

---

**GUÍA PARA LA ELABORACIÓN  
DE INFORMES  
DE EVALUACIÓN  
DE TECNOLOGÍAS SANITARIAS**

---

Informe de Evaluación de  
Tecnologías Sanitarias Nº 19  
Madrid, Junio de 1999

**Agencia de Evaluación de Tecnologías Sanitarias (AETS)  
Instituto de Salud Carlos III  
Ministerio de Sanidad y Consumo**

*Sinesio Delgado, 6 - Pabellón 3  
28029 MADRID (SPAIN)  
Tels.: 91 387 78 40 - 91 387 78 00  
Fax: 91 387 78 41*

1.ª edición, Junio 1999  
2.ª edición, Noviembre 1999

Edita: AGENCIA DE EVALUACIÓN DE TECNOLOGÍAS SANITARIAS  
Instituto de Salud Carlos III - Ministerio de Sanidad y Consumo

N.I.P.O.: 354-99-006-X  
I.S.B.N.: 84-930379-2-3  
Depósito Legal: M-45380-1999

Imprime: Rumagraf, S.A. Avda. Pedro Díez, 25. 28019 Madrid

Este documento es un Informe Técnico de la Agencia de Evaluación de Tecnologías Sanitarias (AETS) del Instituto de Salud Carlos III del Ministerio de Sanidad y Consumo.

**Dirección y coordinación:**  
*José L. Conde Olasagasti*

**Elaboración y redacción:**  
*Iñaki Imaz Iglesia*  
*Jesús González Enríquez*  
*Juan Fco. Alcaide Jiménez*

**Edición y disseminación:**  
*Antonio Hernández Torres*

**Para citar este informe:**

Agencia de Evaluación de Tecnologías Sanitarias (AETS)  
Instituto de Salud Carlos III - Ministerio de Sanidad y Consumo  
«*Guía para la Elaboración de Informes de Evaluación de Tecnologías Sanitarias*»  
Madrid: AETS - Instituto de Salud Carlos III, *Junio de 1999*

Este texto puede ser reproducido siempre que se cite su procedencia.



---

# Presentación

El propósito de este trabajo es servir de orientación para la elaboración de informes de Evaluación de Tecnologías Sanitarias (ETS), tratando de clarificar y hacer explícitas las distintas etapas de su elaboración y los criterios utilizados para la valoración y selección de la información válida. A su vez, pretende describir las herramientas metodológicas más empleadas en la síntesis de resultados y ofrecer una recopilación de los recursos más eficientes para realizar un trabajo de evaluación sanitaria. El documento adopta una aproximación general, sintética y eminentemente práctica, haciendo énfasis en lo esencial y prioritario de cada una de las etapas de elaboración, en el contexto local y la adaptación de la respuesta a la necesidad del solicitante.

Se abordan primero los conceptos generales que definen la ETS. A continuación se refieren los tipos, métodos y proceso de elaboración de los informes, teniendo en cuenta que éstos no están completamente establecidos y sus contenidos, estructura y formato de presentación no han sido consensuados. El propósito es estimular la discusión acerca de cuáles son los elementos que deben conformar un informe de ETS y la posible unificación de un esquema de referencia para este tipo de producto. El último apartado presenta en forma de anexos una serie de herramientas útiles para la realización de informes: una propuesta de estructura o formato; una selección de fuentes de información útiles para la ETS, y varios esquemas para la valoración y gradación de la evidencia científica.

La realización de revisiones sistemáticas que promueve la Colaboración Cochrane, así como el fenómeno de la Medicina Basada en la Evidencia, buscan la mejora de la asistencia sanitaria a través del aumento de la cantidad y calidad de las evidencias científicas, basándose siempre en la búsqueda exhaustiva de la mejor información. La ETS se sitúa en ese mismo escenario y se nutre del trabajo de esas líneas de investigación, además de otras como la investigación en resultados, la investigación en servicios sanitarios, la evaluación económica o la epidemiología clínica.

Todos estos campos han producido en los últimos años un gran volumen de publicaciones sobre metodología aplicable a la ETS, pero no sobre la elaboración de informes o el proceso de realización de un trabajo evaluativo. El presente trabajo pretende aportar ese elemento diferencial. No es una revisión exhaustiva de los procedimientos y métodos de la ETS, sino que, como el término *guía* indica, pretende ser un documento sencillo, que facilite el empleo de la metodología y recursos propios de la ETS.

Su preparación, edición y difusión se enmarcan en el conjunto de actividades de promoción de la ETS que desarrollan el Instituto de Salud «Carlos III» del Ministerio de Sanidad y Consumo de España y la Organización Panamericana de la Salud (OPS) a través de un Convenio Marco de colaboración firmado entre ambas instituciones.



# Índice

	Pág.
INAHTA STRUCTURED ABSTRACT .....	9
INTRODUCCIÓN .....	11
PROCESO DE ELABORACIÓN DE INFORMES DE ETS .....	15
1. Tipos de documentos de ETS .....	16
1.1. Informes de ETS .....	16
1.2. Guías de Práctica Clínica .....	16
1.3. Dictámenes técnicos .....	16
2. Solicitud del informe .....	16
2.1. Priorización en la elaboración de informes de ETS .....	17
3. Diseño de la evaluación .....	17
3.1. Definición de las preguntas de evaluación .....	17
3.2. Estudio de la factibilidad .....	20
3.3. La elaboración del proyecto de evaluación .....	20
4. Búsqueda de la información .....	22
5. Evaluación crítica y selección de los estudios identificados .....	24
5.1. La evaluación crítica de los estudios identificados .....	24
5.2. Selección de los estudios .....	26
5.3. Extracción de datos de los estudios seleccionados .....	27
6. Estudio del contexto local .....	28
7. Técnicas de síntesis de la información y otros métodos utilizados en ETS .....	29
7.1. Síntesis cualitativa y narrativa .....	29
7.2. Meta-análisis .....	30
7.3. Análisis de decisión .....	34
7.4. Estudios de evaluación económica .....	37
7.5. El juicio de expertos .....	40
8. Valoración de los efectos sociales, organizativos, psicológicos y éticos .....	41
9. Elaboración de las conclusiones y recomendaciones .....	42
10. Revisión del documento provisional .....	44
11. Difusión y aplicación de los resultados del informe .....	44
ANEXOS .....	47
Anexo I. Formato y estructura de un informe de ETS .....	47
Anexo II. Recopilación de fuentes de información útiles en ETS .....	50
Anexo III. Esquemas para la interpretación de la evidencia científica .....	55
Anexo IV. Clasificaciones del grado de evidencia científica .....	59
REFERENCIAS .....	63



---

# INAHTA Structured Abstract

**GUIDE FOR PERFORMING HEALTH TECHNOLOGIES ASSESSMENT REPORTS (GUÍA PARA LA ELABORACIÓN DE INFORMES DE EVALUACIÓN DE TECNOLOGÍAS SANITARIAS).**

**Authors:** Imaz-Iglesia I, González-Enríquez J, Alcaide-Jiménez JF, Conde-Olasagasti JL. **Organization:** AETS (Agencia de Evaluación de Tecnologías Sanitarias) (Spanish Health Technology Assessment Agency). **Correspondence Address:** Imaz-Iglesia, Iñaki. Agencia de Evaluación de Tecnologías Sanitarias, Instituto de Salud "Carlos III", c/ Sinesio Delgado, 6 Pab. 3, 28029 Madrid, Spain. **Date:** June of 1999. **Language:** Spanish. **Summary:** No **Publication type:** Methodological guidelines. **Pages:** 71 **References:** 232. **ISBN:** 84-930379-2-3. **Subject index terms:** Health Technology Assessment-methods-guidelines.

**Objectives:** To describe the process of performing Health Technologies Assessment (HTA) reports. To offer a selection of useful sources of information, checklist for critical appraisal of medical literature and other tools for evaluating technologies in health care. To propose a standardized structure for HTA reports. This guide intends to offer tools and proposals for improving the quality and impact of HTA reports, as well as enhancing discussions among the organizations involved in HTA in order to obtain agreements about the structure and practice of elaborating HTA reports.

**Results/Content of the report:** The guide begins with an introduction of the justifications and concepts of HTA. The description of the steps involved in the creation of HTA reports follows. This section sets forth the methods used in HTA with practical appraisal and emphasizes the essential aspects of each method and components of HTA. The process of elaborating a HTA report starts with the

formal request or with the identification of priorities.

The design of the evaluation project includes the study of feasibility, and the formulation of the scientific question. The next step is a systematic and exhaustive bibliographic search and critical reading of the identified articles. The study of the local context and the applicability of the results are essential contents in all HTA reports. The guide explains the methods for synthesis and analysis of information and other methods useful for HTA. Qualitative synthesis, Meta-analysis, Decision Analysis, Economic Analysis and the Consensus Methodology are the main evaluative methods discussed. Moreover, the HTA must assess the social, psychological, ethical, and organizational effects.

This guide offers tools to make HTA reports, including a selection of useful sources of information, checklist for critical appraisal of medical literature, several scales of quality of medical evidence, a format of the HTA reports, a description of the types of HTA documents and other tools for evaluating technologies in health care.

In the last few years many papers concerning methodology applied to HTA have been published, but few or none have been published regarding the elaboration of reports or the process of HTA work. This document tries to cover this subject. It is not an exhaustive, nor in depth description of the HTA procedures and methods, but rather, it seeks to be a guide to aid in the use of HTA methodology and resources. On the other hand, taking into account that HTA reports are not well defined publications, this could contribute to discuss their contents, methodology and format.



# Introducción

Los proveedores y financiadores de servicios sanitarios, en la búsqueda de mayor efectividad y eficiencia de dichos servicios, están cada vez más interesados en disponer de evidencias explícitas sobre la efectividad y coste-efectividad de las tecnologías y procedimientos que proveen o financian. Por otra parte, los ciudadanos de los países desarrollados se ven sometidos en la actualidad a dos elementos en conflicto: la oferta de un número creciente de nuevos procedimientos médicos cada vez más sofisticados y prometedores, y el aviso continuado de las autoridades económicas acerca del crecimiento del gasto sanitario y las dificultades para afrontarlo. Esta situación ha sido analizada por profesionales sanitarios y políticos, llegando a constatar tres fenómenos que configuran un panorama de incertidumbres, que en última instancia pueden ser considerados generadores de la ETS o del movimiento llamado "Medicina Basada en la Evidencia"<sup>1</sup>.

El primero de estos fenómenos es *la ausencia de una buena correlación entre el nivel de gasto sanitario de un país y la salud colectiva de sus ciudadanos*. Así, por ejemplo, Grecia, con el gasto sanitario *per capita* más bajo de la Unión Europea, logra una mayor esperanza de vida en hombres que Alemania o Francia<sup>2</sup>. Los resultados agregados de salud y bienestar de una población son poco sensibles a los cambios en el gasto sanitario a partir de unos niveles mínimos de cobertura de riesgos, alcanzados por la mayoría de los países desarrollados. El sistema sanitario y la inversión en tecnología sanitaria constituyen sólo uno de los determinantes de la salud y el bienestar (junto a factores de la biología humana, medio ambientales, protección social o educación, entre otros).

El segundo es *la variabilidad observada en la práctica clínica*. Fenómeno descrito en los años treinta<sup>3</sup>, y estudiado con detalle a partir de la década de los setenta<sup>4</sup>. Se han realizado multitud de estudios. Unos han descrito diferencias importantes en las tasas de intervenciones realizadas a pacientes según el lugar de residencia. Por ejemplo, Mark Chassin describió variaciones en las tasas de implantación de prótesis de cadera de hasta 6 veces entre 13 Estados de los EE.UU. La diferencia aumen-

taba a 26 veces para el caso de las esclerosis hemorroidales<sup>5</sup>. Otros han descrito diferencias en las tasas de intervención entre hospitales de la misma zona geográfica. Por ejemplo, en 1984 se realizó un estudio que medía la probabilidad de que a un niño residente en Vermont (EE.UU.) le fuera practicada una amigdalectomía durante su infancia, y observó que ésta difería hasta en 9 veces en dependencia del hospital en el que le correspondiera la atención sanitaria, siempre dentro del mismo Estado<sup>6</sup>.

A pesar de los estudios realizados, los determinantes analizados (razones médicas, sociales, de necesidad) no llegan a explicar toda la variabilidad observada en la práctica clínica. Otra de las posibles explicaciones es el uso inadecuado de recursos. Se han detectado diferencias entre países atribuidas no tanto a los criterios de indicación de la técnica, sino a la conducta del profesional en caso de incertidumbre, al estilo de práctica clínica y a cuestiones organizativas<sup>7,8</sup>. En numerosas áreas asistenciales se detecta una infrautilización de procedimientos altamente eficaces y apropiados y distintos tipos de barreras que limitan el acceso a dichos procedimientos.

Por otra parte, los cambios en la práctica clínica no obedecen simplemente a la disponibilidad de información válida sobre efectividad y eficiencia de las técnicas utilizadas. Las fuentes de variación en la práctica clínica superan en mucho a la mera incertidumbre profesional sobre la efectividad de la práctica médica, si bien el juicio de los médicos puede ser el principal factor de variabilidad. Sin embargo, es probable que las recomendaciones basadas en evidencias de alta calidad puedan modificar dicha práctica más que las recomendaciones basadas en otros procedimientos. Todo ello hace necesarios los estudios evaluativos que, basados en la mejor evidencia disponible, contribuyan a disminuir la variabilidad inexplicable e inaceptable en la práctica clínica<sup>9</sup>.

El tercer y último fenómeno es *la incertidumbre acerca del impacto real que sobre la salud de los ciudadanos tiene el uso de muchas prácticas médicas*. Algunas estimaciones sitúan el gasto sanitario dedicado a procedimientos ineficaces

o perjudiciales en un 20% del gasto total. También se ha señalado que un 20-25% de los procedimientos médicos realizados en los EE.UU. están basados en indicaciones inapropiadas o dudosas<sup>10</sup>. La constatación de la falta de eficacia y efectividad de muchas intervenciones médicas consideradas clásicas y ampliamente extendidas en la mayoría de los países desarrollados ha generado un importante debate, que, para algunos autores, ha producido una auténtica revolución en la investigación biomédica<sup>11</sup>.

En conclusión, el rápido incremento del gasto sanitario no acompañado de mejoras proporcionales en el nivel de salud, el alto grado de variabilidad no explicable en la práctica médica y la incertidumbre acerca del impacto real de muchos procedimientos médicos sobre la salud, conducen a admitir la existencia de importantes lagunas en el conocimiento que dificultan la adopción de decisiones juiciosas sobre asignación de recursos y uso y aceptación de prácticas médicas. Esta realidad es la que ha llevado al nacimiento y desarrollo del concepto de ETS<sup>1</sup>.

---

## Concepto de Tecnología Sanitaria

El término Tecnología Sanitaria (TS) se identifica en muchos casos con equipamientos o técnicas complejas, pesadas y sofisticadas. Sin embargo, el Diccionario de la Real Academia Española define tecnología como el "conjunto de conocimientos propios de un oficio mecánico o industrial". Definición que, si se hace extensiva a la práctica de la medicina, lleva a considerar las TS como algo mucho más amplio que los equipamientos pesados.

En general, hoy se acepta la definición que en su día propuso el Institute of Medicine de los EE.UU.: "conjunto de medicamentos, dispositivos y procedimientos médicos o quirúrgicos usados en la atención sanitaria, así como los sistemas organizativos y de soporte dentro de los cuales se proporciona dicha atención"<sup>12</sup>.

---

## Concepto de Evaluación de Tecnologías Sanitarias

El concepto de ETS fue definido inicialmente por la Office of Technology Assessment de los EE.UU. (OTA) como aquella "forma de investi-

gación que examina las consecuencias clínicas, económicas y sociales derivadas del uso de la tecnología, incluyendo el corto y medio plazo, así como los efectos directos e indirectos, deseados e indeseados"<sup>12</sup>. Esta definición utiliza el término *investigación* como casi sinónimo de *evaluación*, lo cual ha sido revisado por algunos autores que consideran más conveniente utilizar términos amplios como "estrategia o proceso de análisis"<sup>13</sup>. Por otro lado, según la Real Academia Española, evaluar es "estimar, apreciar, calcular el valor de una cosa". Así, se propone la siguiente definición de ETS: *proceso de análisis e investigación, dirigido a estimar el valor y contribución relativos de cada tecnología sanitaria a la mejora de la salud individual y colectiva, teniendo además en cuenta su impacto económico y social*.

---

## Requisitos y atributos de los informes de ETS

La propia naturaleza y funciones de los organismos dedicados a la Evaluación de Tecnologías Sanitarias obliga a que sus productos reúnan ciertas características que los hagan acordes y coherentes con su misión. Los atributos básicos deben ser la fundamentación en la evidencia científica, la utilidad para la toma de decisiones, la claridad en la redacción, la facilidad de comprensión de los documentos, y la oportunidad y adaptación al tiempo y ámbito de decisión.

1. *Fundamentación en la evidencia científica*. Las afirmaciones y contenidos de un informe de ETS han de apoyarse siempre y esencialmente en la evidencia científica, incorporando, en determinados casos, la aportación del juicio de expertos. Sólo la revisión exhaustiva y el análisis crítico de la evidencia disponible proporcionan la base sobre la que se pueden sustentar las conclusiones y recomendaciones de un informe de ETS. Dicho de otro modo, un informe de ETS ha de ser incuestionable en su rigor científico. Se podrán compartir o no sus conclusiones pero no rechazar las bases en que se apoya<sup>14</sup>.

2. *Utilidad para la toma de decisiones*. Tal vez sea éste el atributo principal y más claramente diferenciado de las publicaciones de ETS frente a otro tipo de textos científicos. En el caso de la comunicación científica, los autores están obligados a exponer la metodología utilizada y los resultados alcanzados en determinado experimento u observación clínica, transmitiendo un conocimiento que puede o

no servir para la toma de decisiones. Sin embargo, los informes de ETS siempre se realizan con la intención de que el resultado sirva a alguien (político, gestor, administrador, clínico, usuario) para tomar decisiones. No se pretende sustituir a quien tiene la responsabilidad de la decisión, sino facilitar el proceso proporcionando elementos de información relevantes<sup>14,15</sup>.

3. *Redacción clara y fácil de comprender.* Este requisito, importante en cualquier texto escrito, adquiere una particular relevancia en este ámbito si se tiene en cuenta el variado perfil de los posibles destinatarios al que se dirigen los informes. Los destinatarios de un informe sobre una TS no son única ni fundamentalmente los especialistas médicos que la utilizan. Otros profesionales sanitarios o no sanitarios pueden ser lectores de un informe de ETS y, por tanto, el lenguaje que se utilice debe adaptarse a las características de los usuarios, que en muchos casos van a ser múltiples y muy variados, lo cual obliga a un esfuerzo de redacción importante. La adecuada definición y precisión en los términos, el orden y la estructuración, la claridad expositiva, el rechazo de lo accesorio y la huida de los sobreentendidos, deben caracterizar el estilo de redacción de cualquier documento de ETS.

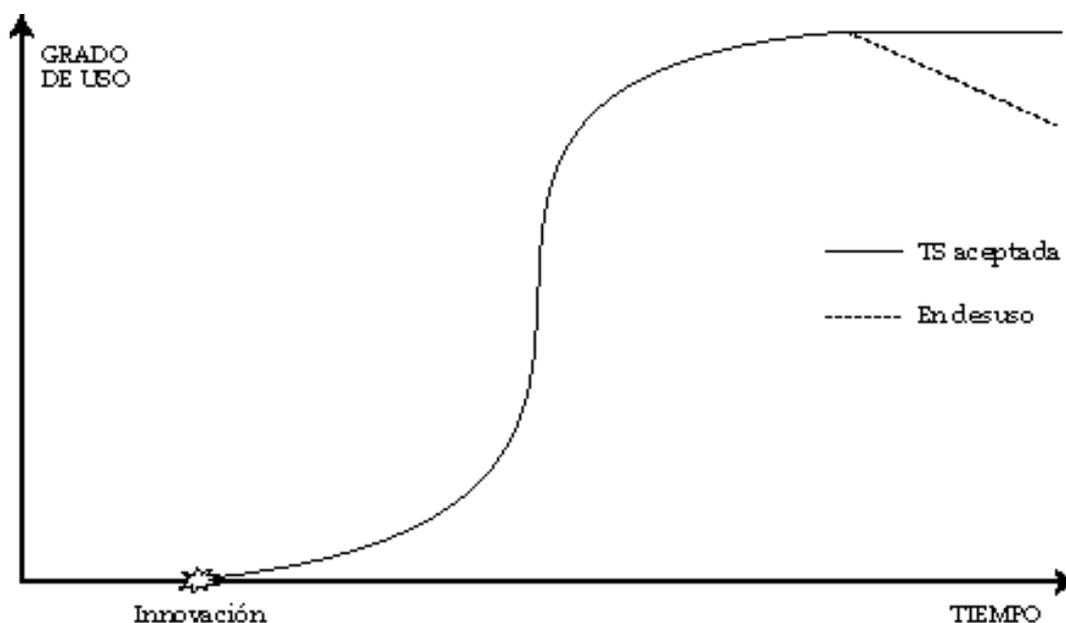
4. *Oportunidad y adaptación al tiempo y ámbito de decisión.* La capacidad de influencia e impacto de la información presentada en un informe de evaluación de TS dependerá del grado en que esta información se adapte al contexto en el que se generó la demanda. El

modo y grado de profundización en la respuesta deben estar también adaptados al marco temporal del proceso de toma de decisiones. Este marco debe estar siempre presente, de forma que exista cierta flexibilidad en la elaboración y formato de la respuesta, con el objetivo de maximizar el grado de oportunidad. El sacrificio de la exhaustividad puede venir justificado por un marco temporal limitado. La búsqueda rápida de información válida y relevante para una toma de decisión inmediata puede constituir el objetivo de un informe de ETS y, por tanto, esto tiene que quedar claro en los primeros momentos tras la solicitud.

## Fases de introducción de una TS

El principal de los atributos mencionados anteriormente, utilidad para la toma de decisiones, tiene un componente importante que se relaciona con la evolución en el tiempo del uso de una tecnología. El proceso de ETS debe considerar el momento y la oportunidad de elaborar un informe, valorando la utilización y el grado de implantación de esa TS. En ese sentido es importante conocer las fases de la introducción de las tecnologías sanitarias en los servicios sanitarios, que se pueden conceptualizar a partir de un esquema aceptado de forma bastante generalizada, como es el de la figura 1.

Figura 1  
Fases de la evolución de una tecnología sanitaria (tomado de Banta<sup>16</sup>)



Las fases de introducción de una TS suelen ser cinco:

- Experimental. Es la fase en la que se suele realizar la innovación, investigación inicial y conceptualización de la TS.
- Introducción. Se inicia la implantación y uso de la TS.
- Expansión. Coincide con la parte más ascendente de la curva.
- Utilización generalizada. La curva comienza a aplanarse y si la TS es aceptada pasa a ser una meseta.
- Declive o sustitución.

Las primeras fases (experimental e introducción) son las más oportunas para la evaluación de la viabilidad técnica, eficacia y seguridad. En estas fases no suelen intervenir los organismos o agencias dedicados a

la ETS. Generalmente, corresponde al investigador o a la industria proporcionar los primeros datos sobre la TS. Sin embargo, durante las fases de introducción y expansión, cuando ya es posible conocer la efectividad, utilidad clínica o impacto económico de una TS, un informe de ETS puede producir un impacto muy importante que logre modificar la curva de introducción de la tecnología. Éste es, por tanto, el momento más eficiente para realizar un trabajo de evaluación.

Por último, durante la fase de utilización generalizada es todavía más útil el estudio de los efectos, al poder medir los resultados e impacto de la tecnología a medio y largo plazo. Las evaluaciones realizadas en esta fase proporcionan datos muy fidedignos y consistentes sobre los efectos reales de la técnica o intervención.

# Proceso de elaboración de informes de ETS

A continuación se estructuran y describen los elementos a considerar en un informe de ETS, estableciendo la secuencia temporal de elaboración, los tipos de informes y los aspectos metodológicos más destacados, a la vez que se ofrecen recursos y bibliografía relacionada. No se pretende realizar una descripción pormenorizada de las técnicas utilizadas en ETS, sino ofrecer una guía de apoyo para la elaboración de informes; por ello, algunos temas sólo son mencionados, aunque se presentan fuentes bibliográficas y otros recursos que permiten profundizar en ellos.

Algunas de las cuestiones planteadas se encuentran actualmente sometidas a debate. En primer lugar, la necesidad de nuevas publicaciones que atiendan a sintetizar la gran cantidad de información disponible, o la influencia creciente del movimiento de Medicina Basada en la Evidencia, línea de colaboración internacional que promueve la mejor utilización de los conocimientos médicos<sup>17,18</sup>. Existen diversos tipos de trabajos que se pueden situar en ese marco común: los *informes de ETS*, las *revisiones sistemáticas* o los *meta-análisis*, entre otros. La descripción somera de las diferencias entre ellos puede servir para caracterizar los informes de ETS y comprender su papel particular.

Las *revisiones sistemáticas* son trabajos que pretenden resumir toda la evidencia que existe sobre un tema concreto, generalmente sobre los efectos de una intervención sanitaria. Deben realizarse con una metodología explícita, ordenada y reproducible, permitiendo establecer las bases científicas sobre las que se sustenta el uso de determinadas técnicas o procedimientos. La utilización de métodos explícitos y sistemáticos limita los sesgos o errores sistemáticos y reduce los errores aleatorios o simples<sup>19</sup>.

El *meta-análisis* es una técnica de síntesis cuantitativa de resultados de investigaciones primarias, cuyo objetivo es proporcionar estimaciones más precisas sobre los efectos de la atención sanitaria que las derivadas de estudios aislados incluidos en una revisión<sup>20</sup>. Las revisiones sistemáticas pueden incorporar un meta-análisis o no hacerlo.

Los *informes de ETS*, sin embargo, son publicaciones que utilizan habitualmente revisiones

sistemáticas (con o sin meta-análisis), entre otros recursos, para determinar la posible contribución de una TS a la mejora de la salud en una población concreta. Su objetivo es responder a solicitudes de terceros (Administración, proveedores de servicios, aseguradores) que precisan información fiable sobre los resultados o impacto de una TS y su posible aplicación en un contexto determinado. En consecuencia, su metodología debe ajustarse al tipo de pregunta planteada y al tiempo requerido para su respuesta. La formulación de conclusiones y recomendaciones aplicables a un ámbito local determinado es otro de los aspectos que diferencia los informes de ETS de las revisiones sistemáticas o los meta-análisis. En los informes de ETS casi siempre es necesario realizar un estudio del contexto local que permita que el solicitante del informe pueda tomar una decisión sobre la incorporación o no de esa TS.

Los responsables de tomar decisiones en los sistemas sanitarios (autoridades, administradores, profesionales) están interesados en disponer de información válida, relevante, a veces en un corto espacio de tiempo y aplicable al contexto en el que actúa. La ETS pretende cubrir esas necesidades, por lo que los informes de evaluación pueden ir consolidándose como publicaciones de gran impacto y utilidad. Esta guía, que describe la metodología, proceso de elaboración y tipos de informes, pretende difundir la práctica de la ETS, así como provocar debate e investigación para el establecimiento de los beneficios y limitaciones de la metodología evaluativa en ciencias de la salud.

**El informe de Evaluación de Tecnologías Sanitarias** se puede definir como un documento estructurado, comprensible, basado en la mejor evidencia disponible, elaborado con metodología explícita y reproducible y que responda a cuestiones relativas a la seguridad, eficacia, efectividad, utilidad, eficiencia o impacto de una tecnología en un contexto previamente definido. Su misión finalista es la ayuda en la toma de decisiones. Su elaboración requiere una metodología acorde con el carácter de la pregunta y el tiempo en el que ésta debe ser contestada.

---

# 1. Tipos de documentos de ETS

## 1.1. Informes de ETS

- *Informes a demanda o internos.* Se realizan en respuesta a solicitudes de parte interesada. Su objeto, extensión, metodología y circulación se ajustan a los términos que previamente se convengan entre las dos partes. La agencia actúa de segunda parte, es decir, elabora el informe para asesorar al solicitante en sus decisiones unilaterales. El producto de esa colaboración, el "informe a demanda", se orienta a satisfacer las cuestiones planteadas, por lo que suele ser un informe interno o de circulación supeditada al solicitante. En general, y debido a la necesidad urgente del solicitante, son informes de ejecución rápida y contenido breve, aunque no se debe asimilar los informes a demanda o internos con documentos breves, ya que en muchos casos no es así.

Algunas organizaciones como la Agencia para la ETS de Alberta, en Canadá, denominan a este tipo de productos consultas técnicas, *Health tech notes*, y las definen como respuestas breves a peticiones de asesoría urgentes que, por su limitación de tiempo, presentan un análisis más limitado y no están necesariamente sujetas a un proceso de revisión externa<sup>21</sup>. Esta organización ha evaluado la calidad de la información que proporcionan y su impacto en la toma de decisiones a través de entrevistas con solicitantes y lectores de los informes. Seis de los once informes considerados tuvieron un impacto elevado en la toma de decisiones, cuatro un impacto moderado pero fueron considerados útiles para la gestión, y únicamente un informe fue valorado como de bajo impacto. La calidad se consideró aceptable en todos ellos. El mismo estudio previene de las limitaciones de este tipo de informes, en el sentido de que la urgencia impuesta por los solicitantes puede perturbar el proceso de búsqueda de información y el análisis ulterior<sup>22</sup>.

- *Informes públicos.* Son los productos más conocidos de las agencias de ETS. En este caso, la agencia actúa de tercera parte, ofreciendo una valoración técnica útil para cualquiera de las partes presentes en el sistema sanitario. Dado que deben ser útiles a cualquiera de los operadores que interactúan en un Servicio de Salud, son públicos y de libre acceso a todo aquel que lo solicite. Su origen y motivación

pueden ser diversos, pero en todo caso la trascendencia de la cuestión que tratan ha aconsejado su difusión pública. Estos informes van a formar parte de la base de datos de informes de ETS de la red internacional INAH-TA y generalmente se elaboran por iniciativa de la Autoridad Sanitaria o de la propia agencia.

## 1.2. Guías de Práctica Clínica

Pretenden ayudar a profesionales sanitarios y usuarios a decidir sobre la atención sanitaria más apropiada en circunstancias clínicas específicas, mediante la formulación de unas recomendaciones desarrolladas sistemáticamente y presentadas de manera estructurada, frecuentemente con algoritmos o árboles de decisión<sup>23</sup>. Constituyen un instrumento para la difusión de los resultados de la evaluación de tecnologías sanitarias.

La elaboración de una Guía de Práctica Clínica debe perseguir la obtención de un texto basado en la evidencia científica y consensado. Sus características básicas deben ser: validez, fiabilidad, aplicabilidad clínica, flexibilidad clínica, multidisciplinariedad, previsión de revisión y documentación adecuada. La calidad de una Guía de Práctica Clínica se basa en la exhaustividad y valoración crítica de la literatura científica y en la adecuada valoración de la evidencia, que debe acompañar a cada recomendación incluida. Su descripción, proceso de elaboración, implantación y evaluación no han sido abordados en esta guía y merecen un tratamiento específico, aunque se pueden recomendar una serie de referencias que permiten conocer sus fundamentos<sup>23,24,25,26,27</sup>.

## 1.3. Dictámenes técnicos

Otra parte de la actividad de las agencias de ETS es la asesoría acerca de lo apropiado de la adopción de una tecnología en un contexto determinado (peritaje, audit). La respuesta a este tipo de consulta se materializa en informes breves o dictámenes sobre aspectos concretos de determinadas técnicas o procedimientos.

---

# 2. Solicitud del informe

El origen del proceso de elaboración de un informe de ETS suele ser la solicitud de evalua-

ción por parte de la administración sanitaria, instituciones u organismos usualmente proveedores o financiadores de TS. Algunos centros han elaborado hojas de solicitud de informe en las que se pide la identificación del solicitante, la formulación de preguntas concretas, información sobre el contexto local de aplicación de la TS, y además se establece el marco temporal y el formato de la respuesta. Este procedimiento puede facilitar y clarificar inicialmente el tipo de trabajo requerido y la previsión de recursos necesarios para su desarrollo.

## 2.1. Priorización en la elaboración de informes de ETS

Con independencia de la respuesta a demandas concretas, es importante que la realización de informes de ETS y su consiguiente planificación de recursos se haga a través de un proceso explícito y meditado. La ETS puede abordar una enorme cantidad de temas, siendo su prioridad, relevancia y oportunidad variables. Establecer prioridades, o sea, escoger los temas o aspectos específicos a estudiar, puede ser el primer paso en el ciclo de la ETS, que se continúa con la realización de la evaluación, la difusión y aplicación de los resultados y la medición de su impacto. Una metodología ideal para este paso de establecimiento de prioridades debería definir el alcance de los temas a priorizar desarrollando una lista de posibles temas a evaluar, los criterios explícitos para establecer prioridades, los métodos de aplicación de esos criterios y la inclusión de un amplio rango de perspectivas en el proceso, desde gestores, financiadores, investigadores, profesionales o pacientes<sup>28</sup>.

Algunas agencias han desarrollado métodos de priorización, que facilitan la planificación de su trabajo y a la vez posibilitan el análisis global de la situación de las TS en su contexto local. Las agencias de ETS españolas han realizado un estudio de priorización de las tecnologías susceptibles de evaluación o de investigación para el Consejo Interterritorial del Sistema Nacional de Salud español<sup>29</sup>. El propósito del estudio era ordenar un amplio grupo de TS en función de la necesidad de elaboración de un informe de ETS o de realización de investigación dirigida a cubrir las lagunas de conocimiento detectadas sobre los efectos de esa técnica o procedimiento.

El listado inicial incluía seiscientos tipos de técnicas o intervenciones sanitarias seleccio-

nadas en función de su relevancia, nivel de concreción y razonable disponibilidad de evidencias acerca de su efectividad y utilidad. Las fuentes de obtención de las TS iniciales fueron la red INAHTA (International Network of Agencies of Health Technology Assessment), las cuatro agencias de evaluación españolas, las propuestas presentadas por Consejerías de Sanidad de las Comunidades Autónomas españolas, el Instituto Nacional de la Salud y el Ministerio de Sanidad y Consumo españoles, y un estudio sobre priorización en TS realizado recientemente en el Reino Unido<sup>30</sup>.

Los criterios que se utilizaron para la priorización realizada por las agencias de ETS españolas fueron<sup>29</sup>:

- Prevalencia de la condición o situación clínica a la que se dirige la tecnología.
- Carga de la enfermedad. Grado de deterioro de la salud que ocasiona la condición clínica implicada. Los datos que pueden definir este concepto son los referentes a mortalidad, discapacidad, años de vida perdidos, jornadas laborales perdidas o calidad de vida, entre otros.
- Grado de incertidumbre. Se refiere a la existencia de información contradictoria, escasa y/o de baja calidad.
- Posible impacto sobre la salud. Efectos de la TS sobre el manejo y resultados clínicos para los pacientes, teniendo en cuenta las alternativas concurrentes.
- Posible impacto económico. Se valora si el uso de la tecnología eleva o reduce el coste agregado.
- Posible impacto organizativo. Se valora si el uso de la TS provoca cambios en las condiciones y organización del trabajo.
- Posible impacto ético, social y/o legal.

---

## 3. Diseño de la evaluación

### 3.1. Definición de las preguntas de evaluación

La elaboración de un informe de ETS requiere determinar desde el inicio el enfoque de la evaluación, estableciendo las preguntas con concreción. En muchas ocasiones, la formulación adecuada de la pregunta no resulta una tarea simple. La petición inicial o solicitud de

evaluación no suele estar planteada en términos suficientemente precisos y específicos; por ello se debe abrir un proceso de discusión con la institución o persona solicitante para delimitar preguntas, objetivos o hipótesis claramente definidos y operativos, que puedan ser objeto de investigación. En esta fase inicial, es útil la realización de una primera reunión con los solicitantes que clarifique los objetivos del informe.

La pregunta de evaluación se compone de un objetivo general o pregunta política y de objetivos específicos o preguntas científicas. El *objetivo general o pregunta política* constituye el motivo inicial que fundamenta la solicitud de un informe de ETS. Su naturaleza puede ir desde un problema o preocupación poco definidos de una autoridad o gestor sanitarios, hasta una pregunta muy concreta sobre su ámbito de gestión. Rara vez se formula inicialmente una cuestión que pueda responderse científicamente<sup>15</sup>.

*Objetivos específicos, preguntas científicas o preguntas de evaluación.* El objetivo general o pregunta política debe descomponerse en cuestiones más delimitadas y que puedan responderse con metodología científica. La formulación de los objetivos específicos va a determinar todo el proceso de evaluación, desde el diseño hasta su posible impacto en la práctica clínica. Durante el proceso de construcción de las preguntas es necesario plantearse cuál es la forma de presentar los resultados más útil y comprensible para los potenciales receptores del informe.

Éste es el momento de realizar una búsqueda y recuperación bibliográfica inicial sobre el tema a evaluar (el anexo II ofrece un listado de fuentes de información para la búsqueda preliminar), con el objeto de identificar los datos o elementos de información más relevantes en relación con las preguntas de investigación. Una posible caracterización y clasificación de los tipos de elementos a considerar en la formulación de la pregunta es<sup>31</sup>: a) datos referidos a las características de los participantes; b) a la tecnología; c) a la metodología, y d) a los resultados.

a) *Datos sobre las características de los participantes.* Las características que tengan influencia sobre el efecto o resultado clínico, o sobre su interpretación, son las que más interesan, y van a determinar en gran parte la validez de las inferencias. Por ejemplo, una característica de los participantes con gran influencia en la efectividad de las intervenciones es el grupo socioeconómico al que pertenecen. Sería el

caso de un informe que evaluara la efectividad de determinadas intervenciones contra la obesidad en Atención Primaria en España. Si estas intervenciones incluyen modificaciones de hábitos alimenticios que conlleven un incremento del gasto habitual en alimentación, los estudios que incorporen una proporción de personas de bajo nivel económico, más numerosa que la media española, obtendrán menor efectividad que aquellos en los que el nivel socioeconómico sea semejante a la media. En ese caso, la evaluación de esas intervenciones sobre hábitos alimenticios precisa que se recojan datos sobre el grupo socioeconómico al que pertenecen los participantes.

Otras características de los participantes con influencia en el efecto o resultado clínico son la naturaleza o determinados aspectos de la enfermedad implicada (por ejemplo, si el informe evalúa un tratamiento antihipertensivo, los niveles iniciales de presión arterial de los participantes antes de la intervención determinarán los resultados del estudio), el ámbito en el que se realiza el estudio (por ejemplo, un hospital de crónicos o un servicio de urgencias), la región geográfica, la edad, el sexo, el nivel de educación o la existencia de co-morbilidad.

b) *Los datos acerca de las características de la TS o intervención* a evaluar también condicionan la validez de los resultados. Por ejemplo, en los ensayos con tratamientos farmacológicos debe considerarse la vía de administración, las dosis y la duración del tratamiento. En los estudios que no comparan con placebo es importante, por ejemplo, registrar la información referente al tratamiento que recibió el grupo control, ya que si los tratamientos de los diferentes estudios son excesivamente heterogéneos no será razonable combinar sus resultados en un meta-análisis. Son múltiples las características de la TS o intervención que podrían enumerarse. Cada técnica tiene aspectos diferenciales que pueden influir en las variables a considerar cuando se redacta una pregunta de investigación. Un estudio exhaustivo de las características de la TS permitirá conocer todos los factores que pueden influir y modificar los resultados, costes y beneficios de la intervención sanitaria que se evalúa.

c) *Los datos sobre la metodología* de los estudios que se suelen incluir en las preguntas de evaluación son:

— Las características de las medidas de resultado que han utilizado los estudios revisados.

- La precisión de los resultados de los estudios.
- La técnica de asignación de pacientes a los grupos.
- La duración del estudio.
- El tipo de diseño, por ejemplo, ensayo aleatorio controlado o estudio observacional de cohortes.
- La existencia de enmascaramiento y en qué grado, ciego simple, doble o triple ciego.
- Las pérdidas de pacientes durante la realización del estudio.
- Las co-intervenciones.

d) *Las medidas de resultado* requeridas por la pregunta de investigación van a determinar totalmente la posibilidad de encontrar una respuesta válida. La elección del tipo de medida más idónea para responder a la pregunta científica es una tarea a la que hay que dedi-

car mucho tiempo y estudio cuando se diseña la evaluación. En la actualidad, el estudio de **las medidas de resultado** en medicina es una importante línea de investigación conocida también como investigación sobre la efectividad y la utilidad (*outcomes research*)<sup>32,33,34</sup>.

Un informe de ETS basado en medidas de resultado que describan efectos sobre la salud o utilidades permitirá establecer conclusiones sólidas, cuya validez difícilmente pueda ser cuestionada. Sin embargo, si se utilizan medidas que describan los resultados clínicos de forma indirecta o intermedia, tales como los cambios producidos en un parámetro bioquímico, pueden surgir dudas sobre la efectividad real de la tecnología evaluada<sup>35,36</sup>. Es importante llamar la atención acerca de la frecuente utilización de evidencias indirectas o *surrogate end-points*, que son, generalmente, parámetros fisiológicos o bioquímicos que forman parte de una cadena causal o de asociación con resultados o efectos clínicos (ver ejemplo en tabla 1).

**Tabla 1**  
**Ejemplo de utilización de una evidencia indirecta o surrogate end-point**

Las conclusiones del ensayo clínico llevado a cabo para evaluar la influencia que la reducción de las arritmias tiene sobre la mortalidad postinfarto (CAST, Cardiac Arrhythmia Suppression Trial)<sup>35</sup> son útiles para explicar el concepto de evidencia indirecta o *surrogate end-point*. La relación encontrada previamente entre las arritmias postinfarto y la muerte súbita<sup>37</sup> y los efectos de la encainida, flecainida y moricizina disminuyendo las arritmias cardíacas durante el año siguiente a un infarto agudo de miocardio<sup>38</sup> conforman una línea tratamiento, resultado intermedio - resultado final, que se inicia con el uso de antiarrítmicos, continúa con la reducción de las arritmias y debería finalizar en la disminución de la mortalidad postinfarto. En este caso, los efectos positivos observados en el paso intermedio o evidencia indirecta, la arritmia cardíaca, estaban justificando la utilización generalizada de antiarrítmicos en los pacientes con infarto de miocardio por una supuesta disminución de la mortalidad postinfarto no demostrada.

El citado ensayo clínico pretendía obtener evidencias que demostraran ese efecto sobre la mortalidad. Sin embargo, diez meses después de iniciarlo, el análisis de los resultados preliminares conllevó su interrupción inmediata. El motivo fue que se observó mayor mortalidad en los grupos que tomaban encainida y flecainida que en los grupos que tomaban placebo. Las curvas de supervivencia que compararon el grupo con tratamiento y el grupo placebo mostraron una mortalidad y una incidencia de infarto mucho mayor en el grupo que tomaba antiarrítmicos, tanto debido a arritmia ( $p=0,0004$ ) como a cualquier causa ( $p=0,001$ ). El ensayo CAST demostró los resultados negativos de una práctica profiláctica extendida, al medir sus efectos con una medida de resultado final (la mortalidad) y no una medida de resultado *indirecta*, como la disminución de arritmias.

La asunción de que las medidas intermedias predicen el resultado final es aventurada, por dos razones fundamentales. En primer lugar, porque la biología es extremadamente sensible a multitud de factores que difícilmente se pueden conocer en su totalidad y, por tanto,

las asociaciones biológicas demostradas en casos concretos no pueden ser utilizadas de manera general para predecir comportamientos clínicos complejos. Las teorías actuales sobre causalidad la definen, en la mayoría de los casos, como multifactorial. Se han descri-

to muchos factores causales necesarios, pero no muchos factores causales suficientes. Además, muchas veces se ha observado que la existencia de asociación, entre un resultado A, resultado intermedio, y uno B, resultado final, en condiciones de laboratorio o ideales, se diluye o desaparece completamente en las condiciones de la práctica habitual<sup>39</sup>.

### 3.2. Estudio de la factibilidad

Antes de iniciar el trabajo de evaluación es necesario valorar la viabilidad del proyecto y la posibilidad de responder a la solicitud. Es

posible que la evaluación esté ya total o parcialmente realizada, que exista información suficiente de fácil acceso para cubrir los objetivos propuestos, o, por el contrario, que los recursos de personal o de tiempo necesarios para la elaboración del informe sean muy superiores a los inicialmente previstos.

A continuación se presenta una propuesta de los aspectos que pueden determinar la factibilidad y oportunidad de realizar un informe de ETS. La Agencia de ETS de Alberta, en Canadá, tiene en cuenta varios criterios para decidir el tipo y grado de la respuesta a las solicitudes de evaluación que, parcialmente modificados, se exponen en la tabla 2.

Tabla 2  
Aspectos que determinan la factibilidad de realizar un informe de ETS  
(modificado de Alberta Heritage Foundation<sup>21</sup>)

- *El propósito de la solicitud.*
- *El impacto de la tecnología.* Prevalencia de la condición clínica implicada, gravedad de ésta e impacto económico, social y organizativo.
- *La pertinencia de realizar la evaluación.* Considerando el momento en el que se presentará el informe. Este criterio depende de aspectos como la sobre o infrautilización percibida o la fase de introducción en que se encuentra la TS.
- *La variabilidad en el uso de esa tecnología* o en el manejo de la enfermedad entre diferentes zonas geográficas o medios clínicos.
- *La disponibilidad de la información* necesaria para realizar la evaluación.
- *Calidad y grado de homogeneidad de la información* que se utilizará para la evaluación.
- *Los recursos* necesarios para realizar la evaluación y la capacidad para obtenerlos. Tiempo, recursos humanos, urgencia u oportunidad política de la solicitud.
- *La idoneidad* de que se responda aportando informes ya elaborados o fuentes de información relacionadas, o la posibilidad de que otros organismos lleven a cabo el proyecto.

En algunas ocasiones el estudio de factibilidad puede concluir con la conveniencia de elaborar un informe preliminar de forma rápida o un informe completo de ETS. Un informe preliminar puede constituir una aproximación inicial al conocimiento científico disponible sobre esa TS, siendo útil para la estimación de los recursos necesarios para culminar el proceso de ETS, o para realizar posteriormente, si así se decide, un informe de evaluación completo.

### 3.3. La elaboración del proyecto de evaluación

En los pasos anteriores se han definido las preguntas de evaluación, se ha estudiado

la factibilidad de realizar el informe y se ha realizado una búsqueda de información preliminar. En este momento ya se tiene información suficiente como para conocer el volumen de literatura que se va a manejar, diseñar la estrategia de revisión de la literatura y considerar la utilización de otros métodos. Se está en disposición de planificar la evaluación presentando un proyecto para la discusión con los solicitantes.

Los contenidos que debe incluir un protocolo o proyecto de evaluación pueden estructurarse tal y como se propone en la tabla 3.

No siempre es posible definir con precisión todos los puntos especificados en la tabla 3,

**Tabla 3**  
**Estructura de un proyecto de evaluación**

- + Definición precisa de las preguntas u objetivos generales y específicos de la evaluación.
- + Metodología que se utilizará para elaborar el informe: revisión sistemática, evaluación económica, métodos de consenso, encuesta sobre el uso actual de una TS, investigación con datos primarios u otros métodos.
- + Fuentes y estrategia de la búsqueda de información.
- + Criterios de inclusión y exclusión de los estudios, y criterios que se van a utilizar para la lectura crítica de los estudios.
- + Datos a extraer de los estudios y métodos que se emplearán para ello.
- + Estudio de los sesgos y limitaciones del proyecto.
- + Equipo de trabajo y, en su caso, métodos de selección de los expertos que van a participar.
- + Necesidades de personal, de tiempo y de presupuesto, para llevar a cabo la evaluación.
- + Estrategia de difusión del informe una vez elaborado.
- + Recopilación bibliográfica preliminar.

ya que son los estadios iniciales del proyecto y no se dispone de información suficiente. Por otra parte, existen otras posibles propuestas de esquema para el proyecto de evaluación. Por ejemplo, el Fondo de Investigación Sanitaria de la Seguridad Social en España exige una serie de requisitos para la financiación de proyectos de investigación que conforman toda una propuesta de protocolo<sup>40</sup>; además, se han publicado muchos trabajos con esquemas sobre cómo planear y conducir una investigación científica que son aplicables a la ETS<sup>41,42,43,44</sup>.

#### **La elección de la metodología a emplear**

Éste es el aspecto del proyecto de evaluación que más va a condicionar el resultado final y las posibles limitaciones del informe. Un informe de ETS puede utilizar un amplio rango de técnicas y métodos de análisis, algunos de los cuales se describen posteriormente en esta guía. Cada uno de esos métodos ha sido ampliamente estudiado y analizado, con sus fortalezas y debilidades. Algunos están en continuo debate y son muy controvertidos.

En muchos casos, los informes de ETS incluirán la utilización de varios métodos complementarios para la obtención de información que permita responder a las preguntas de evaluación. Por ejemplo, si el objetivo es comparar la eficiencia en un Servicio de Salud de una técnica diagnóstica nueva, A, con respecto a la antigua, o B, se puede realizar una *revisión sistemática* acerca de su eficacia que incluya *meta-análisis*. Las medidas resumen obtenidas se pueden incluir en un *análisis de decisión* que cuantifique el valor de las dos estrategias diagnósticas. Además, con los datos sobre los costes que ocasionan las dos técnicas se puede realizar un *estudio de coste-efectividad* que permita comparar su eficiencia. Todo ello puede ser completado con un estudio del contexto local en el que se van a aplicar esas tecnologías mediante una encuesta que valore la práctica habitual y disposición al cambio por parte de los profesionales.

El estudio y consideración de la metodología más adecuada para responder a la pregunta y elaborar unas conclusiones basadas en información de alta calidad requieren, generalmente, la mayor parte del esfuerzo que se emplea en elaborar el proyecto de evaluación.

Por otra parte, uno de los aspectos clave de la planificación de los recursos necesarios para elaborar el informe es el grado de exhaustividad de la búsqueda bibliográfica. Ésta es una tarea que siempre parece inacabada y, además, tiene rendimientos decrecientes. A partir de una determinada cantidad de trabajo, cada nueva unidad de resultado obtenida requiere recursos crecientes. La decisión sobre el tiempo y recursos a emplear en la elaboración del informe debe ser valorada en función del tipo de pregunta, los objetivos del trabajo, la exhaustividad necesaria para responder a la pregunta o los recursos disponibles.

Tras la presentación del proyecto de evaluación se debe convocar una reunión entre los solicitantes y los evaluadores para discutir y

aprobar el proyecto. En cualquier caso, la realización del proyecto debe ser un proceso flexible y abierto a la modificación y valoración continuas<sup>45</sup>.

## 4. Búsqueda de la información

La búsqueda exhaustiva y sistemática de la información requerida por las preguntas de evaluación constituye la base de todo informe de ETS. En los últimos tiempos, a la vez que han progresado los sistemas de identificación bibliográfica, acercándose a la exhaustividad a través de las búsquedas sistemáticas, ha aumentado exponencialmente la cantidad de información médica, de tal manera que estar actualizado cada vez cuesta más esfuerzo; es lo que se ha denominado sobrecarga de información, *information overload*<sup>46</sup>. Hace trece años ya se estimaba que la cantidad de artículos médicos que se publicaban al año rondaba los 2 millones en unas 2.000 revistas médicas<sup>47</sup>.

La búsqueda a través de bases de datos bibliográficas es actualmente imprescindible para realizar una buena revisión de la literatura médica. Las bases de datos electrónicas de amplia cobertura más utilizadas son MEDLINE y EMBASE, si bien el uso exclusivo de estas bases presenta limitaciones. Se ha estimado que MEDLINE sólo identifica el 30% de los ensayos clínicos randomizados y controlados en algunos temas científicos<sup>48</sup>. Otra limitación de las bases de datos proviene, paradójicamente, de su principal ventaja, ya que el sistema de palabras clave posibilita la identificación rápida de algunos artículos de un tema determinado pero no ofrece precisión. Existen múltiples interpretaciones sobre qué palabras clave definen mejor un tema de investigación y puede haber errores o inexactitudes en el proceso de elección de las palabras clave de un artículo por parte de los editores o autores. Además, hay que considerar la posibilidad de que la búsqueda esté afectada por el sesgo de publicación o por el sesgo debido al idioma del artículo.

La búsqueda a través de bases de datos bibliográficas es el mejor sistema para encontrar rápidamente unos pocos trabajos sobre un tema determinado, pero no permite realizar una búsqueda exhaustiva si es utilizado en solitario. Algunos estudios han estimado la capacidad de las bases de datos MEDLINE y EMBASE, concluyendo que su utilización con-

junto mejora mucho la capacidad de identificación de literatura y la precisión de la búsqueda. La base de datos americana, MEDLINE, es más accesible y menos costosa, pero se ha estimado que el solapamiento, o sea, los resultados comunes entre los dos sistemas de búsqueda, es de un 34%, con lo que la mayoría de los artículos sólo están en una de las dos bases de datos<sup>49</sup>.

La estrategia de búsqueda es un paso clave, y a la vez complejo, que merece un tiempo importante que asegure que la revisión bibliográfica realizada para el informe sea sistemática, exhaustiva y precisa. La utilización de bases de datos electrónicas permite diseñar una estrategia de búsqueda sistemática y fácilmente reproducible, disminuyendo los sesgos. La exhaustividad y la precisión dependen en gran parte de la selección de las palabras clave.

Estas dos cualidades, asimilables a la sensibilidad y especificidad, generalmente son inversamente proporcionales; es decir, a medida que aumentamos la exhaustividad de una búsqueda, reducimos su precisión. El intento de obtener un número elevado de trabajos sobre un tema puede producir el denominado *ruido de información*, o aumento del volumen de publicaciones no relevantes para el usuario. Por el contrario, una búsqueda que tienda a la precisión producirá el efecto contrario, *silencio de información*, los resultados de la búsqueda pueden omitir documentos relevantes que no se ajustan a la estrategia de búsqueda, que resulta excesivamente selectiva<sup>50</sup>. Por tanto, una buena búsqueda bibliográfica debe combinar estrategias que consigan un resultado equilibrado en exhaustividad y precisión.

Asimismo, el período de tiempo cubierto por la búsqueda y la fecha en que se realiza influyen considerablemente en el resultado final y, por tanto, deben describirse en la metodología del informe. Un sesgo habitual es el de selección de referencias o *reference bias*, que consiste en que habitualmente se presta mayor atención a las referencias recuperadas que apoyan las tesis de la persona que realiza la búsqueda. Para evitarlo se debe utilizar una estrategia de selección y análisis de los artículos lo más explícita posible<sup>51</sup>.

En el anexo II se presenta una relación de las principales fuentes de información útiles en ETS con su dirección o modo de acceso. Por otra parte, a continuación se clasifican las formas de búsqueda de información que deben considerarse en el proceso de elaboración de un informe de ETS<sup>31,52</sup>.

- *Bases de datos electrónicas.* Existen multitud de bases de datos de investigaciones originales, además de MEDLINE y EMBASE, que facilitan la identificación de trabajos sobre tecnologías sanitarias. Por otra parte, en los últimos años han proliferado las publicaciones secundarias o las bases de datos sobre revisiones, entre las que destacan las que ofrece la Cochrane Library, que en el proceso de elaboración de informes de ETS son de consulta obligada, y a veces previa, a las bases de datos de investigaciones originales.
- *Informes de otras agencias de ETS.* Muchas agencias se encuentran agrupadas en la red INAHTA, que promueve una base de datos con los informes de ETS realizados por sus miembros.
- *Fuentes de información estadística y/o administrativa.* La aplicabilidad de los resultados del proceso de ETS va a depender en gran parte de los contenidos del estudio del contexto local que todo informe debe incluir. Este estudio se nutre de este tipo de fuentes de información. La disponibilidad de registros con los que poder realizar un ajuste al medio local, permitirá dotar de mayor validez a las conclusiones del informe.
- *Consulta de libros o revisiones sobre el tema.* En muchas ocasiones, la mejor manera de iniciar un trabajo de investigación y localizar la información más relevante y aceptada es la lectura de uno o varios libros de texto generales, continuando por un buen artículo de revisión.
- *Consulta en revistas especializadas.* Cuando sea posible, la revisión manual de los índices de revistas relacionadas con el tema estudiado durante un período de tiempo determinado, suele contribuir a la localización de originales. Este método permite la recuperación de artículos no identificados a través de una búsqueda con palabras clave. La inevitable subjetividad o posibles errores durante el proceso de asignación de palabras clave a los artículos hacen que las búsquedas manuales sean una valiosa herramienta. Este método, además, permite localizar las publicaciones más recientes que aún no hayan sido incluidas en las bases de datos.
- *Búsqueda a través de las referencias de los artículos encontrados.* Es un método que suele resultar fructífero, aunque también laborioso. Una buena fuente para realizar este trabajo son las revisiones, ya que suelen presentar un gran número de citas.
- *Búsqueda de artículos publicados por un autor relevante.* Se puede realizar fácilmente mediante el empleo de bases de datos electrónicas. Es útil cuando se trata de autores muy significativos en determinadas áreas o disciplinas, o también cuando se estudian técnicas muy novedosas que aún no se han difundido ampliamente entre la comunidad científica.
- *Búsqueda en actas de congresos.* Pueden ser una fuente de información muy temprana y útil sobre investigaciones recién terminadas o aún en progreso, cuya localización se ve facilitada por el desarrollo actual de bases de datos electrónicas que las incluyen. Si bien los resúmenes de las actas de congresos generalmente son una fuente poco fiable de información, al carecer casi siempre de datos importantes, son una buena fuente de localización de autores y trabajos.
- *Búsqueda de literatura gris.* El concepto de literatura gris es aún algo impreciso, a pesar de que su uso está cada vez más extendido. En general, se entiende por tal un conjunto de documentos de tirada limitada y circulación restringida, que no pueden obtenerse por los canales habituales de distribución. Son materiales que pueden aportar resultados de un considerable valor, tales como libros, folletos, informes técnicos de administraciones y empresas públicas o privadas, actas de congresos, tesis doctorales y otros<sup>53,54</sup>. Aunque su localización exhaustiva es difícil, actualmente existen bases de datos que facilitan su localización. Las sociedades profesionales y los organismos de investigación pueden ser otra interesante fuente de "información gris".
- *Consulta a personas clave.* La consulta a investigadores o profesionales destacados es un método de obtención de información muy útil para conseguir estudios recientes que todavía no han sido publicados o estudios antiguos que nunca llegaron a publicarse. Además, los documentos de literatura gris son conocidos por muchos profesionales que trabajan en ese campo, pero son muy difíciles de identificar mediante los sistemas de búsqueda clásicos. La consulta a personas clave es una manera muy eficaz de disminuir el sesgo de publicación. Dirigiéndose por carta a los autores de algunos de los estudios relevantes recuperados o a

un grupo de expertos, enviándoles una lista de los estudios identificados y preguntándoles si conocen otros artículos o trabajos diferentes a éstos.

- *Solicitud de información a las casas comerciales relacionadas con la tecnología.* Generalmente, los primeros datos y estudios sobre una TS los realiza y conoce la industria o empresa interesada en su utilización. Las casas comerciales disponen de información que, en muchos casos, no se hará pública o es de difícil acceso y que determina la forma en la que la tecnología se introduce en el sistema.
- *Internet.* Recientemente, el espectacular desarrollo de Internet ha posibilitado un cómodo y rápido acceso a muchas de las fuentes de información citadas. Los buscadores de Internet facilitan la posibilidad de encontrar rápidamente información de cualquier tema, aunque en ocasiones los resultados de la búsqueda no ofrecen la calidad suficiente.

El proceso de búsqueda de información en general se compone de tres elementos o fases sucesivas, útiles para su análisis. Son la necesidad de información, la petición de información y la estrategia de búsqueda que finalmente se diseña. Muy a menudo, la información que en verdad se necesita no coincide exactamente con la que se pide en una biblioteca o a través de otro medio. Ello se debe a que el solicitante no conoce con precisión sus necesidades, o tiene dificultades para expresarlas de una manera eficaz. La ayuda de un especialista en documentación puede ser muy útil, tanto en la definición de la pregunta como en la selección de las palabras clave y de las bases de datos o fuentes de información que se deben consultar. Existen abundantes experiencias que demuestran las diferencias, a veces muy acusadas, entre una solicitud de información según se presenta originariamente y tal como queda después de una discusión a fondo con un especialista en documentación<sup>55</sup>.

Para la selección de las palabras clave es útil la consulta de artículos relevantes sobre el tema<sup>56</sup> o la utilización de estrategias prediseñadas, filtros, para la búsqueda bibliográfica (ver anexo II). Se han publicado numerosos trabajos sobre cómo manejar MEDLINE y otras fuentes bibliográficas en la búsqueda de evidencias médicas<sup>57,58,59,60,61,62,63,64</sup>. Además, existen programas informáticos que facilitan la gestión de las referencias bibliográficas, como REFERENCE MANAGER<sup>65</sup>, PRO-CITE<sup>66</sup>, ENDNOTE<sup>67</sup>, SHAPIRE<sup>68</sup>.

## 5. Evaluación crítica y selección de los estudios identificados

### 5.1. La evaluación crítica de los estudios identificados

Es un paso fundamental en la elaboración de un informe de ETS, que contribuye a la identificación de sesgos y sirve de guía en la interpretación de los hallazgos. Los criterios que requieren evaluación crítica se pueden dividir en aquellos relacionados con: *la validez, la precisión y las medidas de resultado de los estudios.*

#### • La validez

Un estudio es válido cuando su diseño y realización garantizan la detección y eliminación de errores sistemáticos y sesgos. Existen dos tipos de validez, interna y externa. La *validez interna* está relacionada con el diseño y los procedimientos de medición de variables y los métodos de análisis. Un estudio tiene validez interna cuando representa o mide aquello para lo que ha sido diseñado. En ese caso, los resultados serán atribuibles a los factores que el estudio ha tenido en cuenta, además de al error aleatorio<sup>69</sup>. El significado del término validez interna se relaciona con los conceptos de exactitud, fiabilidad y reproducibilidad.

La *validez externa* de un estudio es la potencialidad de generalización de sus resultados a personas o poblaciones distintas a los sujetos del estudio. Depende de aspectos relacionados con el diseño del estudio, como el muestreo o las características de la población estudiada, los cuales presentan mayor dificultad de control que los condicionantes de la validez interna<sup>69</sup>.

Los puntos a considerar en el análisis de la validez de un estudio son diversos. Por ejemplo, puede analizarse la asignación aleatoria, por si se detecta el empleo de métodos de *seudoaleatorización*. Éstos se dan cuando el investigador conoce previamente la probabilidad de asignación de un tratamiento u otro, y disminuye de forma considerable la calidad del estudio. Entre estos métodos se encuentran, por ejemplo, la asignación alterna, la asignación según el día de nacimiento o según el orden numérico que supone cada paciente. También pueden analizarse las características de los grupos de comparación. La comparación entre grupos separados en el tiempo, por ejem-

plo con controles históricos, tiene menor validez que la comparación entre grupos simultáneos. También, los estudios retrospectivos y los no controlados, en general, tienen menos validez que los prospectivos y los controlados. En resumen, son variados los aspectos relacionados con la validez que deben analizarse, y se pueden clasificar en aquellos relativos a la metodología del estudio, a los participantes y a la tecnología<sup>45</sup>.

El análisis de la validez es el aspecto que mejor valora la calidad de los estudios y, por tanto, de las evidencias científicas. Éstas deben fundamentar las recomendaciones de todo informe de ETS, relacionando la fuerza de la recomendación con el grado de calidad de la evidencia científica. No es posible establecer una recomendación muy relevante y/o trascendente a partir de evidencias de baja calidad. Por ello, es importante que después de la valoración crítica de los estudios se clasifiquen las evidencias proporcionadas por éstos según el grado de su calidad científica (ver anexo IV).

- **La precisión**

Un concepto importante que no debe ser confundido con la validez es la precisión. Ésta es una medida del error aleatorio que puede afectar al resultado de un estudio, y que generalmente se representa a través del intervalo de confianza de la estimación del efecto. En algunos casos se incluye entre los criterios que ponderan los estudios sintetizados en un meta-análisis, dando más peso a los estudios más precisos<sup>70</sup>.

- **Las medidas de resultado**

El último de los tres criterios, aunque no el menos importante, abordado anteriormente en la sección "**Definición de las preguntas de evaluación**", se convierte, muchas veces, en el criterio más determinante para la selección de los artículos identificados.

- **Los cuestionarios o esquemas de análisis crítico**

Estos esquemas, o *checklists*, son útiles para ordenar los aspectos a incluir en la evaluación crítica de los estudios identificados. Proporcionan un método más objetivo de análisis que la simple lectura crítica, aunque nunca evitan ésta. Presentan los problemas típi-

cos de cualquier cuestionario, la variabilidad inter e intraobservador y otras limitaciones que pueden disminuir su fiabilidad y reproducibilidad, aspectos que deben ser evaluados si se utilizan<sup>71,72</sup>. En ese caso, es importante que se describan claramente en el informe final; esto mejora la fiabilidad y reproducibilidad de las conclusiones. Existen multitud de cuestionarios o tablas de análisis para la evaluación crítica de estudios, algunos de los cuales se han recopilado y resumido en el anexo III.

- **Limitaciones**

Una de las limitaciones más relevantes y de difícil manejo en todo proceso de revisión de la literatura científica es el sesgo de publicación. Consiste en la tendencia a publicar más los estudios con datos positivos, estudios considerados relevantes por el autor o por el editor de la revista. En este último caso el sesgo se denomina sesgo del archivero. En otros casos, se observa que el tipo de revista determina la publicación o no de determinados artículos con datos negativos, o sea, datos que demuestran la ausencia de efectividad de una tecnología. Este tipo de artículos son excluidos de algunas revistas muy especializadas que prefieren divulgar sobre todo investigaciones sobre los avances en determinadas técnicas.

Se han realizado varios trabajos que exploran la repercusión potencial de los diferentes sesgos de publicación que existen, y concluyen que este sesgo puede limitar gravemente la validez de las revisiones bibliográficas<sup>73,74,75,76,77</sup>. El registro prospectivo de ensayos clínicos es, por el momento, la única estrategia difundida que puede ser considerada efectiva en la lucha contra este sesgo<sup>78,79,80</sup>.

Otra de las dificultades que surgen en la evaluación crítica de los estudios es la presentación inadecuada de los métodos y resultados. A veces, se observa la ausencia de datos imprescindibles para la lectura crítica del artículo; por ejemplo, cuando sólo se describe el número de pacientes intervenidos pero no el número de intervenciones totales que se han realizado. Se podría suponer que si no se ofrecen datos adicionales es porque los datos del número de pacientes coinciden con los del número de intervenciones, pero no siempre es así.

Existen métodos para estimar esos datos inexistentes, aunque es necesario realizar algunas asunciones cuyo impacto en los resulta-

dos debería ser estudiado mediante un análisis de sensibilidad (ver el capítulo “**Meta-análisis**”) <sup>81</sup>. En resumen, no existe ninguna solución sencilla para el problema de los datos ausentes, pero se aconseja adoptar una serie de precauciones: siempre se deben describir las asunciones que se realicen, debe realizarse un análisis de sensibilidad que valore los cambios que distintas asunciones producen en los resultados, y, siempre que sea posible, se debe intentar obtener datos adicionales a los ofrecidos en los artículos, contactando directamente con el autor del trabajo <sup>82,83,84</sup>.

## 5.2. Selección de los estudios

Las decisiones sobre cuáles son los estudios que se incluirán en el informe de ETS son difíciles y muy relevantes. No es fácil discriminar todos los artículos con la misma objetividad; además, la exigencia de una descripción clara en el informe de los criterios utilizados para esa selección, aunque mejora la reproducibilidad y rigor de las conclusiones, en algunos casos produce una valoración demasiado rígida e incluso ilógica. Se han propuesto diversas formas para llevar a cabo el proceso de selección de estudios <sup>31</sup>:

- Estableciendo unos criterios umbrales para la inclusión o exclusión de estudios, que puede seguirse de un análisis de sensibilidad de los criterios considerados.
- Asignando pesos a cada estudio en función de criterios explícitos, para posteriormente incluir esos pesos en el análisis cuantitativo de agregación de estudios.
- Analizando y valorando los estudios de manera cualitativa, considerando la calidad el principal criterio de evaluación.

Al igual que para los parámetros de evaluación crítica de los estudios, los criterios de inclusión y exclusión pueden atender a: valoración de la validez (metodología, características de los participantes y de la tecnología), la precisión y las medidas de resultado <sup>85</sup>.

Es conveniente que los criterios de selección de los estudios estén especificados ya en el proyecto inicial. De esta manera se evitarán sesgos de selección como el de “selección de referencias” <sup>51</sup>. Es aconsejable que la selección de estudios, al igual que la utilización de cuestionarios de valoración crítica, se haga por varias personas y que los desacuerdos sean discutidos y resueltos. El grado de concordancia entre los revisores puede ser

valorado mediante el índice Kappa u otros índices de concordancia <sup>72,86,87</sup>.

Si son varias las personas que realizan el estudio o selección de la información, es útil que alguno de ellos sea experto en el tema o tecnología evaluada. La presencia de especialistas en la materia aporta conocimiento y criterios valiosos y consistentes, pero puede introducir a su vez sesgos, dado que los expertos generalmente tienen una opinión preestablecida.

Otro aspecto a decidir es si el proceso de selección de artículos se realiza con un método ciego para los revisores; este enmascaramiento puede ser para los nombres de los autores, instituciones y publicaciones, y para los resultados de los estudios. Algunos estudios empíricos indican que la evaluación a ciegas puede producir puntuaciones inferiores y más uniformes que las evaluaciones abiertas <sup>88</sup>. No obstante, tales evaluaciones requieren un tiempo considerable, por lo que se deben sopesar los beneficios potenciales de las evaluaciones con enmascaramiento frente a los costes ocasionados.

El proceso de decisión acerca del punto de corte entre la inclusión y exclusión de estudios puede representarse como una línea continua que va desde “libre de sesgo” a “sin duda sesgado”. Existen diversos mecanismos para incluir los resultados de la evaluación crítica en esa decisión <sup>89</sup>. Se pueden utilizar gráficos visuales de los resultados ordenados según su grado de validez. Una segunda aproximación es el análisis de subgrupos, que consiste en realizar un análisis de sensibilidad en el que se observa si cambian los resultados al incluir determinados grupos de estudios en un lado u otro del punto de corte <sup>90,91</sup>. Un tercer enfoque consiste en ordenar los estudios según su calidad y combinar secuencialmente los resultados de cada estudio; es un tipo de meta-análisis “acumulativo”, mediante el cual es posible examinar el impacto en los resultados globales a medida que se van incluyendo ensayos de menor calidad. Un cuarto enfoque consiste en utilizar métodos estadísticos para ponderar los estudios según su calidad o utilizar una meta-regresión para explorar la relación entre calidad y magnitud del efecto entre estudios.

El manejo de los criterios de calidad en la selección de estudios presenta otras dificultades, ya que se sabe la elevada influencia de los criterios de calidad en los resultados de los estudios, pero no se ha medido ni se han establecido relaciones que permitan ponderar y ajustar con seguridad esos criterios.

Aunque hay estudios que han demostrado que, en general, tanto la ocultación inadecuada de las asignaciones como la ausencia de doble ciego producen sobreestimaciones de los efectos del tratamiento, en general, no es recomendable el uso amplio de las "puntuaciones de calidad" y la excesiva confianza en las evaluaciones detalladas de calidad de los estudios. Su uso no está sustentado por la evidencia empírica, puede ser lento y es potencialmente engañoso<sup>31</sup>.

Un problema frecuente que se plantea a la hora de seleccionar los estudios es la duplicación total o parcial de trabajos. Para evitarlo es necesario ser cuidadoso en la revisión de los artículos, ya que es frecuente que algunos trabajos se publiquen de nuevo un tiempo después con pequeños cambios o un incremento en el número de pacientes incluidos, y que el autor no indique qué parte de ese trabajo ya ha sido publicada previamente. Por último, es recomendable que el informe final presente tablas con los estudios seleccionados y otras con los excluidos, con el objeto de que los lectores conozcan con claridad cuáles son los estudios que se han quedado fuera del análisis y por qué<sup>45</sup>.

### 5.3. Extracción de datos de los estudios seleccionados

La ayuda de un formulario de recogida de datos, que contenga los criterios utilizados en la evaluación crítica de los estudios y los datos o resultados seleccionados cuando se definió la pregunta, suele ser beneficiosa. Los datos a extraer de los estudios se clasifican en: relativos a los resultados, a su metodología, a los participantes y a la tecnología. La definición de las unidades y medidas de resultado para el formulario de recogida de datos es una de las tareas más claves. No sólo es importante la decisión de incluir una variable de resultado, sino su tipo, y cómo se mide y trata estadísticamente.

El tipo de variable va a condicionar la técnica de síntesis de resultados a utilizar. En el caso de variables dicotómicas, los datos que se deben incluir en el formulario serán los números de personas que experimentaron el suceso en cada grupo de comparación y los números totales de personas en cada grupo. En el caso de variables continuas, los datos que deben incluirse son el número de personas en cada grupo y la media y desviación estándar de la medida de resultado en cada grupo. A veces, los resultados de una variable continua en algunos estudios se presentan como

cambios (media de las diferencias entre mediciones pre y postintervención), y en otros estudios como resultados postintervención<sup>31</sup>.

En ocasiones, los datos de una misma medida de resultado se presentan en forma dicotómica en algunos estudios y continua en otros. Por ejemplo, los resultados en las escalas de depresión pueden presentarse como una media o como el porcentaje de pacientes que estaban deprimidos en cierto período después de una intervención (una puntuación superior a un punto de corte determinado). A menudo, la información de este tipo es más fácil de comprender, y más útil para los que toman decisiones, si se expresa en forma dicotómica. Sin embargo, la decisión de establecer un punto de corte es arbitraria y siempre se pierde información cuando las variables continuas se transforman en dicotómicas.

Hay varias opciones para tratar de combinar datos dicotómicos y continuos. En general, es útil resumir los resultados de todos los estudios válidos y relevantes de manera similar, pero esto no siempre es posible. En ese caso es recomendable resumir los datos en tres grupos: continuos, dicotómicos y otros datos. Otra opción es presentar para todos los datos la media de las diferencias entre los grupos que se comparan o el *log-odds-ratio*<sup>31</sup>. Además, es aconsejable que las unidades se estandaricen para que la comparación entre los estudios sea fácil y que se presenten en unidades naturales que son de comprensión más intuitiva.

Un formulario de recogida de datos debe ser completo pero manejable. Es fácil caer en un exceso por asegurar que la información que se recoge es suficiente. La elaboración de unas instrucciones para la utilización del cuestionario, el entrenamiento de los recopiladores de información y la realización de una prueba piloto que aseguren la validez y precisión de los datos obtenidos son convenientes. La prueba piloto se realizará en una muestra representativa de los estudios identificados y sus resultados pueden aconsejar la modificación de alguna parte del cuestionario.

En todo proceso de obtención de datos se deben tomar una serie de medidas encaminadas a asegurar la fiabilidad y precisión, así como estudiar la concordancia entre las interpretaciones de los diversos recopiladores de datos, asegurar una codificación discriminante pero a la vez manejable o realizar un control de calidad de los posibles errores en cualquiera de las fases del proceso de extracción de datos. La extracción de datos puede

ser aconsejable que se realice por varias personas, estableciendo previamente el sistema de consenso que se utilizará y analizando posteriormente la concordancia entre observadores.

El grado en que diferentes personas coinciden en los datos extraídos de un estudio se corresponde con la fiabilidad del cuestionario. No hay un estándar fijo sobre el nivel de fiabilidad adecuado o sobre cómo evaluar la fiabilidad. Sin embargo, es importante que se analice en todo el proceso de recogida de datos. Por ejemplo, si después de alcanzar el consenso en los primeros estudios los observadores detectan o muestran frecuentes discrepancias en datos específicos, probablemente sea necesario modificar las instrucciones de codificación. Si aparecen cambios frecuentes en la forma de codificar por parte del mismo observador, denominado *coder drift* o cambio temporal en la forma de codificar la información, posiblemente sea necesario volver a codificar o incrementar el entrenamiento<sup>31</sup>.

---

## 6. Estudio del contexto local

Entre la información fundamental que debe incluir un informe de evaluación está la referida a la situación del medio en el que se aplica o se va a aplicar la tecnología. La utilidad de un informe de ETS se maximiza si se ofrece una valoración adaptada a las condiciones locales en las que se encuentra el usuario de la tecnología o receptor del informe. Si los datos y conclusiones ofrecidos por el informe son aplicables al medio del que forma parte el lector, el informe ofrecerá confianza e interés a éste y será más fácil conseguir el efecto final que busca la ETS: *influir en la toma de decisiones relacionadas con la TS*. El estudio del contexto local es especialmente necesario si el carácter de la pregunta va orientado a conocer el uso apropiado de una tecnología o su impacto en una determinada organización o ámbito.

Evaluar el impacto local de una TS u ofrecer información que sitúe la evaluación en un contexto determinado es un elemento específico de las organizaciones que hacen ETS, a diferencia, por ejemplo, de aquellas dedicadas a la realización de revisiones sistemáticas. La valoración de la validez externa o generalizabilidad de las conclusiones de los informes de ETS está íntimamente relacionada con el análisis del contexto local que incluya el informe.

El marco de desarrollo de los estudios (atención primaria, especializada, población general), los criterios de inclusión y exclusión de sujetos en los estudios, las características de la intervención, pueden orientarnos sobre el grupo de población y condiciones locales de aplicación de la tecnología. Las características de la población, la prevalencia y distribución de la enfermedad, el nivel de recursos y tipo de sistema sanitario, la percepción y preferencia de los pacientes y de los profesionales, condicionan la aplicabilidad y efectividad de una determinada tecnología que ha demostrado ser eficaz en determinadas condiciones ideales. Por ello, resulta fundamental añadir a la revisión y búsqueda de evidencias un análisis contextualizado que facilite la toma de decisiones sobre la adquisición o desarrollo de la TS.

El estudio del contexto local puede enfocarse de dos formas, descriptiva o analítica. El **enfoque descriptivo** consiste en ofrecer información relevante sobre la utilización actual o futura de la tecnología o sobre las posibles consecuencias de la implantación de esa tecnología basada en datos de otros estudios o registros sanitarios diseñados para otro fin. El **enfoque analítico** consiste en realizar un estudio con datos primarios que evalúe la utilización o implantación de la tecnología en el medio elegido. Este enfoque ofrece mayor calidad que el descriptivo, ya que la información se recoge *in situ* en el mismo ámbito local de la evaluación, y representa una metodología evaluativa muy potente, dada la enorme validez y aplicabilidad de sus conclusiones.

Otros aspectos a considerar cuando se estudia el contexto son los aspectos legales y organizativos. Conocer la normativa legal que afecta al uso de la TS es esencial para valorar la introducción de ésta en el sistema sanitario o para la elaboración de recomendaciones sobre su uso apropiado o su regulación. En otros casos resulta muy útil realizar un estudio de la normativa legal o de la situación organizativa comparando diversos países, con el objeto de evaluar la conveniencia de que el informe realice propuestas que afecten a la legislación u organización del sistema sanitario.

Por otra parte, el proceso de difusión y aplicación en la práctica de las conclusiones y recomendaciones del informe incluye la realización de actividades que produzcan efectos o motiven cambios en un medio o ámbito determinado. El éxito de esta fase de la ETS dependerá en parte del estudio del contexto local que se haya realizado. La tabla 4 ofrece una de las posibles clasificaciones de la infor-

**Tabla 4**  
Tipos de fuentes de información útil para el estudio del contexto local

- Datos epidemiológicos de la enfermedad o condición clínica implicada, referidos fundamentalmente a la magnitud y distribución de ésta.
- Datos que describan la utilización actual de la tecnología y la práctica clínica habitual, a través de la información contenida en registros sanitarios, encuestas o estudios *ad hoc*.
- Datos sobre los costes que conllevará la utilización de la tecnología.
- Datos económicos y/o sociales de la población afectada.
- Otros datos que ayuden a responder a las preguntas de investigación en el contexto local.

mación que se puede utilizar en un estudio del contexto local. Además, en el anexo II se proporciona una selección de las fuentes de información estadísticas y administrativas, útiles para la valoración del impacto local de una TS.

## 7. Técnicas de síntesis de la información y otros métodos utilizados en ETS

### 7.1. Síntesis cualitativa y narrativa

El empleo apropiado de la metodología cuantitativa es un instrumento muy potente que contribuye a evitar muchos errores y a ajustar las interpretaciones, pero muchas veces se observan importantes limitaciones para el empleo de técnicas de integración cuantitativa de resultados. En contraste, las valoraciones cualitativas deben estar incluidas siempre en los informes de ETS. Es más, muchos de los informes sólo van a incluir una valoración cualitativa o narrativa que, situada en el contexto de la ETS, tiene que ser sistemática, exhaustiva y explícita<sup>19</sup>.

A veces los resultados del empleo de una técnica estadística, por ejemplo un meta-análisis, no son discutidos con la amplitud y profundidad necesarias y no se valoran suficientemente las consideraciones referentes a la calidad de los estudios, las limitaciones de la técnica estadística o la relevancia clínica de los resultados<sup>89,92</sup>. La proliferación de estudios que incluyen un meta-análisis está aumentando, lo que ha llevado a que se analicen los problemas y ventajas de su utilización en medicina, advirtiendo de la peligrosidad de

una excesiva confianza en ellos<sup>93,94,95</sup>. Algunos autores han mostrado que, en muchos casos, los resultados de un meta-análisis no son consistentes con los de ensayos aleatorizados, controlados y de gran tamaño realizados con posterioridad<sup>96</sup>. Los riesgos, complejidad, así como las limitaciones e incertidumbres acerca de las consecuencias de su uso para la obtención de evidencias científicas, han llevado a algunos autores a recomendar restricciones en su empleo<sup>92,95,96</sup>.

Las razones para no emplear técnicas estadísticas de síntesis cuantitativa de resultados en medicina pueden resumirse en las siguientes<sup>45</sup>:

- La excesiva heterogeneidad de los estudios y/o sus resultados.
- La falta de datos válidos y/o relevantes.
- La escasa precisión de los datos disponibles.
- La baja calidad metodológica de los estudios.
- La inseguridad de que la agregación cuantitativa de resultados pueda contribuir a aclarar las dudas que ya existen.

Si los datos seleccionados no tienen la calidad suficiente como para ser considerados válidos, su agregación puede dar lugar a error, cuyo efecto se incrementa por la confianza que pueda tener el lector en un documento que le ofrece la mejor evidencia disponible. La difusión de información sesgada tiene peores consecuencias en muchas ocasiones que el hecho de no difundir ninguna información. El hecho de conocer que no existen evidencias científicas sobre la contribución de una tecnología a la mejora de la salud puede ayudar en la toma de decisiones.

Todo informe de ETS debe contener, al menos, una síntesis cualitativa o narrativa que inclu-

ya una revisión bibliográfica exhaustiva y sistemática con una buena valoración crítica de los estudios identificados, un examen y descripción minuciosos de los datos de los artículos, la elaboración de tablas que resuman y comparen los trabajos, y un estudio del impacto de la tecnología en el contexto local del solicitante.

Una evaluación desarrollada únicamente mediante una síntesis cualitativa debe evitar formas de síntesis *pseudocuantitativas* que no presentan las suficientes garantías de calidad metodológica. Por ejemplo, se pueden cometer errores como comparar el número de estudios positivos y negativos, enfoque del "recuento de votos", o excluir estudios con efectos pequeños pero clínicamente relevantes o estudios que presentan efectos sin significación estadística<sup>94</sup>; a veces, se ponderan los estudios inadecuadamente. La valoración de la importancia de los estudios no puede hacerse subjetivamente, es necesario explicitar una serie de criterios razonados. El recuento de votos concede igual peso a los estudios, ignorando las diferencias en su tamaño y calidad. Todas estas formas de síntesis son metodológicamente deficientes y pueden introducir importantes sesgos, agravados por la confianza que suele producir la utilización de metodología cuantitativa.

A continuación se abordan los enfoques analíticos cuantitativos más útiles para la síntesis de resultados en ETS. El meta-análisis, el análisis de decisión y la evaluación económica son instrumentos metodológicos de gran valor en ETS que pueden usarse de forma combinada, constituyendo una herramienta de gran potencia<sup>97</sup>. Algunos estudios han integrado estas técnicas en la búsqueda de medidas resumen que obtuvieran el máximo consenso<sup>98</sup>. Los métodos citados, junto con las técnicas de consenso a través fundamentalmente del juicio de expertos, deben conocerse con profundidad para la práctica de la ETS; en algunos casos se emplearán para la elaboración del informe; en otros, su uso se limitará a la lectura crítica e integración de sus resultados en la evaluación.

## 7.2. Meta-análisis

Es un método cuantitativo para la agregación de resultados de diferentes estudios basado en la utilización de técnicas estadísticas que posibilitan la combinación y síntesis de los datos de las diversas investigaciones, y en las que las unidades primarias de análisis son los

diferentes estudios. El objetivo del meta-análisis no es únicamente la obtención de una medida que sintetice el efecto de una intervención, sino también el análisis y la valoración de las diferencias entre los resultados de los diferentes estudios. Ya en 1904 aparecen estudios que emplean este tipo de técnica<sup>99</sup>, pero no fue hasta la década de los setenta en que el meta-análisis se comenzó a usar en medicina. El término meta-análisis fue acuñado por Glass<sup>100</sup> y posteriormente redefinido por la Biblioteca Nacional de Medicina de Estados Unidos como sigue: "método cuantitativo de combinación y síntesis de resultados de estudios independientes (obtenidos generalmente de trabajos publicados) con el objeto de conseguir unas conclusiones o un resumen que pueda ser utilizado para la valoración de la efectividad de un tratamiento o la planificación de nuevos estudios, entre otros usos"<sup>20</sup>.

Un meta-análisis estaría indicado en aquellas ocasiones en las que se desee realizar una revisión crítica con el rigor de un protocolo científico; cuando los resultados de los estudios existentes son conflictivos o no concluyentes, por ejemplo, si incluyen un número reducido de pacientes; e incluso cuando se desea detectar la evidencia existente y la calidad de la misma, identificando las lagunas en el conocimiento del área investigada. El meta-análisis se considera útil si los estudios son de alta calidad, existe una heterogeneidad explicable y de pequeño tamaño o la obtención de una medida estimadora de un efecto clínico resulta de alta relevancia. Sin embargo, todavía no puede decirse que los beneficios de su utilización compensen los riesgos y limitaciones que presenta<sup>92,94,95,96</sup>.

Los pasos de un meta-análisis son los siguientes<sup>97</sup>:

- 1) Formulación de los objetivos o preguntas de investigación y establecimiento de un protocolo o proyecto que planifique el estudio.
- 2) Búsqueda y recuperación de estudios, análisis de su calidad y selección, de acuerdo con los criterios de inclusión y exclusión especificados previamente.
- 3) Selección y extracción de los datos de los estudios (fundamentalmente características metodológicas y resultados).
- 4) Test de homogeneidad, que determina si es adecuado el cálculo de una medida resumen a partir de los resultados de los diferentes estudios.

5) Análisis estadístico mediante el cual se obtiene un resultado que sintetice los estudios seleccionados.

6) Análisis de sensibilidad para conocer la solidez de la medida final obtenida.

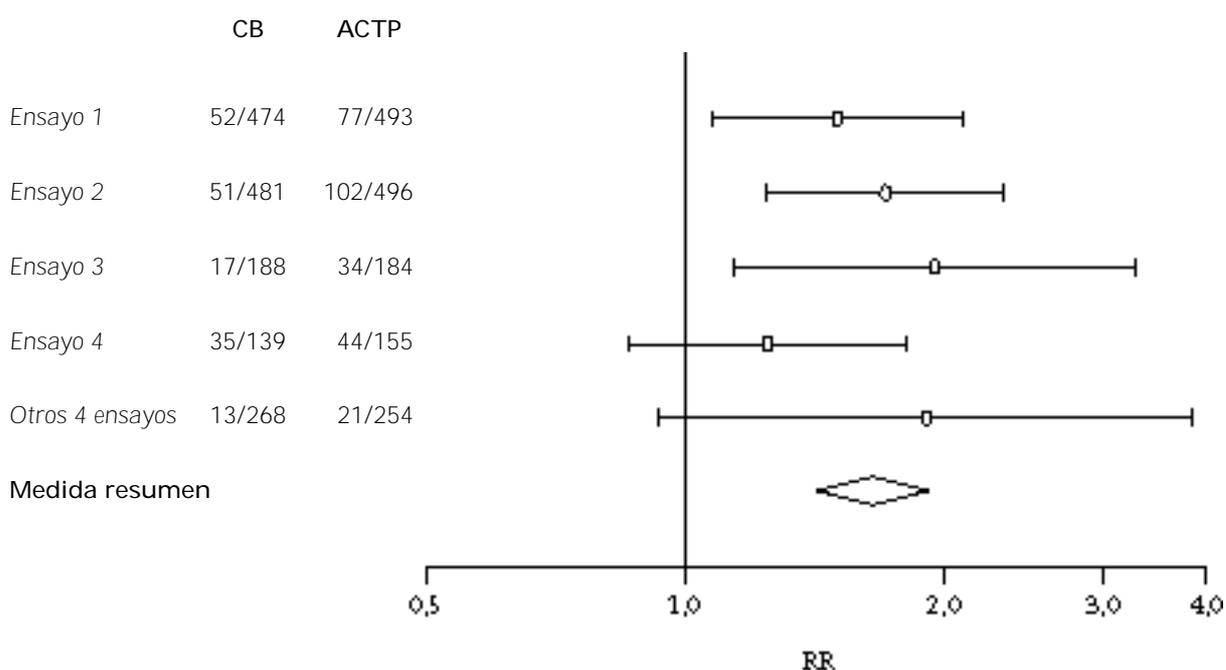
Antes de realizar un meta-análisis se deben establecer con precisión *las medidas de resultados y el estadístico* que se utilizarán para representar los efectos y agregarlos. Las medidas de resultados deben ser coherentes con los criterios de inclusión y exclusión de los estudios. Los estadísticos descriptores de los efectos que más se utilizan son las diferencias de tasas (mediante *odds-ratio*, riesgos relativos, diferencias de riesgo, reducción de riesgo y el número necesario de pacientes intervenidos para prevenir un caso, *Number Needed to Treat*)<sup>101,102,103</sup> o las diferencias en variables continuas (mediante comparaciones de medias o comparaciones de medias estandarizadas).

La *representación gráfica* de los resultados de los estudios incluidos en el meta-análisis es de gran utilidad. Se han usado varios métodos gráficos, el clásico suele tener forma de pirámide o embudo. En él se representa la medida

de efecto en el eje de abscisas (eje x) y a lo largo del eje de ordenadas (eje y) se sitúan los diferentes estudios ordenados por el peso que se les haya dado. Generalmente se ponderan por el inverso de la varianza de la medida de efecto elegida o por el tamaño de la muestra<sup>104</sup>. El siguiente ejemplo (figura 2) forma parte de un meta-análisis que pretende integrar los ensayos clínicos aleatorios que comparan los efectos de la angioplastia coronaria y de la cirugía con *bypass* en pacientes con angina severa<sup>105</sup>.

La figura 2 representa los logaritmos de los riesgos relativos (log RR) del grupo tratado con ACTP frente al grupo tratado con CB, de tal manera que si un RR es mayor de 1 indica un riesgo mayor en el grupo tratado con ACTP. El meta-análisis de los ocho ensayos clínicos (cuatro forman un único grupo de análisis) fue realizado con el método de efectos fijos, obteniendo una medida sumaria de la media ponderada de los log RR. La ponderación, y por tanto el orden de los estudios en el "eje y", se realizó por el inverso de la varianza del RR. Los estudios con mayor varianza aparecen en la parte inferior del gráfico.

**Figura 2**  
Prevalencia de angina (de grado >1) tras un año de seguimiento en el grupo tratado con Angioplastia Coronaria Transluminal Percutánea (ACTP) frente al tratado con Cirugía de Bypass (CB) (tomado de Pocock<sup>105</sup>)



Se han utilizado otros métodos gráficos, como el llamado *odd man out*, en el cual la superposición de los intervalos de confianza facilita la obtención de una región que resuma los intervalos<sup>106</sup>. También se utilizan la representación de la "tendencia cronológica del efecto"<sup>104</sup>, los gráficos de cajas (*box plot*)<sup>107</sup> o gráficos bivariados, radiales o circulares<sup>108</sup>.

Antes de realizar un meta-análisis debemos comprobar que la heterogeneidad entre los diferentes estudios que vamos a combinar se debe al azar, en vez de a factores no aleatorios, mediante la realización de un *test de homogeneidad* (ver tabla 5).

El siguiente paso es la realización de la *técnica estadística* de combinación de estudios, con lo que se obtiene una medida resumen, que generalmente se presenta acompañada de su intervalo de confianza<sup>111</sup>. Los estudios son ponderados generalmente por su varianza, aunque también se aplican otros criterios: calidad del estudio, diseño y validez de los datos<sup>70,89,97</sup>.

La heterogeneidad entre los estudios puede incorporarse en el análisis, a través de los métodos de efectos aleatorios, o no ser incluida, si se utiliza un método de efectos fijos<sup>20</sup>. Los modelos de efectos fijos asumen que existe un único efecto en la población y no tienen en cuenta la heterogeneidad (varianza) entre los distintos estudios. Así, el tamaño del estudio y su propia varianza son los únicos determi-

nantes de su peso ponderado en el meta-análisis. El modelo de "efectos fijos" es válido para comprobar la "hipótesis nula global", la cual defiende que ningún estudio tiene efecto. Un resultado que permita rechazar estadísticamente la hipótesis nula global indica que existe efecto en al menos uno de los estudios.

Los modelos de efectos aleatorios, por el contrario, tienen en cuenta la heterogeneidad entre los estudios al considerar que los efectos de la intervención en la población son diversos y que los estudios incluidos en la revisión son sólo una muestra aleatoria de todos los posibles efectos. La ponderación de los estudios se realiza, además de por la varianza o tamaño de éstos, por la varianza entre los estudios. Los resultados de este método tienden a ser más conservadores, con mayores intervalos de confianza, y generalizables al posible universo de estudios. Deben utilizarse siempre que se determine tras un test de homogeneidad que los estudios son heterogéneos y no puedan identificarse claramente las causas, lo cual es muy habitual. Se puede decir que existe consenso general sobre la utilización prioritaria del modelo de «efectos aleatorios».

Los métodos estadísticos utilizados en el meta-análisis son múltiples y no están totalmente establecidos. Para una visión general pueden ser útiles las siguientes referencias<sup>82,97,104,112,113,114</sup>. La elección de la técnica depende fundamentalmente del tipo de datos analizados (ver tabla 6).

Tabla 5  
Análisis de homogeneidad

El propósito es comprobar si los factores que motivan la heterogeneidad entre los estudios no son aleatorios o, por el contrario, se deben al azar. Una manera de hacerlo es a través del estudio de la representación gráfica conjunta de los resultados de los estudios. Por ejemplo, con la utilización de los gráficos *odd man out*<sup>106</sup>, pero la metodología más utilizada es el Análisis de Homogeneidad<sup>82,109,110</sup>.

Si el resultado de la prueba es que las diferencias observadas no pueden ser explicadas por el azar, se concluye que hay heterogeneidad. En ese caso es importante explorar y explicar en la medida de lo posible los motivos de heterogeneidad. Las características diferenciales que se utilizan más habitualmente para definir la heterogeneidad entre los estudios son las características de los pacientes, de la intervención sanitaria (por ejemplo, la dosis administrada) y de las medidas de resultado utilizadas<sup>45</sup>.

Hay que ser muy cautos a la hora de buscar explicaciones a esa heterogeneidad. Se pueden adoptar varias opciones, de las cuales la más rigurosa metodológicamente es no resumir los estudios en una medida sumaria. En cualquier caso, si se realiza es necesario presentar los resultados con claras advertencias a sus limitaciones analíticas y utilizar el modelo de "efectos aleatorios". En caso de que se encuentren motivos convincentes que expliquen la heterogeneidad entre los estudios, la opción más recomendable es realizar un Análisis de Subgrupos (ver tabla 7).

**Tabla 6**  
Métodos estadísticos para meta-análisis disponibles en el programa informático  
REVMAN 3.0 (tomado de Mulrow<sup>31</sup>)

Tipo de datos	Estadístico descriptor	Modelo	Método
Dicotómicos	<i>Odds-ratio</i>	Efectos fijos	Peto <sup>115,116</sup> Mantel-Haenszel <sup>117,118</sup>
		Efectos aleatorios	Dersimonian-Laird <sup>110</sup>
	Riesgo relativo	Efectos fijos	Mantel-Haenszel
		Efectos aleatorios	Dersimonian-Laird
	Diferencia de riesgos	Efectos fijos	Mantel-Haenszel
		Efectos aleatorios	Dersimonian-Laird
Continuos <sup>119,120,121</sup>	Diferencia de medias ponderada	Efectos fijos	Inverso de la varianza o método de Woolf <sup>109,122</sup>
		Efectos aleatorios	Dersimonian-Laird
	Diferencia estandarizada de medias	Efectos fijos	Inverso de la varianza
		Efectos aleatorios	Dersimonian-Laird
Datos de pacientes individuales	<i>Odds-ratio</i>	Efectos fijos	Peto

Un meta-análisis basado en la agregación de datos de pacientes individuales, *mega-análisis*, es más válido y fiable que el basado en la síntesis de los resultados ofrecidos por cada estudio. El aumento del poder estadístico, la posibilidad de comprobar la calidad de los datos originales y de corregir ciertos errores, la posibilidad de adaptar el análisis a la pregunta que nos interesa, y de realizar análisis de subgrupos (ver tabla 7), son algunas de las ventaj

as de este método. Sin embargo, en la mayoría de los casos es muy difícil obtener los datos individualizados originales de cada estudio, debido a la pérdida de datos, acceso limitado o a la ausencia de colaboración de los investigadores.

La combinación de estudios de validez muy diferente en un meta-análisis debe realizarse con una metodología específica que evite sesgos. Se han utilizado hasta el momento tres

**Tabla 7**  
Análisis de subgrupos

El estudio de los resultados de una serie de trabajos puede enfocarse hacia la búsqueda de explicaciones en las diferencias entre los grupos de participantes. Este tipo de metodología pretende contrastar la significación de una diferencia aparentemente importante en un subgrupo, por ejemplo en mujeres, o en una determinada franja de edad.

La utilización o interpretación de un análisis de subgrupos debe hacerse con mucha cautela, ya que puede haber muchos factores que pueden explicar las diferencias entre los estudios. Podemos estar introduciendo fácilmente errores sistemáticos o heterogeneidad y las conclusiones o recomendaciones en ese caso provocan equivocaciones que afectan a la asignación de recursos a todo un grupo de población.

Los factores que influyen en la fuerza de las conclusiones que un análisis de subgrupos puede proponer son la magnitud y la significación estadística de las diferencias, si la hipótesis fue anterior o posterior al análisis, si el análisis de subgrupos fue una entre las varias hipótesis que se estudiaron, si la diferencia surgió por las comparaciones intra o interestudios, si la hipótesis es consistente y si existen evidencias indirectas que la apoyan.<sup>90,91</sup>

métodos diferentes para ese objetivo. El primero consiste en realizar un meta-análisis para cada grupo de estudios, clasificando éstos por su calidad, y posteriormente comparar los resultados. El segundo consiste en realizar algo parecido a un análisis de sensibilidad, incluyendo en el meta-análisis paso a paso estudios de menor validez que van modificando los resultados; es lo que se llama *meta-análisis acumulativo*<sup>123</sup>. El tercer método introduce la calidad de los estudios como una variable más en un análisis multivariante; es la técnica denominada *meta-regresión*<sup>124,125,126</sup>.

El último paso de un meta-análisis es el *análisis de sensibilidad*. Es una prueba de estabilidad de los resultados de una técnica de síntesis cuantitativa, sea meta-análisis, evaluación económica o análisis de decisión. Se utiliza cuando hay incertidumbre en aspectos importantes que condicionen la validez de los estudios. Sirve, por tanto, para valorar la consistencia de la medida resumen obtenida tras el proceso de combinación cuantitativa de resultados. Con el análisis de sensibilidad se intenta responder a la pregunta ¿serían diferentes los resultados finales de la técnica estadística si se modifican los datos dudosos que se han utilizado en el análisis? Por tanto, se está comprobando la robustez de los resultados con relación a las asunciones y supuestos que han surgido en el proceso de realización del análisis.

En general, los tipos de asunciones o aspectos condicionantes de la validez que se examinan en los análisis de sensibilidad incluyen<sup>31</sup>:

- Cambio de los criterios de inclusión.
- Inclusión o exclusión de estudios en los que persiste ambigüedad sobre si reúnen o no los requisitos de inclusión.
- Exclusión de estudios no publicados.
- Exclusión de los estudios de poca calidad metodológica.
- Cambio en el intervalo de resultados para los estudios que plantean incertidumbres. Por ejemplo, debido a incongruencias en la manera de describir los datos o a diferencias en la definición o medida de los resultados.
- Imputar un intervalo de valores razonable a los datos inexistentes.
- Cambio en el enfoque analítico. Por ejemplo, usando un modelo de efectos aleatorios.

El análisis de sensibilidad se realiza repitiendo la técnica estadística realizada con dife-

rentes valores para determinados datos, incluyendo, por ejemplo, un estudio excluido previamente, observando posteriormente si la modificación cambia sustancialmente el efecto final. Si ocurre así, los resultados no tienen la solidez suficiente y, por tanto, es necesario ser muy cauto en la valoración de los datos y en la elaboración de las conclusiones<sup>97</sup>. El análisis de sensibilidad es una de las mejores maneras de depurar o detectar errores y de realizar una evaluación preliminar del resultado final de una síntesis cuantitativa<sup>127</sup>.

Existen diversos programas informáticos útiles para la realización de un meta-análisis o una revisión sistemática como: REVMAN<sup>128</sup>, FAST\*PRO<sup>129,130</sup>, STATA<sup>129</sup>, TRUE EPISTAT<sup>129</sup>, DSTAT<sup>129</sup>, DESCARTES<sup>129</sup>, EASY MA<sup>129</sup> y META-ANALYST<sup>129</sup>.

### 7.3. Análisis de decisión

Es una herramienta cuantitativa que evalúa el valor relativo de cada una de las opciones existentes en una decisión<sup>97</sup>. Divide los problemas complejos en partes más sencillas, los analiza en detalle y los combina de nuevo de manera que se pueda identificar la mejor estrategia. Trata de evitar los errores sistemáticos producidos al establecer juicios clínicos basados fundamentalmente en la intuición clínica o en la experiencia personal. El análisis de decisión es un procedimiento de apoyo a la toma de decisiones, basado en la utilidad esperada de cada opción desde una determinada perspectiva. A la vez, constituye un método sistemático de identificación de opciones, cuantificación de expectativas o probabilidades, asignación de valores a resultados y comparación entre opciones.

En medicina se utiliza tanto para facilitar la toma de decisiones en la práctica clínica, ante pacientes concretos, como para diseñar estrategias globales aplicables a un amplio grupo de pacientes o a la política sanitaria<sup>97</sup>. Su origen proviene de la "Teoría de la Estrategia de Juegos", descrita en los años veinte y aplicada en economía y política a partir de la década de los cuarenta<sup>131</sup>. Existen varios textos relevantes que desarrollan con amplitud el análisis de decisión<sup>97,132,133</sup>.

¿Cuándo está indicado el abordaje de un problema mediante el análisis de decisión? Se han propuesto las siguientes situaciones<sup>127</sup>:

- 1) Existe un conjunto de alternativas reales bien definidas entre las que escoger.

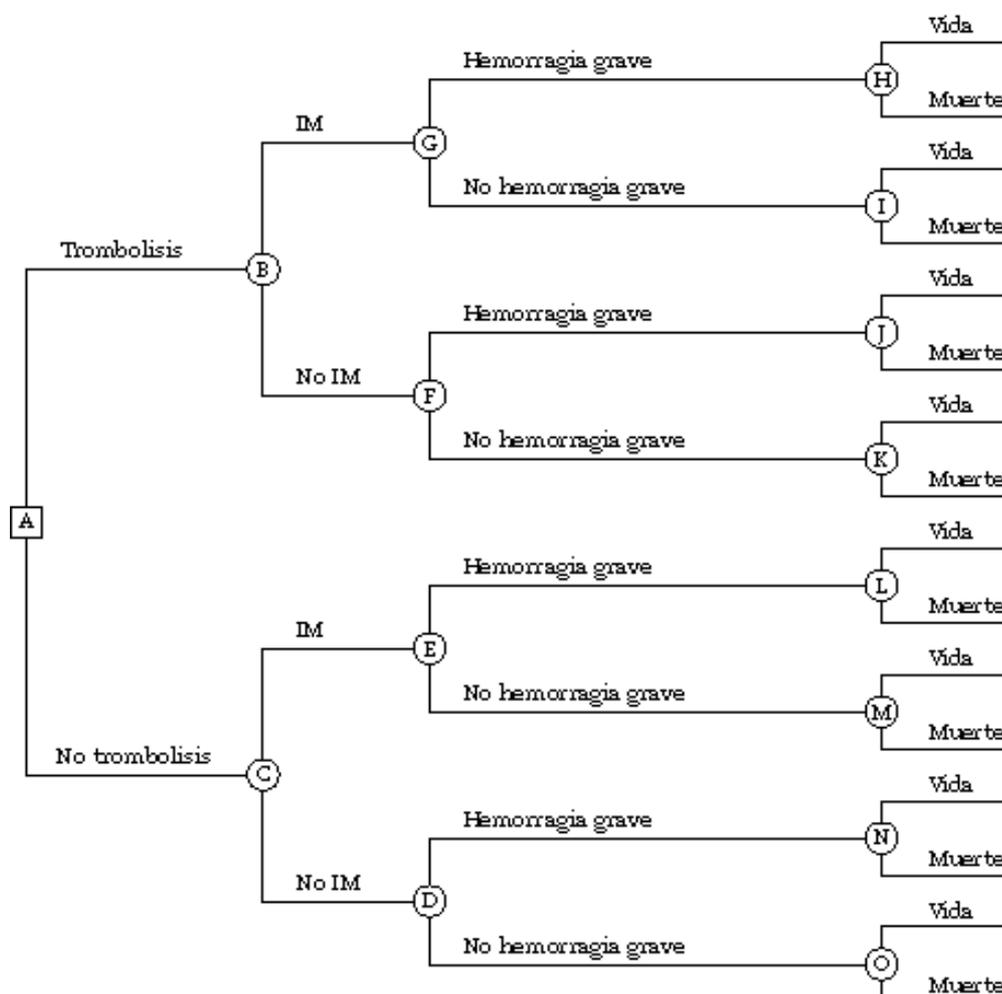
- 2) Se detecta incertidumbre sobre las consecuencias de las alternativas.
- 3) La persona o personas que deciden tienen preferencias acerca de los posibles resultados.
- 4) Existen una o más restricciones que limitan el conjunto de alternativas. Por ejemplo, el coste de oportunidad de vacunar a toda la población escolar de hepatitis B es tan grande que nos impide realizar otras actuaciones sanitarias con una mejor relación coste-efectividad.
- 5) El problema precisa que se tome una decisión.

En algunos estudios que abordan una situación clínica complicada, el análisis de decisión no proporciona más información que la necesidad de datos adicionales imprescindibles y de nuevos estudios en ese campo. En otros casos puede indicar que las diferencias entre los resultados de las distintas estrate-

gias son muy pequeñas, lo que suele deberse a errores del método. Incluso cuando resulta potencialmente útil, puede no ser factible completar los cálculos o el análisis de sensibilidad precisos por limitaciones de tiempo o de recursos. No obstante, el valor del enfoque analítico del proceso de decisión radica en que integra los datos disponibles, obliga a un razonamiento riguroso y expone las áreas de incertidumbre o ignorancia<sup>134</sup>.

Los árboles de decisión son las técnicas más usadas entre la amplia gama de análisis de decisión existentes. Son útiles para explicitar la secuencia temporal y lógica de las diferentes alternativas de un problema, de tal manera que sea fácil comprender todas sus posibles decisiones y consecuencias. La figura 3 muestra un ejemplo de árbol de decisión para decidir si tratar con un determinado fármaco trombolítico intravenoso a los pacientes en los que se sospecha infarto agudo de miocardio.

Figura 3  
Árbol de decisión sobre la trombolisis en pacientes con sospecha de infarto agudo de miocardio (tomado de Goldman<sup>134</sup>)



El cuadrado o "núcleo" marcado con la letra A indica la decisión que el médico debe tomar. Los círculos, marcados de B a O, indican los diferentes resultados. Cada uno de ellos proviene de una rama en la que se debe escribir la probabilidad conocida que tiene de producirse ese resultado. Cada una de las evoluciones finales posibles recibirá una *utilidad* o preferencia, o sea, una valoración cuantitativa de lo apreciado que resulta ese resultado en relación con los demás. El valor de cada resultado final se obtiene multiplicando todas las probabilidades de cada una de las ramas por la utilidad asignada a ese final de rama. Por último, para obtener el valor final de decidir tratar con trombolisis frente al valor de no hacerlo, se suman los valores relativos de todas las puntas de rama resultado de esas dos decisiones, y posteriormente se elegirá la opción que muestre el mayor valor. La tabla 8 describe los pasos del análisis basado en árboles de decisión.

Respecto al diseño de árboles de decisión se han realizado recomendaciones generales para evitar errores e incorrecciones en su construcción<sup>127</sup>. Entre las recomendaciones

se incluyen las siguientes: procurar el balance entre riesgos y beneficios de la intervención en las distintas ramas; situar preferentemente dos ramas en cada nodo de decisión; ramas ligadas a través de los cálculos de probabilidades de resultados o de las utilidades; asegurar la simetría, o incluir todas las condiciones determinantes que pueden afectar a los resultados.

Las utilidades pueden ser estimadas por diversos procedimientos: juicio de expertos; opinión de las personas que realizan el estudio; búsqueda en la literatura científica o estudio *ad hoc*. Existen varias publicaciones que describen las utilidades de un amplio rango de estados de salud<sup>135,136,137</sup>. Existen varias publicaciones útiles para profundizar en los métodos de elaboración y medición de utilidades en salud<sup>135,138,139,140,141,142</sup>.

El árbol de decisión es la técnica principal y más empleada en los análisis de decisión. Otras técnicas son el análisis de umbrales de decisión y el proceso de Markov, que se describen a continuación; los modelos de simulación de Monte-Carlo<sup>143,144</sup>, y los métodos de

**Tabla 8**  
**Fases del diseño de un análisis basado en un árbol de decisión (modificado de Pettiti<sup>97</sup>)**

1. IDENTIFICAR Y LIMITAR EL PROBLEMA, de tal manera que se establezcan todas las alternativas de decisión y los resultados a que pueden dar lugar. Se deben incluir los aspectos más relevantes que puedan determinar los resultados y la utilidad de los cursos de acción, definir claramente la perspectiva del análisis, identificar todas las alternativas y el horizonte temporal de análisis.
2. CONSTRUIR UN ÁRBOL DE DECISIÓN QUE ESTRUCTURE EL PROBLEMA. Selección de las opciones, nodos de decisión y nodos finales. Debe mantenerse el equilibrio entre complejidad y simplicidad, logrando que los elementos esenciales, tanto en lo que se refiere a opciones de acción como a valores, estén incluidos en el modelo.
3. ESTIMACIÓN DE PROBABILIDADES Y ASIGNACIÓN A LOS CURSOS DE ACCIÓN. Recopilar información mediante una revisión sistemática de la literatura, consulta a expertos o búsqueda de datos primarios, para estimar las probabilidades que van a ser asignadas a cada uno de los componentes del árbol. La probabilidad final de que ocurra el suceso definido por una rama se halla multiplicando todas las probabilidades de todos los trayectos de rama necesarios para llegar a ese resultado clínico final.
4. ASIGNACIÓN DE VALORES Y UTILIDADES A LAS CONSECUENCIAS O RESULTADOS FINALES. Los resultados clínicos finales pueden expresarse de diversas formas: años de vida, años de vida ajustados por calidad, esperanza de vida ajustada por calidad, casos de enfermedad evitados o complicaciones evitadas, o utilidades. La <i>utilidad</i> es una medida de la preferencia relativa, perspectiva, que tiene alguien respecto de un determinado resultado. Se expresa en una escala entre 0 y 1, de tal manera que al resultado de menor utilidad (la muerte sería el caso extremo) se le asigna el valor 0; y al de mayor utilidad (la salud plena sería el otro extremo), el valor 1.
5. INTERPRETACIÓN DE LOS RESULTADOS Y VALORACIÓN DE LA INCERTIDUMBRE. Conociendo la probabilidad de que ocurra cada suceso, o rama del árbol, y su utilidad se dispone de todos los datos para interpretar los resultados y poder elegir el curso de acción más favorable. Por último, es aconsejable realizar un análisis de sensibilidad que nos confirme la estabilidad de nuestros resultados en diferentes circunstancias.

medición de preferencias entre estados de salud *standard gamble*<sup>140,145,146</sup>, *time trade-off* y *category scaling*<sup>142,146,147</sup>.

El análisis de umbrales de decisión es un tipo de análisis de decisión que utiliza la teoría de Bayes para valorar, por ejemplo, la conveniencia de realizar una prueba diagnóstica a un paciente. Parte del concepto de que realizar un test a un paciente es inútil si no sirve para modificar la estrategia terapéutica o el manejo clínico global de ese paciente. El umbral es el nivel de probabilidad en el que el beneficio previsible de tomar una decisión iguala al de no tomarla o al de tomar una alternativa<sup>133,148</sup>.

Las cadenas de Markov, o el proceso de Markov, son un método cuantitativo dinámico que toma en consideración la probabilidad de que un paciente cambie de un estado de salud a otro en un determinado período de tiempo, es decir, la probabilidad de transición para ese determinado estado. En este modelo es necesario definir por adelantado cuáles son los estados de salud de interés y especificar todas las posibles transiciones o caminos entre estados<sup>127,149,150</sup>.

Existe una variada oferta de programas informáticos útiles para los análisis de decisión disponibles en el mercado. Destacan DATA 3.5<sup>151</sup>, SMLTREE 2.9<sup>152</sup>, DECISION MAKER<sup>153</sup> y FAST\*PRO<sup>154</sup>. Estos programas facilitan la construcción y análisis de árboles de decisión, cadenas de Markov, modelos de simulación de Monte-Carlo, análisis de sensibilidad y aplicaciones para análisis de coste-efectividad. Anne Haddix, en su libro *Prevention Effectiveness*, ofrece un anexo con la descripción de estos y otros programas informáticos útiles<sup>130</sup>.

#### 7.4. Estudios de evaluación económica

La pertinencia de realizar un estudio de evaluación económica para la elaboración de un informe de ETS depende fundamentalmente de la formulación de la pregunta planteada en la solicitud de evaluación. Cuando los objetivos del informe vayan encaminados a la ponderación de costes y consecuencias, y pueda conseguirse información de la suficiente calidad y validez, se pueden usar las técnicas que a continuación se describen.

El diseño de un estudio de evaluación económica debe partir de la identificación del tipo de información económica que será más útil para el destinatario del estudio o informe y definir de forma precisa las preguntas o problemas de investigación. Estos primeros pa-

sos ofrecen las claves para orientar sobre el tipo de costes y resultados que van a ser analizados y sobre el diseño más apropiado para responder a las preguntas formuladas<sup>130</sup>.

Una vez aclarados estos aspectos iniciales, deben ser identificadas de forma precisa las posibles alternativas razonables de intervención y bases de referencia o comparación, incluida la ausencia de intervención. La perspectiva de análisis debe ser definida a continuación. El tipo de perspectiva adoptada dependerá del problema planteado, del interés del solicitante del estudio y del contexto de la demanda. En general, la preferencia de los estudios de evaluación económica es la adopción de la perspectiva social o general, en la que se incluyen costes indirectos y de oportunidad si bien; otra perspectiva puede ser igual o más relevante para un estudio concreto: la perspectiva individual, la del proveedor o la del financiador.

El marco temporal de cada estudio, es decir, el período en el que se desarrolla la intervención que se evalúa, tiene que adaptarse al tipo de intervenciones estudiadas y quedar perfectamente establecido en el diseño. El horizonte temporal de análisis, o período en el que se van a medir costes y resultados relacionados con la intervención, debe ser de suficiente duración como para incluir los costes y resultados relevantes producidos durante la aplicación de la intervención, aun cuando se produzcan en el futuro, una vez finalizada ésta.

*Tipos de estudios.* Las dos características definitorias del análisis económico en sanidad son<sup>155</sup>:

- *Los costes y beneficios de una actividad sanitaria son los dos pasos de una cadena de producción.* Michael Drummond establece que al igual que se acepta la relación entre factor productivo y producto en la industria, de forma similar se puede enunciar una relación similar en el ámbito sanitario.
- *El análisis económico está íntimamente ligado a la elección.* Pretende identificar criterios útiles para decidir cómo asignar recursos estableciendo comparaciones entre alternativas que ofrezcan información discriminatoria.

Las dos características anteriores formuladas en forma de preguntas sirven para clasificar los estudios de evaluación económica sanitaria en una matriz de seis celdas, tal y como se muestra en la tabla 9. ¿Se examinan tanto los costes como los beneficios de las alternativas? ¿Hay comparación entre dos o más alternativas?

Tabla 9  
Clasificación de los estudios de evaluación económica (tomado de Drummond<sup>155</sup>)

		¿Se examinan costes y consecuencias simultáneamente?	
¿Se comparan dos o más alternativas?	NO		Sí
	Sólo consecuencias	Sólo costes	
NO	Descripción de consecuencias	Descripción de costes	Descripción de costes y consecuencias
Sí	Evaluación de eficacia/efectividad	Análisis de costes	Minimización de costes Análisis coste-efectividad Análisis coste-utilidad Análisis coste-beneficio

La celdilla inferior izquierda corresponde a los estudios que evalúan la eficacia/efectividad, como los ensayos clínicos, que comparan los resultados o consecuencias de las alternativas pero no comparan los costes o efectos secundarios. La celdilla inferior central se refiere a los estudios llamados análisis de costes. Éstos comparan tecnologías cuya efectividad se considera semejante, sólo pretenden estimar el coste de las alternativas. La única de las seis celdillas que cumple las dos características de los estudios de evaluación económica es la inferior derecha. Al resto de los estudios se les denominan *evaluaciones económicas parciales*. Lo que no implica que estos estudios sean inútiles, ya que pueden representar estadios intermedios vitales para la comprensión de los costes y resultados de los servicios y programas sanitarios.

A continuación se describen cada uno de los cuatro tipos de *estudios económicos completos*:

**MINIMIZACIÓN DE COSTES:** Los resultados no se miden en ninguna unidad porque se parte de la asunción de que las intervenciones que se comparan producen aproximadamente los mismos resultados. Se usa cuando, una vez identificado un beneficio común, las distintas opciones permiten alcanzar el mismo objetivo en el mismo grado, aunque previsiblemente con costes diferentes, que son los que se trata de minimizar. El interés también puede estar en el estudio de la distribución de los costes<sup>155</sup>.

**ANÁLISIS DE COSTE-EFECTIVIDAD:** Ha sido, desde los años setenta, el método más utilizado de evaluación económica en el campo de la salud. Las unidades en que se miden los resultados son unidades naturales, unidades de efecto sanitario como las unidades de presión arterial, la incidencia de enfermedad, los ca-

sos evitados, la mortalidad o los años de vida ganados. El propósito de este tipo de análisis es comparar el coste de dos o más alternativas de intervención sanitaria que persiguen el mismo objetivo. Permite identificar el procedimiento que presenta una relación coste-efectividad más beneficiosa para el logro de un objetivo específico en una determinada población o contexto de análisis, o identificar aquel cuya relación coste-efectividad es relativamente alta o baja en comparación con otras alternativas de intervención. Los costes se miden en unidades monetarias y los beneficios en unidades de efectividad<sup>156</sup>. La utilidad de los análisis de coste-efectividad para la toma de decisiones sobre asignación de recursos y priorización ha generado un importante debate y propuestas de estandarización de los métodos de análisis, de forma que se disminuya la heterogeneidad observada entre los estudios publicados y permita avanzar en la comparabilidad entre los estudios<sup>157,158,159,160,161</sup>.

**ANÁLISIS COSTE-UTILIDAD:** Es una variedad del análisis de coste-efectividad en la que un componente importante de los resultados medidos es la calidad de vida. Los resultados se miden generalmente en Años de Vida Ajustados por Calidad (AVAC o, en inglés, QALY, Quality Adjusted Life Year), lo que exige disponer de instrumentos específicos de medida y de ajuste de medidas cuantitativas de salud por calidad de vida, a menudo complejos y de difícil interpretación. Al proveer un método estandarizado de medida y valoración de resultados permite comparar la utilidad relativa de numerosas intervenciones que compiten por los recursos disponibles. Es un tipo de estudio económico más amplio que el análisis coste-efectividad, apropiado cuando la calidad de vida es relevante como resultado de

las intervenciones comparadas, existen efectos de morbilidad o incapacidad relevantes o efectos múltiples que afectan de forma compleja a la calidad de vida. Introduce las preferencias y valoraciones de estados de salud realizadas por individuos, profesionales o la sociedad<sup>135,162</sup>.

**ANÁLISIS COSTE-BENEFICIO:** Es el único tipo de evaluación económica en el que tanto los costes como los resultados se miden en unidades monetarias, de esta manera se facilita la comparación entre costes y beneficios. Este tipo de estudio viene a ser el método más comprensivo de análisis económico y más concluyente para la decisión de hacer o no una intervención, ya que valora todos los efectos en términos económicos, aunque la valoración monetaria de los efectos sanitarios presenta importantes limitaciones. El objetivo es determinar si los beneficios exceden a los costes, utilizando el "valor actualizado neto" (VAN) como expresión de la diferencia entre beneficios y costes. Para la valoración monetaria de la salud y de la vida humana se utilizan el enfoque de capital humano y el de disponibilidad a pagar. El primero se basa en la valoración económica de prevenir la muerte de un sujeto de una determinada edad; el segundo valora la salud y la vida basándose en el dinero que los sujetos estarían dispuestos a pagar por una mejora de la salud. Es, quizás, la modalidad de análisis económico menos utilizada dadas las implicaciones éticas y la dificultad de atribuir valor económico a la vida humana<sup>157,158,163</sup>.

*La medición de los costes y resultados.* Inicialmente, en todo análisis económico se deben estudiar cuáles van a ser los costes a considerar, fijos o variables, directos o indirectos, tangibles o intangibles, pagados por los pacientes, la administración sanitaria o la industria. La adecuada identificación de los costes y resultados es el aspecto más importante que se debe considerar cuando se diseña un estudio de evaluación económica en sanidad<sup>15</sup>.

La identificación, definición y recogida de datos referentes a cada uno de los costes que formarán parte del análisis es una de las tareas más complejas y difíciles de realizar. Los factores que influyen en la inclusión de unos costes u otros son: la perspectiva adoptada para el análisis, proveedor, financiador, sociedad; la adecuada definición del período de tiempo en el que los costes han de ser cuantificados (marco temporal y horizonte de análisis); las unidades de intervención (coste por persona intervenida, costes medios, netos, marginales); las diferencias en la estructura

de costes de las alternativas comparadas, y la disponibilidad de información.

Posteriormente hay que introducir otros factores en la evaluación como<sup>164</sup>:

- La actualización de los valores económicos en el tiempo (tasa de descuento o conversión de costes y beneficios a su valor actual según la preferencia temporal, ajustes según la inflación, amortización de costes iniciales).
- La valoración de la validez y robustez del estudio económico y de las asunciones realizadas mediante un *análisis de sensibilidad*.
- Las *externalidades* (son efectos sobre unidades no relacionadas directamente con la intervención o actividad; por ejemplo, el efecto positivo que produce una inmunización sobre las personas no vacunadas).
- Los posibles *efectos distributivos diferenciales* de costes y beneficios en distintas poblaciones o contextos.
- Las *limitaciones presupuestarias*.
- Los *costes intangibles*.

Después de medir los costes y los resultados en salud de cada tecnología se construyen indicadores que permitan comparar entre las alternativas evaluadas (valor actualizado neto, razón entre beneficios y costes, razón coste/efectividad). Se pueden utilizar medidas sumariadas como la *razón coste/efectividad*, pero es más aconsejable presentar los resultados como *razones incrementales o marginales*, comparando distintas TS alternativas o distintos niveles de intervención de una misma TS<sup>164</sup>. El enfoque marginal expresa el cambio en la relación coste/efectividad que se produce al aumentar una unidad de resultado en cada TS. Por ejemplo, un día más de trabajo, o una unidad más de una droga, o una unidad monetaria más de gasto en un programa de prevención. En contraste, el enfoque incremental analiza el cambio en la relación coste/efectividad pero siempre de una TS en comparación con otra. Generalmente, se calcula la razón que divide el incremento de coste por el incremento de efectividad de varias TS frente a la considerada más efectiva<sup>165</sup>.

La utilización de estos estudios en el área sanitaria es relativamente reciente y está alcanzando un gran desarrollo y difusión. Las técnicas de análisis económico están muy consolidadas, pero no tanto su aplicación en sanidad, ya que es un área donde resulta especialmente difícil la valoración económica

de costes y beneficios. Además existen varios problemas metodológicos sin resolver, como por ejemplo, la utilización de las tasas de descuento o la valoración económica de los años de vida ajustados por calidad<sup>130</sup>. Con respecto a su utilización para la elaboración de un informe de ETS es preciso indicar que el estudio del contexto local, que todo informe de ETS debe contener, conlleva siempre la realización de una valoración económica. Ésta, aunque no consista en un estudio de evaluación económica estándar, debe incluir un estudio de los costes y el impacto económico del procedimiento estudiado para lo que es necesario considerar la metodología descrita anteriormente.

### 7.5. El juicio de expertos

La necesidad de adopción de decisiones sobre una TS pese a la ausencia de evidencias científicas de alta calidad justifica en ocasiones el uso de métodos de consenso basados en el juicio de expertos. En otros casos, el problema analizado o las preguntas más relevantes a las que se pretende dar respuesta presentan numerosos componentes subjetivos, como la valoración de la calidad de vida, el interés social, determinados aspectos éticos o posiciones políticas. En estos y otros casos puede ser relevante añadir a la síntesis de la evidencia la valoración subjetiva y la opinión de especialistas experimentados o de representantes de las personas implicadas e intereses afectados por el uso de la TS. Es deseable también obtener un cierto grado de consenso y aceptación entre los que van a ofrecer o utilizar esa TS, de forma que se facilite su diseminación e implantación<sup>15</sup>.

Los procedimientos de juicio de expertos utilizan técnicas con grados diferentes de formalización y estructuración, y emplean métodos cualitativos de investigación y de decisión. Existe una enorme variedad de aproximaciones metodológicas para la elaboración de productos de acuerdo y consenso. Estos métodos varían en el grado en el que los paneles de expertos adoptan normas formales estrictas y explícitas de interacción y comunicación para llegar al consenso; en el grado en el que cada recomendación se sustenta en la evidencia científica; en la diversidad o heterogeneidad de los expertos representados en el panel, y en el nivel de estructuración explícita e integración en el informe de factores como la evaluación económica, la aceptabilidad, el valor de utilidad y las preferencias de los pacientes<sup>166</sup>.

La búsqueda de respuestas a determinadas preguntas planteadas ha sido abordada con distintos métodos que generan también distintas líneas de productos: Conferencias de Consenso<sup>167</sup>; Método de Uso Apropriado, diseñado por la RAND Corporation<sup>10,168</sup>; Método DELPHI<sup>169</sup>; Grupos Nominales, y otros. Estos procedimientos suelen implicar la participación de expertos en un proceso de valoración de problemas o situaciones clínicas con el apoyo de la literatura biomédica disponible y la elaboración de un documento consensuado en el que se propone una pauta de decisión frente al problema planteado. Así, en los procesos de elaboración de Guías de Práctica Clínica se llega a un acuerdo sobre un conjunto de recomendaciones, diseñadas para ayudar a los profesionales y a los usuarios en la selección de las mejores opciones diagnósticas y/o terapéuticas para el manejo de una condición clínica específica en circunstancias concretas.

El método RAND de uso apropiado constituye un medio de detección de acuerdos sobre las indicaciones de una TS determinada mediante el uso de paneles de expertos. El procedimiento estándar consiste en la valoración de los riesgos y beneficios de cada una de las posibles indicaciones identificadas para una técnica médica o quirúrgica, por parte de un grupo de expertos de distintas especialidades relacionadas con la tecnología analizada. Los miembros del grupo revisan la literatura científica que se les proporciona y valoran individualmente, en función de ésta y de su propio juicio, cada indicación, otorgando una puntuación. El grupo discute después las puntuaciones alcanzadas por cada indicación y se vuelven a otorgar puntuaciones individualmente. En el análisis final de los resultados se establece el grado de acuerdo o desacuerdo para la indicación en cada situación clínica y se clasifican las indicaciones como apropiadas, inciertas o inapropiadas<sup>28</sup>.

En las Conferencias de Consenso se alcanza un acuerdo escrito en un proceso formalizado semejante a un juicio, en el que participan expertos y representantes de grupos implicados en la decisión clínica objeto de la pregunta. El modelo más difundido es el establecido por los Institutos Nacionales de la Salud de los EE.UU. a través de la *Office for Medical Applications of Research*<sup>167,170</sup>. Al final de un largo proceso de preparación, un tribunal revisa la documentación elaborada para la conferencia y escucha a testigos y expertos, para después deliberar y redactar un acuerdo basado en las pruebas y testimonios presentados. Esta metodología es relativamente compleja y

consume considerables esfuerzos, tiempo y recursos. Se ha utilizado preferentemente en los EE.UU. y se ha difundido en algunos países europeos, Canadá y Australia. Permanecen numerosas incertidumbres sobre aspectos de organización y contenidos que determinan la validez y utilidad del procedimiento; como los promotores, los patrocinadores, la representatividad, el proceso de valoración de las evidencias o los criterios y requisitos para su realización.

La principal debilidad de los procedimientos basados en el juicio de expertos es la incertidumbre sobre su validez y reproducibilidad. Por otro lado, estos procedimientos hacen un uso desigual y no protocolizado de la evidencia científica disponible. Son procesos en muchos casos de larga duración que consumen elevados recursos y sus productos requieren una actualización continua. Existen escasos trabajos originales que hayan abordado las limitaciones de estos procedimientos. Quedan por resolver otros interrogantes sobre la adecuación de los sistemas de selección de participantes, la composición de los paneles, la idoneidad de los promotores, financiadores y organizadores, así como la perspectiva de evaluación de estos procedimientos y sus productos. En contrapartida, permiten obtener recomendaciones consensuadas sobre problemas importantes objeto de decisiones clínicas, o de otro tipo, que se han de tomar cuando no hay evidencias científicas disponibles<sup>171</sup>.

Estos métodos permiten, a su vez, la implicación de profesionales sanitarios de prestigio y de representantes de otros estamentos involucrados o afectados por el proceso de toma de decisiones (usuarios, pacientes, Administración, financiadores), lo que facilita la aceptación del proceso de elaboración, la colaboración en la estrategia de difusión y la aceptación de las recomendaciones. Los resultados finales de estos métodos de consenso dependerán en gran medida del rigor metodológico, el grado de preparación, la transparencia de todo el procedimiento y el prestigio de la organización promotora.

## 8. Valoración de los efectos sociales, organizativos, psicológicos y éticos

La cantidad de factores que pueden influir y la diversidad de parámetros que se pueden

utilizar para medir este tipo de efectos son algunas de las múltiples dificultades que presenta su valoración e inclusión en un estudio de evaluación. Las expectativas acerca de la calidad de vida o la salud percibida, por ejemplo, pueden matizarse enormemente de tal manera que cada individuo considere variables o factores diferentes a los demás. Sin embargo, muchas TS ocasionan efectos psicológicos importantes en los pacientes, tienen un elevado impacto social o plantean problemas éticos relevantes. También su impacto organizativo puede influir en el uso y resultados finales de la intervención sanitaria. Por ello, su inclusión en un informe de ETS es importante y merece un apartado específico.

Se ha avanzado algo en la identificación de este tipo de efectos y en el establecimiento de escalas de medida para valorarlos<sup>172,173,174,175,176,177</sup>. Así, la realización de pruebas diagnósticas produce *efectos psicológicos* como el "etiquetado" (clasificar a una persona de enferma tiene consecuencias psicológicas), el miedo a la enfermedad, la ansiedad por el contacto con los servicios sanitarios, la satisfacción con la atención recibida. Los anteriores son aspectos que valoran los resultados. Algunas variables que miden el proceso también pueden ser relevantes psicológicamente, como las creencias y percepciones sobre la salud o la sensación de control de la enfermedad que tiene el paciente durante un tratamiento continuado. Se ha observado, por ejemplo, la influencia determinante del mensaje emitido por el médico en la aceptación de un tratamiento y, por tanto, en la efectividad de la práctica clínica<sup>176</sup>.

La valoración de los *efectos sociales* incluye aspectos como la equidad en el acceso a la TS<sup>28</sup>, los condicionantes de la toma de decisiones colectiva sobre inclusión de prestaciones o realización de intervenciones, la asignación de valores económicos a los resultados clínicos, las necesidades expresadas o percibidas en la comunidad y otros<sup>178</sup>.

A continuación se describen algunas de las variables que pueden medir los *efectos de la introducción de una tecnología en una organización sanitaria*<sup>15</sup>:

- Grado de flexibilidad/vulnerabilidad de la organización a los efectos de la introducción de la TS sobre las limitaciones presupuestarias, el mantenimiento general, los momentos críticos, la falta de personal entrenado en la nueva tecnología o la carga de trabajo.

- Grado de centralización/descentralización de la información, del desarrollo de los procedimientos técnicos, de las decisiones clínicas y de las decisiones económicas o sobre los recursos de la tecnología.
- Nivel de diferenciación que ocasiona la tecnología en la estructura de la organización (creación de un nuevo departamento) y en los niveles del personal (nuevos expertos).
- Requerimientos de personal, de formación o de incentivos.
- Satisfacción en el trabajo (medida mediante diversos aspectos como la seguridad, adaptación al medioambiente, buenas relaciones personales, efectividad y motivación con el trabajo).
- Grado de efectividad de los canales de comunicación, destacando los canales de datos clínicos y administrativos.

La *dimensión ética* en la ETS se relaciona con la evaluación socioeconómica, en el marco de un contexto de recursos limitados. El acceso equitativo a la provisión de servicios sanitarios es un derecho de los ciudadanos. A la vez, el objetivo de los servicios de salud es maximizar el impacto sobre la salud con los recursos que la sociedad establece para ello. Así, la búsqueda de la eficiencia se puede definir como un objetivo ético<sup>28</sup>.

Otro posible enfoque es la identificación de las implicaciones éticas de la utilización de una TS a través de la detección de los puntos de encuentro y acuerdo entre los diferentes grupos de personas implicadas, o *evaluación interactiva*, pero requiere habilidades específicas alejadas de las actividades fundamentales relacionadas con la ETS y más relacionadas con las técnicas de marketing o la investigación cualitativa. Existen más enfoques que relacionan la ética con la ETS: la búsqueda de la ética en la buena práctica clínica; los incentivos a los profesionales sanitarios como análisis de la ética y valor social de una organización; las connotaciones éticas de los resultados de las intervenciones sanitarias; la falta de justificación ética de un consumo excesivo de recursos, o el uso apropiado de las TS como planteamiento ético<sup>28</sup>.

## 9. Elaboración de las conclusiones y recomendaciones

Las conclusiones y recomendaciones son los apartados más leídos de los informes, junto

con el resumen. En las conclusiones deben destacarse aquellos contenidos del informe que tengan mayor relevancia y capacidad de influir en la toma de decisiones. La realización de recomendaciones debe ser adecuadamente valorada, en función de los objetivos, contenidos y las necesidades del solicitante. Este apartado debe ser breve, preciso y consistente con los datos y evidencias aportados. Se deben evitar las conclusiones o recomendaciones obvias o poco relevantes, y las que se basen en asunciones o decisiones arbitrarias realizadas durante el proceso de análisis y síntesis de la información. Los aspectos que deben ser reseñados habitualmente son<sup>15,45</sup>:

- Valoración de la calidad de la información obtenida, así como de las incertidumbres detectadas.
- Valoración de los costes (tanto económicos como relativos a riesgos, complicaciones y efectos adversos) y beneficios del uso de la tecnología. En algunos casos, concluir si la aplicación de la TS es factible en el medio estudiado.
- Valoración de las posibles implicaciones del informe para el Servicio Sanitario en el que se aplica la tecnología.
- Grado en que se han logrado cumplir los objetivos propuestos.
- Formulación de la respuesta o respuestas a las preguntas de evaluación u objetivos específicos.
- Incertidumbres o lagunas de conocimiento que no han podido ser cubiertas con la información obtenida.
- Líneas de investigación que deberían desarrollarse para avanzar en el conocimiento del tema que se ha evaluado.
- Implicaciones para el desarrollo o modificación de regulaciones referidas a la tecnología evaluada.

Para la elaboración de conclusiones sobre la consistencia de una relación causa-efecto entre una intervención y un resultado clínico se utilizan frecuentemente *pruebas indirectas*: unidades de resultado intermedias o *surrogate end-points*, relaciones establecidas a partir de tecnologías o intervenciones análogas u obtenidas en poblaciones similares. Aunque muchas veces las pruebas indirectas aportan conocimientos útiles, es necesario advertir y valorar pormenorizadamente las limitaciones de su utilización.

La calidad de las evidencias y de las recomendaciones que surgen de ellas es un aspecto

que debe ser mencionado siempre en las conclusiones y recomendaciones. Un informe de ETS debe informar del grado de evidencia que apoya cualquier conclusión o recomendación, estableciendo claramente el orden de preferencia que se da a los diferentes resultados. En cualquier caso, un informe de ETS no sólo debe valorar en sus conclusiones los efectos establecidos a partir de datos válidos, sino que debe considerar todos los estudios con alguna aportación importante sobre la tecnología a evaluar. En un gran número de casos, los estudios sobre las tecnologías sanitarias implantadas, no sólo sobre las que están introduciéndose, no van a tener la validez suficiente como para establecer conclusiones, por lo que el informe de ETS deberá basarse en los datos disponibles, aunque no sean de la calidad deseada.

La clasificación y jerarquía de las evidencias científicas es un tema ampliamente abordado en la literatura reciente. Se han propuesto muchos criterios y métodos de clasificación, algunos de los cuales se han seleccionado en el anexo IV.

Un aspecto que debe analizarse en las conclusiones es la aplicabilidad de los resultados finales del informe. Se debe ofrecer una ayuda a los lectores para identificar si las conclusiones son aplicables a una población concreta, de tal manera que se describan las posibles variaciones de la validez o efectos si va-

rían a su vez determinadas circunstancias, tales como las características biológicas de los participantes o del factor de exposición, por ejemplo; las diferentes cepas de una bacteria; determinados componentes culturales que afecten a los resultados de la intervención sanitaria; la adherencia al tratamiento, o el riesgo basal previo a la intervención<sup>31</sup>. Los factores enumerados pueden producir enormes variaciones, por lo que no se puede exigir que un informe de evaluación aborde todos esos posibles escenarios, pero sí se deben analizar las diferencias más importantes e incluir estas variaciones dentro de las posibles limitaciones de la evaluación realizada.

Finalmente, es interesante que las conclusiones establezcan el grado de apoyo o confianza que merece la TS a juicio de los autores del informe. Una clasificación que puede ser útil para ese propósito es la elaborada por el Grupo de Embarazo y Parto de la Colaboración Cochrane (ver tabla 10), que clasifica las intervenciones sanitarias en dos grupos: aquellas para las que hay información disponible como para extraer conclusiones firmes y aquellas para las que se necesita investigación adicional.

Entre los errores comunes en el proceso de elaboración de conclusiones y recomendaciones está la confusión entre la calificación "No hay evidencia de efecto" y "Evidencia de ningún efecto". El primer caso se refiere al

**Tabla 10**  
**Clasificación de las intervenciones sanitarias según el grado de recomendación que se puede establecer para su uso basado en evidencias científicas (modificado de Enkin<sup>179</sup>, Mulrow<sup>31</sup>)**

<p>1. Intervenciones sanitarias (IS) sobre las que hay evidencia suficiente para establecer recomendaciones sólidas y aplicables a la práctica:</p> <p>1.1. IS cuyos beneficios compensan a los efectos adversos o costes (por ejemplo, la continuidad del personal sanitario durante el parto).</p> <p>1.2. IS que deben ser abandonadas según la evidencia disponible (por ejemplo, cesárea electiva ante una diabetes gestacional).</p> <p>1.3. IS en las que hay equilibrio entre beneficios y efectos adversos conocidos (por ejemplo, analgesia epidural durante el parto frente a otras formas de alivio farmacológico).</p> <p>2. IS para las cuales las evidencias disponibles son insuficientes si se desea proporcionar recomendaciones sólidas y aplicables a la práctica, pero que deberían incluirse entre las prioridades de la investigación sanitaria:</p> <p>2.1. IS que parecen prometedoras, pero que requieren una evaluación adicional (por ejemplo, apoyo al personal sanitario durante un embarazo de riesgo).</p> <p>2.2. IS en las que no se han demostrado los efectos esperados, pero que merecen atención adicional (por ejemplo, suplemento férrico continuado durante el embarazo).</p> <p>2.3. IS con evidencia razonable de que no son eficaces para la finalidad perseguida (por ejemplo, regulación dietética ante una diabetes gestacional).</p>
--

punto 2 de la clasificación anterior, y la proposición "Evidencia de ningún efecto" se refiere al punto 1.2. Otra equivocación habitual es presentar conclusiones o establecer recomendaciones que van más allá de los datos que contiene el informe. También deben evitarse afirmaciones del tipo de "Se necesita más investigación". El informe de evaluación debe manifestar exactamente qué investigación se necesita y por qué, y cómo se puede mejorar la evaluación con datos o recursos adicionales.

Las conclusiones y las recomendaciones deben responder a los objetivos y preguntas específicos planteados en el informe, sobre todo de la pregunta política, y a las funciones que tenga asignada la agencia en la regulación legal o normativa de la tecnología a evaluar. A su vez, deben ser útiles a los solicitantes y aplicables en el sistema sanitario implicado.

---

## 10. Revisión del documento provisional

Se debe reservar un período de tiempo suficiente para realizar una revisión externa y completa del texto antes de su publicación. La revisión por pares o *peer review* es una práctica asumida en la actualidad por toda la comunidad científica y considerada actualmente uno de los elementos clave que aseguran la calidad de una publicación<sup>180,181</sup>. Los informes de ETS, aunque en ocasiones no utilizan de forma estricta ese control de calidad, deben incluir un proceso de revisión externa.

El proyecto de evaluación debe prever quiénes van a ser los revisores, además de los recursos materiales y del tiempo necesarios para ello. Es interesante que en el grupo de revisores se encuentren tanto expertos en el tema y en la metodología de evaluación como posibles receptores del informe para que estos últimos puedan valorar la utilidad, claridad y relevancia del documento. Muy probablemente serán necesarias varias revisiones con sus correspondientes correcciones e inclusión de nuevos puntos en el informe, aunque será necesario establecer un calendario firme, que fije sobre todo la fecha final, ya que en caso contrario el proceso de revisión puede prolongarse mucho en el tiempo.

La última reunión entre los revisores y redactores debe decidir los aspectos relacionados con la edición y la estrategia de difusión. Esta

última, debido a su relevancia, se discute a continuación en una sección específica. Además, la edición va a requerir un tiempo adicional para que a través de una o varias pruebas de imprenta se asegure la publicación sin errores y con la presentación deseada.

---

## 11. Difusión y aplicación de los resultados del informe

Esta fase es fundamental para asegurar que un informe de ETS sea efectivo, consiga cambiar la práctica clínica o influir en las decisiones de las autoridades relacionadas con las TS. No basta con elaborar informes basados en la mejor evidencia disponible, también es importante diseñar una estrategia que consiga la difusión y aplicación de sus conclusiones en la práctica diaria. Los informes de ETS deben basarse en la información científica más actualizada y de mayor calidad, pero esa información debe tener el impacto suficiente como para influir en los comportamientos y actitudes de los profesionales o los gestores sanitarios.

En esta fase del proceso de elaboración de informes de ETS los conceptos *dissemination* e *implementation*, que se podrían traducir como *diseminación* e *implementación*, pueden considerarse objetivos importantes en la búsqueda del mayor impacto posible del informe de ETS. El concepto *dissemination* se refiere a un conjunto de actividades muy amplias: la difusión del documento, la estrategia de búsqueda de los lectores clave, el diseño de actividades encaminadas a la discusión e interpretación correcta de las conclusiones del informe, e incluso la participación en procesos de formación y entrenamiento de los profesionales sanitarios que utilicen la TS que se evalúe. Es un concepto que pretende añadir un nuevo valor o componente a los términos *difusión* o *divulgación*. La *diseminación* se corresponde con una difusión más activa, en la que a través de intervenciones o actividades se promueve, no sólo se difunde, una información o una recomendación.

El concepto *implementation* se refiere a actividades que van más allá de la extensión o difusión de un documento, como la aplicación y evaluación en la práctica de las conclusiones y/o recomendaciones. Incluye intervenciones que reduzcan las barreras al cambio, actividades para conseguir la asimilación y posterior aplicación de los resultados del informe,

el seguimiento y evaluación del uso de la tecnología, la medición del impacto del informe y la valoración a largo plazo de los efectos que produce la TS sobre la organización sanitaria o la población<sup>182</sup>.

La evaluación del impacto de los informes de ETS es una actividad que todavía está poco desarrollada. La mayoría de los estudios diseñados con ese objetivo analizan la difusión, comprensión y asimilación de la información resultante de la evaluación, pero no llegan a medir los cambios y el impacto que ésta produce en la práctica clínica. Las limitaciones metodológicas de los ensayos de intervención comunitaria para atribuir cambios a una determinada intervención influyen en la falta de desarrollo de esta línea de evaluación. El diseño de estudios de evaluación antes-después es el diseño más apropiado dentro de los ensayos de intervención comunitaria<sup>182</sup>.

La forma tradicional de comunicación entre los investigadores y los médicos clínicos es la revista científica, pero en muchos casos no ha demostrado efectividad para influir en la práctica diaria de la medicina<sup>183</sup>. El análisis

de los factores que influyen y cambian la práctica clínica, la política sanitaria o la actitud de los pacientes forma parte de las materias de las ciencias del comportamiento, como la sociología y la psicología. Diversos estudios han intentado diseñar modelos explicativos del proceso y los factores que influyen en la atención sanitaria, pero todos ellos han llegado a la conclusión común a todos ellos: la complejidad del tema de análisis<sup>45,184</sup>.

El éxito del proceso de transferencia de conocimientos de la evaluación e investigación a la práctica clínica precisa el desarrollo de una estrategia de difusión y aplicación de los resultados apoyada en actividades e incentivos organizativos, que estimulen el cambio. Esta estrategia debe ir dirigida no sólo a los profesionales sanitarios, sino también a los investigadores, directivos sanitarios, profesionales de los medios de comunicación y ciudadanos. Además, es importante que se planifique y se haga explícita en las primeras fases de la elaboración del informe de ETS, cuando se diseña el proyecto de evaluación.



## Anexo I. Formato y estructura de un informe de ETS

El formato y estructura de los informes no ha sido consensuado entre las agencias u organizaciones que se dedican a la ETS. Lo que se pretende, en suma, es trasladar toda la información científica identificada a un texto comprensible y útil, de tal manera que cada tipo de usuario pueda apoyarse en la rigurosidad y fiabilidad del informe para tomar decisiones. A continuación se propone un formato básico aplicable a los informes públicos; en el caso de los informes a demanda o internos no es posible proponer un formato homogéneo, dado que sus características básicas son la adaptabilidad y flexibilidad.

— **TÍTULO.** Debe ser preciso y describir con claridad los objetivos del informe.

— **RESUMEN.** Un primer resumen en el idioma original y un segundo en inglés, que deben ser breves. Es lo que se identifica como *Summary* o *Executive Summary*. Es aconsejable que se divida en secciones, relacionadas con la estructura del informe final, como: Objetivos, Fuentes de información, Selección de la información, Métodos de análisis y síntesis de la información, Resultados, Conclusiones y Recomendaciones.

— **RESUMEN O ABSTRACT ESTRUCTURADO EN INGLÉS.** Además del resumen descrito en el punto anterior, es necesario escribir también un resumen estructurado para su inclusión en el registro de informes de la red INAHTA, cuya dirección aparece en el anexo II. En la tabla 11 se propone un abstract estructurado, que modifica el que utilizaba anteriormente la citada red y se asemeja a la propuesta que realizó un grupo de trabajo de INAHTA en 1997<sup>185,186</sup>. Esta asociación no tiene actualmente un formato establecido para esta sección.

— **INTRODUCCIÓN.** En esta sección se presenta el documento, indicando las razones que motivaron el inicio del trabajo, profundi-

zando en la información requerida, las lagunas de información que pretende cubrir y los posibles beneficios que puede suponer. Indicar también la entidad solicitante o financiadora del informe, los posibles destinatarios de éste, algunos datos sobre la situación actual y evolución de la tecnología que describan ésta en su contexto internacional y local, y otros antecedentes que sitúen al lector permitiéndole profundizar a través de las referencias ofrecidas.

Algunos informes en lengua inglesa distinguen entre *introduction* y *background*. En el primer apartado incluyen los contenidos de una presentación, y en el segundo describen los antecedentes y el marco de referencia. En algunos informes puede ser útil incluir algún otro apartado de entrada como prólogo, presentación o prefacio.

— **OBJETIVOS O PREGUNTAS.** Pueden situarse tanto en la introducción como en un apartado específico. Dos cualidades esenciales son precisión y brevedad. Cada objetivo debe iniciarse con un verbo que describa la acción a realizar; por ejemplo, "Comparar la eficacia de las tecnologías A y B". Generalmente existe una cuestión de ámbito político o relacionada con la gestión sanitaria que motiva la realización del informe; es el *objetivo general*, que a su vez se desarrolla mediante uno o varios *objetivos específicos* o preguntas científicas.

— **METODOLOGÍA.** Se debe incluir una descripción de los métodos empleados en la realización del informe, con la claridad y los detalles suficientes como para que puedan reproducirse todos los pasos. Las fuentes de información, la estrategia de búsqueda, los criterios de inclusión y exclusión de los estudios y la metodología utilizada para analizar y sintetizar la información son los aspectos más importantes que deben especificarse en esta sección del informe.

— **DESARROLLO DE LOS CONTENIDOS DE LA EVALUACIÓN.** Este apartado muestra el análisis realizado (juicio de expertos, técnica cuantitativa) y sus resultados, así como la síntesis de los estudios considerados. También se incluye aquí la descripción de la tecnología y enfermedad (ver tabla 12), el estudio de la validez y calidad de los datos de los

**Tabla 11**  
**Datos que se deben incluir en los abstracts estructurados en inglés**

1.	<b>Title.</b> Primero el título en inglés y luego el título en el idioma original entre paréntesis.
2.	<b>Author/s.</b> Debe usarse el estándar de Vancouver <sup>187</sup> .
3.	<b>Organization.</b> Por ejemplo, Agencia de Evaluación de Tecnologías Sanitarias.
4.	<b>Contact person.</b> Dirección y nombre de la persona que la organización seleccione.
5.	<b>Date.</b> Mes y año de publicación.
6.	<b>Language.</b> Idioma o idiomas de publicación.
7.	<b>Abstract:</b> Especificar si incluye otro resumen además del abstract estructurado INAHTA y el idioma o idiomas en que están escritos.
8.	<b>Publication type.</b> Por ejemplo: informe, guía de práctica clínica.
9.	<b>Pages.</b> Número de páginas total del informe.
10.	<b>References.</b> Número de referencias que cita el documento.
11.	<b>ISBN.</b> International Standard Book Number.
12.	<b>Technology type.</b> Categoría de la técnica o intervención sanitaria que se evalúa: screening, preventive, diagnostic, therapeutical, palliative, rehabilitating.
13.	<b>Subject index terms.</b> Deben utilizarse las palabras clave que proporciona el Index Medicus en su thesaurus. Indicar primero los Major Descriptors marcándolos con asterisco y después los Minor Descriptors. Si la palabra no está incluida en el thesaurus del Index Medicus, poner "(Non MeSH)" después de esa palabra. Es interesante poner descriptores secundarios que maticen, separados con un guión. Ejemplo: *Aortic Aneurysm – epidemiology; *Stents; Blood Vessel Prothesis; Khar-kov Stent (Non MeSH).
14.	<b>Objectives.</b> Objetivos generales y específicos.
15.	<b>Methods: Data sources:</b> Datos utilizados y su origen. <b>Criteria for study inclusion:</b> Pautas utilizadas para la inclusión y exclusión de estudios. <b>Primary data collection:</b> Especificar si se realizó un estudio con datos primarios <i>ad hoc</i> . <b>Secondary data analysis:</b> Especificar si se realizó un estudio basado en datos secundarios como los obtenidos en registros o historias clínicas. <b>Literature review and integration of evidence:</b> Bases de datos bibliográficas, expertos u otras fuentes de información consultadas. Método de revisión o técnica de síntesis utilizada (revisión sistemática, meta-análisis, evaluación económica, análisis de decisión).
16.	<b>Results.</b> Contenido del informe y resumen de los hallazgos más importantes.
17.	<b>Recommendations.</b>
18.	<b>Peer review process:</b> Especificar: Yes / No / Internal / External / Both.

estudios seleccionados, la descripción de los estudios incluidos y excluidos, el estudio del contexto local, la valoración del impacto social, organizativo, psicológico y ético de la TS, y finalmente, una discusión donde se analicen las limitaciones de la metodología empleada en la evaluación y se valore de manera integral el conjunto de los contenidos definidos en esta sección.

— **TABLAS Y/O GRÁFICOS.** En la mayor parte de los informes son un elemento fundamental de la síntesis de información. Generalmente van a servir para describir los estudios incluidos y excluidos del análisis. A modo de esquema, los datos que se incluirán van a ser fundamentalmente: características de los pa-

cientes receptores de la tecnología, tipo de tecnología, diseño del estudio, calidad de las evidencias y resultados de los estudios<sup>45</sup>.

— **CONCLUSIONES.** Los hallazgos del estudio evaluativo tienen una serie de implicaciones o pueden plantear nuevas preguntas. Éste es el apartado adecuado para explicitar estas cuestiones realizando un compendio de los resultados más importantes, incluyendo las limitaciones y lagunas de conocimiento detectadas. Cada conclusión debe ser breve, precisa, relevante y consistente con las evidencias que el informe aporta.

— **RECOMENDACIONES.** No todos los informes incluyen recomendaciones, aunque son importantes. De la misma forma que para las

**Tabla 12**  
**Descripción de la tecnología y de la enfermedad**

Es evidente que un informe de ETS debe describir el objeto de su evaluación, aunque esta información no siempre es necesaria en el mismo grado. Su extensión debe ser apropiada, de manera que aporte elementos esenciales para la comprensión del documento pero que no se extienda en aspectos demasiado técnicos que no añadan valor a la información relacionada con la pregunta de evaluación.

El tipo de información sobre la enfermedad o condición clínica que puede ser útil incluye los siguientes conceptos: fisiopatología, historia natural, datos epidemiológicos sobre morbilidad y mortalidad, incluyendo incidencia y prevalencia, estimaciones del impacto de la enfermedad en la población y características demográficas de ésta.

La información técnica sobre la TS se puede clasificar de la siguiente manera:

- Características relevantes del mecanismo de funcionamiento de la TS, tanto físicas como biológicas o funcionales. Por ejemplo, en el caso de un fármaco se debe aportar información básica sobre sus características farmacocinéticas y su interacción con otros fármacos. En una TS diagnóstica se precisa conocer datos sobre su validez, precisión, reproducibilidad y valores predictivos.
- Cumplimiento de estándares o criterios de calidad establecidos.
- Uso en la práctica diaria. Pueden incluirse aspectos como el tiempo necesario para realizar el procedimiento, riesgo que asumen los trabajadores y usuarios y nivel de entrenamiento necesario.
- Estrategias de control de calidad y mantenimiento.

conclusiones, las recomendaciones deben obtenerse a partir de evidencias científicas y basarse exclusivamente en los contenidos o resultados del informe<sup>15</sup>.

— **BIBLIOGRAFÍA.** Lo ideal es que se transcriban cumpliendo las normas de Vancouver<sup>187</sup>. Las referencias de la metodología utilizada para la elaboración del informe son también

relevantes. Es importante que no se olvide incluirlas. Asimismo, la bibliografía puede presentarse en tres apartados: los estudios incluidos en el análisis final, los excluidos y aquellos estudios que se han utilizado en la metodología o en el estudio general del tema a evaluar pero que no son objeto de análisis<sup>188</sup>.

---

## Anexo II. Recopilación de fuentes de información útiles en ETS

El presente anexo se configura como una ayuda para la localización de diversos tipos de información. En su elaboración no se ha pretendido alcanzar una gran exhaustividad, que podría dificultar su manejo y concreción, sino ofrecer un listado de las fuentes de información de mayor relevancia que sea de manejo fácil. Por otro lado, hay que reconocer la existencia de un claro sesgo hacia la presencia de fuentes de origen español, especialmente en el apartado de las fuentes de información administrativa. Sin embargo, consideramos importante su inclusión ya que, en cualquier caso, permite orientar sobre el tipo de fuentes que pueden ser de utilidad en otros contextos.

Al final del anexo se ha incluido un listado que recoge una selección de fuentes, excluyendo las guías de práctica clínica, que pueden ser exploradas como paso preliminar de la búsqueda de información y para localizar rápidamente revisiones ya existentes.

### Índice del anexo:

1. Selección de fuentes de información para la búsqueda preliminar .....	50
2. Fuentes bibliográficas y documentales ...	50
2.1. Bases de datos de publicaciones e investigaciones originales .....	50
— Bases de datos generales .....	50
— Bases de datos específicas o temáticas.....	51
2.2. Fuentes de información secundaria .....	51
— Bases de datos .....	51
— Publicaciones secundarias .....	52
— Guías de práctica clínica .....	52
— Literatura gris .....	52
2.3. Productores o distribuidores de bases de datos .....	53
3. Fuentes estadístico-administrativas .....	53
3.1. En España .....	53
4. Buscadores de información sanitaria en Internet .....	54
5. Estrategias prediseñadas de búsqueda bibliográfica .....	54

### 1. SELECCIÓN DE FUENTES DE INFORMACIÓN PARA LA BÚSQUDA PRELIMINAR

Listado de fuentes de información útiles para la búsqueda de información inicial. Ésta se realiza fundamentalmente cuando se diseña la evaluación.

- **Red INAHTA** (<http://www.inahta.org>).
- **Base de datos de revisiones sistemáticas de la Colaboración Cochrane** (<http://www.update-software.com/ccweb/cochrane/cdsr.htm>).
- **HSTAT** (Health Services/Technology Assessment Text) (<http://text.nlm.nih.gov/>).
- **TRIP database** (<http://www.gwent.nhs.gov.uk/trip/test-search.html>).
- **SchARR-Lock's Guide to the Evidence** (<http://www.shef.ac.uk/uni/academic/R-Z/scharr/ir/sceb.htm>).
- **IDEA Topics List** ([http://www.ohsu.edu/bicc-informatics/ebm/ebm\\_topics.htm](http://www.ohsu.edu/bicc-informatics/ebm/ebm_topics.htm)).
- **MEDLINE**: Buscar fundamentalmente artículos de revisión o de síntesis.  
PubMed: (<http://www.ncbi.nlm.nih.gov/PubMed/>)  
GratefulMed: (<http://igm.nlm.nih.gov/>)

### 2. FUENTES BIBLIOGRÁFICAS Y DOCUMENTALES

#### 2.1. Bases de datos de publicaciones e investigaciones originales

*Bases de datos generales*

**Medline**: Es la principal base de datos de la National Library of Medicine (NLM) de EE.UU., y abarca los campos de medicina, enfermería, odontología

y veterinaria. Contiene citas bibliográficas de más de 3.900 revistas biomédicas y más de 9 millones de registros desde 1966. LA NLM ofrece actualmente dos sistemas de acceso gratuito a Medline vía Internet:

- PubMed. Internet (<http://www.ncbi.nlm.nih.gov/PubMed/>).
- GratefulMed. Internet (<http://igm.nlm.nih.gov/>).

Accesible también en CD-ROM.

**Embase:** Es la base de datos electrónica de Excerpta Médica. Comprende información farmacológica y biomédica. Producida por Elsevier Science, Amsterdam.

Accesible vía Internet (<http://www.elsevier.nl/>) y en CD-ROM.

**Uncover:** Base de datos que recoge más de 17.000 revistas en inglés de contenido multidisciplinar, no exclusivamente biomédico. Producido por CARL Corporation (Colorado Alliance of Research Libraries).

Accesible vía Internet (<http://uncweb.carl.org>).

**HSRProj:** Base de datos de la NLM. Contiene citas de proyectos de investigación en curso acerca de Servicios Sanitarios, incluyendo Evaluación de Tecnologías Sanitarias, así como el desarrollo y uso de guías de práctica clínica.

Accesible en Internet vía GratefulMed (<http://igm.nlm.nih.gov/>).

**LILACS (Literatura Latinoamericana y del Caribe en Ciencias de la Salud):** Referencias bibliográficas desde 1982. Producida por BIREME. Se encuentra accesible en CD-ROM, el cual contiene, además, otras bases de datos bibliográficas con contenidos sociosanitarios y legislativos de Latinoamérica.

Accesible vía Internet (<http://www.bireme.br/iahlii/homepage.htm>).

**National Research Register (NRR):** Registro electrónico de proyectos de investigación y desarrollo, en curso o recién terminados, financiados por el Servicio Nacional de Salud del Reino Unido (NHS). Contiene proyectos de investigación, ensayos clínicos y revisiones sistemáticas. Publicado por Update Software Ltd.

Accesible vía Internet (<http://www.update-software.com/default.htm>) o en CD-ROM.

**IME (Índice médico español):** Contiene la producción científica española en el área de biomedicina. CD-ROM distribuido por el Centro de Información y Documentación Científica (CINDOC).

Accesible vía Internet (<http://www.cindoc.csic.es/principal.html>) y en CD-ROM.

*Bases de datos específicas o temáticas*

**Bases de datos de la NLM (National Library of Medicine):** La NLM ofrece acceso a muchas bases de datos de contenidos más específicos, muchas de ellas

son también accesibles a través de GratefulMed. Brevemente, pueden destacarse las siguientes:

- **AIDSLINE:** Citas bibliográficas sobre SIDA y materias relacionadas.
- **AIDSTRIALS:** Ensayos clínicos relacionados con SIDA.
- **CANCERLIT:** Citas bibliográficas cuyo principal tema es el cáncer.
- **DIRLINE:** Directorio de fuentes de información sobre temas sanitarios, principalmente de EE.UU., pero también de algunas organizaciones internacionales.
- **HealthSTAR:** Citas bibliográficas sobre temas clínicos y de administración sanitaria. Incide principalmente en la evaluación de resultados y la efectividad de los procedimientos. Actualmente está siendo incorporada a PubMed.
- **PDQ:** Información actualizada mensualmente de tratamientos y ensayos clínicos sobre cáncer.
- **POPLINE:** Citas bibliográficas acerca de políticas poblacionales, planificación familiar, atención primaria y problemas específicos de los países en vías de desarrollo.

Internet ([http://www.nlm.nih.gov/pubs/factsheets/online\\_databases.html](http://www.nlm.nih.gov/pubs/factsheets/online_databases.html)).

**PsycINFO (Psychological Abstracts):** Recoge la literatura internacional en psicología y disciplinas relacionadas tales como medicina, psiquiatría, educación, trabajo social o ciencias sociales. Producida por la Asociación Americana de Psicología.

Accesible vía Internet (<http://www.apa.org/psycinfo/contact.html>) y en CD-ROM.

**ASSIA (Applied Social Sciences Index and Abstracts):** Contiene resúmenes y referencias bibliográficas de ciencias sociales aplicadas a diversos temas como geriatría, abusos infantiles, desempleo o reformas sanitarias. Un 40% de sus contenidos son del Reino Unido. Distribuida por Bowker-Saur Ltd.

Accesible vía Internet ([http://www.bowker-saur.co.uk/products/catalog/a\\_and\\_i/assia\\_plus\\_c.htm](http://www.bowker-saur.co.uk/products/catalog/a_and_i/assia_plus_c.htm)) y en CD-ROM.

**CINAHL (Cumulative Index to Nursing and Allied Health Literature):** Recoge literatura e información sobre enfermería y otras profesiones sanitarias relacionadas, como fisioterapia o terapia ocupacional, desde 1982.

Accesible vía Internet (<http://www.CINAHL.com>) y en CD-ROM.

## 2.2. Fuentes de información secundaria

*Bases de datos*

**The Cochrane Library:**

Contiene las siguientes bases de datos:

- Base de datos de revisiones sistemáticas (**CSR**D).
- Base de datos de resúmenes de revisiones de efectividad (**DARE**) (Accesible también vía CRD).
- Registro Cochrane de ensayos controlados (**CCTR**).
- Base de datos sobre metodología de las revisiones (**CRMD**).

Accesible vía Internet (<http://www.update-software.com/ccweb/cochrane/cdsr.htm>) y en CD-ROM.

**Red INAHTA** (International Network of Agencies for Health Technology Assessment): Referencias de los informes de evaluación de tecnologías sanitarias, realizados por las agencias pertenecientes a la red INAHTA. Pueden localizarse a través de Internet (<http://www.inahta.org>).

**ISTAHC Database** (International Society of Technology Assessment in Health Care): Recopilación del trabajo de los miembros de ISTAHC desde 1992. Contiene resúmenes de congresos y de publicaciones periódicas.

Accesible vía Internet (<http://www.istahc.org/live/viewdata>).

**HSTAT (Health Services / Technology Assessment Text)**: Producida por la NLM. Proporciona acceso a documentos completos de diversos tipos, tales como guías de práctica clínica y publicaciones de la AHCPR (Agency for Health Care Policy and Research), de los NIH (National Institutes of Health) y de los CDC (Centers for Disease Control and Prevention).

Accesible vía Internet (<http://text.nlm.nih.gov/>).

**CRD (Centre for Reviews and Dissemination)**: Dependiente del Servicio Nacional de Salud (NHS) británico. Produce las siguientes bases de datos:

- **DARE** (Base de datos de resúmenes de revisiones de efectividad). Incluida también en The Cochrane Library.
- **NEED** (NHS Economic Evaluation Database). Base de datos de resúmenes de estudios de evaluación económica de las intervenciones sanitarias.
- **HTA** (Health Technology Assessment). Base de datos de informes de Evaluación de Tecnologías Sanitarias de la Red INAHTA.

Accesible vía Internet (<http://nhscrd.york.ac.uk/welcome.html>).

**TRIP database**: Base de datos producida por TRIP (Turning Research Into Practice), grupo formado por profesionales sanitarios del Reino Unido. Permite explorar 18 bases de datos de materias relacionadas con la Medicina basada en la evidencia y acceder a los documentos contenidos en muchas de ellas.

Accesible vía Internet (<http://www.gwent.nhs.gov.uk/trip/test-search.html>).

**Programa británico de Evaluación de Tecnologías Sanitarias**: Permite acceder a informes de ETS y proyec-

tos de investigación financiados por el NHS y a las bases de datos de las divisiones regionales del NHS.

Accesible vía Internet (<http://www.soton.ac.uk/~hta/>).

**SCHARR-Lock's Guide to the Evidence**: Producida por la Sección de Fuentes de Información de la Escuela de Salud e Investigación Sanitaria (School of Health and Related Research, SCHARR) de la Universidad de Sheffield (UK). Es una guía de fuentes de evidencia impresas, ordenada por descriptores MeSH (Medical Subject Heading). Contiene principalmente literatura gris académica u oficial del Reino Unido.

Accesible vía Internet (<http://www.shef.ac.uk/uni/academic/R-Z/scharr/ir/sceb.html>).

**IDEA Topic List** (Internet Database of Evidence-Based Abstracts and Articles): Base de datos que recoge estudios seleccionados. Sus contenidos están ordenados por términos MeSH.

Accesible vía Internet ([http://www.ohsu.edu/bicc-informatics/ebm/ebm\\_topics.htm](http://www.ohsu.edu/bicc-informatics/ebm/ebm_topics.htm)).

#### Publicaciones secundarias

Son revistas que publican estudios basados en revisiones de otros ya publicados. Recogen trabajos que han demostrado una adecuada calidad con arreglo a los criterios de la medicina basada en la evidencia. Entre las muchas existentes en la actualidad, pueden destacarse:

- **Evidence-based medicine journal** (<http://www.acponline.org/journals/ebm/ebmmenu.htm>).
- **ACP journal club** (<http://www.acponline.org/>).
- **Bandolier** (<http://www.jr2.ox.ac.uk/Bandolier>).

#### Guías de práctica clínica

**NGC** (National Guideline Clearinghouse): Base de datos de guías de práctica clínica y documentos relacionados. Producida por la Agency for Health Care Policy and Research (AHCPR).

Accesible vía Internet ([http://www.guidelines.gov/body\\_home\\_nf.asp?view=home](http://www.guidelines.gov/body_home_nf.asp?view=home)).

**CPG Infobase**: Base de datos de guías de práctica clínica producidas por la Asociación Médica de Canadá (Canadian Medical Association).

Accesible vía Internet (<http://www.cma.ca/cpgs/>).

**SIGN** (Scottish Intercollegiate Guidelines Network): Guías de práctica clínica producidas por los colegios médicos de Escocia.

Accesible vía Internet (<http://pc47.cee.hw.ac.uk/sign/home.htm>).

#### Literatura gris

**SIGLE** (System for Information on Grey Literature): Recoge literatura gris producida en Europa. Es

una base de datos de contenido multidisciplinar, no exclusivamente biomédico, con 364.000 registros aproximadamente. Producida por EAGLE (European Association for Grey Literature Exploitation).

Accesible vía Internet (<http://www.konbib.nl/infolev/signle/ea/home.htm>) y en CD-ROM a través de SilverPlatter (<http://www.silverplatter.com/>).

**Conference Papers Index:** Información de los programas finales, libros de resúmenes y actas publicadas en las principales reuniones científicas de todo el mundo. Desde 1995 predominan los contenidos relativos a ciencias de la vida y del medio ambiente. Producido por Cambridge Scientific Abstracts.

Accesible vía Internet (<http://www.csa1.co.uk/>) y en CD-ROM.

**NTIS** (National Technical Information Service): Base de datos producida por el Ministerio de Comercio de EE.UU. Es una importante fuente para localizar investigaciones, estudios e informes de múltiples campos científicos, patrocinados por el gobierno federal y desarrollados dentro o fuera de EE.UU.

Accesible vía Internet (<http://www.ntis.gov/index.html>).

**TESEO:** Contiene información sobre tesis doctorales leídas en las Universidades españolas desde 1976.

Accesible vía Internet (<http://www.mec.es/teseo/>).

### 2.3. Productores o distribuidores de bases de datos

Se recogen aquí una serie de compañías que producen o distribuyen, entre otras, las bases de datos mencionadas en este anexo. Suelen ofrecer acceso a las bases de datos tanto on-line como a través de CD-ROM.

**The Dialog Corporation.** Ofrece acceso a más de 700 bases de datos. (<http://www.dialog.com>).

**STN International.** Ofrece acceso a más de 200 bases de datos (<http://www.fiz-karlsruhe.de/stn.html>).

**SilverPlatter.** Ofrece acceso a más de 250 bases de datos (<http://www.silverplatter.com/>).

**Ovid Technologies.** Ofrece acceso a más de 80 bases de datos (<http://www.ovid.com>).

## 3. FUENTES ESTADÍSTICO-ADMINISTRATIVAS

**OMS/WHO** (Organización Mundial de la Salud): Ofrece acceso a múltiples estadísticas a nivel mundial y permite buscar en sus bases de datos

de informes técnicos y otras publicaciones (<http://www.who.org>).

**OPS/PAHO** (Organización Panamericana de la Salud): Organismo internacional de salud pública dedicado a mejorar la salud y las condiciones de vida en América. Ofrece acceso a interesantes documentos, programas sanitarios e importante información técnica (<http://www.paho.org/>).

**FDA** (Food and Drug Administration): Organismo que registra y autoriza todos los productos alimenticios y medicamentos en los EE.UU. (<http://www.fda.gov/hometext.htm>).

**OCDE** (Organización para la Cooperación y el Desarrollo Económico): Ofrece información valiosa sobre estadísticas y publicaciones clasificadas por temas, entre las que destaca por su utilidad para el estudio del contexto la importante base de datos "OECD Health Data", que presenta estadísticas que permiten comparar la situación sanitaria de 29 países (<http://www.oecd.org>).

**CORDIS** (Community Research and Development Information Service): Dependiente de la Comisión Europea. Proporciona información sobre actividades de investigación y desarrollo realizadas en la Unión Europea (<http://www.cordis.lu/>).

**EUROSTAT** (Oficina de Estadística de la Unión Europea): Dedicada a proporcionar un servicio de información estadística de alta calidad a la Unión Europea. Recoge datos de los Institutos Nacionales de Estadística de los Estados miembros (<http://europa.eu.int/en/comm/eurostat/eurostat.html>).

**Health Atlases on the Internet:** Colección de atlas y censos sobre información sanitaria. Incluye una selección de mapas, tablas, gráficos e informes de la página web de la OMS ([http://members.xoom.com/mgdigest/medgpy\\_atlases.html](http://members.xoom.com/mgdigest/medgpy_atlases.html)).

### 3.1. En España

**CNE** (Centro Nacional de Epidemiología): Ofrece datos relacionados con enfermedades sometidas a vigilancia epidemiológica, estadísticas sobre el cáncer e indicadores de mortalidad y otros recursos (<http://www.isciii.es/unidad/sge/cne/ccne.html>).

**MSC** (Ministerio de Sanidad y Consumo): Ofrece acceso a bases de datos como el Conjunto Mínimo Básico de Datos (CMBD) y a la información ofrecida mediante el uso de los Grupos Relacionados con el Diagnóstico (GRD), además de otra información valiosa (<http://www.msc.es>).

**INE** (Instituto Nacional de Estadística): Ofrece la posibilidad de consultar el banco de datos de series estadísticas TEMPUS, además de poder conocer la información estadística nacional actualizada (<http://www.ine.es>).

**CIS** (Centro de Investigaciones Sociológicas): Ofrece acceso a más de 1.200 encuestas y estudios de opinión pública y situación de la sociedad española (<http://www.cis.es/>).

**Información legislativa del BOE:** Este servicio del Boletín Oficial del Estado de España ofrece acceso a la información de carácter jurídico-administrativo en el ámbito nacional, autonómico, de la Unión Europea, Latinoamérica y otros países (<http://www.boe.es/tablon/otrasdir.htm>).

**DOCUMED:** Base de datos que contiene resúmenes seleccionados de trabajos publicados desde 1994 en más de 200 revistas. Elaborada por el Instituto de Estudios Documentales e Históricos sobre la Ciencia de la Universidad de Valencia (<http://www.uv.es/~docmed/documed/documed.html>)

#### 4. BUSCADORES DE INFORMACIÓN SANITARIA EN INTERNET

Se incluyen una serie de buscadores de información sanitaria en Internet. También este listado podría haber sido de gran extensión pues el número de buscadores de información en Internet crece continuamente. Se recogen aquí sólo algunos de los más conocidos y que se consideran de buena calidad.

**Medical Matrix.** Permite acceso a recursos clínicos disponibles en Internet clasificados y comentados. Dirigido principalmente a médicos y trabajadores sanitarios de Estados Unidos (<http://www.medmatrix.org/>).

**Altavista.** Uno de los grandes buscadores de Internet. Aunque no es específico de temas relacionados con la salud, recoge un gran número de recursos sanitarios (<http://www.altavista.digital.com/>).

**Yahoo Health.** Recursos sanitarios del que, probablemente, sea el buscador de Internet más popular (<http://www.yahoo.com/health>).

**OMNI** (Organising Medical Networked Information). Recurso británico de fuentes biomédicas de alta calidad en Internet (<http://www.omni.ac.uk>).

#### 5. ESTRATEGIAS PREDISEÑADAS DE BÚSQUEDA BIBLIOGRÁFICA

Se denominan también filtros y se utilizan para optimizar los resultados de las búsquedas en bases de datos bibliográficas. Entre las diversas instituciones que ofrecen estos filtros, pueden citarse las siguientes:

**IHS** (Institute of Health Sciences). Dependiente de la Universidad de Oxford, Reino Unido (<http://www.ihs.ox.ac.uk/library/filters.html>).

**EBM.** Centro para la Medicina Basada en la Evidencia del NHS británico. Ubicado en Oxford (<http://cebm.jr2.ox.ac.uk/docs/searching.html>).

**CRD** (Centre for Reviews and Dissemination). Ubicado en la Universidad de York (Reino Unido) (<http://www.york.ac.uk/inst/crd/search.htm>).

---

## Anexo III. Esquemas para la interpretación de la evidencia científica

Existen múltiples cuestionarios o tablas para el análisis crítico de publicaciones científicas. Se han incluido en este documento los que, a nuestro juicio, son más útiles y están más difundidos. El conjunto de cuestionarios más completo y coherente publicado hasta ahora ha sido la serie que la revista JAMA ha ofrecido desde 1993<sup>17,189,190</sup>. Además, se han incluido otras tablas de análisis con el propósito de incrementar el rango de tipos de estudios que proporciona la serie citada y una lista de referencias que permitan ampliar la lista ofrecida.

Las tablas 13 a 22 fueron publicadas como una serie en la revista JAMA y se componen de tres secciones: la primera analiza la validez interna, la segunda los resultados y la tercera la validez externa. Sólo en caso de que se concluya que los resultados del estudio tienen una validez interna suficiente tiene sentido responder al resto de las preguntas. En cualquier caso, la decisión última sobre la validez del estudio la tiene el lector, que no puede decidir la inclusión o no del estudio únicamente por las preguntas enunciadas en los esquemas, sino que debe adaptar esa decisión a las características de la investigación que se está realizando. El valor de un artículo va a depender del valor otorgado al resto de los artículos identificados, o sea, que la decisión acerca de la inclusión de un artículo o no se realizará mediante una valoración de la importancia relativa de éste en relación con la información disponible.

Los esquemas presentados a continuación no pueden usarse como una lista de criterios exhaustiva para la valoración de artículos. La mayoría de estos esquemas son listados de criterios mínimos que deben cumplir todos los artículos del mismo tipo, pero que en ningún caso pueden ser considerados una guía con la que se evita la lectura de cada artículo

en su totalidad y el estudio metodológico riguroso que cada trabajo precise.

Cada artículo es diferente y no es posible redactar un esquema rígido que sirva para la evaluación de todos los artículos de un determinado tipo; por ejemplo, para los artículos sobre pruebas diagnósticas. El mayor beneficio de la utilización de cuestionarios es que se explicitan los criterios que se han utilizado para el análisis crítico de los artículos y que, por tanto, es posible que un lector reproduzca íntegramente y reinterprete la evaluación realizada valorando la importancia relativa de cada artículo en relación con la información existente.

Además de las tablas que se incluyen a continuación, existen muchas propuestas de cuestionarios para el análisis crítico de la literatura científica, entre los que destacan:

- Otros cuestionarios para ensayos clínicos<sup>83,191,192,193,194,195</sup>.
- Otros cuestionarios para estudios de evaluación económica<sup>161,164</sup>.
- Revisiones sistemáticas de ensayos clínicos controlados y aleatorizados<sup>196</sup>.
- Meta-análisis<sup>197,198,199,200,201</sup>.
- Estudios observacionales<sup>202,203,204,205,206</sup>.
- Análisis de decisión que incorporan utilidades<sup>138</sup>.
- Análisis de subgrupos<sup>90</sup>.
- Artículos que analizan la probabilidad de enfermar y el diagnóstico diferencial<sup>207</sup>.
- Auditorías clínicas o estudios que valoran la calidad de la práctica clínica<sup>208</sup>.
- Libros que ofrecen un compendio de esquemas para la lectura crítica<sup>209,210,211</sup>.

**Tabla 13**  
**Artículos sobre intervenciones terapéuticas y/o preventivas\***

<p><b>¿Son válidos los resultados del estudio?</b></p> <p>Criterios primarios:</p> <ul style="list-style-type: none"> <li>— ¿Se ha realizado de forma aleatoria la asignación de los tratamientos a los pacientes?</li> <li>— Inclusión de todos los pacientes que entraron en el estudio en el análisis: ¿Se ha realizado un seguimiento completo de los pacientes? ¿Se ha mantenido la asignación inicial de pacientes durante todo el estudio?</li> </ul> <p>Criterios secundarios:</p> <ul style="list-style-type: none"> <li>— ¿Hubo enmascaramiento respecto al tratamiento, a los pacientes, los clínicos y los investigadores?</li> <li>— ¿Eran similares los grupos al inicio del estudio?</li> <li>— Aparte de la intervención experimental, ¿se ha tratado de forma idéntica a los grupos de pacientes?</li> </ul> <p><b>¿Cuáles han sido los resultados del estudio?</b></p> <ul style="list-style-type: none"> <li>— ¿Cuál ha sido la magnitud del efecto del tratamiento?</li> <li>— ¿Con qué precisión se ha estimado el efecto del mismo?</li> </ul> <p><b>¿Son útiles los resultados en la práctica clínica?</b></p> <ul style="list-style-type: none"> <li>— ¿Son aplicables los resultados a nuestros pacientes?</li> <li>— ¿Se han considerado todos los posibles resultados clínicamente importantes?</li> <li>— ¿Compensan los beneficios del tratamiento, los costes y riesgos del mismo?</li> </ul>
---

\* Tomado de Guyatt<sup>212</sup> y Guyatt<sup>213</sup>.

**Tabla 14**  
**Artículos sobre pruebas diagnósticas\***

<p><b>¿Son válidos los resultados del estudio?</b></p> <p>Criterios primarios:</p> <ul style="list-style-type: none"> <li>— ¿Se ha comparado de forma independiente, ciega, con un patrón de referencia?</li> <li>— ¿La muestra de pacientes estudiados ha recogido todo el espectro de pacientes a los que se aplicará el test en la práctica clínica?</li> </ul> <p>Criterios secundarios:</p> <ul style="list-style-type: none"> <li>— ¿Hay sesgo de verificación? Existe cuando el test estándar se realiza más en los casos que resultaron dudosos o positivos en el test a evaluar.</li> <li>— ¿Se ha descrito la técnica con el detalle suficiente como para permitir su reproducción?</li> </ul> <p><b>¿Cuáles han sido los resultados del estudio?</b></p> <ul style="list-style-type: none"> <li>— ¿Se han presentado los cocientes de probabilidad, <i>likelihood ratios</i>, o los datos sobre sensibilidad y especificidad del test o los datos para su cálculo?</li> </ul> <p><b>¿Son útiles los resultados en la práctica clínica?</b></p> <ul style="list-style-type: none"> <li>— ¿Es reproducible el test, tanto en la realización como en la interpretación, para unos determinados pacientes?</li> <li>— ¿Son aplicables los resultados a nuestros pacientes?</li> <li>— ¿Modificará el test el manejo de los pacientes?</li> <li>— ¿Mejorarán los pacientes por la utilización del test?</li> </ul>
--

\* Tomado de Jaeschke<sup>214</sup> y Jaeschke<sup>215</sup>.

**Tabla 15**  
**Artículos sobre efectos de intervenciones y/o exposiciones\***

<p><b>¿Son válidos los resultados del estudio?</b></p> <p>Criterios primarios:</p> <ul style="list-style-type: none"> <li>— ¿Se han utilizado grupos de comparación similares en cuanto a otros factores determinantes del efecto diferentes a la exposición?</li> <li>— ¿Se han valorado de la misma forma en todos los grupos de comparación las exposiciones y los efectos?</li> <li>— ¿El seguimiento ha sido suficientemente largo y completo?</li> </ul> <p>Criterios secundarios:</p> <ul style="list-style-type: none"> <li>— ¿Precede la exposición al efecto?</li> <li>— ¿Existe gradiente dosis-respuesta?</li> <li>— Aparte de la intervención experimental, ¿se ha tratado de forma idéntica a los grupos de pacientes?</li> </ul> <p><b>¿Cuáles han sido los resultados del estudio?</b></p> <ul style="list-style-type: none"> <li>— ¿Cuál es la fuerza de la asociación exposición-resultado?</li> <li>— ¿Qué precisión tiene la estimación del riesgo?</li> </ul> <p><b>¿Son útiles los resultados en la práctica clínica?</b></p> <ul style="list-style-type: none"> <li>— ¿Son aplicables los resultados a nuestros pacientes?</li> <li>— ¿Cuál es la magnitud del riesgo?</li> <li>— ¿Se debe evitar la exposición?</li> </ul>
--

\* Tomado de Levine<sup>203</sup>.

**Tabla 16**  
**Artículos sobre pronósticos\***

<p><b>¿Son válidos los resultados del estudio?</b></p> <p>Criterios primarios:</p> <ul style="list-style-type: none"> <li>— ¿La muestra de pacientes está bien definida y es representativa de pacientes en un momento similar del curso de la enfermedad?</li> <li>— ¿El seguimiento ha sido suficientemente largo y completo?</li> </ul> <p>Criterios secundarios:</p> <ul style="list-style-type: none"> <li>— ¿Los parámetros utilizados para medir los resultados han sido objetivos y no sesgados?</li> <li>— ¿Se ha ajustado por todos los factores pronósticos importantes?</li> </ul> <p><b>¿Cuáles han sido los resultados del estudio?</b></p> <ul style="list-style-type: none"> <li>— ¿Se ha especificado la probabilidad de que ocurran los eventos en períodos de tiempo determinados?</li> <li>— ¿Cuál es la precisión de las estimaciones de probabilidad?</li> </ul> <p><b>¿Son útiles los resultados en la práctica clínica?</b></p> <ul style="list-style-type: none"> <li>— ¿Los pacientes del estudio son similares a los nuestros?</li> <li>— ¿Son útiles los resultados para decidir si dar un tratamiento?</li> <li>— ¿Son útiles los resultados para aconsejar o tranquilizar a los pacientes o a sus familiares?</li> </ul>
--

\* Tomado de Laupacis<sup>204</sup>.

**Tabla 17**  
**Artículos de revisión\***

<p><b>¿Son válidos los resultados de la revisión?</b></p> <p>Criterios primarios:</p> <ul style="list-style-type: none"> <li>— ¿Es suficientemente concreto y ajustado a nuestros objetivos el tema de la revisión?</li> <li>— ¿Los criterios utilizados para la inclusión de estudios en la revisión fueron apropiados?</li> </ul> <p>Criterios secundarios:</p> <ul style="list-style-type: none"> <li>— ¿Es baja la probabilidad de que falte algún estudio relevante?</li> <li>— ¿Se ha evaluado la validez de los estudios incluidos?</li> <li>— ¿Es reproducible la evaluación realizada para la inclusión y exclusión de estudios en la revisión?</li> <li>— ¿Existe suficiente homogeneidad entre los estudios como para sintetizarlos en una revisión?</li> </ul> <p><b>¿Cuáles han sido los resultados del estudio?</b></p> <ul style="list-style-type: none"> <li>— ¿Cuáles son los resultados de la revisión en su conjunto?</li> <li>— ¿Hasta qué punto fueron precisos los resultados?</li> </ul> <p><b>¿Son útiles los resultados en la práctica clínica?</b></p> <ul style="list-style-type: none"> <li>— ¿Son aplicables los resultados a nuestros pacientes?</li> <li>— ¿Se han considerado todos los resultados clínicamente importantes?</li> <li>— ¿Compensan los beneficios previsible, los costes y los riesgos?</li> </ul>
--

\* Tomado de Oxman<sup>216</sup>.

**Tabla 18**  
**Análisis de decisión clínica\***

<p><b>¿Son válidos los resultados del estudio?</b></p> <ul style="list-style-type: none"> <li>— ¿Se han incluido todos los resultados y estrategias importantes?</li> <li>— ¿Se ha utilizado un método adecuado para identificar, seleccionar y combinar las evidencias y establecer probabilidades?</li> <li>— ¿Se han obtenido las utilidades de forma adecuada y a partir de fuentes fiables?</li> <li>— ¿Se ha estimado el impacto de las posibles incertidumbres existentes?</li> <li>— ¿Se ha ajustado por todos los factores pronósticos importantes?</li> </ul> <p><b>¿Cuáles han sido los resultados del estudio?</b></p> <ul style="list-style-type: none"> <li>— ¿Hay una estrategia claramente superior a las otras?</li> <li>— ¿Cuál es el nivel de calidad de las evidencias de los estudios incluidos en análisis?</li> <li>— ¿Podría modificarse el resultado a causa de las incertidumbres existentes?</li> </ul> <p><b>¿Son útiles los resultados en la práctica clínica?</b></p> <ul style="list-style-type: none"> <li>— ¿Se ajustan las estimaciones de probabilidad a las características clínicas de nuestros pacientes?</li> <li>— ¿Valorarían nuestros pacientes las utilidades de forma similar?</li> </ul>
---

\* Tomado de Richardson<sup>217</sup> y Richardson<sup>218</sup>.

**Tabla 19**  
**Guías de práctica clínica\***

<p><b>¿Son válidos los resultados de la revisión?</b></p> <p>Criterios primarios:</p> <ul style="list-style-type: none"> <li>— ¿Se han considerado todas las opciones y resultados clínicos importantes?</li> <li>— ¿Existe un procedimiento adecuado que se haya utilizado para identificar, seleccionar y combinar las evidencias?</li> </ul> <p>Criterios secundarios:</p> <ul style="list-style-type: none"> <li>— ¿Se ha utilizado un procedimiento adecuado para considerar el valor relativo de los distintos pronósticos?</li> <li>— ¿La información que ofrece está actualizada?</li> <li>— ¿Ha sido valorada por expertos, <i>peer review</i>?</li> <li>— ¿Se ha comprobado que la aplicación de la guía en la práctica produzca beneficios?</li> </ul> <p><b>¿Cuáles son sus recomendaciones?</b></p> <ul style="list-style-type: none"> <li>— ¿Las recomendaciones son aplicables y clínicamente relevantes?</li> <li>— ¿Cuál es la fuerza de las recomendaciones?</li> <li>— ¿Cuál es el impacto de las incertidumbres de las evidencias en las que se basa la guía?</li> </ul> <p><b>¿Son útiles los resultados en la práctica clínica?</b></p> <ul style="list-style-type: none"> <li>— ¿Las recomendaciones son aplicables a nuestros pacientes?</li> <li>— ¿El objetivo de la guía es consistente con sus objetivos?</li> </ul>
--

\* Tomado de Hayward<sup>219</sup> y Wilson<sup>220</sup>.

**Tabla 20**  
**Artículos sobre calidad de vida relacionada con la salud\***

<p><b>¿Son válidos los resultados del estudio?</b></p> <p>Criterios primarios:</p> <ul style="list-style-type: none"> <li>— ¿Las variables incluidas miden aspectos considerados importantes por los pacientes?</li> <li>— ¿Los instrumentos utilizados miden la calidad de vida de forma válida, fiable y precisa?</li> </ul> <p>Criterios secundarios:</p> <ul style="list-style-type: none"> <li>— ¿Se incluyen todos los aspectos importantes relacionados con la calidad de vida?</li> <li>— En el caso de que existan estudios que hayan realizado una evaluación económica o hayan medido la relación calidad-cantidad de vida, ¿se han empleado en el artículo las medidas adecuadas?</li> </ul> <p><b>¿Cuáles han sido los resultados?</b></p> <ul style="list-style-type: none"> <li>— ¿Cuál fue la magnitud del efecto sobre la calidad de vida?</li> </ul> <p><b>¿Son útiles los resultados en la práctica clínica?</b></p> <ul style="list-style-type: none"> <li>— ¿Son útiles los resultados en el proceso de información a nuestros pacientes?</li> <li>— ¿Consigue el estudio una buena simulación de la práctica clínica real?</li> <li>— ¿Se ajustan las estimaciones de probabilidad a las características clínicas de nuestros pacientes?</li> </ul>
---

\* Tomado de Guyatt<sup>221</sup>.

**Tabla 21**  
**Ensayos clínicos aleatorios\***

<p>1. Asignación de tratamientos:</p> <ul style="list-style-type: none"> <li>— ¿La asignación fue aleatoria?</li> <li>— ¿Hubo enmascaramiento en la asignación?</li> </ul> <p>2. ¿Los diferentes grupos fueron similares en cuanto a los más importantes factores pronósticos?</p> <p>3. ¿Se especificaron los criterios de inclusión y exclusión establecidos?</p> <p>4. ¿Hubo enmascaramiento para los investigadores?</p> <p>5. ¿Hubo enmascaramiento para los médicos que atendieron a los pacientes participantes?</p> <p>6. ¿Hubo enmascaramiento para los pacientes?</p> <p>7. ¿Se presentaron las medidas de resultado junto con sus medidas de dispersión?</p> <p>8. ¿Se realizó un análisis por intención de tratar?</p>
--

\* Tomado de Verhagen<sup>222</sup>.

**Tabla 22**  
**Estudios de evaluación económica de la atención médica\***

<p><b>¿Son válidos los resultados del estudio?</b></p> <ul style="list-style-type: none"> <li>— ¿Se han considerado todas las estrategias clínicas relevantes?</li> <li>— ¿Se han considerado los costes y resultados desde los puntos de vista más importantes?</li> <li>— ¿Se han medido correctamente los costes y resultados?</li> <li>— ¿Se ha analizado el impacto de las incertidumbres que existen en los datos?</li> <li>— ¿Se ha considerado la situación clínica basal de la población en los datos incluidos en el análisis?</li> </ul> <p><b>¿Cuáles han sido los resultados?</b></p> <ul style="list-style-type: none"> <li>— ¿Cuáles fueron los costes y resultados incrementales de cada una de las estrategias consideradas?</li> <li>— ¿Se han observado diferencias en los costes y resultados incrementales entre subgrupos?</li> <li>— ¿Cuál es el impacto de las incertidumbres de los datos incluidos?</li> </ul> <p><b>¿Son útiles los resultados en la práctica clínica?</b></p> <ul style="list-style-type: none"> <li>— ¿Los beneficios compensan los costes y efectos adversos?</li> <li>— ¿La valoración de los resultados clínicos y de los costes es aplicable a nuestros pacientes?</li> </ul>
--

\* Tomado de Drummond<sup>223</sup> y O'Brien<sup>224</sup>.

## Anexo IV.

# Clasificaciones del grado de evidencia científica

La ETS supone la formulación de recomendaciones basadas en la mejor evidencia científica disponible y en el análisis del contexto. Las escalas de clasificación de la evidencia científica contribuyen a determinar la calidad y el rigor científico de la información disponible, a la vez que permiten establecer la fuerza de cada una de las recomendaciones del informe, tanto en el sentido de aconsejar la adopción de la tecnología como de que ésta se deje de utilizar si las evidencias indican la carencia o insuficiencia de datos positivos sobre su efectividad.

### 1. NIVELES DE CALIDAD DE LA EVIDENCIA CIENTÍFICA

Las siguientes clasificaciones diferencian de forma jerárquica los distintos niveles de evidencia en función del rigor científico del diseño del estudio, aunque una clasificación adecuada de la importancia de los estudios debería incluir también el análisis crítico de su calidad (desarrollo, pérdidas de sujetos, calidad de las fuentes de datos y otros aspectos). A continuación se presentan las clasificaciones más aceptadas y difundidas.

**Tabla 23**  
**Clasificación de la evidencia científica según el diseño del estudio<sup>a</sup>**  
**(tomado de US Preventive Task Force<sup>225</sup>)**

- |       |   |
|-------|---|
| I.    | Evidencia obtenida a partir de al menos un ensayo aleatorizado y controlado diseñado de forma apropiada.  |
| II.1. | Evidencia obtenida de ensayos controlados bien diseñados, sin randomización.  |
| II.2. | Evidencia obtenida a partir de estudios de cohorte o caso-control bien diseñados, realizados preferentemente en más de un centro o por un grupo de investigación. |
| II.3. | Evidencia obtenida a partir de múltiples series comparadas en el tiempo con o sin intervención <sup>b</sup> .   |
| III.  | Opiniones basadas en experiencias clínicas, estudios descriptivos o informes de comités de expertos.  |

a: De mayor (I) a menor (III) calidad.

b: Este tipo de evidencia también incluye resultados muy llamativos en experimentos sin grupo control, como los derivados de la introducción de la penicilina en los años cuarenta.

**Tabla 24**  
**Calidad de la evidencia según el diseño del estudio (modificado de Bertram<sup>226</sup> y Goodman<sup>56</sup>)**

- |  |
|--|
| <ul style="list-style-type: none"><li>• Ensayos aleatorios controlados de gran tamaño, revisiones sistemáticas o meta-análisis de ensayos aleatorios controlados.</li><li>• Ensayos aleatorios controlados de pequeño tamaño.</li><li>• Ensayos no aleatorios con controles coincidentes en el tiempo.</li><li>• Ensayos no aleatorios con controles históricos.</li><li>• Estudios observacionales de cohortes.</li><li>• Estudios observacionales casos-controles.</li><li>• Vigilancia epidemiológica, Estudios descriptivos, Información basada en registros.</li><li>• Estudio de series de casos multicéntrico.</li><li>• Estudio de series de casos unicéntrico.</li><li>• Estudio de un caso o anécdota.</li></ul> |
|--|

Tabla 25  
Clasificación de la evidencia científica (modificado de Jovell<sup>227</sup>)

Nivel	Tipo de diseño	Condiciones de rigurosidad científica
I	Meta-análisis de ensayos controlados y aleatorizados	No heterogeneidad, calidad de los estudios
II	Ensayo controlado y aleatorizado de muestra grande	Evaluación del poder estadístico, multicéntrico, calidad del estudio
III	Ensayo controlado y aleatorizado de muestra pequeña	Evaluación del poder estadístico, calidad del estudio
IV	Ensayo prospectivo controlado no aleatorizado	Controles coincidentes en el tiempo, multicéntrico, calidad del estudio
V	Ensayo prospectivo controlado no aleatorizado	Controles históricos, calidad del estudio
VI	Estudios de cohortes	Multicéntrico, apareamiento, calidad del estudio
VII	Estudios de casos y controles	Multicéntrico, calidad del estudio
VIII	Series clínicas sin grupo control Estudios descriptivos Juicio de expertos	Multicéntrico
IX	Anécdotas o casos únicos	

## 2. RELACIÓN ENTRE LA CALIDAD DE LA EVIDENCIA CIENTÍFICA Y LA FUERZA DE LAS RECOMENDACIONES

Las recomendaciones acerca de la adopción o no de una nueva tecnología sanitaria formuladas por las agencias de ETS deben surgir del

análisis y síntesis de la evidencia disponible en la literatura médica, para lo cual puede ser útil presentar las evidencias clasificadas según el rigor científico. En ese sentido, la tabla 26 establece la correspondencia entre el nivel de calidad científica y la fuerza de la recomendación que surge de la evidencia. Se ha seleccionado esa propuesta, aunque se han publicado otras interesantes<sup>228,229</sup>.

Tabla 26  
Relación entre los niveles de calidad de la evidencia científica y el grado de recomendaciones (modificado de Sackett<sup>230</sup>)

Nivel de calidad	Grado de recomendación <sup>a</sup>
I. Ensayos aleatorizados con una muestra grande y resultados bien definidos (y un riesgo bajo de error estadístico tipo $\alpha$ y $\beta$ )	A Adecuada evidencia científica para recomendar la adopción de la tecnología E Adecuada evidencia para recomendar la no adopción
II. Ensayos aleatorizados con una muestra pequeña (y un riesgo moderado o alto de error estadístico tipo $\alpha$ y $\beta$ )	B Cierta evidencia científica para recomendar la adopción de la tecnología D Cierta evidencia para recomendar la NO adopción
III. Estudios no aleatorizados, cohortes concurrentes en el tiempo	C Insuficiente evidencia científica, por lo que la decisión de adoptar la tecnología debe basarse en otros criterios
IV. Estudios no aleatorizados, cohortes históricas	C
V. Estudios no controlados, series clínicas	C

a: El grado de recomendaciones se define en US Preventive Task Force<sup>225</sup> y en Canadian Task Force<sup>231</sup>.

### 3. CLASIFICACIÓN DE LOS ESTUDIOS SOBRE EFICACIA DIAGNÓSTICA

Tabla 27  
Clasificación de los niveles de estudios sobre eficacia diagnóstica con las medidas de resultado más habituales (tomada de Fryback<sup>232</sup>)

Nivel	Medidas de resultado
<b>I. Eficacia Técnica</b>	<ul style="list-style-type: none"> <li>— Resolución de pares de líneas</li> <li>— Cambio en la función de transferencia de modulación</li> <li>— Rango en la escala de grises</li> <li>— Cantidad de abigarramiento</li> <li>— Nitidez</li> </ul>
<b>II. Eficacia en la Exactitud Diagnóstica</b>	<ul style="list-style-type: none"> <li>— Cantidad de diagnósticos normales o anormales en una serie de casos</li> <li>— Exactitud diagnóstica (porcentaje de diagnósticos correctos en una serie de casos)</li> <li>— Sensibilidad y Especificidad en una situación clínica definida</li> <li>— Medidas del área de la curva ROC</li> </ul>
<b>III. Eficacia Presumible Diagnóstica</b>	<ul style="list-style-type: none"> <li>— Porcentaje de casos en los que la imagen se juzgó útil para realizar el diagnóstico.</li> <li>— Cambios en la tendencia de la distribución de probabilidades del diagnóstico diferencial</li> <li>— Diferencias subjetivas clínicas en las probabilidades diagnósticas antes y después de la prueba</li> <li>— Cociente de probabilidades estimado de forma empírica y subjetiva para una prueba positiva y negativa</li> </ul>
<b>IV. Eficacia Terapéutica</b>	<ul style="list-style-type: none"> <li>— Porcentaje de veces en que la imagen se consideró útil para planificar el manejo de los enfermos en series de casos.</li> <li>— Porcentaje de veces en que procedimientos médicos (terapéuticos o diagnósticos) fueron evitados por la información suministrada por la imagen</li> <li>— Porcentaje de veces en que la terapia pre-prueba cambió tras la información de la imagen (inferido retrospectivamente de los informes clínicos)</li> <li>— Porcentaje de veces que cambiaron los clínicos su terapia después de obtener la información de la prueba.</li> </ul>
<b>V. Eficacia en Resultados para los Enfermos</b>	<ul style="list-style-type: none"> <li>— Porcentaje de enfermos que mejoraron con la prueba diagnóstica en comparación con los que no la usaron</li> <li>— Procedimientos médicos o morbilidad evitados tras la información suministrada por la imagen</li> <li>— Cambios en AVACs (Años de Vida Ajustados por Calidad)</li> <li>— Coste por AVACs ahorrado por la información suministrada por la prueba diagnóstica</li> </ul>
<b>VI. Eficacia Social</b>	<ul style="list-style-type: none"> <li>— Análisis coste/beneficio, coste/efectividad o coste/utilidad realizado desde una perspectiva social.</li> </ul>

**Tabla 28**  
**Criterios de clasificación de estudios que evalúan pruebas diagnósticas por imagen**  
**(tomado de Flynn<sup>188</sup>)**

Grado	Criterio fundamental	Requisitos
I	Estudios con resultados generalizables a una amplia variedad de enfermos y que no presentan defectos metodológicos significativos	<ul style="list-style-type: none"> <li>— <math>\geq 35</math> enfermos y <math>\geq 35</math> sujetos sin la enfermedad analizada (ya que tales números dan un intervalo de confianza del 95% cuyo límite inferior excluye 0,90 si el error estándar es de 1).</li> <li>— Pacientes escogidos de una muestra clínica representativa (no filtrados para incluir sólo casos graves) cuyos síntomas sean completamente descritos.</li> <li>— Los diagnósticos han de ser definidos por un estándar de referencia apropiado.</li> <li>— Los estudios han de ser de alta calidad técnica y evaluados independientemente del diagnóstico de referencia.</li> </ul>
II	Estudios con un espectro más estrecho de generalización, y con pocos defectos metodológicos que están bien descritos (su impacto en los resultados puede ser evaluado)	<ul style="list-style-type: none"> <li>— <math>\geq 35</math> enfermos con y sin patología.</li> <li>— Espectro más limitado de pacientes, presencia de sesgo de selección (por ejemplo: pacientes con patología más grave por estar ingresados en Centros Universitarios).</li> <li>— Libre de otros defectos metodológicos que promuevan interacción entre los resultados de la prueba y el diagnóstico de la enfermedad.</li> <li>— Estudios prospectivos.</li> </ul>
III	Estudios con varios defectos metodológicos	<ul style="list-style-type: none"> <li>— Muestras pequeñas.</li> <li>— Información incompleta.</li> <li>— Estudios de exactitud diagnóstica retrospectivos</li> </ul>
IV	Estudios con múltiples defectos metodológicos	<ul style="list-style-type: none"> <li>— No existe un estándar de referencia fiable para el diagnóstico.</li> <li>— Dependencia entre los resultados de la prueba estudiada y la determinación del diagnóstico final.</li> <li>— No se pudo determinar la fuente de la cohorte de pacientes o estaba obviamente influida por los resultados de la prueba (sesgo de verificación).</li> <li>— Opiniones no basadas en los datos.</li> </ul>

# Referencias

1. Conde JL. Evaluación de Tecnologías Sanitarias: Racionalización frente a racionamiento. *Revista de Administración Sanitaria*. 1998; vol: 2, n.º 5: 139-147.
2. OCDE Health Data. *A software package for the international comparison of health care systems*. [archivo de ordenador] Version 1.5. París: OCDE, 1993.
3. Glover JA. The incidence of Tonsillectomy in School Children. *Proc Roy Soc Med* 1938; 31: 1219-36.
4. Wennberg JE, Gittelsohn A. Small area variations in health care delivery. *Science* 1973; 182: 1102-1108.
5. Chassin MR, Kosecoff J, Park ER, et al. Variations in the use of medical and surgical services by the Medicare population. *New Engl J Med*. 1986; 314: 285-90.
6. Wennberg JE. Dealing with medical practice variations: a proposal for action. *Health Affairs*. 1984; 3: 6-32.
7. McGlynn EA, Naylor CD, Anderson GM et al. Comparison of the appropriateness of coronary angiography and coronary artery bypass graft surgery between Canada and New York State. *JAMA* 1994; 272: 934-940.
8. Bernstein SJ, Kosecoff J, Gray D, Hampton JR, Brook RH. The appropriateness of the use of cardiovascular procedures. British versus US perspectives. *Int J Technol Assess Health Care* 1993; 9: 3-10.
9. US Congress, Office of Technology Assessment. *Identifying health technologies that work. Searching for Evidence*. OTA-H-608. Washington DC: US Government Printing Office, 1994.
10. Brook RH, Chassin MR, Fink A, and cols. A method for the detailed assessment of the appropriateness of medical technologies. *Int J of Technology Assessment in Health Care*. 1986; 2: 52-63.
11. Relman AS. Assessment and accountability: the third revolution in medical care [editorial]. *N Engl J Med* 1988; 319: 1.220-2.
12. US Congress, Office of Technology Assessment, *Assessing the efficacy and safety of medical technologies*. Publication n.º OTA-75. Washington DC: US Government Printing Office, 1978.
13. Granados A. La evaluación de las Tecnologías Médicas. *Med Clin (Barc)* 1995; 104: 581-588.
14. Ministerio de Sanidad y Consumo - Instituto de Salud Carlos III - Agencia de Evaluación de Tecnologías Sanitarias (AETS). *Evaluación Epidemiológica de Tecnologías de Salud*. Madrid: AETS - Instituto de Salud Carlos III, 1995.
15. Liberati A, Sheldon TA, Banta HD. EUR-ASSESS Project Subgroup Report on Methodology. Methodological Guidance for the Conduct of Health Technology Assessment. *Intl J of Technology Assessment in Health Care*. 1997; 13, 2: 186-219.
16. Banta HD, Luce B. *Health Care Technology and its Assessment. An international perspective*. New York: Oxford University Press, 1993.
17. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA* 1992; 268: 2.420-5. Edición española 1997.
18. Gol JM. Bienvenidos a la medicina basada en la evidencia. *JAMA* (edición española) 1997.
19. Cook DJ, Mulrow CD, Haynes RB. Systematic Reviews: Synthesis of Best Evidence for Clinical Decisions. *Ann Intern Med* 1997; 126: 376-380.
20. Thompson SG, Pocock SJ. Can Meta-Analyses Be Trusted?. *Lancet* 1991; 338 (8775): 1127-1130.
21. Alberta Heritage Foundation for Medical Research. *Health Technology Assessment in Alberta* (folleto divulgativo). Edmonton, Alberta, Canada, 1998.
22. Hailey D, Harstall C, Corabian P, Swist F. The Impact of rapid technology assessments. *Proceedings of the 14<sup>th</sup> Annual Meeting of the International Society of Technology Assessment in Health Care*; 1998, June 7-10; Ottawa, Ontario, Canada.
23. Field MJ, Lohr KN, editores. *Guidelines for clinical practice. From Development to Use*. Institute of Medicine. Washington, DC: National Academy Press, 1992.
24. US Department of Health and Human Services, Agency for Health Care Policy and Research. *Using Clinical Practice Guidelines to Evaluate Quality of Care (volume I)*. Publication n.º 95-0045. Rockville, MD: Executive Office Center, 1995.

25. US Department of Health and Human Services, Agency for Health Care Policy and Research. *Using Clinical Practice Guidelines to Evaluate Quality of Care (volume II)*. Publication n.º 95-0046. Rockville, MD: Executive Office Center, 1995.
26. Audet A, Greenfield S, Field M. Medical practice guidelines current activities and future directions. *Ann Intern Med* 1990; 113: 709-714.
27. Woolf SH, Battista RN, Anderson GM, Logan AG, Wang E, and the Canadian Task Force on the Periodic Health Examination. A report by the Canadian Task Force on the Periodic Health Examination. *J Clin Epidemiol* 1990; 43: 891-905.
28. Lázaro P. Evaluación de Tecnología Médica. *Papeles de Gestión Sanitaria*. Monografía n.º II-94. Valencia: M/C/Q Ediciones, Valencia, 1994.
29. AETS (Agencia de Evaluación de Tecnologías Sanitarias). *Fichas técnicas de Evaluación de Tecnologías Sanitarias*. Consejo Interterritorial del Sistema Nacional de Salud. Madrid, octubre de 1998.
30. Stevens RG. Identifying new health care technologies in the United Kingdom. *International Journal of Technology Assessment in Health Care* 1997; 13 (1): 59-67.
31. Manual de la Colaboración Cochrane (versión española de "Mulrow CD, Oxman AD (eds). *The Cochrane Collaboration Handbook*") [actualización de Septiembre 1997]. Sabadell: Centro Cochrane Español; 1998.
32. Wennberg JE, Bunker JP, Barnes B. The need for assessing the outcome of common medical practices. *Annual Review of Public Health* 1980; 1: 277-295.
33. Wennberg JE. What is outcomes research? En: Gelijns AC, ed. *Modern methods of clinical investigation*. Washington DC: National Academy Press, 1990, pp. 33-46.
34. Greenfield S. The state of outcomes research: are we on target? *N Engl J Med* 1989; 320: 1142-1143.
35. Echt DS, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. *N Engl J Med* 1991; 324: 781-788.
36. Miller FW, Leitman SF, Cronin ME et al. Controlled trial of plasma exchange and leukapheresis and dermatomyositis. *N Engl J Med* 1992; 326: 1380-1384.
37. Ruberman W, Weinblatt E, Goldberg JD, Frank CW, Shapiro S. Ventricular premature beats and mortality after myocardial infarction. *N Engl J Med* 1977; 297: 750-7.
38. The CAPS (Cardiac Arrhythmia Pilot Study) Investigators. Effects of encainide, flecainide, imipramine and moricizine on ventricular arrhythmia during the year after acute myocardial infarction: the CAPS. *Am J Cardiol* 1988; 61: 501-9.
39. Gotzsche PC, Liberati A, Torri V, Rosetti L. Beware surrogate endpoints. *International Journal of Technology Assessment in Health Care* 1996; 12: 238-246.
40. Guallar E, Conde JL, de la Cal A, Martín-Moreno JM. Guía para la evaluación de proyectos de investigación en ciencias de la salud. *Med Clin (Barc)* 1997; 108: 460-471.
41. Friedman GD. *Principios de epidemiología*. México: Panamericana, 1975.
42. Hulley SB, Cummings SR, editores. *Diseño de la investigación clínica. Un enfoque epidemiológico*. Barcelona: Doyma, 1993.
43. Bakke OM, Carné X, García-Alonso F. *Ensayos clínicos con medicamentos: Fundamentos básicos, metodología y práctica*. Barcelona: Doyma, 1994.
44. Roca J. Cómo y para qué hacer un protocolo. *Med Clin (Barc)* 1996; 106: 257-262.
45. CRD (The National Health Service Centre for Reviews and Dissemination). *Undertaking Systematic Reviews of Research on Effectiveness. CRD Guidelines for Those Carrying Out or Commissioning Reviews*. York: University of York, 1996.
46. Faber RG. Information overload [carta]. *BMJ* 1993; 307: 383.
47. Mulrow CD. Rationale for systematic reviews. *BMJ* 1994; 309: 597-599.
48. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *British Medical Journal* 1994; 309: 1286-91.
49. Smith BJ, Darzins PJ, Quinn M, Heller RF. Modern methods of searching the medical literature. *Med J Aust* 1992; 157(9): 603-11.
50. Pérez JR. *Introducción a la Información y Documentación Científica*. Madrid: Alhambra, 1987.
51. Gotzsche, P.C. Reference bias in reports of drug trials. *BMJ* 1987; 295: 654-6.
52. US Congress, Office of Technology Assessment. *Tools for Evaluating Health Technologies: Five Background Papers*. Publication n.º BP-H-142. Washington DC: US Government Printing Office, 1995.

53. Torres I. La llamada literatura gris. Reflexiones sobre su naturaleza y desarrollo de instrumentos bibliográficos para identificación y localización. *Boletín de la Asociación Andaluza de Bibliotecarios* 1994, 34: 41-59.
54. Torres I. La llamada literatura gris. Reflexiones sobre su naturaleza y desarrollo de instrumentos bibliográficos para identificación y localización. *Boletín de la Asociación Andaluza de Bibliotecarios* 1994, 35: 9-26.
55. Martín W. Medicina Basada en la Evidencia: el papel de las bibliotecas médicas. *Anuari SOCADI de documentació i informació* 1998, pp. 241-4.
56. Goodman C. *Literature Searching and evidence interpretation for assessing health care practices*. Estocolmo: The Swedish Council on Technology Assessment in Health Care. 1993.
57. McKibbin KA, Walker-Dilks CJ. Beyond ACP Journal Club: how to harness MEDLINE to solve clinical problems [editorial]. *ACP Journal Club* 1994 mar-apr: A10-A12 (*Ann Intern Med* 120 suppl 2).
58. KA, Walker CJ. Beyond ACP Journal Club: how to harness MEDLINE for therapy problems [editorial]. *ACP Journal Club* 1994 jul-aug: A10 (*Ann Intern Med* 121 suppl 1).
59. McKibbin KA, Walker CJ. Beyond ACP Journal Club: how to harness MEDLINE for diagnosis problems [editorial]. *ACP Journal Club* 1994 sep-oct: A10 (*Ann Intern Med* 121 suppl 2).
60. McKibbin KA, Walker-Dilks CJ, Wilczynski NL et al. Beyond ACP Journal Club: how to harness MEDLINE for review articles [editorial]. *ACP Journal Club* 1996 may-jun: A12-13.
61. Walker-Dilks CJ, McKibbin KA and Haines RB. Beyond ACP Journal Club: how to harness MEDLINE for etiology problems [editorial]. *ACP Journal Club* 1994 nov-dec: A10-A11 (*Ann Intern Med* 121 suppl 3).
62. Walker-Dilks CJ, McKibbin KA and Haines RB. Beyond ACP Journal Club: how to harness MEDLINE for prognosis problems [editorial]. *ACP Journal Club* 1995 jul-aug: A12-A14 (*Ann Intern Med* 123 suppl 1).
63. Hewitt P, Chalmers TC. Using MEDLINE to Peruse the Literature. *Controlled Clinical Trials* 1985; 6: 75-83.
64. Greenhalgh T. How to read a paper. The Medline database. *BMJ* 1997; 315: 180-3.
65. Reference Manager [archivo de ordenador]. Windows version 8.5. Carlsbad: Research Information Systems Inc, 1997.
66. Pro-Cite [archivo de ordenador]. DOS version 2.1. Ann Arbor: Personal Bibliographic Software, 1993.
67. End-Note version 3.0. [archivo de ordenador]. Berkeley, California: Niles Software, (<http://www.software-cientifico.com/endnote.htm>) 1999.
68. Hersh W, Hickam DH, Haynes RB, McKibbin KA. Evaluation of SAPHIRE: an automated approach to indexing and retrieving medical literature. *Proc Annu Symp Comput Appl Med Care* 1991; 808-812.
69. Last JM. *A Dictionary of Epidemiology*. New York: Oxford University Press, 1995.
70. Laird NM, Mosteller F. Some statistical methods for combining experimental results. *Int J Tech Assess in Health Care* 1990; 6: 5-30.
71. Stone DH. Design a questionnaire. *BMJ* 1993; 307: 1264-6.
72. Hernández I, Porta M, Miralles M, García F, Bolúmar F. La cuantificación de la variabilidad en las observaciones clínicas. *Med Clin (Barc)* 1990; 95: 424-429.
73. Begg CB, Berlin JA. Publication bias and dissemination of clinical research. *J Natl Cancer Inst* 1989; 81: 107-15.
74. Berlin JA, Begg CB, Louis TA. An assessment of publication bias using a sample of published clinical trials. *J Am Stat Assoc* 1989; 84: 381-92.
75. Hedges LV. Modelling publication selection effects in meta-analysis. *Stat Sci* 1992; 7: 246-55.
76. Iyengar S, Greenhouse JB. Selection models and the file drawer problem. *Stat Sci* 1988; 3: 109-35.
77. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991; 337: 867-72.
78. Chalmers I. A register of controlled trials in perinatal medicine. *WHO Chron* 1986; 40: 61-65.
79. Dickersin K. Report from the panel on the Case for Registers of Clinical Trials at the Eighth Annual Meeting of the Society for Clinical Trials. *Controlled Clin Trials* 1988; 9: 76-81.
80. Dickersin K. Keeping posted. Why register clinical trials? - revisited. *Controlled Clin Trials* 1992; 13: 170-7.
81. Follmann D, Elliott P, Suh I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol* 1992; 45: 769-73.
82. Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiology Reviews* 1987; 9: 1-30.
83. The Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. *JAMA* 1994; 272: 1926-31.

84. Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. Call for comments on a proposal to improve reporting of clinical trials in the biomedical literature. *Ann Intern Med* 1994; 121: 894-5.
85. ECRI. *ECRI's Methods of Healthcare Technology Assessment*. ECRI Executive Briefings. Plymouth: ECRI 1996.
86. Sackett DL, Haynes RB, Tugwell P. *Epidemiología Clínica. Una ciencia básica*. Madrid: Díaz de Santos, 37.
87. Fleiss JL. *Statistical methods for rates and proportions*. Nueva York: John Wiley and Sons, 1981, pp: 212-225.
88. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clin Trials* 1996; 17: 1-12.
89. Detsky AS, Naylor CD, O'Rourke K, McGreer AJ, L'Abbe KA. Incorporating variations in the quality of individual randomized trials into meta-analysis. *J Clin Epidemiol* 1992; 45: 255-65.
90. Oxman AD, Guyatt GH. A consumer's guide to subgroup analyses. *Ann Intern Med* 1992; 116: 78-84.
91. Yusuf S, Wittes J, Probstfield J, Herman AT. Analysis and Interpretation of Treatment Effects in Subgroups of Patients in Randomized Clinical Trials. *JAMA* 1991; 266: 93-8.
92. Bailar JC III. Passive smoking, coronary heart disease, and meta-analysis. *N Engl J Med* 1999; 340: 958-9.
93. L'Abbé KA, Detsky AS, O'Rourke K. Meta-Analysis in Clinical Research. *Annals of Internal Medicine* 1987; 107: 224-233.
94. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *JAMA* 1992; 268: 240-8.
95. Bailar JC III. The practice of meta-analysis. *J Clin Epidemiol* 1995; 48: 149-57.
96. Le Lorier J, Grégoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med* 1997; 337: 536-42.
97. Petitti D. *Meta-Analysis, Decision Analysis and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine*. New York: Oxford University Press, 1994.
98. Hillner BE, Smith TJ. Efficacy and cost effectiveness of adjuvant chemotherapy in women with node-negative breast cancer: a decision-analysis model. *New England Journal of Medicine* 1991; 324: 160-168.
99. Pearson K. Report on Certain Enteric Fever Inoculation Statistics. *British Medical Journal* 1904; 2: 1243-1246.
100. Glass GV. Primary, Secondary and Meta-Analysis of Research. *Educational Researcher* 1976; 5: 3-8.
101. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988; 318: 1728-33.
102. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *Br Med J* 1995; 310: 452-454.
103. Figueredo E, Canosa L. Limitaciones y falacias de la fórmula del "número de pacientes que será necesario tratar". *Med Clin (Barc)* 1999; 112, 1: 36.
104. Jenicek M. Meta-análisis en medicina: reunión de experiencias. En: Jenicek M. *Epidemiología: la lógica de la epidemiología moderna*. Barcelona: Masson, 1996: pp. 285-317.
105. Pocock SJ et al. Meta-analysis of randomised trials comparing coronary angioplasty with bypass surgery. *Lancet* 1995; 346: 1184-89.
106. Walker AM, Martin-Moreno JM, Artalejo FR. Odd man out: a graphical approach to meta-analysis. *Am J Public Health* 1988; 78:961-6.
107. Williamson DF, Parker RA, Kendrick JS. The Box Plot: A simple visual method to interpret data. *Ann Int Med* 1989; 110: 916-21.
108. Galbraith RF. A note on graphical presentation of estimated odds ratios from several clinical trials. *Stat Med* 1988; 7: 889-94.
109. Woolf B. On estimating the relationship between blood group and disease. *Ann Human Genet*, 1955; 19: 251-3.
110. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986; 7: 177-88.
111. Altman DG, Gardner MJ. Confidence intervals for research findings. *Br J Obstet Gynaecol* 1992; 99: 90-1.
112. Light RJ, Pillemer DB. *Summing up. The science of reviewing research*. Cambridge (MA): Harvard University Press, 1984.
113. Delgado M, Sillero M, Gálvez R. Metaanálisis en epidemiología (1.ª parte): características generales. *Gaceta Sanitaria* 1991; 5 (27): 265-72.

114. Delgado M, Sillero M, Gálvez R. Metaanálisis en epidemiología (2.ª parte): métodos cuantitativos. *Gaceta Sanitaria* 1992; 6 (28): 30-39.
115. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, Mantel N, McPherson K, Peto J, Smith PG. Design and analysis of randomized clinical trials requiring prolonged observations of each patient. II. Analysis and examples. *Br J Cancer* 1977; 35: 1-39.
116. Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta blockade during and after myocardial infarction, an overview of the randomized trials. *Prog Cardiovas Dis* 1985B; 27: 335-71.
117. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiology Research*. California, Belmont: Lifetime Learning Publications, 1982.
118. Rothman KJ. *Epidemiología Moderna*. Madrid: Díaz de Santos, 1987.
119. Dales RE, Stark RM, Raman S. Computed tomography to stage lung cancer. *Am Rev Resp Dis* 1990; 141: 1096-1101.
120. Gilbert ES, Fry SA, Friggs LD, Volez GL, Cragle DL, Petersen GR. Methods for analyzing combined data from studies of workers exposed to low doses of radiation. *Am J Epidemiol* 1990; 131: 917-27.
121. Longnecker MP, Berlin JA, Orza MJ, Chalmers TC. A meta-analysis of alcohol consumption in relation to risk of breast cancer. *JAMA* 1988; 260: 652-6.
122. Sillero M, Delgado M, Rodríguez R, Bueno A, Gálvez R. Hormone Replacement therapy and breast cancer: A meta-analysis. *Obstet Gynecol* 1992; 79: 276-304.
123. Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992; 327: 248-54.
124. McIntosh M. The population risk as an explanatory variable in research synthesis of clinical trials. *Harvard Medical School*: 1995 1-35.
125. Sowden AJ, Deeks JJ, Sheldon TA. Volume and outcome in coronary artery bypass surgery: true association or artefact? *British Medical Journal* 1995; 311: 151-155.
126. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled clinical trials. *JAMA* 1995; 273: 408-412.
127. Detsky AS, Naglie G, Krahn MD, Naimark D, Redelmeier DA. Primer on Medical Decision Analysis. *Medical Decision Making* 1997; 17 (2): 123-159.
128. Review Manager [archivo de ordenador]. Windows version 3.0.1. Oxford: Cochrane Collaboration, Update Software, 1997.
129. Egger M, Sterne JAC, Davey Smith G. Meta-analysis software. [en línea] *eBMJ* 1998 (<http://www.bmj.com/collections/ma.htm>)
130. Haddix AC, Teutsch SM, Shaffer PA, Duñet DO. *Prevention Effectiveness. A Guide to Decision Analysis and Economic Evaluation*. Centers for Disease Control and Prevention. New York: Oxford University Press, 1996.
131. Von Neumann J, Morgenstern O. *Theory of Games and Economic Theory*. New York Wiley, 1947.
132. Weinstein MC, Fineberg HV. *Clinical Decision Analysis*. Philadelphia, WB Saunders Company, 1980.
133. Sox H, Blatt MA, Higgings MC, Marton KI. *Medical Decision Making*. London: Butterwood & Co., 1988.
134. Goldman L. Aspectos cuantitativos del juicio clínico. En: *Harrison. Principios de Medicina Interna*, 14.ª edición en castellano. Fauci AS et al eds. Madrid: Mac Graw-Hill, 1998.
135. Torrance GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 1986; 5: 1-30.
136. Fryback DG, Dasbach EJ, Klein R, et al. The Beaver Dam health outcomes study: initial catalog of health-state quality factors. *Medical Decision Making* 1993; 13: 89-102.
137. Kaplan RM, Anderson JP. A general health policy model: update and applications. *Health Services Research* 1988; 23: 203-35.
138. Redelmeier DA, Detsky AS. A clinician's guide to utility measurement. *Primary Care* 1995; 22: 271-8.
139. Froberg DG, Kane RL. Methodology for measuring health-state preferences -I. Measurement strategies. *Journal of Clinical Epidemiology* 1989; 42: 345-54.
140. Froberg DG, Kane RL. Methodology for measuring health-state preferences -II: Scaling methods. *Journal of Clinical Epidemiology* 1989; 42: 459-71.
141. Froberg DG, Kane RL. Methodology for measuring health-state preferences -III: Population and context issues. *Journal of Clinical Epidemiology* 1989; 42: 585-92.

142. Froberg DG, Kane RL. Methodology for measuring health-state preferences -IV: Progress and a research agenda. *Journal of Clinical Epidemiology* 1989; 42: 675-85.
143. Rué M. Las técnicas de ayuda a la decisión en Salud Pública. En: Martínez F, Antó JM, Castellanos PL, Gili M, Marset P, Navarro V. *Salud Pública*. Madrid: McGraw-Hill, 1997.
144. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation: a practical approach. *Med Decis Making* 1985; 5: 157-177.
145. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. Describing health states: methodologic issues in obtaining values for health states. *Med Care* 1984; 22: 543-552.
146. Read JL, Quinn RJ, Berwick DM, Fineberg HV, Weinstein MC. Preferences for health outcomes: comparison of assessment methods. *Med Decis Making* 1984; 4: 315-329.
147. Torrance GW, Thomas WH, Sackett DL. A utility maximization model for evaluation of health care programs. *Health Services Res* 1972; 7: 118-133.
148. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med* 1980; 302: 1109-17.
149. Beck JR, Pauker SG. The Markov process in medical prognosis. *Medical Decision Making* 1983; 3: 419-458.
150. Ransohoff DF, Gracie WA, Wolfenson LB, Neuhauser D. Prophylactic Cholecystectomy of Expectant Management for Silent Gallstones. *Annals of Internal Medicine* 1983; 99: 199-204.
151. DATA 3.5 [archivo de ordenador]. TreeAge Software, Inc. 1075 Main St. Williamstown, MA 01267 USA. Phone: 413-458-0104. Fax: 413-458-0105. February 1998.
152. Hollenberg JP. SMLTREE [archivo de ordenador]. New York: Cornell University, 1987.
153. Pauker SG. DECISION MAKER [archivo de ordenador]. Boston: Division of Clinical Decision Making, New England Medical Center, 1998.
154. Eddy DM, Hasselblad V. FAST\*PRO [archivo de ordenador]. Durham: Center for Health Policy Research and Education, Duke University.
155. Drummond MF, Stoddart GL, Torrance GW. *Métodos para la evaluación económica de los programas de atención de la salud*. Madrid, Díaz de Santos, 1991.
156. Gold MR, Siegel JE, Russell LB, Weinstein MC (eds.) *Cost-effectiveness in health and medicine*. New York: Oxford University Press, 1996.
157. Warner KE, Luce BR. *Cost-benefit and cost-effectiveness analysis in health care: principles, practice and potential*. Ann Arbor: Health Administration Press, 1982.
158. Luce BR, Elixhauser A. *Standards for socioeconomic evaluation of health care products and services*. New York: Springer-Verlag, 1990.
159. Russell LB, Gold MR, Siegel JE, Daniels N, Weinstein MC. The role of cost-effectiveness analysis in health and medicine. *JAMA* 1996; 276:1172-1177.
160. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the panel on cost-effectiveness in health and medicine. *JAMA* 1996; 276: 1253-1258.
161. Siegel JE, Weinstein MC, Russell LB, Gold MR. Recommendations for reporting cost-effectiveness analyses. *JAMA* 1996; 276: 1339-1341.
162. Torrance GW, Feeny DH. Utilities and quality-adjusted life years. *Int J Technology Assessment in Health Care* 1989; 5: 559-575.
163. Gramlich EM. *A guide to benefit-cost analysis*, 2<sup>nd</sup> ed. New Jersey: Englewood Cliffs, Prentice-Hall, 1990.
164. Eisenberg JM. Clinical Economics. A Guide to the Economic Analysis of Clinical Practices. *JAMA* 1989; 262: 2879-86.
165. Rovira J. *Glosario de términos y conceptos de uso frecuente en la evaluación económica de medicamentos y programas sanitarios*. Barcelona: SOIKOS SL. Química Farmacéutica Bayer, 1996.
166. Peiró S, Portella E. No todo es acuerdo en el consenso: limitaciones de los métodos de consenso en los servicios de salud. *Gac Sanit* 1993; 7: 294-300.
167. Mullan F, Jacoby I. The Town Meeting for Technology. The maturation of Consensus Conferences. *JAMA* 1985; 254 (8): 1068-1072.
168. Ministerio de Sanidad y Consumo - Instituto de Salud Carlos III - Agencia de Evaluación de Tecnologías Sanitarias (AETS). *Cirugía de la hipertrofia benigna de próstata. Estándares de uso apropiado*. Madrid: AETS - Instituto de Salud Carlos III, 1996.
169. Dalkey NC. The Delphi method: an experimental study of group opinion. RM-5888-PR. Santa Monica, California: Rand Corp, 1969.

170. Vinokur A, Burnstein E, Sechrest L, Wortman PM. Group decision making by experts: field study of panels evaluating medical technologies. *J Personality S Psych* 1985; 49 (1): 70-84.
171. Granados A. El proceso de evaluación de tecnologías médicas. La evidencia científica y la complejidad del contexto. *JANO*. 18-24 Octubre 1996; LI, n.º 1184: 1141-1143.
172. Guyatt GH, Mitchell A, Molloy DW, Capretta R, Horsman J, Griffith L. Measuring patient and relative satisfaction with level or aggressiveness of care and involvement in care decisions in the context of life threatening illness. *J Clin Epidemiol* 1995; 48: 1215-1224.
173. Levine RA, Shanaman RH. Translating clinical outcomes to patient value: an evidence-based treatment approach. *Int J Periodontics Restorative Dent* 1995; 15: 186-200.
174. O'Meara JJI, McNutt RA, Evans AT, Moore SW, Downs SM. A decision analysis of streptokinase plus heparin as compared to heparin alone for deep-vein thrombosis. *N Engl J Med* 1994; 330: 1864-1869.
175. Kassirer JP. Incorporating patients preferences into medical decisions. *N Engl J Med* 1994; 330: 1895-1896.
176. Hux JE, Naylor CD. Communicating the benefits of chronic preventive therapy: does the format of efficacy data determine patients' acceptance of treatment? *Med Decis Making* 1995; 15: 152-157.
177. Naylor CD, Llewellyn Thomas HA. Can there be a more patient-centred approach to determining clinically important effect sizes for randomized treatment trials? *J Clin Epidemiol* 1994; 47: 787-795.
178. Keeney RL. Decisions about life-threatening risks. *N Engl J Med* 1994; 331: 193-196.
179. Enkin M, Keirse MJNC, Renfrew MJ, Neilson JP. *A Guide to Effective Care in Pregnancy and Childbirth*. 2nd ed. Oxford: Oxford University Press, 1994.
180. Relman AS. Peer review in scientific journals-what good is it? *West J Med* 1990; 153: 520-522.
181. Feinstein AR. Construction, consent and condemnation in research on peer review [editorial]. *J Clin Epidemiol* 1991; 44: 339-341.
182. Granados A, Jonsson E. EURASSESS Subgroup on Dissemination and Evaluation of Impact. *International Journal of Technology Assessment in Health Care* 1997; 13 (2): 220-287.
183. Kanouse D, Kallich D, Kahan JP. Dissemination of effectiveness and outcomes research. *Health Policy* 1995; 34: 167-192.
184. Lomas J. Retailing research: increasing the role of evidence in clinical services for child birth. *The Milbank Quarterly* 1993; 71: 439-475.
185. Banta D, Granados A, Hailey D, Menon D, Parada A. *Document I: Review of the current format of the INAHTA structured abstracts*. INAHTA Work Group, October 1997.
186. Banta D, Granados A, Hailey D, Menon D, Parada A. *Document II: New proposal of the INAHTA structured abstracts*. INAHTA Work Group, October 1997.
187. International Committee of Medical Journal Editors. Uniform Requirements for Manuscripts Submitted to Biomedical Journals. *N Engl J Med* 1997; 336: 309-15.
188. Flynn K, Anderson D. *Positron Emission Tomography*. Boston: Management Decision and Research Center, Department of Veterans Affairs, 1996.
189. Oxman AD, Sackett DL, Guyatt GH, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, I: how to get started. *JAMA*. 1993; 270: 2093-2095. Ed. española 1997.
190. Guyatt GH, Rennie D. User's guides to the medical literature. *JAMA* 1993; 270: 2096-7. Edición española 1997.
191. Chalmers TC, Smith H, Blackburn B, Silverman B, Schoeder B, Reitman D, Ambroz A. A method for assessing the quality of a randomized control trial. *Controlled-Clin-Trials* 1981, Vol. 2, p. 31-49.
192. Moher D, Jadad AR, Nichol G et al. Assessing the quality of randomised controlled trials: an annotated bibliography of scales and checklists. *Controlled Clinical Trials* 1995; 16: 62-73.
193. Moher D, Jadad AR, Tugwell P. Assessing the quality of Randomized Controlled Trials – current issues and future directions. *International Journal of Technology Assessment in Health Care* 1996; 12 (2): 195-208.
194. Sacristán JA, Soto J, Galende I. Evaluación crítica de ensayos clínicos. *Med Clin (Barc)* 1993; 100: 780-787.
195. Jadad AR. Assessing the quality of RCTs: why, what, how and by whom? En: Jadad AR. *Randomised Clinical Trials. A user's guide*. London, BMJ Publishing Group, 1998, p. 45-60.
196. Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomised control trials in health care from the Postdam Consultation on Meta-analysis. *Journal of Clinical Epidemiology* 1995; 48 (1): 167-171.
197. Muir Gray JA. Atención sanitaria basada en la evidencia. Cómo tomar decisiones en gestión y política sanitaria. Madrid: Churchill. Livingstone, 1997.

198. Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. *Journal of Clinical Epidemiology* 1996; 49 (2): 235-243.
199. Lau J, Chalmers TC. The rational use of therapeutic drugs in the 21<sup>st</sup> century. Important lessons from cumulative meta-analyses of randomized control trials. *International Journal of Technology Assessment in Health Care* 1995; 11 (3): 509-522.
200. Moher D, Olkin I. Meta-analysis fo randomized controlled trials. A concern for standards; [comment]. Comment on: *JAMA* 1995; 274 (24): 1935-8. Comment on: *JAMA* 1995; 274 (24): 1942-8. *JAMA* 1995; 274 (24): 1962-4.
201. Nony P, Cucherat M, Haugh MC et al. Critical reading of the meta-analysis of clinical trials. *Therapie* 1995; 50 (4): 339-351.
202. Naylor CD, Guyatt GH, for the Evidence-Based Medicine Working Group. User's guides to the medical literature. X: how to use an article reporting variations in the outcomes of health services. *JAMA* 1996; 275: 554-558.
203. Levine M, Walter S, Hui Lee, Haines T, Holbrook A, Moyer V, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, IV: How to Use An Article About Harm. *JAMA*. 1994; 271: 1615-1619. Ed. española 1997.
204. Laupacis A, Wells G, Scott Richardson W, Tugwell P, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, V: How to Use an Article About Prognosis. *JAMA*. 1994; 272: 234-237. Ed. española 1997.
205. Fleiss JL, Gross AJ. Meta-analysis in epidemiology, with special reference to studies of the association between exposure to environmental tobacco smoke and lung cancer: a critique. *Journal of Clinical Epidemiology* 1991; 44: 127-139.
206. McMaster University, Department of Clinical Epidemiology and Biostatistics. How to read clinical journals, IV: to determine etiology or causation. *Canadian Medical Association Journal* 1981; 124: 985-990.
207. Richarson WS, Wilson MC, Guyatt GH, Cook DJ, Nishikawa J, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, XV. How to Use an Article About Disease Probability for Differential Diagnosis. *JAMA* 1999; 281 (13): 1214-19.
208. Naylor CD, Guyatt GH, for the Evidence-Based Medicine Working Group. User's guides to the medical literature. XI: how to use an article about a clinical utilization review. *JAMA* 1996; 275: 1.435-1439.
209. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. *Medicina Basada en la Evidencia. Cómo ejercer y enseñar la MBE*. Madrid, Churchill Livingstone España, 1997.
210. Salmi LR. *Lecture Critique et Rédaction Médicale Scientifique. Comment lire, rédiger et publier une étude clinique ou épidémiologique*. Paris, Elsevier, 1998.
211. Crombie IK. *The Pocket Guide to Critical Appraisal*. London, BMJ Publishing Group, 1996.
212. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, II: How to Use an Article About Therapy or Prevention. A. Are the results of the study valid?. *JAMA*. 1993; 270: 2598-2601. Edición española 1997.
213. Guyatt GH, Sackett DL, Cook DJ, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, II: How to Use an Article About Therapy or Prevention. B. What Were the Results and Will They Help Me in Caring for My Patients?. *JAMA*. 1994; 271: 59-63. Edición española 1997.
214. Jaeschke R, Guyatt GH, Sakett DL, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, III: How to Use an Article About a Diagnostic Test. A. Are the Results of the Study Valid?. *JAMA*. 1994; 271: 389-391. Ed. española 1997.
215. Jaeschke R, Guyatt GH, Sakett DL, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, III: How to Use an Article About a Diagnostic Test. B. What Are the Results and Will They Help Me in Caring for My Patients. *JAMA*. 1994; 271: 703-707. Ed. española 1997.
216. Oxman AD, Sackett DL, Guyatt GH, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, VI: How to Use an Overview. *JAMA*. 1994; 272: 1367-1371. Ed. española 1997.
217. Richardson WS, Detsky AS, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, VII: How to Use an article on clinical decision analysis. A: are the results of the study valid?. *JAMA* 1995; 273: 1292-1295. Ed. española 1997.
218. Richardson WS, Detsky AS, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, VII: How to Use an article on clinical decision analysis. B: What Are the Results and Will They Help Me in Caring for My Patients?. *JAMA* 1995; 273: 1610-1613. Ed. española 1997.
219. Hayward RSA, Wilson MC, Tunis SR, Bass EB, Guyatt GH, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, VIII: How to Use Clinical Practice Guidelines. A. Are the recommendations valid?. *JAMA*. 1995; 274: 570-574. Edición española 1997.

220. Wilson MC, Hayward RSA, Tunis SR, Bass EB, Guyatt G, for the Evidence-Based Medicine Working Group. User's guides to the medical literature. VIII: how to use clinical practice guidelines, B: what are the recommendations and will they help you in caring for your patients? *JAMA* 1995; 274: 1630-1632.
221. Guyatt GH, Naylor CD, Juniper E, Heyland DK, Jaeschke R, Cook DJ, for the Evidence-Based Medicine Working Group. User's guides to the Medical Literature. XII: how to use articles about health-related quality of life. *JAMA* 1997; 277: 1232-1236.
222. Verhagen AP, de Vet HCW, de Bie RA, Kessels AGH, Boers M, Bouter LM, Knipschild PG. The Delphi List: A Criteria List for Quality Assessment of Randomized Clinical Trials for Conducting Systematic Reviews Developed by Delphi Consensus. *J Clin Epidemiol* 1998; 51 (12): 1235-1241.
223. Drummond MF, Richardson WF, O'Brien BJ, Levine M, Heyland D, for the Evidence-Based Medicine Working Group. User's Guides to the Medical Literature. XIII: How to use an Article on Economic Analysis of Clinical Practice. A. Are the Results of the Study Valid? *JAMA* 1997; 277: 1552-1557.
224. O'Brien BJ, Heyland D, Richardson WS, Levine M, Drummond MF, for the Evidence-Based Medicine Working Group. User's Guides to the Medical Literature. XIII. How to use an Article on Economic Analysis of Clinical Practice. B. What Are the Results and Will They Help Me in Caring for My Patients?. *JAMA* 1997; 277: 1802-1806.
225. US Preventive Task Force. *Guide to clinical preventive services: an assessment of the effectiveness of 169 interventions*. Baltimore: Williams and Wilkins, 1989.
226. Bertram DA, Flynn K, Alligood E. *Endovascular Placed Grafts for Infrarenal Abdominal Aortic Aneurysms: A Systematic Review of Published Studies of Effectiveness*. Boston, Health Services Research & Development Service, Veteran Affairs Medical Center. Technology Assessment Program, Report n.º 9. 1998.
227. Jovell AJ, Navarro-Rubio MD. Evaluación de la evidencia científica. *Med Clin (Barc)* 1995; 105: 740-743.
228. Guyatt GH, Sackett DL, Sinclair JC, Hayward RSA, Cook DJ, Cook RJ, for the Evidence-Based Medicine Working Group. User's guides to the medical literature, IX: A method for grading health care recommendations. *JAMA* 1995; 274: 1800-1804. Edición española 1997.
229. Laupacis A, Feeny D, Detsky AS, Tugwell PX. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *Can Med Assoc J* 1992; 146 (4): 473-81.
230. Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* 1989; 95 (Supl): 2-3.
231. Canadian Task Force on the Periodic Health Examination. The periodic health examination. *Can Med Assoc J* 1979; 121: 1193-254.
232. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Medical Decision Making*. 1991; 11: 88-94.