente. Es decir, tiene que trasladar la práctica clínica recomendable en un grupo de pacientes más o menos seleccionados a la atención óptima de un paciente individual (tiene que ir desde la certeza general a la especificidad particular). De hecho, es bien sabido que la práctica clínica no depende sólo de la eficacia rigurosa y científicamente demostrada de un determinado tratamiento (evidencia científica), sino también de la pericia profesional, los recursos al alcance, las características (antecedentes) del paciente al que se le aplica y sus preferencias (fig. 1). Por tanto, el profesional ha de saber encontrar el equilibrio entre estas 5 vertientes de la práctica clínica y poder convertir cada aplicación de la recomendación en un acto único. Así pues, la intención de una GPC no es, en absoluto, suplantarse el juicio clínico (como quizá se haya podido pensar), sino ayudar al profesional, como mínimo, en la vertiente del conocimiento o la evidencia científica. De esta manera, las GPC nunca podrán ser un libro o un recetario de cocina, sino un instrumento cuya utilidad será mayor cuanto mejor sea utilizado por el profesional sanitario.

Esta dificultad que implica la traslación de lo ideal a lo real, junto a la dificultad de extrapolación contextual, ha hecho que, en algunas ocasiones, los grupos de trabajo que han elaborado GPC hayan redefinido las

Figura 1. Elementos que intervienen en la práctica clínica.



escalas de niveles de evidencia científica y grados de recomendación para mejorar su aplicabilidad a un entorno concreto. Así, por ejemplo, el Scottish Intercollegiate Guidelines Network (SIGN) de reconocido rigor y prestigio en la elaboración de GPC, incluyó en el grado de recomendación la capacidad de aplicar los resultados de la literatura médica a la población a quien va dirigida la GPC<sup>21</sup>. Además, también consideró una nueva categoría para las situaciones donde no existen estudios, pero que son práctica establecida (tabla 2).

La nueva clasificación del SIGN también recogía un aspecto adicional que otras escalas de evidencia científica ya habían incorporado, la calidad de los estudios publicados, pero además consideraba aspectos de apli-

cación. En este sentido, algunos grupos de elaboración de GPC, como los grupos del programa OncoGuías de Cataluña, se han encontrado, en algún momento, con este callejón sin salida: existen uno o más ECA que apoyan una recomendación, pero estos ECA han sido muy criticados o los criterios de inclusión son tan restrictivos que hacen difícil asignarle a la recomendación el máximo grado, por más que el nivel de evidencia científica sea —según el diseño del estudio— elevado. El grupo de las OncoGuías superó el escollo asignando a cada recomendación, además del nivel de evidencia científica, un grado de consenso (unánime, mayoritario o inexistente)<sup>22</sup>. Este grado de consenso se alcanzó, de manera formal y abierta, durante las distintas reu-

**Tabla 2. Cambio en los niveles de evidencia científica y grados de recomendación del Scottish Intercollegiate Guidelines Network (SIGN)**

Niveles de evidencia científica hasta otoño de 2000 <sup>a</sup>	Niveles de evidencia científica a partir de otoño de 2000
<p>Ia: La evidencia científica procede de metaanálisis de ECA Ib: La evidencia científica procede de al menos un ECA</p> <p>Ila: La evidencia científica procede de al menos un estudio prospectivo controlado bien diseñado sin aleatorizar Ilb: La evidencia científica procede de al menos un estudio cuasiexperimental bien diseñado</p> <p>III: La evidencia científica procede de estudios observacionales bien diseñados IV: La evidencia científica procede de documentos u opiniones de comités de expertos y/o experiencias clínicas de autoridades de prestigio</p>	<p>1++: Metaanálisis de ECA y revisiones sistemáticas de ECA de alta calidad o ECA con muy bajo riesgo de sesgo 1+: Metaanálisis de ECA y revisiones sistemáticas de ECA bien hechas, o ECA con bajo riesgo de sesgo 1-: Metaanálisis de ECA y revisiones sistemáticas de ECA, o ECA con alto riesgo de sesgo 2++: Revisiones sistemáticas de alta calidad de estudios de casos y controles o estudios de cohortes, o estudios de casos y controles y estudios de cohortes de alta calidad con muy bajo riesgo de confusión, sesgo o azar y una alta probabilidad de que la relación sea causal 2+: Estudios de casos y controles y estudios de cohortes bien hechos con bajo riesgo de confusión, sesgo o azar y una probabilidad moderada de que la relación sea causal 2-: Estudios de casos y controles y estudios de cohortes con alto riesgo de confusión, sesgo o azar y un riesgo significativo de que la relación no sea causal 3: Estudios no analíticos, por ejemplo, series de casos o descripción de casos 4: Opinión de expertos</p>
Grados de recomendación hasta otoño de 2000	Grados de recomendación a partir de otoño de 2000
<p>A: Requiere al menos un ECA como parte de un conjunto de evidencia científica de buena calidad globalmente y consistencia con relación a la recomendación específica. (Niveles de evidencia científica Ia y Ib)</p> <p>B: Requiere disponer de estudios clínicos metodológicamente correctos que no sean ECA sobre el tema de la recomendación. (Niveles de evidencia científica II y III)</p> <p>C: Requiere disponer de documentos u opiniones de comités de expertos y/o experiencias clínicas de autoridades reconocidas. Indica la ausencia de estudios clínicos directamente aplicables y de alta calidad. (Nivel de evidencia científica IV)</p>	<p>A: Al menos un metaanálisis, revisión sistemática o ECA considerados 1++, directamente aplicables a la población diana, o una revisión sistemática de ECA o un conjunto de estudios considerados principalmente 1+, directamente aplicables a la población diana y que demuestren una consistencia global de los resultados B: Un conjunto de estudios considerados 2++, directamente aplicables a la población diana y que demuestren una consistencia global de los resultados, o evidencia científica extrapolada de estudios considerados 1++ o 1+ C: Un conjunto de estudios considerados 2+, directamente aplicables a la población diana y que demuestren una consistencia global de los resultados, o evidencia científica extrapolada de estudios considerados 2++ D: Evidencia científica de nivel 3 o 4; o evidencia científica extrapolada de estudios considerados 2+ Good practice point: recomendada como la mejor práctica clínica basada en la experiencia clínica del grupo que elabora la GPC</p>

ECA: ensayo clínico aleatorizado. <sup>a</sup>Son los que utiliza la Agency for Healthcare Research and Quality (AHRQ) descritos también en la tabla 1.



niones de trabajo que mantuvieron los grupos bajo la tutela de un coordinador que, a su vez, era el responsable de ir incorporando la clasificación de la evidencia científica disponible y del grado de consenso alcanzado. Por tanto, el consenso clínico no debería considerarse como una fuerza paralela respecto al nivel de evidencia científica, sino como una modulación (un tipo de factor de corrección), puesto que venía a traducir discrepancias en el grupo respecto a la calidad o aplicabilidad de los estudios originales.

Así pues, se ha visto que los problemas relacionados con la calidad metodológica de los estudios originales pueden superarse modificando y refinando las escalas de niveles de evidencia científica y los grados de recomendación y que parte de la aplicabilidad podría solucionarse mediante la realización de ECA con criterios de inclusión que se acercasen más a la práctica clínica habitual. La segunda dificultad, la de extrapolación contextual, es decir, la que depende, además de las características y preferencias del paciente, de los recursos disponibles en un contexto determinado, también sería parcialmente superable llevando a cabo estudios en el mismo contexto donde hay que aplicar la GPC.

Mientras no se realicen los estudios suficientes, los metaanálisis de ECA pueden conferir parte de la validez externa (capacidad de generalización) que se espera de los resultados de la investigación clínica. El metaanálisis, al incluir diferentes ECA, está en cierto modo ampliando —aunque sea artificialmente— criterios de inclusión y valorando diferentes contextos. Sin embargo, cuando además de la eficacia de una intervención, interesa también recomendar en función de la efectividad o del coste-efectividad —ambos con una relación directa con el contexto— los estudios originales producidos en el mismo contexto de aplicación se hacen mucho más necesarios.

---

### De las 2 dimensiones (papel) a las 3 dimensiones (realidad)

Suponiendo que se llegara a producir tanta investigación aplicada de calidad como la necesaria para escribir una recomendación en una GPC que admitiese sólo una incertidumbre residual, aún quedaría una tercera dificultad por superar: cómo hacer que la recomendación se aplique a la práctica clínica, es decir, cómo conseguir la implementación de las GPC. En este punto es cuando entran en juego las estrategias de diseminación e implementación. La difusión de la GPC a través del correo, la formación de los profesionales y la discusión pública sobre su contenido; las técnicas de recordatorio de las recomendaciones, las de *audit* o de *feedback* sobre los resultados conseguidos con la apli-

cación de la GPC o las visitas personales entregando y explicando la GPC son las estrategias que pueden utilizarse para diseminar las GPC y conseguir que se implementen. Éste, sin embargo, es un área de la investigación en ciencias de la salud que aún está en las fases más iniciales. De momento, la escasa evidencia científica al respecto apunta que la máxima implementación que se puede conseguir con la mejor de las estrategias (visitas personales) es sólo del 15%<sup>23</sup>, pero aún se está lejos de tener estudios de calidad que puedan evaluarlo correctamente (en este sentido, recientemente, bajo el auspicio de la Agencia de Evaluación de Tecnologías Sanitarias de Andalucía (AETSA) se ha traducido un libro que describe las ventajas y los inconvenientes de los diseños para llevar a cabo estudios que evalúen la implementación<sup>24</sup>).

Mientras tanto, habría que esforzarse, sobre todo, en elaborar buenas GPC y esto implica cumplir toda una serie de requisitos, ya casi clásicos, definidos a principios de los años noventa por el Institute of Medicine (IOM) de los Estados Unidos: validez, fiabilidad, reproducibilidad, aplicabilidad clínica, claridad, multidisciplinariedad en la elaboración, revisión planificada y documentación<sup>25</sup>. Cabe destacar que alguno de estos atributos, como la claridad, puede parecer fácil de alcanzar y, en cambio, a menudo puede ser uno de los más difíciles de cumplir. La claridad significa que las recomendaciones tienen que estar perfectamente diferenciadas del resto del texto que las justifica, que los algoritmos de actuación clínica deben de ser inteligibles, que la GPC ha de tener una versión reducida que permita al clínico consultarla rápidamente y que la edición tiene que ser atrayente. De otro modo, la GPC sería difícil de distinguir de un libro o de un artículo de revisión. La documentación es el otro atributo que parece sencillo de conseguir, pero puede no serlo, puesto que significa no sólo documentar todas las recomendaciones y afirmaciones de la GPC con las citas de los estudios de donde provienen, sino también determinar todo el proceso que se ha seguido para elaborarla: los componentes del grupo de trabajo y los conflictos de intereses, los revisores externos y la manera de considerar sus sugerencias, la forma de selección y evaluación de la evidencia científica o la financiación que se ha recibido para elaborar la GPC.

De hecho, todos estos atributos son los que tiene en cuenta el cuestionario AGREE de evaluación de calidad de las GPC (Appraisal of Guidelines for Research and Evaluation)<sup>26</sup>. Este instrumento fue desarrollado mediante un proyecto de colaboración financiado por la Comisión Europea y, entre todos los instrumentos de evaluación de la GPC, es el único que contiene, como mínimo, una pregunta para cada uno de los atributos definidos por el IOM<sup>27</sup>. A pesar del gran número de GPC existentes, no todas reúnen los criterios mínimos exigibles que demuestran su calidad y éste es un buen

instrumento para «separar el grano de la paja», ya sea para seleccionar una GPC con el propósito de utilizarla en la práctica clínica o para desarrollarla (puesto que el primer paso que se debe dar es buscar si ya han sido publicadas otras GPC con el mismo objetivo que la que se pretende desarrollar).

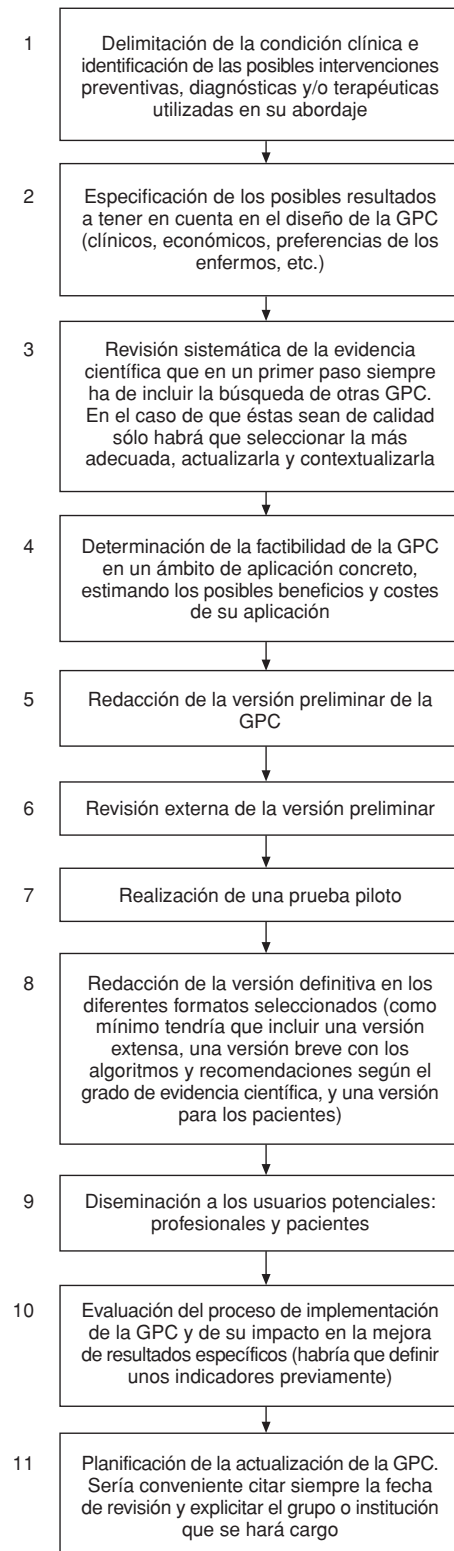
Teniendo en cuenta los atributos o criterios mínimos de calidad de las GPC, serán los propios profesionales quienes buscarán la GPC que les ayude en su práctica clínica y no será la institución sanitaria que las impulsa quien haya de convencerlos para que las utilicen. Por tanto, de momento, ese 15% de aceptación y uso que conseguiría una buena estrategia de disseminación e implementación sería suficiente si la parte más importante de ese éxito dependiera del buen diseño, desarrollo y contenido (atributos) de la GPC. Así pues, los que elaboran GPC deben tener muy presentes todos estos criterios y ser sistemáticos, siguiendo una etapa detrás de otra (fig. 2)<sup>28</sup> para poder elaborar las GPC que garanticen una calidad adecuada.

Aun así, lo que quizá sería preciso es que desde las instituciones sanitarias se fomentaran y pusieran al alcance los medios para poder desarrollar GPC para las condiciones clínicas en las que, por su variabilidad, prevalencia u otras circunstancias, fuera necesario. Además, sería interesante que hubiese un programa de GPC en el que colaborasen diferentes instituciones y periódicamente se editaran GPC de buena calidad al estilo del SIGN escocés<sup>29</sup>, o un organismo del Sistema Nacional de Salud al estilo del National Institute for Clinical Excellence (NICE)<sup>30</sup> en Inglaterra y Gales.

### Para finalizar: el estándar y la realidad

Las GPC contribuyen a disminuir la variabilidad de la práctica clínica mediante la definición de unos estándares que deben tender a mejorar la atención sanitaria, gracias a la incorporación del conocimiento científico, a pesar de las dificultades que entraña traspasarlo del papel a la realidad. Se debe destacar que seleccionar el conocimiento científico de calidad requiere que quien desarrolla las GPC tenga las herramientas y la destreza para leer críticamente la literatura médica y también capacidad de síntesis y claridad en la exposición (para elaborar los algoritmos necesarios y las versiones reducidas). De todos modos, la tarea de elaborar una GPC exige también cierta dosis de honestidad, puesto que el proceso consiste en ver qué dice la investigación científica y recomendar en relación con ésta y no a la inversa (escribir la recomendación y buscar la investigación que la apoya). La ventaja de hacerlo así es que se recomienda con más o menos fortaleza (según los niveles de evidencia científica) lo que está más demostrado. No se toma con la misma seguridad

Figura 2. Etapas en el desarrollo de una guía de práctica clínica (GPC).



una decisión clínica sabiendo que está avalada por una buena evidencia científica, que si lo está por una evidencia científica limitada o sólo por la opinión (respetable, pero que no proviene de los resultados de la investigación científica).

Después, por supuesto, hay que dedicar esfuerzos al diseño de publicación de la GPC (en papel y en formato electrónico), su difusión y las estrategias de implementación que exige la realidad clínica; pero, en primer lugar, hay que tener un producto (un estándar) de calidad al que poder aplicar todas estas herramientas de edición y *marketing*.

Sirva, por último, esta conclusión para reclamar, una vez más, la necesidad de llevar a cabo más investiga-

ción. Son necesarios más estudios de efectividad y coste-efectividad (que tengan en cuenta nuestro contexto) y también más estudios de eficacia (aquí y en cualquier lugar) que confieran más validez externa a sus conclusiones.

---

## Agradecimientos

Las autoras quieren agradecer a Joan M.V. Pons la lectura crítica y la aportación de ideas a una versión previa de este trabajo, y a Cristina Jaldón la traducción al castellano del manuscrito.

---

## Bibliografía

1. Wennberg J, Gittelsohn A. Variations in medical care among small areas. *Sci Am* 1982;246:120-34.
2. Jané E, Barba G, Salvador X, Salas T, Sánchez E, Bustins M. Variaciones en la tasa de hospitalización por procedimientos quirúrgicos seleccionados. Aplicación del análisis de áreas pequeñas. *Gac Sanit* 1996;10:211-9.
3. Pla R, Pons JMV, González JR, Borràs JM. ¿Hay niveles asistenciales en cirugía oncológica? ¿Los que más hacen, lo hacen mejor? Informe sobre la relación entre volumen de procedimientos y resultados en cirugía oncológica. Barcelona: Agència d'Avaluació de Tecnologia i Recerca Mèdiques, CatSalut. Departamento de Sanidad y Seguridad Social, Generalitat de Catalunya, 2003.
4. Buen JM, Peiró S, Márquez Calderón S, Meneu de Guillerma R. Variaciones en la práctica médica: importancia, causas e implicaciones. *Med Clin (Barc)* 1998;110:382-90.
5. Pons JMV. Gestión clínica basada en las fuentes de información: las guías de práctica clínica. Congreso Gestión Clínica, una nueva forma de trabajar que mejora las decisiones. Programa, ponencias, experiencias y comunicaciones. Barcelona, 2003; febrero 22; p. 31-32.
6. Canadian Task Force on the Periodic Health Examination. The periodic health examination. *CMAJ* 1979;121:1193-254.
7. The Swedish Council on Technology Assessment in Health Care. Literature searching and evidence interpretation for assessing health care practices. (SBU Report N.º 119E). Stockholm: SBU, 1996. Disponible en: <http://www.sbu.se/admin/Docs/PUB/570/indexsid1.html>
8. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, et al. Systems to rate the strength of scientific evidence. Evidence Report/Technology Assessment N.º 47 (Prepared by the Research Triangle Institute—University of North Carolina Evidence-based Practice Center under contract n.º 290-97-0011). AHRQ Publication n.º 02-E016. Rockville: Agency for Healthcare Research and Quality, 2002.
9. Jovell AJ, Navarro-Rubio MD. Evaluación de la evidencia científica. *Med Clin (Barc)* 1995;105:740-3.
10. Aymerich M, Jovell AJ, Estrada MD. Revisión sistemática de la evidencia científica. En: Jovell AJ, Aymerich M. Evidencia científica y toma de decisiones en sanidad. Barcelona: Monografies Mèdiques de l'Acadèmia de Ciències Mèdiques de Catalunya i de Balears, 1999; p. 93-105.
11. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1993;270:2598-601.
12. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1994;271:389-91.
13. Levine M, Walter S, Lee H, Haines T, Holbrook A, Moyer V. Users' guides to the medical literature. IV. How to use an article about harm. Evidence-Based Medicine Working Group. *JAMA* 1994;271:1615-9.
14. Laupacis A, Wells G, Richardson WS, Tugwell P. Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-Based Medicine Working Group. *JAMA* 1994;272:234-7.
15. Sacristán JA, Soto J, Galende I. Evaluación crítica de ensayos clínicos. *Med Clin (Barc)* 1993;100:780-7.
16. Chalmers TC, Smith H, Blackburn B, Silverman B, Schroeder B, Reitman D, et al. A method for assessing the quality of a randomized control trial. *Controlled Clin Trials* 1981;2:31-49.
17. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds JM, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clin Trials* 1996;17:1-12.
18. Sackett DL, Wennberg JE. Choosing the best research design for each question. *Bred Med J* 1997;315:1636.
19. Starfield B. Quality-of-care research: internal elegance and external relevance. *JAMA* 1998;16;280:1006-8.
20. Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C. Threats to applicability of randomised trials: exclusions and selective participation. *J Health Serv Res Policy* 1999;4:112-21.
21. Harbour R, Miller J, for the Scottish Intercollegiate Guidelines Network Grading Review Group. A new system for grading recommendations in evidence based guidelines. *BMJ* 2001;323:334-6.
22. OncoGuies de limfoma B difús de cèl·lules grans, de pulmó, de colon i recte, i de mama. Agència d'Avaluació de Tecnologia i Recerca Mèdiques. CatSalut. Barcelona: Departament de Sanitat i Seguretat Social. Generalitat de Catalunya, novembre 2003 (OG01/2003, OG02/2003, OG03/2003 i OG04/2003).
23. Freemantle N. Implementation strategies. *Family Practice* 2000; 17:S7-S11.
24. Thorsen T, Mäkelä M, editors. Cambiar la práctica de los profesionales. Teoría y práctica de la implementación de la Guías



- de Práctica Clínica. Sevilla: Agencia de Evaluación de Tecnologías Sanitarias de Andalucía, 2002.
25. Field MJ, Lohr KN, editors. Guidelines for Clinical Practice: from development to use. Committee on Clinical Practice Guidelines, Institute of Medicine. Washington; DC. National Academy Press, 1992. Disponible en: <http://www.nap.edu/books/0309045894/html/index.html>
  26. The Appraisal of Guidelines, Research, and Evaluation in Europe (AGREE) Collaborative Group. Guideline Development in Europe: an international comparison. Int J Technol Assess Health Care 2000;16:1039-49. Disponible en: <http://www.agreecollaboration.org>
  27. Graham ID, Calder LA, Hebert PC, Carter AO, Tetroe JM. A comparison of clinical practice guideline appraisal instruments. Int J Technol Assess Health Care 2000;16:1024-38.
  28. 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;20:259-66.
  29. Disponible en: <http://www.sign.ac.uk>
  30. Disponible en: <http://www.nice.org.uk>
- 

La revista GACETA SANITARIA ha concedido sendas becas de inscripción para el Curso Get Your Paper Written and Published, impartido por Tim Albert y organizado por el Departamento de Salud Pública de la Universidad d'Alacant, a Dña. Laura Inés González Zapata y a Dña. Pamela Eliana Pereyra Zamora. Nuestra enhorabuena a las candidatas elegidas y nuestro agradecimiento a todos los que se han interesado por esta convocatoria.