

LAS CONFERENCIAS DE CONSENSO. ¿UN MÉTODO PARA LA EVALUACIÓN DE LAS TECNOLOGÍAS MÉDICAS?

Alicia Granados i Navarrete

Oficina Tècnica d'Avaluació de Tecnologia Mèdica / Departament de Sanitat i Seguretat Social / Generalitat de Catalunya

La evaluación de la tecnología es una forma de investigación que permite potencialmente una valoración de los riesgos, los beneficios y los costes de la utilización de las tecnologías médicas (TM), tanto nuevas como ya establecidas. Es, a su vez, un método de análisis que permite una valoración sistemática de las consecuencias deseables, indeseables o inciertas, de la introducción y difusión de las TM.

Uno de los diversos métodos para llevar a cabo la evaluación de las TM son las llamadas conferencias de consenso (CC), cuyo origen se remonta a 1977, en los National Institutes of Health (NIH) de los Estados Unidos de Norteamérica¹, siendo concebidas para mejorar la práctica clínica a través de un evaluación pública de los hallazgos en la investigación de tecnología biomédica.

Desde entonces y con una producción media de 7 CC por año, los métodos organizativos y el proceso de las conferencias de los NIH, así como las estrategias de difusión de sus resultados, han evolucionado. Muchos países han iniciado sus propios programas de desarrollo de consenso, preservando los principios básicos de los NIH que son, por lo general, comunes a todos ellos y adaptando tan sólo el proceso organizativo a su contexto nacional.

El proceso de consenso

Las bases fundamentales de una CC incluyen los siguientes elementos²:

1. *Jurado independiente*: Lo compone

un grupo de alrededor de 12 expertos independientes, cuya composición es variable según el tema objeto de cada CC. Dicho grupo debe juzgar de forma objetiva la información científica publicada y la expuesta tanto por los ponentes invitados como por la audiencia que asiste al acto.

2. *Preguntas previas*: El cometido principal del grupo de consenso consiste en responder específicamente a un número de preguntas (generalmente de 4 a 6), que son determinantes para definir el perfil y dirección de la conferencia.

3. *Sesiones*: El grupo se reúne, primero, en una sesión abierta (pública) en la que se procede a la presentación y discusión de todos los datos y, luego, en una sesión cerrada (ejecutiva) para la preparación del informe en el que se expone el resultado de su juicio consensuado, que posteriormente es presentado a continuación en sesión plenaria.

4. *Informe consensuado*: Al cierre de la sesión pública, el grupo prepara durante toda la noche el borrador del informe que será presentado por la mañana en sesión plenaria. Después de otra discusión abierta y tras recoger las enmiendas o comentarios del auditorio que el grupo de expertos considere apropiadas, se realiza la versión definitiva que constituirá la publicación formal de la conferencia.

5. *Difusión del informe consensuado*: Para lograr el máximo impacto en la práctica médica habitual, es imprescindible una amplia difusión del informe, tanto en la literatura científica especializada, como en los medios de difusión.

Los criterios para la elección del tema objeto de la conferencia son diversos, pero es condición *sine qua non* que exista suficiente evidencia científica para dar respuestas objetivas a las preguntas previamente planteadas, con tal de lograr que los resultados de la conferencia no dependan de las impresiones y juicios de valor de los miembros del jurado².

Propósitos y realidad de las CC

Las CC fueron diseñadas por los NIH para mejorar la comunicación entre los investigadores, los responsables de la labor clínica y la población en general. Las modificaciones en la composición del grupo de consenso introducidas por algunos países, como es el caso del Reino Unido (RU), que suele incorporar otros especialistas, como economistas y epidemiólogos, pretenden, a su vez, sensibilizar a los responsables de la política sanitaria³. El elemento clave para ello es el proceso de consenso en sí mismo que permite reunir en una discusión abierta a investigadores de diversas disciplinas junto a los usuarios para lograr un acuerdo sobre aspectos concretos de una tecnología dada.

Ante este perfil resulta tentador contemplar este método de evaluación como una posible panacea. Bien dirigida, una CC puede constituirse en un instrumento válido de docencia pluridisciplinaria y de educación sanitaria, así como un foro único donde se generen las mejores recomendaciones para la toma de decisiones en materia tecnológica. Sin embargo,

a lo largo de sus 14 años de vida, las CC también han evidenciado sus limitaciones y deficiencias. La más importante y frecuente de éstas es su impacto irregular en la práctica médica y en las estrategias de introducción y difusión de las TM (4-10). Uno de los defectos metodológicos que influye más directamente en la credibilidad de las conclusiones del informe emitido es que dicho informe casi nunca va avalado por referencias bibliográficas⁶, con lo que se desconocen las bases científicas (ensayos clínicos, estudios epidemiológicos, observacionales u opiniones) en las que se apoyan los expertos para emitir un juicio, lo cual impide, a su vez, valorar los grados de certeza en los que se basa dicho informe.

Una característica común de las CC que yo definiría como peculiar y que tiene, a mi entender, un efecto negativo sobre el resultado del informe final, es que éste se escribe en una última sesión cerrada y nocturna que puede llegar a durar toda la noche, con lo cual, el observador tiene la impresión de que las recomendaciones del informe son más el fruto de la capacidad de resistencia al sueño de unos pocos miembros del grupo que del grado de consenso alcanzado por la totalidad de los componentes del grupo.

Sin embargo, la limitación más importante detectada hasta ahora es, la práctica ausencia de una sistemática y exhaustiva revisión y síntesis de la literatura científica en la mayoría de programas de desarrollo de consenso. Casi en ningún caso se realiza un metaanálisis¹¹ ni ningún otro tipo de análisis de la información cuantitativa y cualitativa.

El consenso en España

En 1989 se reunieron en nuestro país un grupo de expertos y responsables de diversas sociedades científicas con el objeto de alcanzar un acuerdo de consenso sobre el control de la colesterolemia^{12, 13}. No se dispone de información publicada sobre cuál fue la metodología utilizada en el proceso seguido para lograr el consenso. No se explicitan tampoco, las preguntas previas a las que el grupo de consenso hubo de responder ni se define claramente cuál fue el grupo que juzgó, defendió o refutó la evidencia. Es decir, no queda claro si aquella reu-

nión fue o no una CC. Sea como fuere, casi simultáneamente, en el RU se realizaba una CC formal sobre el mismo tema¹⁴.

Es interesante observar que a pesar de que el resultado de ambos informes (el inglés y el español) debería haberse basado en la misma evidencia científica disponible, dada la proximidad de fechas en las que se realizaron ambas reuniones las recomendaciones de ambos grupos fueron sin embargo, distintas. Así pues, la presunta objetividad que proporciona la evidencia científica obtenida mediante la investigación rigurosa, queda en entredicho. Dicha objetividad se puede tornar subjetiva cuando intervienen diferentes observadores en contextos distintos, lo cual, como puede observarse en este caso, da lugar a resultados desiguales probablemente alejados de la rigurosidad inicial.

Comentario final

¿Están las CC diseñadas con un enfoque de modelo científico inadecuado? Es más que probable que cuando a la subjetividad de los observadores se le añade una ausencia de rigor metodológico en la revisión y síntesis del material a observar y juzgar, el grado de incertidumbre alcanzado puede llegar a anular completamente la credibilidad de las recomendaciones.

Sin embargo, ante la rápida e incuestionada introducción y difusión de las TM se hace necesario un foro de debate científico, es decir, crítico en esencia, que analice y disemine el conocimiento o desconocimiento acerca de las TM.

Para aumentar la verosimilitud e impacto de este foro quizás sea preciso rediseñar ya el proceso de consenso y clasificar, mediante una escala de grados posibles de certeza, las recomendaciones del grupo de expertos, dependiendo del nivel de evidencia de las fuentes consultadas. Así las recomendaciones basadas en los resultados de extensos ensayos clínicos con escaso margen de error, deberían tener mayor valor que aquellas basadas en el resultado del análisis de casos, o de series de casos.

Cuando se decide iniciar un programa de desarrollo de consenso que culmine en una CC, otro aspecto a valorar es la capacidad de producción científica que tiene una comunidad dada. Así ante una

comunidad científica poco productiva como la nuestra, habrá que tener en cuenta que la mayor parte de las fuentes en las que debe basarse la conferencia serán extranjeras, así como los expertos que expongan o juzguen la evidencia. Ello va a encarecer mucho el proceso de consenso; así pues se deberá decidir si vale la pena iniciarlo, es decir, debe valorarse si una CC en nuestro país puede resultar coste-efectiva, sobre todo si no se rediseña el proceso en sí mismo, que como hemos podido observar puede dar lugar a un escaso, si alguno, impacto en la práctica médica y en las estrategias de introducción y difusión de las TM.

Bibliografía

1. Perry S, Kalberer JT (Jr.), The N.I.H. Consensus Development Program and the Assessment of Health-Care Technologies. *N Eng J Med* 1980; 303: 169-72.
2. Office of Medical Applications of Research Methodes National Institutes of Health 1986. *Guidelines for the Selection and Management of Consensus Development Conferences*.
3. Stocking B. First Consensus Development Conference in United Kingdom: on coronary artery by pass grafting. *Br Med J* 1985; 291: 713-6.
4. Smith T. Consensus on cabbage. *Br Med J* 1984; 289: 1477.
5. Vang J. Who Regional Office for Europe. The Consensus Development Conference and the European Experience. *Int J Technol Assessment Health Care* 1986; 2: 65-76.
6. Perry S. The N.I.H. Consensus Development Program; A Decade Later. *N Eng J Med* 1987; 317: 485-8.
7. Kosecoff J, Kanouse DE, Rogers WH, McCloskey L, Winslow CM, Brook RH. Effects of the National Institute of Health Consensus Development Program on Physician Practice. *JAMA* 1987; 258: 2708-13.
8. Kanouse DE, Jacoby I. When Does Information Change Practitioner's Behavior? *Int J Technol Assessment Health Care* 1988; 4: 27-33.
9. Wortman PM, Vinokur A, Sechrest, L. Do Consensus Conferences Work? A Process Evaluation of the National Institute of Health. Consensus Development Program. *J Health Politics, Policy Law* 1988; 13: 469-98.
10. Calltorp J. Consensus Development Conferences in Sweden. Effects on Health Policy and Administration. *Int J Technol Assessment Health Care* 1988; 4: 75-88.
11. McGlynn E, Kosecoff J, Brook RH. Format and conduct of Consensus Development Conferences. Multination Comparison. *Int J Technol Assessment Health Care* 1990; 450-69.
12. Banegas JR. Las conferencias de consenso llegan a nuestro país. Un instrumento racional para la evaluación y difusión de las tecnologías médicas. *Rev San Hig Pub* 1989; 63: 7-14.
13. Acuerdo para el control de la Colesterolemia en España. *Rev San Hig Pub* 1989; 63: 109-21.
14. *Consensus Statement on Blood Cholesterol Measurement in the Prevention of Coronary Heart Disease*. The Sixth King's Fund Forum, 1989. London: King's Fund Center, 1990.