



TESIS DOCTORAL

MEJORA DE UN MÉTODO DE CONSENSO EXPERTO PARA REDUCIR LA VARIABILIDAD DE LA PRÁCTICA CLÍNICA

APLICABILIDAD Y RENDIMIENTO DE UNA ADAPTACIÓN DEL
MÉTODO DELPHI MODIFICADO, COMO INSTRUMENTO DE
CONSENSO EXPERTO PARA ELABORAR RECOMENDACIONES
EN SITUACIONES DE INCERTIDUMBRE CLÍNICA

FERNANDO CABALLERO MARTINEZ

MADRID, 2014

DIRECTOR: LUIS CABALLERO MARTÍNEZ

DEPARTAMENTO DE PSIQUIATRIA. FACULTAD DE MEDICINA

© Fernando Caballero Martínez, 2014 - f.caballero@ufv.es

INFORME DEL DIRECTOR DE TESIS PARA LA AUTORIZACIÓN DE DEFENSA DE TESIS DOCTORAL

D. LUIS CABALLERO MARTÍNEZ.

Director/a de la tesis doctoral de **D. FERNANDO CABALLERO MARTÍNEZ.**

Informa favorablemente la solicitud de autorización de defensa de la tesis doctoral con el título:

MODIFICACIÓN DE UN MÉTODO DE CONSENSO EXPERTO PARA REDUCIR LA VARIABILIDAD DE LA PRÁCTICA CLÍNICA (APLICABILIDAD Y RENDIMIENTO DE UNA ADAPTACIÓN DEL MÉTODO DELPHI MODIFICADO, COMO INSTRUMENTO DE CONSENSO EXPERTO PARA ELABORAR RECOMENDACIONES EN SITUACIONES DE INCERTIDUMBRE CLÍNICA).

presentada por dicho/a doctorando/a.

Programa de Doctorado: RD 185/1985 (PSIQUIATRIA)

La tesis está sometida a procesos de confidencialidad: Sí NO X

La tesis se presenta como compendio de publicaciones: Sí X NO

Fecha: Madrid a 1 de septiembre de 2014

Fdo. LUIS CABALLERO MARTÍNEZ

Índice

ÍNDICE GENERAL

	Pág.
AGRADECIMIENTOS Y DEDICATORIA.....	11
RESUMEN.....	15
LISTA DE TABLAS.....	23
LISTA DE FIGURAS.....	27
ABREVIATURAS.....	31
TITULO Y SUBTÍTULO.....	35
1. INTRODUCCIÓN.....	39
1.1. Sobre la incertidumbre clínica, la variabilidad injustificada de la práctica médica y la necesidad de una medicina basada en pruebas científicas.	41
1.2. Sobre algunos límites, equívocos y dificultades para la práctica de una <i>Medicina Basada en la Evidencia</i> .	47
1.3. Reivindicación del criterio experto: el papel de la experiencia en la toma de decisiones clínicas.	56
1.4. Conceptos generales sobre el consenso profesional y comparación crítica de los principales métodos.	61
1.5. El método Delphi de consenso profesional y el método Delphi modificado.	69
1.6. Necesidad de desarrollar una nueva adaptación del método Delphi modificado para su uso en el ámbito clínico.	81
2. HIPÓTESIS Y OBJETIVOS.....	87
2.1. Hipótesis que pretende demostrar esta tesis	89
2.2. Objetivos	90
3. MÉTODOS.....	93
3.1. Diseño metodológico de la tesis por compendio de publicaciones.	95
3.2. Aspectos generales de la adaptación clínica de método Delphi Modificado (DEMOCLI) usada en los estudios presentados.	98
3.3. Elaboración de los cuestionarios y procedimiento de síntesis de la evidencia científica utilizados en los estudios presentados.	101
3.4. Escalas de medida de los cuestionarios usadas en los estudios presentados.	103
3.5. Particularidades de las rondas de encuestas en la adaptación clínica del método Delphi modificado (DEMOCLI) propuestas en los estudios presentados.	105

3.6. Criterios de análisis estadístico e interpretación de resultados de los consensos.	109
3.7. Contexto clínico de los estudios presentados: tema, ámbito de investigación y fin estratégico del consenso pretendido.	113
3.8. Selección de panel de expertos participantes en los consensos.	117
4. RESULTADOS	123
4.1. Referencia general de las publicaciones científicas derivadas de los estudios presentados en esta tesis.	126
4.2. Breve resumen del marco teórico justificativo, finalidad y resultados principales de los estudios compilados.	129
4.3. Separatas originales en facsímil de las publicaciones en revistas científicas de los estudios presentados.	143
5. DISCUSIÓN	145
5.1. Reflexiones sobre el origen de la hipótesis de estudio, sobre los fines de esta tesis y sobre la pertinencia de una tesis por compendio de publicaciones.	147
5.2. Reflexiones sobre las aportaciones y la efectividad de la adaptación del método Delphi modificado (DEMOCLI) para el diseño apropiado de las encuestas.	150
5.3. Reflexiones sobre los atributos identitarios del experto-clínico y sobre la adecuación del tamaño y composición de los paneles expertos empleados en los estudios.	153
5.4. Reflexiones sobre la idoneidad de las distintas escalas de respuesta empleadas en las encuestas y sobre la validez del criterio estadístico de consenso empleado en los estudios presentados.	159
5.5. Reflexiones sobre el impacto de la sustitución de la reunión presencial de panelistas entre rondas por el proceso interacción no presencial propuesto en el método DEMOCLI.	163
5.6. Reflexiones sobre el valor del uso alternativo de la adaptación clínica del método Delphi (DEMOCLI) para la validación y refrendo de recomendaciones preexistentes (otras Guías).	165
5.7. Reflexiones sobre posibles limitaciones de esta tesis, sugerencias sobre nuevas líneas de investigación sobre la materia y algunas secuelas científicas de esta tesis.	167
6. CONCLUSIONES	171
EPÍLOGO	177
BIBLIOGRAFÍA	181

ANEXOS	195
ANEXO 1: Instrucciones al comité científico para la elaboración del cuestionario Delphi.	101
ANEXO 2: Niveles de evidencia. Oxford Centre for Evidence-Based Medicine.	203
ANEXO 3: Declaración aclaratoria del autor principal de los estudios en los que el doctorando no es primer o último firmante (autor senior).	205
ANEXO 4. Originales en facsímil de las publicaciones compendiadas en esta tesis, en orden cronológico de publicación (2009-2013).	207
• Artículo 1	211
• Artículo 2	223
• Artículo 3	235
• Artículo 4	249
• Artículo 5	265
• Artículo 6	279
• Artículo 7	297
• Artículo 8	307
• Artículo 9	321

Agradecimientos y dedicatoria



*Where is the wisdom we have lost in knowledge?
Where is the knowledge we have lost in information?*

The Rock, 1934. T. S. Eliot (1888-1965).

¿Dónde está la sabiduría que hemos perdido con el conocimiento?

¿Dónde está el conocimiento que hemos perdido con la información?

DEDICATORIA

- A mi familia, por el tiempo y la energía robados.
- A mi padre, médico de cabecera durante 48 años, por mostrarme día a día el valor de la experiencia clínica.

AGRADECIMIENTOS

- A los 56 coautores del doctorando en los estudios que articulan esta tesis, por su rigor y su generosidad.
- A los 513 expertos que aportaron su opinión profesional en los consensos.
- A todos quienes han revisado y criticado los estudios, por su ayuda imprescindible.
- A Luis Caballero, Director de esta tesis, por su acogida y su orientación constante.
- A Patricia, colaboradora imprescindible en los estudios, por su paciencia y por suplir con orden mis carencias.
- A Cristina, a Diana y a todos mis compañeros de trabajo, por multiplicar mi capacidad y permitirme encontrar el tiempo para acabar esta tesis.

Resumen

RESUMEN

Marco teórico:

La práctica de una auténtica “medicina basada en pruebas”, en su concepción actual, es la integración juiciosa de las mejores pruebas de la investigación clínica, con la maestría clínica (experiencia) del profesional y las preferencias del paciente, para tomar las decisiones clínicas más apropiadas en el cuidado sanitario de cada sujeto concreto.

En este proceso, el clínico afronta en ocasiones situaciones respecto a las que no tiene experiencia o conocimiento suficientes. La necesidad de tomar decisiones en condiciones de incertidumbre es un atributo inherente de la tarea clínica, y un determinante principal de la alta variabilidad de prácticas apreciable en todos los ámbitos asistenciales. Esta variabilidad no es reflejo de un estilo de práctica arbitraria sino que, en gran medida, es inducida por la ausencia de pruebas científicas que avalen cuál es la opción, diagnóstica o terapéutica, de la que más se beneficiaría el paciente.

El devenir del paciente depende, en ocasiones críticamente, de conseguir la mejor respuesta posible a estas dudas clínicas. La mayoría de estas cuestiones se resuelven, en la práctica, consultando con compañeros expertos ante la dificultad de acceder a la información científica en el punto de atención (o no existen evidencias o no resultan fáciles de localizar, interpretar o aplicar justo cuando se necesitan).

En este contexto, las recomendaciones basadas en el criterio de los expertos siguen siendo una ayuda estimable para tomar decisiones en la práctica rutinaria, en particular cuando el clínico afronta escenarios de incertidumbre por ausencia de evidencias sólidas, por controversia científica entre las pruebas disponibles, o por una limitada posibilidad de aplicación directa de estas evidencias en su contexto clínico real (ancianidad, comorbilidad, riesgo de efectos adversos o interacciones medicamentosas, baja adherencia a recomendaciones, ...).

Por ello, aunque los consensos de expertos se localizan en rango inferior de nivel de prueba que puede avalar una recomendación clínica, si están bien

desarrollados, con frecuencia resultan la única opción viable para disponer de un criterio clínico unificador que reduzca la variabilidad de prácticas y optimice la asistencia.

Entre las diversas opciones técnicas para unificar criterio entre expertos, el método Delphi a distancia tiene una trayectoria consolidada de 70 años de aplicación y desarrollo en diversos ámbitos. Sus señas de identidad nucleares son:

- Proceso iterativo mediante un número indeterminado de rondas de encuesta hasta conseguir estabilidad en las opiniones del grupo.
- Anonimato riguroso de los participantes, a fin de reducir el efecto pernicioso que pueden ejercer los posibles elementos dominantes del grupo.
- *Feedback* informativo sobre los resultados del grupo en cada ronda para permitir su interacción controlada.
- Valoración estadística de las respuestas y del consenso alcanzado.

El “método Delphi modificado” en dos rondas, una variante del procedimiento original introducida hace 20 años, ha logrado una enorme proyección gracias a ciertas ventajas de aplicabilidad y rendimiento frente a otras alternativas de consenso presencial (grupos nominales, conferencias de consenso, etc.) y frente al propio Delphi tradicional. Se trata de un método híbrido entre el método Delphi a distancia (sucesivas iteraciones de encuesta) y el grupo nominal, ya que requiere una sesión de debate presencial entre sus dos únicas rondas de encuesta.

Aunque el método se ha usado con éxito en muchos ámbitos profesionales, y en la propia medicina para la evaluación experta de usos apropiados de la tecnología sanitaria, su aplicación para consensuar la opinión de clínicos expertos y ayudar a responder dudas clínicas, presenta ciertos inconvenientes insalvables.

Una modificación técnica de aspectos no sustanciales del procedimiento podría transformarlo en una alternativa viable y efectiva para ayudar a resolver situaciones inciertas (diagnósticas, terapéuticas, pronósticas,...) y unificar las opiniones de grandes paneles de expertos sobre la actuación clínica más recomendable.

Hipótesis:

Una nueva adaptación clínica del método Delphi modificado (DEMOCLI) permite alcanzar un consenso rápido y efectivo en grupos de clínicos expertos sobre recomendaciones de actuación en escenarios de incertidumbre clínica, por ausencia de una evidencia sólidamente establecida, por controversias científicas entre las pruebas disponibles o por inaplicabilidad estas en la práctica rutinaria.

Objetivos:

1. Presentar y discutir las características metodológicas innovadoras de una nueva adaptación del método Delphi modificado (DEMOCLI) especialmente diseñado para el debate y consenso de expertos clínicos sobre dudas de la práctica asistencial.
2. Comprobar que DEMOCLI es una metodología de consenso válida (viable, eficiente, aceptable y versátil) para la construcción de recomendaciones clínicas expertas en respuesta a situaciones de incertidumbre.

Metodología:

Se efectúa una revisión bibliográfica exhaustiva previa, justificadora de la tesis e introductoria al método Delphi y sus variantes, a sus usos, ventajas y limitaciones.

Se describe el diseño, y se argumenta la innovación que representa, una nueva variante clínica de diseño propio (**DEMOCLI**) de la técnica Delphi modificada. Respecto al método Delphi modificado en dos rondas previo, las características principales de la adaptación metodológica propuesta incluyen:

- La responsabilidad exclusiva del Comité Científico (líderes de referencia en la especialidad y materia de estudio) en la elaboración de los cuestionarios de estudio, a partir de preguntas clínicas, sin implicar al pleno del panel de expertos en esta fase.
- El uso de paneles de expertos clínicos más numerosos (hasta 90 miembros) que la composición habitual usada en el Delphi modificado (en general 8-12).
- El hecho de que ambas rondas de encuesta completen sus tareas sin necesidad de reunión presencial de debate, para evitar sesgos de influencia y garantizar viabilidad logística del procedimiento con paneles numerosos.

- Se suple esa reunión por el ofrecimiento de un *feedback* informativo detallado entre rondas (resultados obtenidos y opiniones libres de panelistas)

Se presentan y analizan críticamente nueve publicaciones recientes del doctorando (2009-2013) que comparten el método DEMOCLI como herramienta para el desarrollo de consensos clínicos expertos en diversas especialidades y temas de interés clínico.

Todos los estudios, consensos expertos multicéntricos de ámbito estatal, han sido dirigidos por el autor de la tesis. Su contribución a los estudios incluyó la consultoría metodológica a los equipos investigadores (formados por líderes clínicos, docentes o investigadores en sus respectivas especialidades) y las labores de conducción técnica, recogida de información, análisis de datos, interpretación de resultados y redacción científica.

Se analizan comparativamente sus resultados y se extraen conclusiones conjuntas que intentarán responder a los objetivos planteados, sobre una experiencia real de investigación constatable, prolongada y en distintas áreas clínicas.

Resultados:

Se presentan nueve estudios multicéntricos de ámbito estatal que han empleado por primera vez el nuevo método **DEMOCLI** para el desarrollo de recomendaciones clínicas por consenso experto.

La adaptación metodológica del nuevo método responde a necesidades propias de consenso profesional entre expertos clínicos, y parece haber resultado determinante de la aceptación de los expertos a la propuesta técnica, de la viabilidad práctica, en tiempo y forma, de los proyectos, y de que sus costes de ejecución hayan resultado asumibles.

En los estudios han participado activamente 513 expertos españoles de 13 especialidades médicas. Se logró reclutar más del 90% de los candidatos seleccionados mediante técnicas de muestreo *ad hoc* (*snow ball sample*). No se registró ningún abandono durante el proceso de consenso.

Los paneles alcanzaron acuerdos satisfactorios en todos los estudios (superior al 85% de las proposiciones analizadas), en los plazos previstos (2-3 meses) y sin conflictos relevantes (el 100% de los participantes refrendaron las conclusiones de los consensos sin expresar veto a ninguna recomendación).

Los estudios, promovidos bajo el auspicio de hasta 14 sociedades científicas, han sido publicados en siete revistas científicas indexadas en bases de datos internacionales, cuyos revisores y comités editoriales han ratificado su interés y rigor técnico.

Todo ello puede considerarse un pilotaje suficiente de la propuesta y una constatación empírica de la aplicabilidad y rendimiento del nuevo método.

Conclusiones:

La tesis presenta, pilota y evalúa, sobre la experiencia objetiva de una amplia serie de 9 estudios ya publicados, los resultados de una innovadora adaptación técnica (“clínica”) del método Delphi Modificado (**DEMOCLI**), como procedimiento de consenso no presencial para elaborar recomendaciones expertas que orienten la actuación de los médicos en situaciones de incertidumbre clínica y/o alta variabilidad de prácticas.

En todos los proyectos en los que se empleó **DEMOCLI** se ha mostrado como una técnica operativa de negociación y consenso profesional que permite alcanzar una alta tasa de consenso experto.

El éxito en el reclutamiento de 513 panelistas y su permanencia sin excepciones hasta la finalización de los proyectos confirma el alto grado de aceptación y participación logrado por este procedimiento entre clínicos expertos. La conformidad final de los participantes con todas las recomendaciones desarrolladas mediante **DEMOCLI**, sin conflictos ni vetos, avalan la aceptabilidad y suficiencia del nuevo método para los clínicos.

DEMOCLI ha mostrado su versatilidad en distintos tipos de ámbitos asistenciales y especialidades médicas (hasta 13). El aval científico de los proyectos por diversas sociedades científicas médico-quirúrgicas y otros grupos especializados de

referencia estatal en diversas disciplinas, respalda la pretensión de propagar la propuesta metodológica a nuevos entornos clínicos.

Al menos tres grupos ajenos al investigador han empleado **DEMOCLI** en nuevos consensos clínicos expertos publicados en el último año, refrendando la generabilidad del método.

Las innovaciones técnicas introducidas por **DEMOCLI** han superado con éxito el criterio editoriales de diversas revistas científicas que siguen los procedimientos normalizados de revisión y aprobación del Comité Internacional de Editores de Revistas Médicas (ICMJE) y que están internacionalmente indexadas.

DESCRIPTORES

- Variabilidad de práctica clínica (Clinical variation)
- Incertidumbre clínica (Clinical uncertainty)
- Toma de decisión clínica (Clinical decision making)
- Consenso Experto (Expert consensus)
- Delphi modificado (Delphi modified)
- Medicina basada en pruebas (Evidence based medicine)

Tablas

LISTA DE TABLAS

Tabla I: Comparativa de los principales métodos de consenso. Ventajas e inconvenientes.	68
Tabla II: Contenidos e implementación de las encuestas en los estudios presentados.	108
Tabla III: Objetivos, contenido temático y ámbito de los consensos.	116
Tabla IV: Características del panel de expertos de los estudios.	122
Tabla V: Referencia bibliográfica y bibliométrica, coautoría, auspicio y patrocinio de los estudios.	128
Tabla VI: Resultados de los estudios. Grado de consenso final alcanzado sobre las recomendaciones analizadas.	130

Figuras

LISTA DE ILUSTRACIONES (FIGURAS)

Figura 1: Dimensiones de la competencia profesional.	42
Figura 2: Estructura y atributos básicos de las rondas de encuesta Delphi.	73
Figura 3: Etapas de implementación de la técnica Delphi original.	74
Figura 4: Etapas de implementación del método Delphi modificado (DEMO).	78
Figura 5: Etapas de implementación de la adaptación clínica del método Delphi modificado (DEMOCLI).	100
Figura 6: Presentación a los panelistas de la escala de Likert de 9 puntos usada en algunos estudios.	104
Figura 7: Presentación a los panelistas de la escala de Likert de 5 puntos usada en algunos estudios.	105
Figura 8: Ejemplo de la retroalimentación informativa tras la 1ª ronda (gráficos e indicadores estadísticos, comentarios y aclaraciones aportadas por participantes) ofrecida como información intermedia a los panelistas para su reflexión.	110
Figura 9: Ejemplo de representación comparativa de puntuaciones medias (e IC95%) entre distintos ítems de una encuesta.	111

Abreviaturas

ABREVIATURAS Y SÍMBOLOS

- **EBM:** *Evidence Based Medicine*.
- **MBP:** Medicina Basada en Pruebas.
- Método **DEMO:** Método Delphi Modificado.
- Método **DEMOCLI:** Adaptación Clínica del método Delphi Modificado.

Nota aclaratoria: Para facilitar la lectura del texto y tablas de este escrito, los nueve estudios que forman el cuerpo de publicaciones compiladas en esta tesis se citarán mediante una única palabra (a modo de descriptor semántico) y su año de publicación, que permitan evocar los aspectos claves del estudio sin tener que acudir a la consulta constante a las llamadas bibliográficas.

- **DISLIPEMIA 2009:** Ampudia Blasco FJ, Baos Vicente V, Caballero Martínez F, Álvarez-Sala Walther LA. Criterios profesionales sobre el papel de ezetimiba en el manejo clínico del paciente dislipémico en el ámbito de atención primaria. *Av Diabetol* 2009; 25(5): 399-407.
- **ASMA 2010:** Caballero Martínez F, Plaza Moral V, Quirce Gancedo S, Fernández Benítez M, Gómez Ruiz F, López Viña A, Molina París J, Quintano Jiménez JA, Soler Vilarrasa R, Villa Asensi JR, Balmes Estrada S. Valoración externa de las recomendaciones de la GEMA 2009 por un panel multiprofesional de expertos en asma. *Arch Bronconeumol* 2010; 46(8): 411-419.
- **EPILEPSIA 2010:** Villanueva Haba V, Sánchez Álvarez JC, Peña Mayor P, Salas Puig J, Caballero Martínez F, Gil-Nagel Rein A. Treatment initiation in epilepsy. An expert consensus in Spain. *Epilepsy and Behaviour* 2010; 19(3): 332-42.
- **PEDIATRIA 2010:** Medrano López C, García-Guereta L, Fernández Pineda L, Malo Concepción P, Maroto Álvaro E, Santos de Soto J, Lirio Casero J, Suárez Cabrera P, Caballero Martínez F. Consenso clínico sobre la profilaxis de la infección por virus respiratorio sincitial (VRS) y el uso del palivizumab en cardiología pediátrica. *An Pediatr (Barc)*. 2010; 72(6): 432.e1-432.e13.
- **DEPRESION 2011:** Martín Carrasco M, Agüera Ortiz L, Caballero Martínez L, Cervilla Ballesteros J, Menchón Magriñá JM, Montejo González AL, Moríñigo Dominguez A, Caballero Martínez F. Consenso clínico sobre la depresión en el paciente anciano. *Actas Esp Psiquiatr* 2011; 39(1): 20-31.

- **ESQUIZOFRENIA 2012:** Bobes García J, Saiz Ruiz J, Bernardo Arroyo M, Caballero Martínez F, Gilaberte Asín I, Ciudad Herrera A, en representación del Grupo de Trabajo sobre la Salud Física del Paciente con Trastorno Mental Grave. Consenso Delphi sobre la salud física del paciente con esquizofrenia: valoración por un panel de expertos de las recomendaciones de las Sociedades Españolas de Psiquiatría y de Psiquiatría Biológica. *Actas Esp Psiquiatr* 2012; 40(3): 114-28.
- **INHALADORES 2012:** Plaza Vicente, Calle M, Molina J, Quirce S, Sanchis J, Viejo JL, Caballero F. Validación, mediante el método Delphi, del Consenso sobre Terapia Inhalada para el tratamiento de las enfermedades respiratorias obstructivas. *Arch Bronconeumol* 2012; 48(6): 189-196.
- **PROSTATA 2012:** Cózar Olmo JM, Hernández Fernández C, Miñana López B, Amón Sesmero JH, Montlleó González M, Rodríguez Antolín A, Caballero Martínez F. Consenso experto sobre el impacto clínico de la nueva evidencia científica disponible sobre Hiperplasia Benigna Prostática. *Actas Urol Esp* 2012; 36(5): 265-275.
- **DOLOR 2013:** Pérez C, Rodríguez MJ, Guerrero A, Margarit C, Martín-Estefanía C, Oteo-Álvaro A Caballero-Martínez F. Consenso experto sobre el uso clínico de los tratamientos por vía tópica en el manejo del dolor neuropático periférico. *Rev. Soc. Esp. Dolor.* 2013; 20(6): 308-323.

Título

TÍTULO

MEJORA DE UN MÉTODO DE CONSENSO EXPERTO PARA REDUCIR LA VARIABILIDAD DE LA PRÁCTICA CLÍNICA

SUBTÍTULO

APLICABILIDAD Y RENDIMIENTO DE UNA ADAPTACIÓN DEL MÉTODO DELPHI MODIFICADO, COMO INSTRUMENTO DE CONSENSO EXPERTO PARA ELABORAR RECOMENDACIONES EN SITUACIONES DE INCERTIDUMBRE CLÍNICA

Introducción



1. INTRODUCCIÓN

1.1. Sobre la incertidumbre clínica, la variabilidad injustificada de la práctica médica y la necesidad de una medicina basada en pruebas científicas.

La **práctica clínica**, el desempeño profesional por antonomasia del médico, es el amplio abanico de actuaciones médicas que se desarrollan durante el proceso de atención a los problemas de salud de un paciente, con fines diagnósticos, pronósticos, terapéuticos, preventivos o rehabilitadores. La tarea requiere la puesta en juego de las diversas dimensiones que integran la competencia profesional del médico (perceptivas, cognitivas, psicomotoras, actitudinales, relacionales, éticas,...).

Ser un clínico competente requiere un amplio conocimiento previo de las ciencias básicas y clínicas, el dominio de ciertas destrezas para la recogida y análisis de información clínica (que incluyen habilidades relacionales, comunicativas y exploratorias, así como el uso eficiente de tecnologías sanitarias), una rigurosa capacidad de razonamiento y buen juicio para integrar conocimiento médico e información clínica, criterio técnico para tomar decisiones apropiadas durante la resolución de problemas, capacidad de aplicar las intervenciones clínico-terapéuticas propias de cada especialidad, etc..., todo ello puesto en acción dentro de un marco de valores éticos y profesionales que garanticen un ejercicio profesional efectivo, seguro, satisfactorio para el paciente, socialmente aceptable y asumible en términos de coste.

La competencia profesional de un médico es, por tanto, un constructo multidimensional complejo que incluye conocimiento científico, habilidades personales variadas (cognitivas, manuales, instrumentales, sociales y relacionales), capacidad de juicio y actitudes profesionales adecuadas, todo ello sustentado en un sólido referente ético.

Figura 1: Dimensiones de la competencia profesional



Una actividad profesional como la práctica de la medicina, que requiere una competencia tan poliédrica y que depende de otras muchas variables externas (modelo asistencial, medios accesibles, contexto socio-económico, entorno cultural, ...) resulta virtualmente imposible de estandarizar. La mayoría de los estudios que han analizado la práctica clínica en los más diversos ámbitos clínicos, constatan que no se trata de un fenómeno exacto y reproducible, sino que está sujeto a una significativa **variabilidad**. En ocasiones, este fenómeno es tan llamativo que algunas decisiones médicas parecen arbitrariamente azarosas y difíciles de explicar (1,2).

Las diferencias en la forma en que los médicos tratan a los pacientes con una condición de enfermedad similar han sido un motivo recurrente de investigación durante las últimas décadas bajo el supuesto de que al reducir la variación, se podrá mejorar la calidad asistencial, se reducirán sus costes o ambas cosas. Hay una gran cantidad de literatura que describe diferencias sustanciales en el uso del sistema sanitario en comunidades aparentemente similares (3-10).

Se han hecho intentos de identificar los factores determinantes de esta variación en cuanto a diferencias de la población (por ejemplo, edad, sexo, tipo de aseguramiento, estado de salud, nivel cultural, etc.) o del sistema sanitario (por ejemplo, ámbito asistencial, accesibilidad geográfica, dotación de medios diagnósticos o terapéuticos, modelo de gestión, recursos humanos o número de camas de hospital disponibles, etc.).

Cuando las diferencias significativas en las tasas de consumo sanitario se mantienen después de controlar las posibles diferencias en las circunstancias externas antedichas, tal situación se puede atribuir a determinantes intrínsecos de la práctica médica, como puedan ser algunos atributos del propio médico (perfil sociodemográfico y clínico, nivel formativo, grado de experiencia, actitudes personales, etc.) o a cierta inseguridad inherente al ejercicio de la medicina, conocida como "incertidumbre clínica".

La esencia de la incertidumbre clínica deriva de la dificultad para disponer de un criterio cierto sobre lo que constituye el manejo clínico óptimo para cada paciente concreto, individualmente considerado. Una temprana y frustrante experiencia del médico novel es constatar que no existe un modelo exacto y reproducible de enfermedad (*"no existen enfermedades sino enfermos"*), que la realidad frecuentemente diverge del arquetipo teórico estudiado en los libros, y que la variabilidad interindividual que observa en su práctica le obliga a tomar continuas decisiones en condiciones inciertas, donde la supuesta seguridad de las reglas y recomendaciones científicas (diagnósticas, pronósticas o terapéuticas) no resulta de aplicación exacta y pierde validez.

Probablemente parte de esta variabilidad está justificada cuando los médicos, durante el proceso de diagnóstico y tratamiento, ven limitado por circunstancias externas el acceso o la calidad de los recursos apropiados a las necesidades clínicas del paciente concreto que atienden (impuestas por el entorno y las condiciones del trabajo). Todo ello se consideraría una "variabilidad **inducida**". Pero, si pudieran identificarse y eliminarse todas las variables que conforman estas influencias, la variabilidad remanente en el ejercicio clínico representaría una "variabilidad **innata**" interprofesional, incluso intraprofesional, debida, en principio, a diferencias injustificadas de criterio y de estilo de práctica.

Una definición sencilla, ampliamente aceptada, de la variabilidad **injustificada** ha sido propuesta por Jack Wennberg como la "variación en la utilización de los servicios de salud que no puede ser explicada por diferencias de enfermedad en los pacientes o por sus preferencias" (11,12).

Si esta última variabilidad, difícil de explicar, fuera consecuencia directa de la arbitrariedad, pondría en riesgo la calidad de los cuidados médicos, produciría un efecto lesivo sobre el crédito profesional y cuestionaría la eficiencia de un sistema en el que se constatan usos no apropiados de procedimientos e intervenciones y donde no parecen emplearse criterios claros sobre adecuación diagnóstica o terapéutica.

Por fortuna, el verdadero el origen de esta variabilidad no explicable recae principalmente en la limitada calidad del conocimiento científico que subyace a la toma de decisiones, la cual puede está condicionada por (13):

1. Ausencia de un conocimiento científico sólidamente establecido. En tal caso, el médico tiende a basar las decisiones en su experiencia y las posibilidades de equívoco aumentan. Ello se debe a que las observaciones personales son escasas y no sistematizadas, la memoria es selectiva, las apreciaciones sesgadas y a que la mente humana no es capaz de elaborar comparaciones aleatorias entre pacientes.
2. Inaccesibilidad de las fuentes documentales y/o falta de destreza en el análisis de la información, lo que impide incorporar el conocimiento científico válido en la práctica clínica.
3. Presencia de información incorrecta o tendenciosa. El afán de comunicación provoca, a pesar del cuidado de los gestores de sistemas de documentación científica, que información no válida ni fiable inunde los medios de comunicación profesionales, llenándolos de “ruido” que confunde, desorienta e induce a error en la toma de decisiones.
4. Introducción de prácticas no contrastadas. En la práctica clínica también hay modas, inercias u otras situaciones que se mantienen o propagan por intereses espurios con gran éxito. El paciente que visite varios médicos puede escuchar opiniones diferentes y, por supuesto, recibir diferentes tratamientos. Como mal menor, estas prácticas pueden implicar mayor carga de ineficiencia en el sistema, más gasto y menos credibilidad, pero en ocasiones ponen en riesgo la salud del paciente.
5. Existencia de “zonas oscuras” en la utilización clínica del conocimiento científico disponible. Aun disponiendo de información de calidad, quedan

preguntas pendientes de resolver sobre cuánta evidencia y de qué tipo es necesaria para tomar decisiones, como por ejemplo:

- ¿Qué puntos finales se deben considerar para valorar efectividad en un ensayo clínico? (¿sólo se deben evaluar resultados finales en términos de salud global, como la supervivencia, la calidad de vida, etc.?, ¿sirven los resultados intermedios, como la normalización de las enzimas séricas en el caso de un infarto, o la reducción tumoral en una prueba diagnóstica?).
 - ¿Todos los estudios que analicen estos resultados tienen que ser estudios experimentales controlados?, ¿sirven los diseños observacionales prospectivos o retrospectivos?
 - ¿Cuántos estudios se necesitan para confirmar el beneficio de una intervención?, ¿basta un solo ensayo clínico para justificar una recomendación a todas las personas posibles?
 - ¿Quién tiene la autoridad para sentar la indicación y hacer recomendaciones?
 - ¿Son los resultados disponibles, aplicables y aceptables en el contexto particular de un paciente concreto?
 - ¿Qué papel tiene el coste económico en la toma de decisiones?
6. Pérdida de la actualización científica: la información médica se produce y renueva permanentemente de una manera inabarcable, influyendo de manera errática o no validada en la práctica. Es difícil el acceso y seguimiento de toda esta información, de forma que el grado de actualización del profesional decae con el tiempo y la formación médica continuada clásica no subsana este déficit (14).

Todas estas cuestiones permanecen abiertas, en mayor o menor medida, y es probable que los médicos tengamos que seguir tomando decisiones sin disponer de todas las respuestas en cada caso, en situaciones complejas, bajo circunstancias difíciles y con muy poco apoyo. Sin embargo, la medicina actual no es una práctica aleatoria o fraudulenta. Aunque no está exenta de riesgos, no puede negarse que el conjunto de la práctica médica es eficaz. Pero lo cierto es que con frecuencia

resulta difícil de dilucidar si el paciente no podría haberse beneficiado de una mejor opción en su atención.

Para complicar este panorama, el médico debe, además, enfrentarse e integrar en su decisión a otras variables alejadas de la biología y del core de sus conocimientos profesionales, como son las creencias, expectativas, necesidades y preferencias del paciente, o los condicionantes socio-sanitarios del sistema o país en el que practica.

Para la mayoría de la población el proceso de toma de decisiones médicas es una caja negra. Los pacientes no son capaces de juzgar directamente las decisiones de los médicos y esto otorga a la profesión médica una condición social y jurídica especial (15). En general, la gente espera que el tratamiento que reciben haya sido indicado por médicos que se adhieren a normas profesionales estrictas. La práctica médica -en contraste con otras prácticas sociales- goza del crédito de estar basada en argumentos sólidos y científicos, y de ser defensora y garante de la preciada salud de la población con conocimientos cada vez más sofisticados (16).

La existencia de variabilidad injustificada en la práctica médica desafía esa creencia general. Si los tratamientos se basaran en el conocimiento teórico y en la condición médica del paciente, los enfermos con condiciones similares recibirían siempre el mismo tratamiento, independientemente del médico, hospital o consulta al que asistieran. Sin embargo, hay numerosos ejemplos de la literatura que muestran una situación diferente.

Iniciado el siglo XXI, y en puertas de la nueva era de la medicina personalizada, sigue en pie el reto de reducir la incertidumbre clínica y mejorar los mecanismos de la toma de decisiones por parte de los médicos. ¿Es la capacitación personal para la práctica autónoma de una “medicina basada en la evidencia” la vía que garantizará estos cambios?. Sin descartarlo, no se deben ignorar las dificultades y limitaciones que tal propuesta conlleva, ya identificadas desde el mismo inicio de este movimiento (17).

1.2. Sobre algunos límites, equívocos y dificultades para la práctica de una Medicina Basada en la Evidencia.

Hace más de 20 años, en noviembre de 1992, se publicó en la revista JAMA el artículo "*Evidence-based medicine. A new approach to teaching the practice of medicine*" (18). En este provocador escrito, el *Evidence-Based Medicine Working Group* proponía un cambio en el modelo o paradigma para el aprendizaje y la práctica de la medicina, y formulaba el ideario de la llamada *Evidence Based Medicine (EBM)*. Aunque un año antes Guyatt, presidente del grupo, ya había usado el concepto de EBM en un editorial en la revista *ACP Journal Club* (19), el artículo de JAMA fue la presentación pública de un movimiento iniciado en la década de los ochenta en la universidad canadiense de McMaster (Ontario).

Durante sus primeros años de difusión, la expresión *Evidence Based Medicine*, se empleó sin una definición explícita del concepto, lo que probablemente contribuyó a interpretaciones dispares sobre el significado y alcance de la propuesta. No fue hasta 1996 cuando Sackett, su más reconocido promotor, definió expresamente la EBM como el uso consciente, explícito y sensato de las mejores pruebas disponibles en la toma de decisiones sobre el cuidado del paciente individual, en su contexto concreto (20). Apenas una década después de aquellos primeros pasos la EBM había alcanzado una gran difusión mundial en todos los sectores sanitarios, revelándose como un auténtico fenómeno social, incluso fuera del ámbito profesional (21).

Hoy se considera este enfoque de la asistencia sanitaria como una condición inexcusable para garantizar la calidad del ejercicio clínico. Su componente principal consiste en el aprovechamiento y uso de la literatura científica como base para la toma de decisiones relativas a la práctica, siempre que a esta información se incorpore la experiencia clínica, una formación básica sólida (conocimiento de la fisiopatología y de los mecanismos de acción) y la consideración de la situación particular de cada paciente (22).

Sin embargo, el desarrollo y formalización de esta visión sobre la EBM no fue inmediata, ni estuvieron exentos de polémica. Como suele ocurrir al inicio de cualquier nuevo movimiento en ciencia, pronto se destacaron grupos de

promotores entusiastas que deseaban cambiar lo antes posible la práctica estándar y los comportamientos previos. Algunos de estos defensores del nuevo modelo se posicionaron entonces en una visión simplista de la realidad que no admitía términos medios y que no se matizó hasta que la propuesta fue incorporada al uso rutinario y se contrastó con la experiencia real. En esos primeros pasos de la historia de la EBM los editores de *The Lancet* llegaron a acusar a sus valedores como defensores de una postura subversiva, intolerante y poco inteligente, identificando en el movimiento ciertas similitudes con los cultos fundamentalistas (23).

La controversia entre los defensores fanáticos de una EBM entendida como un nuevo paradigma exclusivo en la forma de hacer medicina y docencia clínica (18,24) (no basta la experiencia, sino que es obligatorio disponer de “evidencia científica” para tomar decisiones clínicas acertadas), y su contrapartida de detractores, defensores de una “medicina tradicional”, que juzgaban tal pretensión exagerada, promotora de una práctica deshumanizada (25,26) e ignorante del valor de la experiencia clínica y de los expertos (repetir tratamientos que han sido útiles a otros enfermos, en la experiencia propia o de otros más antiguos o más experimentados, puede beneficiar a los próximos pacientes similares y es una base suficientemente sólida para la toma de decisiones), no se ha extinguido completamente aun hoy.

Ambas posturas resultan igualmente extremas. Probablemente, la pretensión de que la EBM sea considerada como un cambio de paradigma en la medicina resulta excesivamente presuntuosa (27). El propio Sackett, sitúa los orígenes filosóficos de la EBM en la Escuela de París, a mitad del siglo XIX: “... *más que constituir una revolución, parece que la medicina basada en la evidencia no es sino un ejemplo de un nuevo avance y profesionalización del paradigma actual...*” (28). Sin embargo, rechazar que la EBM sea un nuevo paradigma, no disminuye en absoluto la importancia y el impacto que este fenómeno está teniendo en la enseñanza y en la práctica de la medicina occidental. Si las decisiones médicas sólo se basaran en la experiencia previa quizá se seguirían aplicando sangrías a los pacientes con neumonía, como preconizó Louis en 1835 (29), o lidocaína profiláctica a los pacientes con infarto al miocardio (30). De esta forma se mantendrían en uso

procedimientos obsoletos, mientras que el enorme volumen de la investigación clínica actual sería un esfuerzo inútil.

Hoy se acepta que más que un cambio de paradigma científico en la clínica (como sí ocurrió en la Epidemiología al superarse el “paradigma determinístico” decimonónico -toda causa tiene un factor determinante- por el actual “paradigma de relación probabilística” -existe una red causal de factores de riesgo que se relacionan en mayor o menor intensidad con un efecto-), estamos asistiendo a la transición de modelos de ejercicio: del modelo clínico tradicional hacia un modelo de ejercicio sustentado en la evidencia científica.

Algunas de las asunciones del modelo tradicional sobre el conocimiento necesario para guiar la práctica clínica que han sido puestas en tela de juicio desde la EBM son (13):

1. Las observaciones no sistematizadas de la experiencia clínica son válidas y suficientes para construir y mantener el conocimiento del médico acerca del pronóstico del paciente, del valor de las pruebas diagnósticas y de la eficacia del tratamiento.
2. La comprensión de los mecanismos básicos y fisiopatológicos de la enfermedad resulta suficiente para validar decisiones clínicas.
3. La enseñanza médica tradicional y el sentido común capacitan suficientemente para entender y evaluar la información científica sobre tratamientos, pruebas y procedimientos.
4. La experiencia y maestría individuales son base suficiente sobre la que generar guías y recomendaciones válidas para la práctica clínica.

En contraposición, el nuevo modelo de Medicina Basada en la Evidencia propone asunciones radicalmente distintas (31):

1. La experiencia clínica y el desarrollo del instinto clínico son cruciales y necesarios para ser un médico competente, pero no suficientes. En ausencia de información recogida con orden, rigor y comprobación científica, se ha de ser cauto en interpretar la información derivada de la experiencia personal y la intuición, que pueden resultar erróneas.

2. El conocimiento de los mecanismos básicos de la enfermedad son necesarios pero no suficientes para guiar la práctica clínica. El razonamiento para el diagnóstico y el tratamiento derivados exclusivamente de mecanismos fisiopatológicos a menudo es incorrecto y da lugar a predicciones inexactas sobre el funcionamiento de las pruebas diagnósticas o sobre la eficacia de los tratamientos.
3. El entendimiento y dominio de ciertos principios, métodos y reglas de comprobación científica resultan necesarios para interpretar correctamente la información científica disponible sobre causalidad, pronóstico, pruebas diagnósticas y estrategias de tratamiento.

Pese a la general aceptación de estas premisas, no han faltado opiniones (en tono más suave, pero con una carga de mayor profundidad que las de los editores de *The Lancet*) que cuestionan si el loable objetivo de tomar decisiones clínicas basadas en la evidencia es alcanzable. Algunas voces de tanto peso como las de Feinstein y Horowitz (32) han señalado algunas limitaciones incuestionables del modelo clínico EBM:

- La primera y quizá más importante, es que cada paciente conlleva una situación particular, unas necesidades propias, unos valores y unas preferencias que pueden condicionar la aplicabilidad y efectos esperables de la evidencia disponible.
- Por otra parte, resulta obvio que la verdad hoy aceptada puede convertirse fácilmente en información obsoleta o inapropiada mañana.
- Además, aunque se disponga de firmes evidencias (provenientes de ensayos clínicos) se debe ser cauto al extrapolar conclusiones procedentes de entornos experimentales a la práctica rutinaria. En esta última, circunstancias como la comorbilidad, las posibles interacciones entre tratamientos concomitantes, la tolerabilidad de los fármacos, su influencia sobre la adherencia terapéutica, o la mayor autonomía del paciente durante su seguimiento clínico, pueden determinar unos resultados terapéuticos radicalmente distintos de los experimentales.
- Finalmente, si se pretendiese fundamentar todo el trabajo clínico sobre evidencias científicas de alta calidad, la gran mayoría de los pacientes

quedarían sin tratamiento, ya que no existen tales pruebas respecto a muchas de las decisiones clínicas que se toman rutinariamente, ni probablemente se dispondrá nunca de ellas.

La ausencia de evaluación de muchas prácticas se debe a veces a restricciones éticas que impiden evaluar actuaciones no comprobadas experimentalmente, pero avaladas por el sentido común etiopatogénico y fisiopatológico (p. ejem., administrar líquidos a lactantes con diarrea,...), cuya no aplicación supondría un riesgo inaceptable. En otras ocasiones son los organismos de financiación (agencias estatales, industria farmacéutica,...) quienes no se interesan en financiar estudios a gran escala sobre cuestiones que permanecen como “áreas grises” de la práctica que no han sido evaluadas concienzudamente (p. ejem., el efecto del uso de antibióticos en ciertas infecciones, ...).

Entender la EBM como una práctica basada en las mejores pruebas disponibles, desafortunadamente no siempre de alta calidad (esto es, sin disponer en ocasiones de una “evidencia” cierta), es la clave para acabar con la polémica estéril entre el beneficio indudable y los riesgos potenciales de esta corriente. En el ámbito hispanohablante, parte de este problema proviene de la inapropiada traducción del término inglés *evidence*. Aunque no pueda considerarse un *faux-ami* lingüístico en sentido estricto (33), el uso del término “evidencia” por *evidence* es impreciso ya que este posee un campo semántico más amplio que aquel, aunque compartan una obvia analogía fonética. El sustantivo español que mejor recoge el significado de *evidence* es “prueba”, bien entendido que una prueba puede tener distinto valor esclarecedor sobre aquello que pretende explicar, desde lo meramente indiciario a lo rigurosamente evidente. Debería reservarse, por tanto, “evidencia” al caso particular de la prueba segura que conlleva una certeza plena, de la que no se puede dudar ni requiere demostración adicional. El Diccionario Panhispánico de Dudas (34), aclara y alerta sobre el error de uso de este anglicismo en el mundo jurídico, desde donde parece haber pasado de forma generalizada al entorno clínico:

Evidencia. ‘Certeza clara y manifiesta de la verdad o realidad de algo’. (...) Además de designar la cualidad abstracta, este sustantivo tiene el significado concreto de ‘cosa evidente, esto es, clara y manifiesta’. (...) Esto

no justifica el empleo indiscriminado en español de la voz *evidencia* como sinónimo de *prueba* o *indicio*, calco censurable del inglés *evidence*: en inglés, *evidence* es toda prueba (circunstancial, testimonial, material, documental, etc.) que se alega en un proceso judicial; en español, solo sería aceptable como sinónimo de *prueba evidente*, esto es, prueba clara y manifiesta; así, no resultan apropiados usos como los siguientes: «*Las evidencias que se han aportado no parecen en todo caso muy convincentes*» (Ninyoles *Idiomas* [Esp. 1977]).

Desafortunadamente no se ha aceptado y extendido el uso general en castellano del término “medicina basada en pruebas” (**MBP**), sugerido con precisión algunos expertos (35-38), como la denominación que delimita mejor las pretensiones de la EBM. En reconocimiento a esta propuesta, este será el término empleado en lo sucesivo en esta tesis.

Resumiendo, practicar una MBP es integrar la competencia o maestría individual, es decir la experiencia clínica, con las mejores pruebas disponibles procedentes de investigación válida y fiable o, lo que es lo mismo, con los últimos hechos científicamente probados. Ambos aspectos, experiencia personal y aplicación de las "evidencias", deben ir unidos. Sin el conocimiento y la aplicación de las últimas pruebas disponibles el ejercicio de la medicina queda rápidamente desfasado y pueden conllevar perjuicio para el paciente. Pero sin la experiencia clínica la medicina puede convertirse en un ejercicio automático no menos peligroso. Una maniobra diagnóstica o terapéutica indicada acríticamente, aunque sea de utilidad científicamente comprobada, puede resultar inapropiada o contraproducente para un paciente concreto.

En cualquier caso, más allá de las dificultades semánticas y epistemológicas, quizá el principal inconveniente intrínseco para generalizar la práctica de una MBP sean las carencias formativas y las dificultades prácticas que encuentra el médico interesado al intentar aplicar las mejores pruebas disponibles a sus decisiones clínicas. Definir adecuadamente las preguntas clínicas inherentes a cada decisión, identificar y recopilar las pruebas accesibles en la literatura de una forma eficiente, evaluar críticamente y ponderar el valor dicho material (lectura y análisis de los estudios localizados), integrar la información con los factores relativos al paciente

y la experiencia del clínico para tomar y poner en práctica la decisión clínica, asegurando su aplicabilidad e impacto, son tareas que exigen una capacitación no fácil de conseguir. Disponer de tiempo y oportunidades de aprendizaje para garantizar las habilidades necesarias de manejo de las bases de datos bibliográficas, de interpretación estadística y de las bases de la epidemiología clínica puede resultar un deseo inaccesible para muchos clínicos.

Se ha apuntado, no sin sarcasmo, que la aplicación sistemática y generalizada de una MBP en nuestro sistema sanitario produciría un movimiento migratorio masivo de los médicos desde la cabecera del enfermo al teclado del ordenador y a la biblioteca (39). En esta línea, es común encontrar entre los médicos ajetreados por la tarea clínica absorbente la opinión de que la MBP, idealmente formulada, no deja de ser una propuesta retórica e ingenua de expertos relacionados con la gestión de la bibliografía biomédica, desconocedores de las circunstancias de la práctica clínica real y perpetuadores de una propuesta que se autojustifica más allá de sus resultados prácticos. No parece una opinión irreflexiva si se tiene en cuenta la baja proporción de cuidados sanitarios que se basan en “evidencias” sea cual sea el entorno en el que este fenómeno haya sido analizado.

Aunque la situación no resulte tan pesimista como la descrita por White en los setenta (apenas el 10%) (40), una guía colaborativa reciente coordinada por Booth ha confirmado datos decepcionantes sobre 18 estudios que analizan la cuestión (41). Entre un cuarto y un tercio de los tratamientos que reciben los pacientes asegurados en Estados Unidos resultan inapropiados (por una mala relación beneficio – riesgo) o dudosos (o por no haber demostrado efectividad) (42).

Así las cosas, si la MBP se considera una vía imprescindible para mejorar la calidad de la asistencia a la población, se impone una revisión de su formulación práctica para convertirla en una estrategia viable, accesible y efectiva.

Por otra parte, quizá la crítica más sustancial a la MBP es que su apariencia técnica irrefutable puede haber llevado a la profesión a un desarrollo que ha ignorado la cuestión de cómo debe ser el médico en el siglo XXI, y de qué necesidades debe satisfacer al hombre enfermo, a su entorno y a la sociedad en general.

El siglo pasado, espoleado en sus últimas décadas por la MBP, contempló un increíble desarrollo de la capacidad técnica de la Medicina para prevenir y solventar enfermedades. Tal proceso se acompañó de una necesaria superespecialización del médico que le obligó a profundizar en disciplinas de extensión cada vez más restringida. Sólo en un limitado ámbito especializado se puede conocer y dominar la ingente “evidencia” científica disponible, exponencialmente creciente.

Los beneficios derivados de este proceso han sido incuestionables en respuesta a los problemas de salud más complejos. Sin embargo, la limitación progresiva de la amplitud de la responsabilidad clínica de los especialistas médicos también ha inducido algunos efectos indeseables en el ejercicio de la medicina actual:

- una parcelación de trabajo clínico que parece dirigirse a tratar fragmentariamente órganos enfermos, síntomas concretos o enfermedades específicas, en detrimento de una atención médica integral que entienda al paciente como un ser humano único, y que intente comprender todas sus necesidades como persona enferma.
- una clasificación simplificadora y reduccionista del enfermar humano en base a la investigación disponible (biológico versus psicológico), obviando que la enfermedad influye en todas las dimensiones en las que se despliega una biografía personal (física, mental, emocional, espiritual, familiar, social, laboral, ...).
- un descuido del rol de apoyo y referencia que el médico debe representar para el paciente en el complejo trance de la enfermedad, con clara insuficiencia de las habilidades relacionales y comunicativas que tal papel requiere, y con abandono de valores y actitudes profesionales imprescindibles para ofrecer cuidados, además de científicamente fundados, humanamente excelentes (confiables, benevolentes, compasivos, abnegados, intelectualmente humildes, justos).
- una cierta creencia irracional en las posibilidades ilimitadas del conocimiento científico para solventar problemas de salud, lo que provoca la decepción de médicos y pacientes cuando la técnica claudica y no puede ofrecer respuestas demostradas. Son frecuentes entonces actitudes de

inseguridad en el profesional y de desesperación en el enfermo particularmente en la atención a pacientes atípicos, pluripatológicos y complejos.

En la última década, la propia profesión, en reacción al riesgo de derivar la práctica clínica hacia una MBP deshumanizada, ha impulsado iniciativas profesionales, de base bioética o legal, que basadas en el derecho del paciente a la información y a que se respete su autonomía ("*Patient Centred Medicine*") (43), intentan reorientar la práctica clínica hacia la toma de decisiones compartidas entre médicos y pacientes en los aspectos relevantes para los últimos, a la vez que intentan dar una respuesta integral a las todas las necesidades de una salud ampliamente entendida, esto es, con dimensión bio-psico-social-espiritual ("*Person Centred Medicine*") (44,45). Y para ello se promueve que se considere lo que las pruebas científicas disponibles sugieran en el marco modulador de la experiencia personal del médico y del criterio y preferencias de su paciente.

Ello es necesario porque el determinante principal de la naturaleza de la relación entre médicos y pacientes es el estilo de práctica clínica que impone el médico en la consulta: "doctor-centred" o "patient-centred". En el modelo *doctor-centred* un médico dominante (autoritario o paternalista) toma las decisiones que cree oportunas en interés del paciente. En el estilo *patient-centred* hay menos control del acto clínico por parte de un médico que anima a la participación del paciente a tomar decisiones compartidas en un marco de colaboración y respeto mutuo. Los estudios muestran que los médicos que practican una "medicina centrada en el paciente" tienden a ser más flexibles, tienen mayor capacidad de respuesta a las necesidades de los enfermos y consiguen resultados clínicos más favorables.

Si no se centra en el paciente, la práctica médica corre el riesgo de convertirse en un ejercicio puramente intelectual, en lugar de una tarea de servicio profundamente humano y moral. Y esta amenaza acecha particularmente a quienes justifican estrictamente su actuación profesional sólo en las pruebas científicas que puedan legitimarla. Esta reflexión tiene implicaciones considerables en el profesionalismo del futuro y en tipo de atención que recibirán los enfermos en las próximas décadas.

1.3. Reivindicación del criterio experto: el papel de la experiencia en la toma de decisiones clínicas.

En el contexto de limitación de auténticas evidencias para la práctica sistemática de una MBP, las recomendaciones basadas en el criterio de los expertos siguen siendo una ayuda estimable para tomar decisiones en la práctica rutinaria, en particular cuando el clínico afronta escenarios de incertidumbre por ausencia de pruebas sólidas, por controversia científica entre distintas pruebas disponibles, o por la limitada posibilidad de aplicar ciertas evidencias en un contexto clínico concreto (ancianidad, comorbilidad, intolerancia medicamentosa, riesgo de interacciones farmacológicas, situaciones de baja adherencia a recomendaciones, ...).

Precisamente la práctica de una genuina “medicina basada en pruebas”, en la concepción original de Sacket (20), es la utilización consciente, explícita y juiciosa de las mejores pruebas que la investigación clínica ofrece, integrada con la maestría clínica del profesional (su experiencia) y las preferencias del paciente, al tomar decisiones clínicas sobre el cuidado de un sujeto concreto.

Durante la práctica habitual, aun los clínicos más experimentados afrontan con frecuencia situaciones sobre las que no disponen de conocimiento o experiencia suficientes. Tales ocasiones plantean dudas clínicas de cuya respuesta, acertada o no, puede depender el devenir del cuidado del paciente. Con frecuencia, muchas de estas cuestiones se resuelven consultando con compañeros expertos, ante la dificultad de acceder a la información científica pertinente en el punto de atención (o no existen las pruebas necesarias, o estas no resultan fáciles de localizar, interpretar o aplicar justo cuando se necesitan).

Sobre esta cuestión hay una amplia bibliografía disponible (46-58), entre la que destacan tres estudios:

Una investigación histórica, previa a la “era MBE”, identificó en la práctica clínica de médicos experimentados lagunas de información, que de solventarse podrían haber modificado sustancialmente sus decisiones clínicas, en la atención de dos de cada tres pacientes. Sólo el 30% de estas necesidades de información se

subsanan durante la visita del paciente, por lo general consultando con otro médico o profesional de la salud (56).

Coincidiendo con los primeros pasos de la MBE otra investigación basada en entrevistas y en observación directa de la asistencia habitual contabilizó el número de dudas que surgen a los médicos en su ejercicio. Los resultados dependieron del entorno asistencial, variando desde una cuestión relevante cada quince pacientes en atención primaria, a una duda cada uno o dos pacientes en ciertas especialidades (57).

Más recientemente, autores de nuestro entorno han analizado con rigor metodológico las necesidades de información de médicos de familia españoles y han descrito sus patrones de búsqueda de información (58). Se analizaron videograbaciones de la actividad asistencial de un gran número de consultas de AP aleatorizadas, para registrar las preguntas clínicas que surgían durante las visitas de pacientes y las fuentes de información utilizadas por el médico dentro de la consulta para responderlas. Los médicos que no consiguieron respuesta a sus dudas fueron seguidos posteriormente por teléfono para determinar si finalmente lo habían logrado y que fuentes de información utilizaron para ello. Sobre un total de 3.511 consultas grabadas (de 7,8 minutos promedio) se identificaron 635 preguntas clínicas (0,18 preguntas por consulta). Las preguntas más frecuentes estaban relacionados con el diagnóstico (53%) y el tratamiento (26%). Sólo el 22,8% de los médicos buscó respuesta a sus dudas (un 9,6% durante las propias consultas). El tiempo empleado y la tasa de éxito en lograr una respuesta durante una consulta y después de ella fueron de 2 minutos (100%) y 32 minutos (75%), respectivamente. Los autores concluyen que los médicos de este ámbito (en consultas de corta duración) solo tienen tiempo para responder 1 de cada 5 dudas clínicas sentidas. Los resultados del estudio ponen en cuestión si, en este medio, es realmente posible el ejercicio de una práctica clínica basada en la evidencia.

La reacción de los médicos que toman conciencia de sus dudas y de que necesitarían más información para poder tomar mejores decisiones clínicas depende de su capacidad de tolerar la incertidumbre. En contextos de alta presión asistencial, esta experiencia puede llegar a generar ansiedad, culpabilidad o, en su grado máximo, de bloqueo cognitivo. Otros sujetos desarrollan estrategias de

adaptación más defensivas como el ignorar estas necesidades de información o el “aparcar” el problema y aplazar la decisión hasta el momento en que se pueda iniciar un proceso de búsqueda para disponer de información. En este último caso, obviamente, la simple intención de buscar información no se sigue necesariamente de la acción de buscarla. Para muchos médicos, el esfuerzo que se requiere para elaborar preguntas clínicas apropiadas y buscar respuestas excede su capacidad por diversos motivos: tiempo escaso para leer literatura científica, dificultades para comprender inglés técnico, poca habilidad informática, desconocimiento de las fuentes bibliográficas más adecuadas para cada necesidad de información, falta de acceso a internet en el puesto de trabajo, poca confianza en su capacidad de lectura crítica,...). Por alguno de ellos, o por todos estos problemas, lo más probable es que las necesidades de información inmediatas no sean finalmente cubiertas y, prácticamente nunca, en tiempo real y en el punto de atención sanitaria donde las carencias informativas surgieron.

En respuesta a este contexto real de limitaciones y dificultades para resolver en la bibliografía las dudas afrontadas en la práctica diaria, cada médico en ejercicio suele incluir en su red personal de contactos profesionales a una serie de colegas que considera disponen de criterio experto en materias que no son de su campo de especialización, y a quienes acude con confianza en caso de necesidad profesional (o personal) (56). En un marco más general, esta situación suele reproducirse dentro de los equipos facultativos de cada centro sanitario, donde la mayoría suele compartir el reconocimiento tácito a ciertos especialistas de referencia, por su particular capacidad de juicio y resolución de situaciones complejas. Estos clínicos son considerados por el resto como expertos en función de su dilatada experiencia o a su particular subespecialización.

El funcionamiento espontáneo de estas redes informales de conocimiento optimiza la gestión conjunta de la experiencia acumulada entre los distintos equipos clínicos de cada centro y promueven una práctica diaria más eficaz y segura. En sentido estricto, esta forma de trabajar está plenamente alineada con el imperativo ético y deontológico de “humildad intelectual” que obliga al médico a solicitar ayuda a compañeros más competentes en situaciones que exceden su capacidad personal (59).

Aunque la mayoría de los clínicos aceptan y promueven de facto, en su funcionamiento diario, esta dinámica de intercambio de información y asistencia mutua, se da la paradoja que bajo el paradigma de la “medicina basada en la evidencia” el criterio del experto ocupa el más bajo escalón de confianza en la calidad de la evidencia en la que se puede soportar una recomendación clínica. Esta fuente de información resultaría menos fiable y las recomendaciones derivadas de ella menos robustas que la procedente de cualquier otro tipo estudio de tipo observacional o, por supuesto, de diseño experimental. Siendo esto cierto, durante algún tiempo, los partidarios más estrictos de la MBE demonizaron el uso de este recurso como aval para la toma de decisiones por dudoso, cambiante y potencialmente manipulable (60).

Es indiscutible que es imposible valorar la motivación última y el peso de las intenciones en la opinión de un experto y, por tanto, caben dudas sobre la confianza que debe prestarse a la misma. Este problema es el mismo que cuestionó y, en buena medida desprestigió, el valor de los artículos de revisión narrativa y no sistematizada sobre el *state of the art* en el abordaje de patologías que en otra época se publicaban frecuentemente por reconocidos expertos en las revistas médicas de prestigio. Estos estudios, en general de encargo, se difundían y se usaban como una fuente de actualización y puesta al día aceptada por la mayor parte de la comunidad científica. Hoy, las publicaciones médicas más reconocidas exigen que este tipo de aportaciones tengan una metodología sistemática y transparente de revisión y evaluación de la bibliografía disponible sobre el tópico, o incluso que apliquen un sistema objetivo de análisis de la información en función del tratamiento estadístico de los resultados rescatables en los estudios previos de calidad (metaanálisis). Pese a ello, no faltan controversias sobre estas consideraciones (61), ni ejemplos donde “el hombre superó a la máquina”, resultando fallida una técnica estructurada de revisión sistemática (62) frente al trabajo narrativo de expertos (63).

En cualquier caso, y en defensa de “criterio experto”, debe aclararse la diferencia sustancial entre el crédito que puede merecer la opinión de un solo experto sobre un tema de su conocimiento y el que debe otorgarse al resultado de un procedimiento riguroso y contrastable de consenso entre un grupo de expertos

acreditados de tamaño suficiente. Cuando las conclusiones expertas se derivan de la interacción transparente y controlada de distintos profesionales que confrontan sus criterios y debaten abiertamente sobre los puntos de discordancia en los que no están preliminarmente de acuerdo, la confianza en las conclusiones de este proceso aumenta de forma incuestionable.

Y esto es así porque en un proceso de debate y negociación abierta sobre cuestiones controvertidas entre un grupo de expertos, cada uno de ellos solo puede argumentar con legitimidad desde el conocimiento o la experiencia personal como aval de sus opiniones. Otros elementos que puedan sesgar o manipular el criterio defendido, aun de forma inconsciente, son difíciles de mantener frente a la opinión divergente de otros profesionales que argumentan, igualmente, desde su propio conocimiento experto o su dilatada experiencia.

Se ha propuesto el símil de un iceberg para explicar la efectividad de los procesos de consenso grupal para alcanzar conclusiones válidas: aunque los determinantes de la opinión de cada experto fuesen mayoritariamente ocultos y poco defendibles (la parte “sumergida” del iceberg), la ocasión de debate igualitario e imperativo con otros colegas condiciona que cada uno de ellos ponga en juego solo el sobrenadante visible que puede ser defendido de las críticas con argumentos de rigor científico-técnico. En circunstancias normales, las propias dinámicas grupales, en las que quedan expuestos la capacidad de juicio, la honradez intelectual y el propio prestigio profesional de cada participante, mantienen sumergidos e inactivos bajo la superficie del debate los argumentos espurios o los intereses personales. Es difícil mantener posturas irrenunciables si no se pueden aportar elementos racionales que las avalen.

Por ello, aunque las taxonomías científicas localizan al consenso de expertos en el escalón inferior del rango de calidad de las pruebas científicas que pueden avalar una recomendación clínica, un consenso bien desarrollado puede resultar la única vía para proponer al colectivo clínico un criterio claro y unificador que promueva la reducción de la variabilidad injustificada de la práctica clínica. Allí donde no se disponga de datos experimentales o derivados de una investigación observacional bien estructurada, el clínico puede acudir legítimamente a la información

procedente de los expertos, que integran su conocimiento teórico, con la intuición y el buen juicio entrenados en su mayor experiencia.

Cada experto participante en un consenso ofrece al grupo su pericia clínica (el *expertise* anglosajón), esto es, su capacidad para establecer un buen juicio clínico, entendido como un proceso de toma de decisiones en ausencia de leyes o reglas explícitas. El poder deductivo derivado de la pericia clínica no se fundamenta solo en sus años de experiencia, sino también en la observación y el estudio continuados, el rigor en la obtención de datos y el sentido común (64).

Disponer de un sistema que sintetice y ordene la pericia de los expertos referentes en quienes se deposita generalizadamente confianza profesional, puede iluminar eficazmente los múltiples ámbitos de incertidumbre que el clínico afronta en su práctica clínica.

1.4. Conceptos generales sobre el consenso profesional y comparación crítica de los principales métodos.

Consensuar es, según la RAE, “adoptar una decisión de común acuerdo entre dos o más partes”. Tal empeño representa el contrapunto conceptual a la opinión individual y unilateral de un solo líder. Aunque el consenso ha sido ampliamente usado en actividades sociales y profesionales desde antiguo (consejos municipales, comunidades de regantes, ...), alcanzar acuerdos o decisiones en grupo puede resultar complejo debido a la diferencia de conocimientos de sus miembros, a la fuerza de convicción de las argumentos aportados o incluso a los intereses particulares (65).

Procedentes de las ciencias sociales y diseñados como procedimientos de negociación (66), desde hace más de 40 años se emplean técnicas de trabajo en grupo en el mundo sanitario (67), con fin de disminuir en lo posible los sesgos o influencias indeseadas en las conclusiones que de ellos se extraigan.

Estas técnicas comenzaron a utilizarse como herramientas de ayuda a la toma de decisión en la planificación y gestión de instituciones o empresas, involucrando a expertos en la predicción de tendencias futuras de escenarios inciertos. Con una finalidad similar, pronto pasaron al ámbito de la salud donde se han usado

profusamente para la realización de predicciones con fines prospectivos, para la detección de necesidades de salud en la comunidad y para el establecimiento de prioridades de actuación o de investigación (68).

Una aplicación más reciente, donde más controversia se ha suscitado sobre su validez y utilidad (69), ha sido el uso de estas técnicas en el ámbito clínico-asistencial, para construir pautas terapéuticas o diagnósticas, en especial, para orientar el manejo de nuevas tecnologías.

En estos casos se recurre al criterio experto cuando no hay suficiente información sobre cómo orientar las actuaciones clínicas para alcanzar la mayor efectividad clínica, o cuando, por el contrario, la información resulta excesiva y no necesariamente de la mejor calidad. Aunque esto ocurre especialmente cuando se necesita responder a cuestiones muy concretas o relativamente novedosas, en la práctica, como se ha comentado, hay muy pocas áreas de la atención sanitaria con suficiente evidencia de buena calidad, basada en una investigación bien diseñada y rigurosa. En tal circunstancia, los métodos de consenso permitirían desarrollar recomendaciones tomando en cuenta la opinión de los expertos y la experiencia de los clínicos.

Esta pretensión se fundamenta en la presunción básica que impulsó el desarrollo de las técnicas grupales, esto es: que el resultado del comportamiento conjunto de un grupo de individuos es superior, bajo determinadas condiciones, a la suma de los resultados individuales de los miembros que lo componen. La veracidad de tal presunción está suficientemente contrastada, desde hace tiempo, en el campo de la psicología social (70, 71).

Pese a ello, PubMed, el repositorio bibliográfico mundialmente más utilizado en biomedicina, mantenido por el Centro Nacional de Información Biotecnológica (NCBI) de la *US National Library of Medicine (National Institutes of Health)*, no incluyó el término “consenso” en su *Medical Subject Headings* (el tesoro de vocabulario controlado empleado para la indización de artículos) hasta 2003. En ese año admitió el descriptor **consenso** [*consensus*], como “acuerdo general u opinión colectiva; juicio que alcanza a la mayoría de los interesados”.

En el campo de la salud, se han empleado tres abordajes principales de consenso experto (72): en los sesenta se introdujo el método Delphi (73-74), seguido por la técnica de Grupo Nominal en los setenta (75) y, a partir de 1977, el método de la Conferencia de Consenso, desarrollado por el *NIH Consensus Development Program*, del *National Institute of Health* de los EEUU) (76).

Simplificando mucho y teniendo en cuenta la existencia de variantes de cada técnica, estos tres métodos de consenso principales (Método Delphi, Grupo Nominal y Conferencia de Consenso) utilizan, de forma sucesiva y siguiendo procedimientos altamente formalizados, técnicas de reflexión individual y estimulación de la creatividad, algún tipo de discusión estructurada y la combinación de juicios individuales para construir el criterio global de grupo (69).

Aunque se ha discutido al respecto (77), parece que las metodologías de consenso no deben ser consideradas técnicas de investigación cualitativa, al menos en la concepción estricta que de ellas se hace en sociología. A diferencia de estas (la observación participante, la entrevista abierta, las historia de vida y los grupos de discusión), los métodos de consenso utilizan herramientas cuantitativas (como son los cuestionarios y la cuantificación de resultados) y los investigadores intervienen con sus propias opiniones en la producción de las de los participantes. Frente a ellas, las técnicas cualitativas se basan en el análisis (lingüístico, semántico, antropológico, sociológico, hermenéutico,...) y la interpretación del lenguaje de los discursos de participantes por un moderador que no ofrece su opinión ni influye sobre las personas que participan en la investigación (su autoridad es simbólica y no reconocida previamente por los participantes).

Quizá las técnicas de consenso no admitan su categorización ni en la metodología epidemiológica ni en la cualitativa, sino que deben considerarse una vía particular con fines propios y estrategias específicas (alcanzar el acuerdo de un conjunto, que no grupo, de personas expertas en un tema científico-técnico, para acordar cuales son las mejores ideas o las mejores decisiones a tomar sobre un problema planteado). Más adelante, en la sección “Discusión” de esta tesis, se aporta un nuevo elemento de debate sobre esta cuestión, acercando las técnicas de consenso a los procedimientos psicométricos.

Las técnicas de consenso, aunque no emplean un diseño muestral estricto en cuanto a la selección de las personas que han de participar en el debate (no se eligen de forma aleatoria), ni en cuanto a la determinación del tamaño del grupo (no es necesariamente representativo de la comunidad de profesionales que podría opinar), sí que utilizan cuestionarios para la recogida de información y técnicas cuantitativas para la agregación estadística de datos y para inferir resultados.

La primera que se utilizó fue la **técnica Delphi**, creada en 1952 por la Rand Corporation, en los Estados Unidos (73). En su formulación original, es una peculiar técnica grupal en la que los participantes del grupo nunca se llegan a reunir. Es más, cada uno de ellos desconoce quien más participa con él, por lo que el grupo, en realidad, no existe. En su formulación original, se articula mediante sucesivas (tres o más) rondas de encuesta escrita vehiculada por vía postal. Con el análisis estadístico de las respuestas a cada cuestionario se construye la siguiente ronda, hasta llegar a una situación de consenso estable que permite elaborar un informe final de conclusiones. Más adelante se explica detalladamente el procedimiento y la variante técnica (método Delphi Modificado) que ha inspirado la técnica usada en los proyectos que sustentan esta tesis. Entre sus ventajas, destaca la comodidad de poder hacerse sin presencia física de los expertos y la posibilidad de poder acceder a un amplio número de ellos. A cambio, la elaboración y envío de cuestionarios puede resultar una tarea laboriosa, prolongada e incierta.

El **Grupo Nominal** (o Panel de Expertos) se desarrolló en 1968 como variante de la técnica Delphi (75). Su nombre procede de que los miembros del grupo, a diferencia del anonimato requerido en el Delphi, sí conocen los nombres de quienes están participando junto a él. Aunque se reúnen presencialmente bajo la coordinación de una persona con autoridad reconocida en el tema de interés, en un principio los participantes no pueden comunicarse libremente de forma verbal. La dinámica es sencilla: el moderador lanza una pregunta al grupo, sobre las que cada participante recoge en anotaciones escritas sus ideas personales. Después se exponen de forma exhaustiva todas las ideas generadas en el grupo, se discute sobre las mismas y cada miembro vota de forma individual y anónima sobre una escala cuantitativa previamente establecida (1 a 5, 0 a 100,...). Tras conocer los

resultados estadísticos obtenidos por cada propuesta, se debate de nuevo sobre los puntos más conflictivos y se plantea una puntuación final de las propuestas para identificar las consensuadas como las más votadas por el panel. Para ello se emplean medidas estadísticas previamente establecidas numéricamente para determinar la centralización de opiniones y el grado de acuerdo requerido.

El Grupo Nominal, adecuadamente conducido, gracias a la presencia de expertos y a las posibilidades de interacción directa entre ellos, permite profundizar más en las cuestiones que el Delphi y es relativamente más rápido que este. Por el contrario, no resulta inmune a sesgos asociados con la presencia de personas con más conocimientos o mayor liderazgo durante la reunión (78).

Una variante menos estructurada del método anterior es la “**tormenta de ideas**” (*brain storming*). En esta técnica, el líder estimula a los componentes del grupo presentando sus propias ideas al inicio. Tras ello, las ideas libremente emergidas del grupo son el material de partida para desarrollar un ejercicio de reestructuración (redefinición, división, síntesis,...) que produce una lista abreviada sobre la que se eligen por puntuación las mejores ideas (79).

Las **Conferencias de Consenso** se diseñaron en respuesta a la complejidad creciente de las cuestiones a resolver planteadas por el mundo sanitario, en el National Institute of Health, de Estados Unidos, en 1977 (76). Desde entonces se han analizado allí más de 150 temas con esta metodología hasta la actualidad. Información detallada sobre el *NIH Consensus Development Program* y los informes técnicos de conclusiones de las conferencias están disponibles de forma abierta en la web <http://prevention.nih.gov/cdp/about.aspx>.

Desde allí la metodología se ha difundido para la elaboración de Guías de Práctica Clínica en otros países, como Francia donde se celebran conferencias de este tipo por la *Agence pour l'Accréditation et l'Evaluation en Santé* (ANAES) desde 1993.

En España, con mayor o menor rigor metodológico, también existe una larga experiencia de uso del método promovido, con distintos objetivos, por instituciones privadas, como la Fundación Sira Carrasco para Ayuda a la Fibrosis Quística (tratamiento antibiótico intravenoso domiciliario, 2003); públicas, como la Agencia de Evaluación de Tecnologías Sanitarias de Andalucía (hemofilia A,

2005; climaterio, 2007); o por sociedades científicas, como la Sociedad Española de Contracepción (contracepción hormonal oral, 1997; anticoncepción intrauterina, 2001), la Sociedad Española de Patología del Aparato Respiratorio (EPOC, 2002), la Asociación Española de Vacunología (gripe, 2005), la Sociedad Española de Cardiología (trasplante cardiaco, 2007), la Asociación Española de Pediatría y 10 sociedades pediátricas más (bronquiolitis, 2010), la Sociedad Española de Medicina Intensiva, Crítica y Unidades Coronarias (pancreatitis aguda, 2012), la Sociedad Española de Medicina Interna y cuatro sociedades más (anciano diabético, 2012), o la más reciente de la Asociación Española de Gastroenterología y la Sociedad Española de Patología Digestiva (*Helicobacter pylori*, 2013).

El objetivo de las Conferencias de Consenso es elaborar recomendaciones para la práctica clínica en una sesión pública donde los expertos presentan la mejor evidencia disponible para responder a las preguntas clínicas planteadas, con base en un análisis crítico previo de la bibliografía. El método remeda el proceso de una vista oral en el ámbito jurídico. Durante la sesión los expertos defienden las conclusiones de las revisiones sistemáticas previamente encargadas a un grupo técnico de revisión bibliográfica e interactúan con el público invitado a la conferencia (expertos, clínicos, investigadores, representantes de pacientes y otros colectivos involucrados o con interés en el tema a debate). Todos ellos pueden expresar opiniones en presencia de un jurado multidisciplinar y multiprofesional que, tras escrutar datos, escuchar opiniones y clarificar los aspectos necesarios, elabora un informe de conclusiones y recomendaciones en una sesión de puesta en común a puerta cerrada. Para ello, se requiere el acuerdo consensuado de al menos el 80% de los 8-16 miembros que componen el dicho jurado, eso es, el acuerdo de una mayoría absoluta cualificada (reforzada) del grupo (80).

La conferencia de Consenso es un sistema válido para elaborar recomendaciones de práctica clínica fundadas en bibliografía científica de calidad bien revisada, con la ventaja añadida de la participación de los expertos en el propio proceso de elaboración, lo puede resultar determinante para que sus conclusiones sean aceptadas y seguidas por los clínicos. No obstante, no faltan críticas (81-84) a los riesgos de esta metodología, que puede resultar sesgada por las opiniones preponderantes de los líderes participantes y expuesta a influencias de los

financiadores. El número y calidad de las opiniones a favor de una postura durante la sesión pueden determinar los resultados finales. Hay conocidos ejemplos de algunos denominados consensos invalidados por falta de imparcialidad o rigor metodológico en el proceso, o por el uso interesado de sus resultados.

A modo de resumen, la tabla I resume las ventajas e inconvenientes de cada uno de los métodos, valorados críticamente por un experto conocedor de estas metodologías (85). Como elemento de comparación basal se proponen las características esperables de una simple reunión informal de expertos donde estos discutiesen de forma no estructurada a la búsqueda de un consenso grupal (con las limitaciones propias derivadas de la interacción directa e incontrolada de la información subjetiva entre los miembros del grupo).

Tabla I. Comparativa de los principales métodos de consenso. Ventajas e inconvenientes.

→→ TÉCNICA DE CONSENSO →→	Reuniones no estructuradas	Grupo Nominal	Método Delphi	Conferencia de Consenso
PARTICIPANTES				
Minimización de influencias entre participantes	No	Alta	Muy alta	No
Da iguales oportunidades de participación	No	Sí	Sí	No
Buen funcionamiento con grupos heterogéneos	No	Sí	Sí	Sí
Garantiza la representatividad de participante	No	No	No	Algo
Necesidad de habilidades de comunicación verbal	Sí	Mínima	No	Sí
Necesidad de habilidades de comunicación escrita	No	No	Elevada	No
Sensación de manipulación	Escasa	Alta	Alta	Media
RESULTADOS				
Capacidad de obtención de consenso	Escasa	Alta	Alta	Alta
Capacidad de generación de nuevas ideas	Escasa	Muy alta	Alta	Media
Aportación de perspectivas distintas	Escasa	Muy alta	Alta	Muy alta
Integración de información diversa	A veces	Muy alta	Muy alta	Muy alta
Exploración de más de un problema a la vez	Sí	No	Sí	No
Dificultad en la interpretación de resultados	Alta	Media	Media	Baja
OPERATIVA				
Pérdida de opiniones extremas	No	Siempre	Siempre	A veces
Fácil realización	Sí	Sí	No	No
Esfuerzo en la preparación	Bajo	Medio	Alto	Alto
Duración	Corta	Corta	Larga	Media
Necesidad de reunir miembros	Sí	Sí	No	Sí

Reproducida de: Peiró S, Portella E. 2003. Referencia completa en bibliografía (cita 85)

1.5. El método Delphi de consenso profesional y el método Dephi modificado

El punto de partida para el desarrollo del método Delphi, a mediados el siglo XX, fue la insuficiencia de las técnicas de previsión económica tradicionales (proyecciones de tendencia, series temporales, modelos probabilísticos, analógicos o econométricos,...) para satisfacer las necesidades de información del decisor industrial en circunstancias de incertidumbre (66).

A partir de los setenta, las primeras crisis económicas globales interrumpieron un largo periodo de calma en los países desarrollados e impusieron la necesidad de una nueva forma de gestión más orientada al futuro. En esta nueva circunstancia, los modelos matemáticos predictivos, más o menos complejos, no podían aplicarse bien por carecer de datos para extrapolar o bien porque la velocidad de los cambios invalidaban continuamente cualquier referencia cuantitativa pasada. Aunque aumentó el reconocimiento de la necesidad de previsión, la variabilidad creciente del entorno conllevó una disminución sistemática en la precisión de las previsiones.

La complejidad y las limitaciones de las técnicas predictivas que la ciencia ofrecía redujeron su aplicación práctica, especialmente cuando se trataba de tomar decisiones importantes: generalmente el decisor se fía entonces, sobre todo, de su intuición y su experiencia, y de los de la gente que le merecen confianza.

Ahora bien, esta intuición individual y colectiva, entendida como el juicio subjetivo de ciertas personas informadas que disponen de un conocimiento privilegiado (expertos), debe ser convenientemente canalizada, estructurada y conjugada con el saber disponible y acumulado hasta hoy, para garantizar que la información final que se ofrece ha sido adecuadamente procesada (66). En definitiva se trata de destilar un juicio de valor útil del experto, a partir de su experiencia y la información que acumula, filtrado través de sus creencias, expectativas y opiniones. Pero todo ello, sin renunciar a una aproximación científica, con una metodología adecuada que garantice unas conclusiones confiables.

Así entendido, un experto es todo aquel individuo que puede aportar información, objetiva o subjetiva, válida para la realización de una previsión (en general, buscada con el fin de ayudar a tomar decisiones en situaciones inciertas). En otras palabras, todo aquel que puede aportar un input relevante a la consecución del fin que motivó la iniciación del trabajo Delphi (74), ya sea un especialista, un afectado o un facilitador (individuos que sin pertenecer a las categorías anteriores tienen capacidad de decidir, estimular u organizar). En este sentido, hay una tendencia actual a replantear el restrictivo concepto inicial de experto, por uno más abierto y, en ocasiones, cercano al de colaborador o participante implicado en el problema en cuestión (86).

El método Delphi inspira su nombre en el antiguo oráculo del santuario de Delfos (Grecia), un templo sagrado dedicado a Apolo y a las Musas, que fue centro religioso del mundo helénico y lugar de consulta a los dioses del Olimpo (87). A través de un oráculo mediado por la pitonisa (la sacerdotisa vitalicia del templo), cualquier mortal, desde los grandes reyes griegos a la gente más menesterosa, podía conocer la previsión de los dioses sobre sus inquietudes de futuro (previo ofrecimiento de sacrificios y un pago de tasas históricamente documentado). El nombre, propuesto por el filósofo Abraham Kaplan, se ha mantenido hasta hoy, aunque algunos autores lo han considerado poco afortunado por relacionar la técnica con algo oculto, profético y fatalista, siendo esto precisamente lo contrario a lo que sus desarrolladores pretendían alcanzar.

El **método Delphi convencional** fue ideado, originalmente con fines militares, a comienzos de los años 50 en el seno del Centro de Investigación estadounidense RAND Corporation por Dalkey y Helmer, como un instrumento para realizar predicciones sobre un posible caso de catástrofe nuclear (88). A inicio de los setenta se desclasificó por las Fuerzas Aéreas de los Estados Unidos como método de uso reservado (89), y su empleo se extendió con rapidez, geográfica y temáticamente, en muy diversas áreas, en especial hacia la previsión tecnológica, la toma de decisiones empresariales en situaciones inciertas y la evaluación de problemas sociales complejos (transporte, sanidad, educación, entorno, ...). Como muestra de esa rápida difusión de la técnica en pocos años, Linstone y Turoff

realizaron una exhaustiva recogida y análisis de 489 artículos y documentos publicados sobre el procedimiento hasta 1974 (90).

Desde entonces, ha sido utilizado frecuentemente como sistema para avanzar previsiones sobre el futuro y se ha desarrollado un notable cuerpo de conocimiento teórico que ha interesado a investigadores de la filosofía, la psicología social, la gestión empresarial, la metodología de investigación y la estadística. De todo ello ha resultado un complejo cuerpo de conocimientos que abarca discusiones sobre la epistemología de las ciencias inexactas en la que se sustenta, análisis sobre los planteamientos científico-técnicos que lo articulan, propuestas sobre nuevos desarrollos y variantes del método, y comprobaciones experimentales de la fiabilidad y estabilidad de sus resultados.

El objeto del método Delphi, en palabras de sus creadores es “obtener el consenso de opinión más fidedigno de un grupo de expertos” (89). Linston y Turoff lo definen como un método de estructuración del proceso de comunicación grupal que es efectivo a la hora de permitir a un grupo de individuos, como un todo, tratar un problema complejo (90).

La capacidad de predicción del Delphi se basa en la utilización sistemática de un juicio intuitivo emitido reiteradamente por el grupo de expertos. El método procede por medio de la interrogación a expertos con la ayuda de cuestionarios sucesivos, a fin de poner de manifiesto los aspectos coincidentes y divergentes en las opiniones del grupo, y de promover el consenso entre ellos, pero con la máxima autonomía por parte de los participantes, esto es, sin que sus opiniones sean determinadas por la posible presencia de líderes del grupo.

Las características definitorias (principios básicos) del método original, esquematizados en la figura 2 son (66):

a) Proceso iterativo mediante un número indeterminado, en su formulación original, de rondas de encuesta hasta conseguir estabilidad en las respuestas del grupo (ausencia de cambios en la variabilidad de opiniones).

b) Anonimato riguroso de los participantes, que desconocen al resto de sus compañeros de trabajo, a fin de reducir el efecto pernicioso que ejercen para la comunicación efectiva los posibles elementos dominantes del grupo. Para ello se usan cuestionarios escritos y un control estricto de las comunicaciones, sin permitir la confrontación directa entre los expertos.

c) Retroalimentación controlada para garantizar los resultados grupales (se ha demostrado experimentalmente que dependen de la interacción de sus individuos). Para ello el Delphi mantiene y promueve la interacción, solicitándola expresamente en cada ronda y facilitándola antes de la iniciación de la siguiente. Se transmite la posición general del grupo en cada momento del proceso frente al problema analizado y, frecuentemente, aportaciones o sugerencias significativas de los expertos (razonamientos discordantes o cualquier otra información adicional que, *motu proprio*, los panelistas consideren de interés). Esta información se filtra por el grupo coordinador, para evitar la aparición de “ruido” (información no relevante a los fines del estudio, redundante o errónea) y para unificar un lenguaje común que impida el anclaje en posiciones por causas psicológicas no racionales.

d) Valoración estadística de las respuestas del grupo, en general a través de la mediana de las respuestas individuales y del rango intercuartílico como indicador del nivel de consenso (o de dispersión de las respuestas). Se ha argumentado amplísimamente sobre la conveniencia de diversos sistemas de medición (escalas de Likert con distinta amplitud -5, 7 o 9 puntos-, respuestas si/no, escalas de priorización entre diversas alternativas, estimación directa de valores o proporciones, etc.), y sobre el procedimiento estadístico más apropiado para analizar e interpretar los resultados.

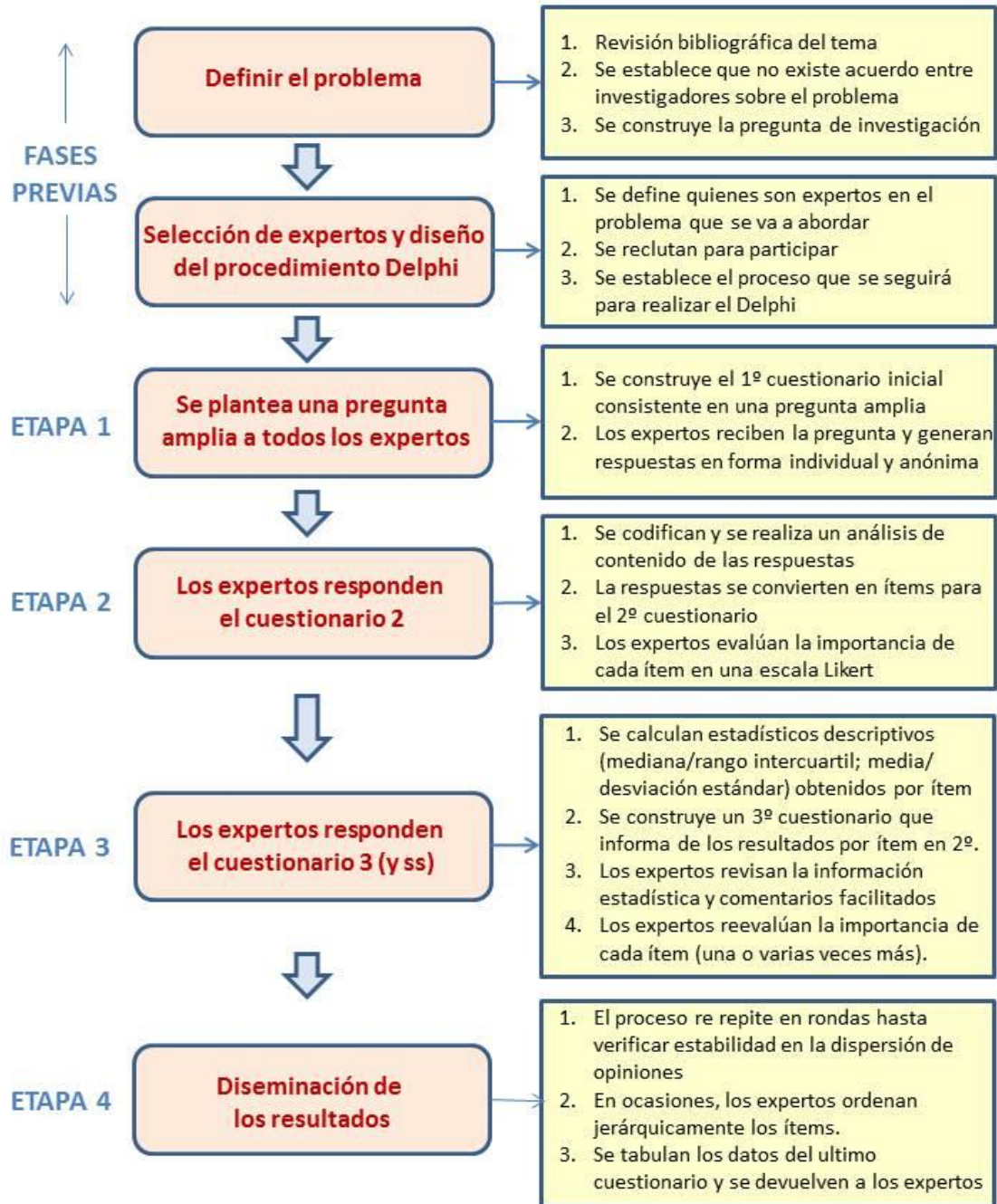
Figura 2: Estructura y atributos básicos de las rondas de encuesta Delphi



Reproducido de Berra A, Marin I, Alvarez R, 1996
(Referencia completa disponible en bibliografía: cita 67)

La figura 3 esquematiza la sistemática de trabajo de las tareas de los distintos participantes en un estudio Delphi convencional.

Figura 3: Etapas de implementación de la técnica Delphi original



Adaptado de Yañez R, Cuadra R, 2008
(Referencia completa disponible en bibliografía: cita 68)

Pese a su larga trayectoria, la formulación tradicional del método Delphi descrito sigue estando sujeta a permanente controversia metodológica y debate técnico, lo que ha motivado el diseño de no pocas variantes que intentan solventar algunas de sus dificultades. Al analizar la literatura sobre la materia, el lector puede sacar la conclusión precipitada de que no hay requisitos estandarizados, aceptados universalmente, para desarrollar un proyecto Delphi. La variabilidad existente en las investigaciones que se valen de este método, en algunas características aparentemente esenciales, resulta llamativa: la selección de expertos, las características del cuestionario, el número de rondas, la forma de retroalimentación, el procesamiento estadístico, la definición de acuerdo general de grupo, entre otras. Para algunos, esto es motivo de confusión y duda sobre el procedimiento (91,92), pero en realidad si existen unos principios básicos (descritos más atrás) y fases de aplicación (66) que se consideran universales, y que deben estar presentes (aunque, en ocasiones, no explícitamente) en todos los proyectos. Respetando tales salvaguardas, la variabilidad técnica no es una insuficiencia del método, sino una fortaleza derivada de su versatilidad y de su adaptabilidad a las condiciones del estudio en el que se emplea.

Se puede encontrar una revisión autorizada sobre el estado actual del desarrollo del método en una fuente documental de prestigio (*Practical Assessment, Research & Evaluation*) (93,94), una dilatada línea de investigación con múltiples estudios sobre el método y sus aplicaciones en *Technological Forecasting and Social Change* (<http://www.referencerepository.com/homes/tagsearch/page:2/jrnl:Technological%20Forecasting%20and%20Social%20Change/tag:delphi>) y un repositorio de acceso libre a documentos fundacionales del Delphi en la propia web de *Rand Corporation* (<http://www.rand.org/>), el laboratorio de ideas norteamericano, cuna del método y de la investigación de sus principales desarrollos. La Public Library of Science (PLOS), la biblioteca de literatura científica bajo licencia de contenido abierto, permite también acceder a una revisión sistemática y otros materiales de interés sobre el uso del Delphi en temas de salud (91,92).

Entre las cuestiones sujetas a variabilidad, destaca que no exista un tamaño predeterminado de panel Delphi que pueda considerarse técnicamente óptimo. En realidad, desde sus inicios se han realizado intentos a partir del algebra booleana, finalmente desechados por simplistas, (95) y esfuerzos de investigación empírica para poner de manifiesto la relación entre el tamaño del plantel de expertos y la

precisión de la estimación grupal obtenida (96). Aunque a partir de aquellos estudios, algunos hayan aconsejado paneles de no menos de 7 (el error disminuye notablemente por cada experto añadido hasta integrar esos siete elementos), y no más de 30 miembros (la mejora de previsión que se obtiene es muy pequeña y no compensa el incremento de coste y trabajo), en general, se observa que el error de previsión tiende a disminuir exponencialmente a medida que se amplía el tamaño de la muestra.

Otra cuestión compleja del Delphi convencional es el proceso de definición de las preguntas y construcción de los cuestionarios que puntuarán los expertos. En el proceso convencional esta tarea no recae exclusivamente bajo la responsabilidad de los investigadores que impulsan el estudio, sino que se solicita la participación activa de los panelistas expertos, generalmente partiendo de una ronda inicial de pregunta abierta. Un tamaño excesivo del grupo resultaría un factor limitante para la viabilidad práctica del proceso.

Los distintos posibles formatos de ítems para componer las encuestas (de jerarquización, de valoración mediante escala o de cuantificación directa) (97) determinan también un sinnúmero de posibilidades de uso de escalas de medida, de análisis de la información resultante y de interpretación de su significado.

Por último, y quizá uno de los elementos más sujetos a debate de todo el proceso, son los criterios que se emplean para la finalización del procedimiento, y el número de rondas de encuestas convenientes a cada caso. Razones relativas a tiempo, coste y precisión exigen rigor en esta decisión, que generalmente se toma bajo dos criterios estadísticos complementarios: el consenso de resultados (obtener un grado de convergencia suficientemente próximo de las opiniones individuales, en general comprobado por una medida de dispersión reducida, como el rango intercuartílico o la desviación típica, reducida), y la estabilidad en los mismos (no variación significativa de las opiniones de los expertos entre rondas sucesivas, medida por la permanencia de los resultados apreciados). La falta de consenso o la inestabilidad de los resultados, aconsejarían nuevas rondas de encuesta para intentar reducirlos (98). El importante consumo de tiempo requerido para la administración de los cuestionarios y para el análisis entre

rondas, así como el riesgo de aumentar los no respondedores con cada nuevo intento, llevan a los impulsores de muchos proyectos a que sean la disponibilidad de tiempo y medios quienes limitan el número de rondas de encuesta a 2 o, a lo sumo, 3 iteraciones, sin tomar en consideración los criterios propuestos por los investigadores.

Por todo ello, en su formulación original, el Delphi resulta una técnica exigente que se ha empleado con distinto nivel de rigor en el mundo biosanitario (68, 69, 99, 100), casi siempre para temas externos al ámbito clínico-asistencial (gestión, formación, planificación de recursos, calidad,...).

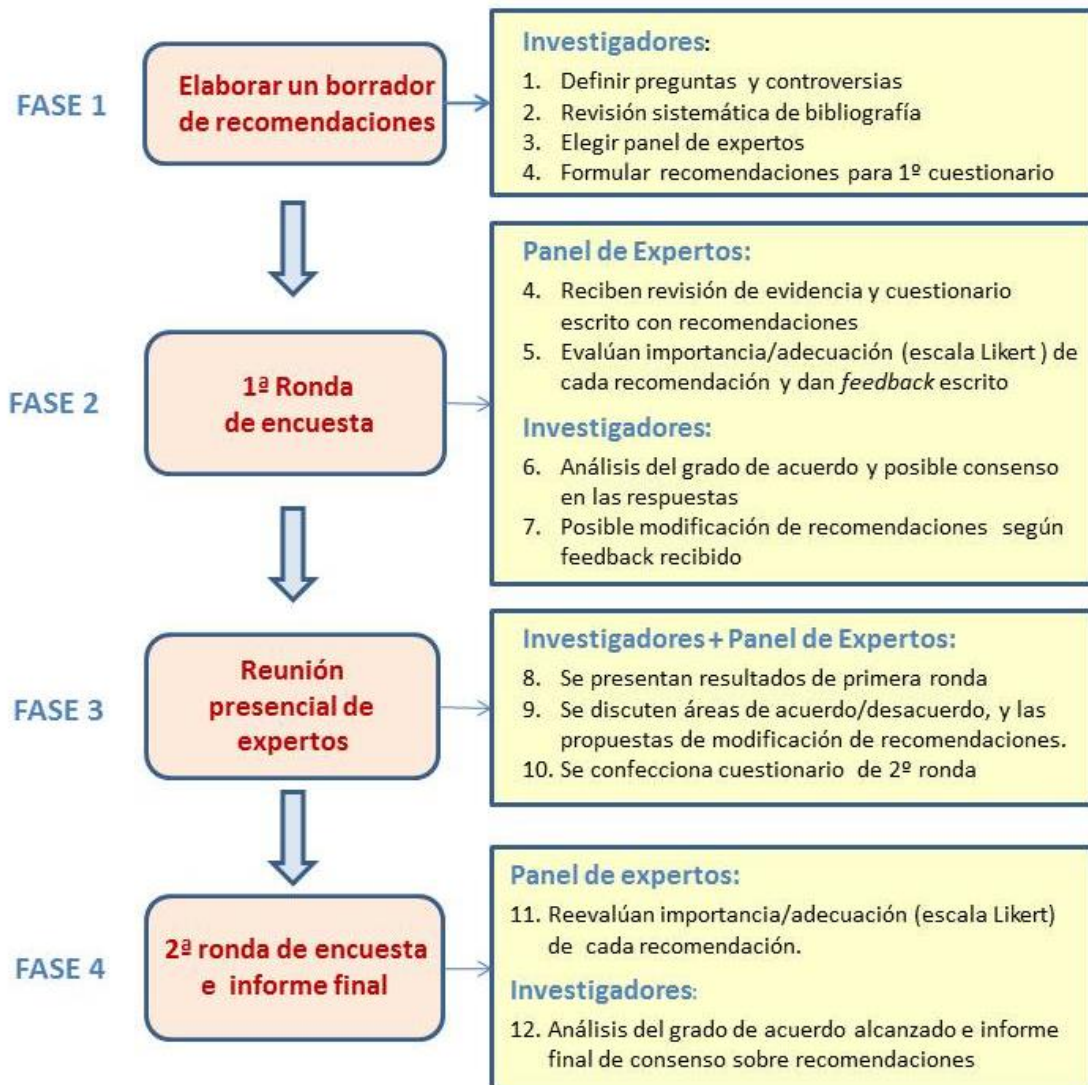
En el tesoro de PubMed, el procedimiento no se incluyó como “Técnica Delphi” hasta 1980. En su acepción revisada y vigente [*Delphi Technique*, 1991], se describe como “un cuestionario iterativo diseñado para medir el consenso entre las respuestas individuales. En el enfoque de Delphi clásico, no hay interacción entre el respondedor y entrevistador”.

Por las exigencias y dificultades del método convencional, en los últimos 25 años se ha difundido ampliamente una variante técnica del mismo conocida como **método Delphi Modificado** (101, 102). La propuesta se introdujo por Nelms y Porter en 1985, quienes desarrollaron una modificación del procedimiento tradicional al que llamaron EFTE (*Estimate, Feedback, Talk, Estimate*). El acrónimo no hizo fortuna por no evocar intuitivamente sus características diferenciales, pero desde entonces se ha empleado extensamente en muchas disciplinas bajo la denominación “método Delphi modificado”. Aunque, en puridad, no es la única variante disponible del método (66, 103, 104) por su exitosa divulgación hoy se le considera el “Delphi modificado” por excelencia (105). En este escrito, y en lo sucesivo se le identificará con la abreviatura **DEMO** para distinguirlo del método Delphi tradicional.

Su principal innovación consiste en permitir una sesión presencial de encuentro y debate experto (*talk*), entre dos únicas rondas de inicio y fin del proceso durante las que los panelistas puntúan de forma anónima e individual los asuntos

debatidos a través de encuestas escritas (*estimate*). La propuesta resulta, así, una formulación híbrida entre el método Delphi y los grupos nominales (figura 4).

Figura 4. Etapas de implementación del método Delphi modificado



Adaptado de Canadian Partnership Against Cancer's Capacity Enhancement Program (CEP). McMaster University 2010 (Referencia completa disponible en bibliografía: cita 106)

Aunque el DEMO tiene similitudes al Delphi tradicional en términos de procedimiento (es decir, una serie de rondas de encuesta a expertos seleccionados) y de intención (es decir, busca obtener un criterio grupal de consenso), se le considera un método más explícito y reproducible por simplificar

su aplicación práctica. Gracias a ello ha tenido una rápida propagación, en particular en el ámbito biomédico por su integración en el procedimiento de valoración del uso de la tecnología sanitaria conocido como “método RAND-UCLA o del uso apropiado” (107-109). Esta metodología, empleada con éxito para el desarrollo de guías de práctica clínica en materias o contextos para los no se dispone de pruebas científicas, ha sido ampliamente refrendada por la comunidad científica (72). Hoy, el DEMO se aplica a otros muchos intereses de investigación en Medicina, hasta el punto de existir estudios sobre su empleo, a través de revisiones sistemáticas y narrativas (91,110).

En el método RAND-UCLA, el proceso comienza con la revisión de la literatura y la creación de una lista exhaustiva de posibles indicaciones para la intervención que se someterá a análisis. Tras ello, un panel de expertos multidisciplinar califica si la intervención es apropiada o no para cada indicación propuesta, usando un cuestionario con escala de Likert (generalmente de 9 puntos). Los resultados de esta valoración individual se presentan en una reunión presencial de los panelistas donde se discuten las áreas de acuerdo y desacuerdo, y donde tienen la oportunidad de reconsiderar sus calificaciones personales. Para finalizar el proceso, las posibles indicaciones se votan confidencialmente de nuevo, en una segunda y definitiva ronda de encuesta, y se categorizan como apropiadas, inapropiadas o inciertas, según las puntuaciones obtenidas (107-109).

Las ventajas procedimentales derivadas de la reducción a dos rondas del DEMO frente a las tres o más requeridas por la técnica original son obvias:

- El Delphi convencional llevado a un extremo puede convertirse en una tarea larga y costosa para ambas partes (investigadores y expertos), ya que cada aplicación consume un tiempo extendido, haciendo cada vez más difícil mantener unas tasas de respuesta aceptable, mientras que con dos rondas es más probable mantener el interés de los panelistas.
- A los expertos se les ofrece una mayor oportunidad de realizar aclaraciones a sus ideas y de recibir esa misma información de sus pares durante el encuentro. Así, la sesión presencial facilita el proceso de clarificación y aproximación de posturas pese a reducir el número de iteraciones sobre los

panelistas (lo que reduce, además del tiempo, el coste de realización del proyecto).

- Por otra parte, el método modificado ha sistematizado en una única vía el procedimiento de análisis estadístico de la información producida, simplificando la objetivación e interpretación de los resultados producidos respecto al método convencional (102), lo que reduce las controversias. La simplicidad de su abordaje estadístico, basado en medidas descriptivas simples, hace que los resultados sean comprensibles para un mayor número de personas.

En contrapartida, algunos inconvenientes del método Delphi modificado son:

- Aunque se preserva la confidencialidad de la decisión final de cada panelista, la sesión presencial desvela el anonimato de los participantes. Pese a que en la reunión se adopten mecanismos de control de la interacción entre los participantes, no se anulan absolutamente los posibles sesgos de influencia. Esto supone perder una de las principales fortalezas del Delphi original y expone los resultados del consenso al riesgo de manipulación por parte del coordinador o de los líderes de opinión.
- El tiempo preparatorio requerido en la elaboración de los cuestionarios estructurados que se someten a debate puede ser muy prolongado, sobre todo si se solicita en ese proceso la participación de todos los panelistas. Además, ello exige que los participantes tengan una alta motivación intrínseca para comprometer su participación en todo el proceso y alcanzar buenas tasas de respuesta.
- En términos prácticos, la necesidad de desplazar a todos los expertos al lugar de reunión conlleva unos costes y unas dificultades logísticas, con frecuencia, inasumibles, si se trata de un grupo relativamente numeroso.

Todo ello, en la opinión autorizada del responsable de una revisión sistemática de un gran número de estudios de calidad que usaron el método entre 1978 y 2009 (91), hace que el método DEMO resulte un procedimiento mejorable, cuya aplicación, con frecuencia no se efectúa con el necesario rigor técnico.

En el apartado de “Métodos” de esta tesis se ofrecen más detalles técnicos del método DEMO y se explican extensamente los cambios que se proponen en la adaptación propia del método.

1.6. Necesidad de desarrollar una nueva adaptación del método Delphi modificado para su uso en el ámbito clínico.

Una búsqueda bibliográfica actual en la plataforma PUBMED que incluya [*“Delphi Technique”*] entre los términos del tesoro [Mesh] empleados, permite localizar el ingente material producido sobre la cuestión en los últimos años, incluyendo los estudios que han usado la herramienta en su desarrollo, y publicaciones que discuten el propio método. En fecha reciente (7 septiembre 2014) la búsqueda devolvía 2760 artículos con el término entre sus descriptores principales. Una estrategia más detallada permite apreciar que el método ha sido profusamente empleado en el ámbito de la gestión sanitaria [*“Health Services Administration”* o términos relacionados], 1778 ítems; en temas de calidad asistencial [*“Quality of Health Care”* o términos relacionados], 460 ítems; en el desarrollo de guías clínicas [*“Practice Guidelines”* or *“Clinical Protocols”*], 232 ítems; y en asuntos de formación médica de todo nivel [*“Medical Education”*], 149 ítems .

En todos estos ámbitos se puede comprobar el uso creciente del método DEMO en dos rondas (aunque con frecuencia no se le denomine “modificado” de forma explícita), por su versatilidad y su mayor facilidad de conducción, respecto a una formulación tradicional, hoy relegada a usos testimoniales. Esta situación parece propia de la Medicina, dada a la persistencia y vitalidad del método convencional en otras disciplinas. Una mirada, aun superficial, a las fuentes bibliográficas autorizadas, ya comentadas, en la materia permiten constatar esta particularidad (*Practical Assessment, Research & Evaluation; The Rand Corporation ; Technological Forecasting and Social Change*).

Seguramente un determinante de esta imposición creciente del DEMO en los estudios biomédicos está relacionado con la adopción del procedimiento modificado dentro del método RAND/UCLA de “uso apropiado”. Este método ha sido profusamente empleado en las Agencias de Evaluación de Tecnología

Sanitaria para emitir informes expertos que tratan de racionalizar, en términos de beneficio/riesgo y coste/oportunidad, el uso de los procedimientos clínicos hoy disponibles, en particular los más recientes, los menos contrastados y los que conllevan un mayor gasto sanitario.

La mayoría de estudios que emplean el método DEMO con tal fin lo hacen a través de paneles multiprofesionales de expertos de pequeño tamaño, en general de entre 8 y 12 miembros (72). Un grupo típicamente empleado, para el que existe incluso una formulación estadística propia para interpretar el grado de acuerdo que alcanzan sus miembros, es de 9 participantes (107). Esta característica no es un requerimiento metodológico que pretende optimizar los resultados del método, bien al contrario, se sabe experimentalmente que a mayor tamaño, los grupos optimizan su capacidad predictiva (66). Se trata, más bien, de una imposición logística: la necesidad de ofrecer a los expertos una sesión de encuentro y debate entre rondas, solo puede garantizarse, en la mayoría de proyectos, si el número de expertos que han de ser movilizados geográficamente es reducido.

Al introducir una sesión de debate “cara a cara” de los expertos, en la que se permite su interacción (en principio, controlada) antes de abordar confidencialmente la segunda y definitiva encuesta, el método DEMO quebranta los principios básicos del método original. Se abandona, así, una de sus señas principales de identidad, como era la exigencia de completar la tarea a distancia y de forma anónima, manteniendo la confidencialidad de los expertos durante todo el proceso.

Aunque, en rigor, tal necesidad se impuso a los desarrolladores de la técnica original por las dificultades (en los años cincuenta) para desplazar con frecuencia grupos de debate en un país tan extenso como los Estados Unidos de América (66), ya se ha comentado su beneficio en términos de evitar los sesgos de influencia. En toda dinámica de grupos existe el riesgo de que uno o varios individuos adopten un papel dominante sobre el grupo (ya sea por su personalidad, su modo de expresarse, su ascendencia sobre los miembros o su autoridad formal), dirigiéndolo e imponiendo sus opiniones. El anonimato en las respuestas de cada miembro del panel previene que esto ocurra. Dalkey llamó “ruido” a estas

influencias dominantes o espurias (interferencias provocadas por la defensa de intereses individuales o de los grupos a los que pertenecen cada uno de los participantes, o a la presión del grupo por homogeneizar las opiniones) y propuso la retroalimentación escrita como solución.

En términos positivos, es de esperar que la interacción directa de los panelistas durante la reunión presencial del DEMO facilite la necesaria aproximación de posturas inherente al consenso y abrevie la necesidad de debate, haciendo razonable la finalización del proceso en solo dos rondas (a diferencia de las 3 o más iteraciones de encuesta que el Delphi original requería).

En cualquier caso, más allá de la controversia (67) sobre los riesgos y beneficios de esta reunión presencial, resulta indiscutible que tal exigencia reduce las posibilidades de aplicar el método DEMO a proyectos que pretendan movilizar grupos de expertos numerosos (si tal pretensión fuera de interés). Las molestias y costes de organización que conllevaría una reunión presencial con muchos asistentes, y las dificultades técnicas para conducir tal sesión sin influencias incontroladas de los investigadores hacia los expertos y entre los propios expertos, la harían, en la práctica, inviable.

Esta situación no es un escenario hipotético: la necesidad de apelar al criterio y experiencia de grupos numerosos de expertos en los nueve estudios que se presentan en esta tesis está ligada a los propios fines de los consensos que se pretende alcanzar en ellos (y en otros posibles proyectos futuros en la misma línea): responder a la necesidad del clínico de disponer de ayuda experta para tomar decisiones clínicas complejas (por inciertas, infrecuentes, peligrosas, caras, etc.), para las que no dispone de pruebas científicas suficientes que avalen una opción ventajosa.

Con tal fin, es imprescindible que el proceso de elaboración de las recomendaciones expertas goce del mayor crédito posible y resulte aceptable para la mayoría de sus destinatarios finales. A todo ello colaboraría sustancialmente que las recomendaciones emitidas fueran avaladas por un alto grado de acuerdo, del mayor número posible de expertos de reconocido prestigio que pudiesen opinar en

la materia. En este sentido, aunque el término experto es ambiguo (111), en el ámbito clínico, puede considerarse experto “todo aquel cuya formación y experiencia previa le ha permitido alcanzar un dominio sobre su especialidad que excede el nivel promedio de sus iguales y que está en disposición de exponer sus opiniones para que sean utilizadas como juicios conclusivos, considerándosele apto para emitir criterios certeros por quienes se lo solicitan” (110).

Si, como se argumentará, este tipo de clínicos expertos son profesionales numerosos y ubicuos en nuestro sistema sanitario, se debería intentar el acceso a la opinión de una parte representativa, en número, de dicho colectivo. Cuantos más, mejor. Pero, con tal fin, el método DEMO presenta restricciones insalvables por los costes y dificultades para reunir en un tiempo y lugar concreto a un grupo de individuos numeroso.

Sin tal posibilidad, resulta imperativo desarrollar una adaptación metodológica al DEMO específica para el fin clínico descrito que no requiera encuentro presencial, sin perder por ello fiabilidad como medio para la recogida, debate y síntesis de las opiniones de los expertos. Si así se consiguiese, los clínicos podrían confiar en las recomendaciones elaboradas por esta vía como procedentes del consenso del mejor criterio disponible en la cuestión, a diferencia del escaso crédito concedido a la opinión procedente de un único (o pocos) expertos, siempre expuesta a sospechas de parcialidad o de visión limitada.

En la sección “Discusión” de esta tesis se argumenta más detalladamente que se entiende por un “clínico experto” y porque se considera que su presencia en nuestro sistema sanitario es prevalente. Quizá esto represente una característica diferencial del mundo clínico respecto al entorno profesional de otras disciplinas en los que se desarrolló y se sigue usando profusamente el método Delphi, por motivos que se razonan allí.

Aunque las características de la “adaptación clínica” del DEMO empleada en los estudios de esta tesis (en adelante identificada por la abreviatura **DEMOCLI**) se comentarán por completo y en detalle en la sección “Métodos”, se anticipa aquí otra innovación sustancial introducida por DEMOCLI, referida al procedimiento de

redacción de los ítems de encuesta. Intentar que un panel numeroso de expertos se implique en las fases iniciales del DEMO (que incluyen, ver fig. 4, la selección de preguntas clínicas objeto de controversia o incertidumbre, la búsqueda y revisión de pruebas sobre la materia, y la formulación explícita de las recomendaciones que se someterán al juicio experto en las encuestas), es una pretensión poco realista. DEMOCLI descarga a los expertos de responsabilidad en esas tareas, eximiéndoles de dominar ciertos procesos técnicos, como la redacción de preguntas clínicas en formato estructurado, las habilidades de búsqueda, de lectura crítica o de capacidad de síntesis bibliográfica, y la formulación apropiada de ítems aptos para un cuestionario Delphi.

En nuestra nueva propuesta técnica, todas ellas son tareas que quedan bajo el gobierno de los investigadores impulsores de cada consenso asistidos, generalmente, por técnicos conocedores de la metodología necesaria. De esta forma, los expertos reaccionan ante un tema en vez de generarlo y su tarea se centra no en la construcción de las propuestas a debatir sino en llegar acuerdos sobre la misma.

Iniciar el proceso con una encuesta que incluye ítems cuidadosamente seleccionados a priori por el equipo investigador (extraídos de diversas fuentes, incluyendo opiniones sintetizadas de la literatura, criterios de los propios investigadores, o consultas con otros expertos líderes en la materia a debate), adecuadamente contruidos y acompañados de una síntesis de calidad de la evidencia disponible, tiene, entre otras ventajas, acelerar el proceso Delphi y mejorar la tasa de respuesta de la ronda inicial. En el apartado “Discusión” se analiza en detalle si esta dinámica de inicio del proceso DEMOCLI puede representar algún tipo de sesgo sobre los contenidos que se debaten en el consenso (pondrían ser parciales o tendenciosamente presentados) y se discuten las precauciones que se toman para evitarlo.

De igual forma, se discutirá allí en detalle las características del proceso de retroalimentación a distancia que se ofrece a los expertos y que suple la reunión presencial entre rondas del DEMO, y en qué medida garantiza la necesaria interacción de los panelistas. Este *feedback* (compuesto de medidas estadísticas,

gráficos descriptivos y un resumen de los comentarios cualitativos recibidos) informa a cada participante de su posición con respecto al resto del grupo, y le ayuda a tomar decisiones sobre sus respuestas en la ronda final.

Hipótesis y Objetivos



2. HIPÓTESIS Y OBJETIVOS

2.1. Hipótesis que se pretenden demostrar en esta tesis:

- El método Delphi Modificado (**DEMO**) en dos rondas es una técnica estandarizada de negociación profesional rápida y efectiva, frente a otras alternativas, para que un grupo heterogéneo de expertos confronte sus opiniones personales y alcance el consenso sobre materias sujetas a incertidumbre. Una adaptación técnica propia del método Delphi Modificado para la clínica (**DEMOCLI**) facilita su uso y garantiza su efectividad para que grupos clínicos-expertos desarrollen por consenso recomendaciones que ayuden a reducir la variabilidad de prácticas al tomar de decisiones en contextos de incertidumbre clínica.
- **DEMOCLI** resulta aplicable en todo tipo de especialidad y ámbito asistencial, y logra un alto grado de aceptación y participación de los expertos como procedimiento para analizar en conjunto situaciones de incertidumbre clínica y desarrollar por consenso recomendaciones que unifiquen su manejo.
- Pese a que **DEMOCLI** es una técnica estructurada de debate no presencial que solo permite una interacción limitada (a distancia) y controlada de los expertos participantes, estos asumen sin conflicto su dinámica técnica y comparten las conclusiones resultantes del proceso como recomendaciones clínicas provenientes de un consenso grupal.
- **DEMOCLI**, adecuadamente diseñado y conducido, es una metodología de rigor científico-técnico suficiente para superar con éxito los procedimientos de *peer-review* y para que los estudios que la apliquen logren su difusión en revistas de impacto científico en la medicina clínica.

2.2 Objetivos:

Objetivo general:

Presentar las características de una nueva adaptación para la clínica del método Delphi Modificado en dos rondas (**DEMOCLI**) y constatar, a partir de una experiencia amplia y objetivable de uso en proyectos reales, que resulta un procedimiento riguroso, eficiente y viable de debate no presencial para elaborar por consenso recomendaciones expertas que orienten la actuación médica en situaciones de incertidumbre clínica y/o alta variabilidad de prácticas.

Objetivos específicos:

1. Comprobar que **DEMOCLI** en 2 rondas es una técnica de negociación profesional eficiente (resolutiva, de rápida ejecución y coste asumible), para que grupos heterogéneos de médicos expertos confronten su criterio personal y desarrollen por consenso recomendaciones clínicas de ayuda a la toma de decisión en escenarios de incertidumbre y/o alta variabilidad de prácticas.
2. Comprobar que **DEMOCLI** logra un alto grado de aceptación y participación entre expertos clínicos como procedimiento de consenso profesional para elaborar recomendaciones clínicas en escenarios de incertidumbre y/o alta variabilidad de prácticas.
3. Comprobar que el tipo de interacción de los expertos (no presencial) que ofrece **DEMOCLI** es suficiente para que estos asuman sin conflicto ni veto las conclusiones resultantes del procedimiento, y las respalden como recomendaciones de consenso grupal.
4. Comprobar que **DEMOCLI** resulta versátil, esto es, igualmente aplicable y eficiente en distintos tipos de ámbitos asistenciales y/o de especialidades médicas.
5. Comprobar la suficiencia de **DEMOCLI**, pese a su escasa tradición de empleo en el ámbito clínico y a la novedad técnica que incluye, para que los estudios que

la utilizan con los fines descritos resulten metodológicamente aptos para su publicación científica en revistas de impacto.

Métodos



3. MÉTODOS

3.1. Diseño metodológico de la tesis por compendio de publicaciones.

Como soporte argumental empírico de esta tesis por compendio de publicaciones se presentan nueve estudios coordinados por el doctorando en los últimos cinco años (entre 2009-2013). Se trata de una serie de proyectos multicéntricos de ámbito estatal, publicados en revistas científicas indexadas en bases de datos internacionales, que comparten un diseño y un procedimiento de ejecución, en lo sustancial, idénticos. Los cinco último estudios, como exige la normativa prevista para la defensa de tesis por compendio de publicaciones fueron publicados en fecha posterior a la aprobación del proyecto de tesis doctoral. Se presentan también las restantes publicaciones que iniciaron la serie, cuyos resultados preliminares inspiraron la hipótesis de estudio.

Para responder a los objetivos de la tesis, se realiza un análisis comparado de las claves metodológicas y los resultados disponibles en los nueve originales publicados. Se pretende con ello aportar una reflexión metodológica al cuerpo de conocimientos establecido sobre las técnicas de consenso experto aplicadas a la medicina (80, 112-115).

En estos estudios, auspiciados por sociedades científicas o promovidos por grupos clínicos expertos, el doctorando se ha responsabilizado personalmente de las tareas de diseño, asesoría metodológica y conducción técnica, como experto en la formulación y desarrollo de procedimientos de consenso profesional.

Los coautores del doctorando en estas publicaciones son, en su mayoría, líderes clínicos de reconocido prestigio en sus respectivas especialidades o áreas de subespecialización, que asumen diversas responsabilidades en centros asistenciales, docentes o investigadores de referencia estatal. En el anexo XX se incluye una declaración personal del autor principal de las publicaciones en las que el doctorando no ha sido el primer o último firmante, que certifica la

responsabilidad directa de este como coordinador técnico del estudio y el que ninguna de las publicaciones presentadas ha sido anteriormente empleada en la defensa de otra tesis doctoral.

Los estudios, en secuencia cronológica, son los siguientes (116-124):

- Ampudia Blasco FJ, Baos Vicente V, **Caballero Martínez F**, Álvarez-Sala Walther LA. Criterios profesionales sobre el papel de ezetimiba en el manejo clínico del paciente dislipémico en el ámbito de atención primaria. *Av Diabetol* 2009; 25(5): 399-407.
- **Caballero Martínez F**, Plaza Moral V, Quirce Gancedo S, Fernández Benítez M, Gómez Ruiz F, López Viña A, Molina París J, Quintano Jiménez JA, Soler Vilarrasa R, Villa Asensi JR, Balmes Estrada S. Valoración externa de las recomendaciones de la GEMA 2009 por un panel multiprofesional de expertos en asma. *Arch Bronconeumol* 2010; 46(8): 411-419.
- Villanueva Haba V, Sánchez Álvarez JC, Peña Mayor P, Salas Puig J, **Caballero Martínez F**, Gil-Nagel Rein A. Treatment initiation in epilepsy. An expert consensus in Spain. *Epilepsy and Behaviour* 2010; 19(3): 332-42.
- Medrano López C, García-Guereta L, Fernández Pineda L, Malo Concepción P, Maroto Álvaro E, Santos de Soto J, Lirio Casero J, Suárez Cabrera P, **Caballero Martínez F**. Consenso clínico sobre la profilaxis de la infección por virus respiratorio sincitial (VRS) y el uso del palivizumab en cardiología pediátrica. *An Pediatr (Barc)*. 2010; 72(6): 432.e1-432.e13.
- Martín Carrasco M, Agüera Ortiz L, Caballero Martínez L, Cervilla Ballesteros J, Menchón Magriñá JM, Montejo González AL, Moríñigo Domínguez A, **Caballero Martínez F**. Consenso clínico sobre la depresión en el paciente anciano. *Actas Esp Psiquiatr* 2011; 39(1): 20-31.

- Bobes García J, Saiz Ruiz J, Bernardo Arroyo M, **Caballero Martínez F**, Gilaberte Asín I, Ciudad Herrera A, en representación del Grupo de Trabajo sobre la Salud Física del Paciente con Trastorno Mental Grave. Consenso Delphi sobre la salud física del paciente con esquizofrenia: valoración por un panel de expertos de las recomendaciones de las Sociedades Españolas de Psiquiatría y de Psiquiatría Biológica. *Actas Esp Psiquiatr* 2012; 40(3): 114-28.
- Plaza Vicente, Calle M, Molina J, Quirce S, Sanchis J, Viejo JL, **Caballero F**. Validación, mediante el método Delphi, del Consenso sobre Terapia Inhalada para el tratamiento de las enfermedades respiratorias obstructivas. *Arch Bronconeumol* 2012; 48(6): 189-196.
- Cózar Olmo JM, Hernández Fernández C, Miñana López B, Amón Sesmero JH, Montlleó González M, Rodríguez Antolín A, **Caballero Martínez F**. Consenso experto sobre el impacto clínico de la nueva evidencia científica disponible sobre Hiperplasia Benigna Prostática. *Actas Urol Esp* 2012; 36(5): 265-275.
- Pérez C, Rodríguez MJ, Guerrero A, Margarit C, Martín-Estefanía C, Oteo-Álvaro A **Caballero-Martínez F**. Consenso experto sobre el uso clínico de los tratamientos por vía tópica en el manejo del dolor neuropático periférico. *Rev. Soc. Esp. Dolor*. 2013; 20(6): 308-323.

En lo sucesivo, en el cuerpo de redacción y tablas de los siguientes apartados de esta tesis, los nueve estudios analizados se citarán mediante una única palabra (a modo de descriptor semántico) y el año de publicación, que permitan al lector evocar con facilidad los aspectos esenciales de cada estudio sin tener que acudir al uso de una referencia bibliográfica convencional, a fin de facilitar una lectura ágil del manuscrito. Los términos elegidos, y siguiendo el orden cronológico comentado de los trabajos, son los siguientes:

- DISLIPEMIA 2009.
- ASMA 2010.

- EPILEPSIA 2010.
- PEDIATRIA 2010.
- DEPRESIÓN 2011.
- ESQUIZOFRENIA 2012.
- INHALADORES 2012.
- PRÓSTATA 2012.
- DOLOR 2013.

Salvo en lo relativo a particularidades de cada estudio derivadas de sus temas y objetivos específicos que se detallan en las tablas II-VI, los nueve estudios comparten un diseño metodológico similar que se describe a continuación (una adaptación propia del autor del “método Delphi modificado” de consenso profesional).

Las tablas citadas permiten, no obstante, comparar los pormenores técnicos de cada uno de ellos. Para indagar con mayor detalle en algún estudio en particular, se sugiere revisar el apartado de metodología del propio texto facsímil de cada original, tal y como fue publicado en su revista correspondiente (disponibles en la sección de “Resultados” de esta tesis).

3.2. Aspectos generales de la adaptación clínica del método Delphi Modificado (DEMOCLI) usada en los estudios presentados:

En los nueve estudios discutidos se ha empleado una adaptación original orientada a la clínica del método Delphi Modificado (DEMOCLI), una técnica estructurada de consenso profesional, a su vez, variante del procedimiento Delphi original, ambos presentados en detalle en la sección “Introducción”.

El desarrollo de la adaptación clínica de método DEMO tiene por objeto desarrollar recomendaciones clínicas de ayuda a la decisión clínica en situaciones de incertidumbre, con la ayuda de grupos de clínicos expertos. Para ello, intenta solventar algunos inconvenientes reconocidos del método Delphi modificado (85), como la laboriosa preparación de los materiales de encuesta; el emplear, en general, paneles expertos de pequeño tamaño, dada la necesidad de garantizar una

alta motivación intrínseca en los participantes con para lograr altas tasas de respuesta; la dificultad de los expertos para matizar o aclarar libremente sus valoraciones personales preservando su confidencialidad; el no ser un método inmune a la posible influencia de los promotores (en particular, en la elección del panel de expertos y en la discusión de los resultados, dada su intervención directa en el encuentro presencial de los expertos durante el que se efectúa la ronda final de encuesta del proyecto); y las propias dificultades organizativas y costes derivados, precisamente, de la organización de esta reunión.

Para controlar el riesgo de influir en los expertos y condicionar su participación, los equipos de investigadores que impulsen estos proyectos (el Comité Científico de nuestros estudios) deberían declarar sus conflictos de interés, garantizar la confidencialidad de opinión durante todo el proceso a los panelistas, y seguir procedimientos sistematizados y objetivables para la selección de panelistas (método “bola de nieve”) (125) y durante el análisis estadístico e interpretación de resultados (126,127).

Como seña de identidad de las publicaciones analizadas y aportación original de esta tesis, en los estudios presentados se ha empleado una adaptación propia del método Delphi Modificado que propone algunas modificaciones técnicas al procedimiento estándar. Sin alterar lo sustancial del método, como se discutirá en adelante, consideramos estos cambios particularmente apropiados para usar el Delphi modificado en entornos clínicos, más en concreto, cuando este método de consenso se destina a la elaboración de recomendaciones expertas de actuación en situaciones de incertidumbre clínica. En este sentido, se justificará:

- La responsabilidad exclusiva del Comité Científico de cada estudio (líderes de referencia en la especialidad y materia de estudio) en la elaboración directa de los cuestionarios de estudio, sin implicar al pleno del panel de expertos en esta fase.
- El uso de paneles de expertos clínicos algo más numerosos que la composición habitual del formato usado en otras disciplinas.
- El hecho de que ambas rondas de encuesta completen sus tareas sin necesidad de reunión presencial alguna.

Como se comentará, estos detalles responden a especificidades propias de consenso profesional entre expertos clínicos, y pueden resultar determinantes de la aceptación de las recomendaciones derivadas de estos consensos entre el colectivo clínico destinatario de sus conclusiones, y de la viabilidad práctica, en tiempo y forma, de los proyectos, y de que sus costes de ejecución resulten asumibles.

La figura 5 esquematiza las etapas de implementación del nuevo método, tal y como ha sido empleado en los estudios presentados.

En los siguientes apartados de esta sección se explican detalladamente todas las particularidades diferenciales de DEMOCLI respecto al DEMO.

Figura 5: Etapas de implementación de la adaptación clínica del método Delphi modificado (DEMOCLI)



Fuente: elaboración propia

3.3. Elaboración de los cuestionarios y procedimiento para la síntesis de la evidencia científica utilizados en los estudios presentados.

En la adaptación del método Delphi propuesta, el Comité Científico impulsor del proyecto, formado por varios médicos especialistas en la patología de estudio (en general, líderes de referencia para la especialidad) son los responsables directos de analizar y discutir las cuestiones objeto de incertidumbre durante el proceso asistencial a la patología de estudio y los problemas de manejo clínico que conllevan, esto es, de formular las preguntas clínicas que los médicos suelen plantearse durante el ejercicio. Para ello se apela a su destacada experiencia personal conjunta, sin necesidad de implicar en este proceso al panel de expertos que protagonizará las fases posteriores del proyecto.

El objetivo de esta tarea inicial es decidir el contenido argumental y estructura general del consenso (los temas que incluirá, las preguntas clínicas que se plantearán y el formato del cuestionario de recomendaciones clínicas que se deseen someter al criterio general de los expertos). Aunque el Comité es libre de abordar el trabajo de acuerdo a sus preferencias, en los estudios presentados el proceso se inició con una reunión presencial. Durante este encuentro, fundacional para el propio Comité Científico y de puesta en marcha del proyecto, se discuten y clarifican las especificaciones técnicas de la tarea a desarrollar y se inicia, preliminarmente, el debate para la selección y la justificación de los contenidos del estudio. Para alcanzar los objetivos de esta sesión, el Comité Científico de cada estudio contó con la asistencia presencial de un asesor técnico (el doctorando de esta tesis).

Tras la reunión inicial, el proceso de construcción del cuestionario avanza mediante una repetida iteración de propuestas y opiniones entre los miembros del Comité, en general, compartidos por vía electrónica (con frecuencia usando algún servicio de alojamiento de archivos en la nube, que permite almacenar y sincronizar información en línea, y compartir documentos entre los ordenadores de los usuarios, tipo Drop-box o similar). En el anexo 1 se transcribe un ejemplo del documento de instrucciones básicas que se facilitó a los miembros de cada Comité Científico para afrontar esta tarea.

A la vez que se avanza en ello, el propio equipo investigador realiza una revisión sistemática de la literatura y elabora una síntesis de la evidencia científica disponible en el momento del estudio. Los artículos se localizan a través de búsquedas sistemáticas en los buscadores en línea de las bases documentales más usadas en medicina clínica (PUBMED - Medline, EMBASE - Excerpta Médica, IME - Índice Médico Español), para un determinado período temporal y se obtienen mediante búsquedas manuales de las citas localizadas. En general, para evaluar y categorizar la calidad de las pruebas científicas recopiladas se indicó a los investigadores el empleo de las recomendaciones del Oxford Centre for Evidence-Based Medicine (128). En el anexo 2, se reproduce la tabla resumen de estas consideraciones en su formulación original.

Salvo el proyecto ASMA 2010, INHALADORES 2013 y ESQUIZOFRENIA 203, cuyas encuestas se configuraron con recomendaciones directamente extraídas, en su literalidad, de los documentos previos que pretendían ser refrendados por el consenso experto, los restantes estudios se iniciaron por un procedimiento reglado, como el descrito, de búsqueda bibliográfica. Aunque este proceso se realizó bajo la responsabilidad de cada comité científico, dispuso, en general, del apoyo de documentalistas que asesoraron externamente en esta materia.

Para completar la selección de los contenidos del cuestionario Delphi, las propuestas preliminares recibidas de cada miembro del comité se sometieron a un proceso de revisión, adaptación y agrupación temática, asistido por el consultor externo (el doctorando de esta tesis). Se intenta garantizar con ello la redacción apropiada de los ítems de encuesta, de forma que resulten conceptualmente válidos (comprensibles, precisos, unívocos, ...) y, en general, aptos para su empleo en un cuestionario Delphi.

Obtenido el primer borrador del cuestionario se hizo circular entre los miembros del comité hasta logra una versión definitiva, técnicamente validada y unánimemente satisfactoria para todos los miembros del comité. Cada ítem finalmente seleccionado es una aseveración (afirmativa o negativa), que recoge un criterio profesional o una recomendación clínica de ayuda a la toma de decisión, en

algún aspecto de interés o controversia posible a lo largo del proceso asistencial de la patología de estudio.

El producto final del trabajo preparatorio del comité científico de cada proyecto fue un cuestionario compuesto por una batería de 40 a 70 recomendaciones clínicas. En todos los casos, tal extensión de encuesta y la carga de trabajo que cumplimentarla requirió, fue considerada razonable y asumible por los expertos. Como excepción singular, el proyecto EPILEPSIA 2010 propuso y consiguió completar el proceso de consenso sobre un listado de recomendaciones individuales para 194 escenarios clínicos. Esta dimensión, en apariencia excesiva, resultó finalmente viable dada la particular formulación del cuestionario en dicho proyecto (que preguntaba repetidamente a los expertos por la idoneidad de usar, de forma independiente, distintas moléculas sobre un mismo escenario).

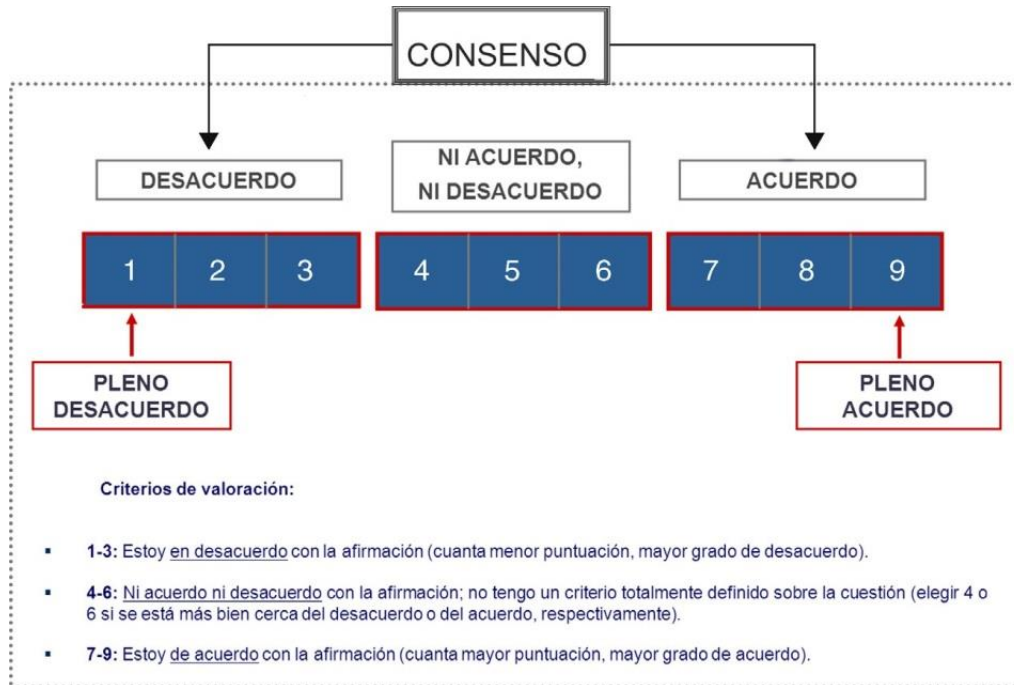
En la tabla 2 se detallan el número de ítems que se incluyeron en la versión final del cuestionario de cada estudio. Pueden consultarse ejemplos de su formulación en los originales de los diferentes estudios.

3.4. Escalas de medida de los cuestionarios utilizadas en los estudios presentados.

En los estudios de publicación más reciente (ESQUIZOFRENIA 2012, INHALADORES 2012, PRÓSTATA 2012, DOLOR 2013), el cuestionario para la valoración de los ítems empleó para todas las cuestiones una escala valorativa ordinal de tipo Likert de nueve puntos (entre un mínimo de 1=pleno desacuerdo y un máximo de 9=pleno acuerdo), según el formato desarrollado en UCLA-Rand Corporation para el método de evaluación del uso apropiado de tecnología sanitaria (107-109).

Las categorías de respuesta en esta escala se describen mediante descriptores semánticos en tres regiones (1-3="desacuerdo"; 4-6="ni acuerdo, ni desacuerdo"; 7-9="acuerdo"). En cada caso, el encuestado pudo discriminar con detalle su opinión particular, eligiendo uno de los tres puntos contenidos en cada región.

Figura 6: Presentación a los panelistas de la escala de Likert de 9 puntos usada en algunos estudios



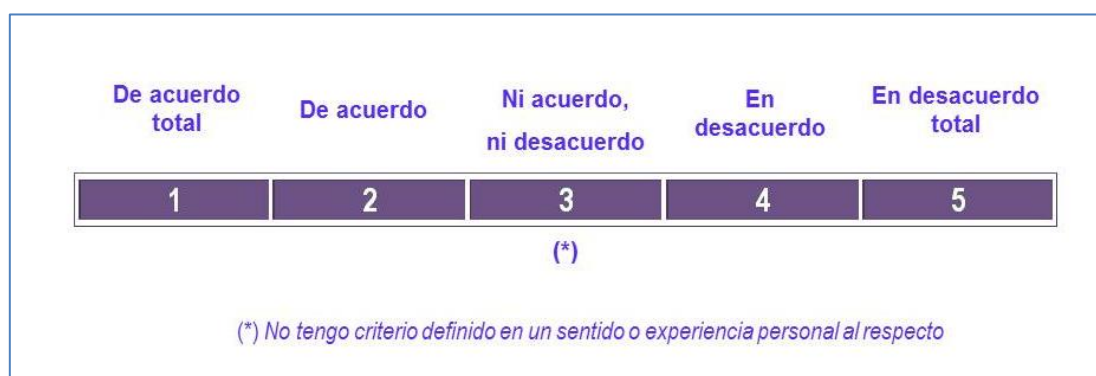
El formulario electrónico de encuesta solicitó, además, que los participantes incluyeran en formato de texto libre observaciones aclaratorias de respaldo a su voto particular y/o comentarios críticos para cada ítem (impugnación, propuesta de cambio,...). Igualmente, ofreció un apartado final para recoger, en caso necesario, propuestas de nuevos ítems ampliatorias del planteamiento inicial del Comité Científico. Si alguna de las cuestiones no se cumplimentó por considerarse el propio panelista no cualificado en la materia, el ítem se trató como caso perdido a efectos estadísticos.

A diferencia de lo descrito *ut supra*, en los cinco primeros estudios de esta serie (DISLIPEMIA 2009, ASMA 2010, EPILEPSIA 2010, PEDIATRIA 2010, DEPRESIÓN 2011) se empleó una escala de Likert reducida de cinco puntos, siguiendo las propuestas originales de una serie de estudios que emplearon por primera vez el

método Delphi modificado para la elaboración de recomendaciones clínicas para el uso de medicamentos en ancianos, hace más de 20 años (127, 129, 130).

En tal caso, las categorías de respuesta numérica fueron de 1 a 5 y sus calificadores lingüísticos los siguientes: 1=pleno acuerdo con el ítem, 2=más bien de acuerdo, 3=ni acuerdo ni desacuerdo (no tengo un criterio definido), 4=más bien en desacuerdo, 5=pleno desacuerdo con ítem. Más adelante se discuten ventajas e inconvenientes de ambas alternativas.

Figura 7: Presentación a los panelistas de la escala de Likert de 5 puntos usada en algunos estudios



3.5. Particularidades de las rondas de encuestas en la adaptación clínica del método Delphi modificado (DEMOCLI) propuestas en los estudios presentados:

Durante el proceso Delphi modificado (DEMO), las indicaciones son puntuadas por los panelistas en dos rondas. En la primera ronda, el informe de síntesis de la evidencia recopilada y la lista de recomendaciones a valorar son enviados a los miembros del panel, junto con instrucciones sobre la forma de puntuar el cuestionario. Su cumplimentación se hace individualmente, sin que los expertos mantengan ningún tipo de comunicación o intercambio de opiniones.

Tras ello, los panelistas envían las puntuaciones al equipo investigador. Una vez analizadas estadísticamente, la información de dichos resultados se les devuelve durante la celebración de una segunda ronda presencial que reúne a los panelistas (durante un día o dos) para discutir y repuntuar las indicaciones. Cada panelista recibe para ello un documento que muestra sus propias puntuaciones identificadas y el resultado de las puntuaciones de los demás panelistas sin identificar, pudiendo así comparar su opinión frente a la del resto de los componentes del panel, sin verse por ello forzado a modificarla por la posible presión del grupo. El moderador de la reunión señala las áreas de desacuerdo o confusión para discusión. Si los panelistas lo desean, se puede modificar la estructura de la lista de indicaciones ampliándola, simplificándola, añadiendo nuevas categorías o eliminando categorías de la primera ronda. Al final de la sesión, los panelistas puntúan de nuevo de forma anónima e individual, cada indicación.

La adaptación DEMOCLI elimina la sesión presencial de la 2ª ronda, supliendo la interacción personal de los panelistas por el envío de un informe de resultados de la primera ronda (estadístico y gráfico) con todos los comentarios críticos y las explicaciones de posturas personales que el resto de panelistas facilitaron durante la 1ª ronda. Disponer de un tiempo tranquilo y confidencial para la reflexión personal, sin recibir presiones ni influencias de otros panelistas, ni de propio Comité Científico, minimiza el riesgo de sufrir algún tipo del “sesgo de autoridad”, tan común en las reuniones presenciales de trabajo grupal (131).

Por otra parte, evitar la convocatoria de una sesión presencial y el desplazamiento de un panel de expertos numeroso a un lugar único de reunión facilita completar el procedimiento en un plazo razonable, reduce costes y mejora la adherencia de los expertos al proyecto hasta su conclusión.

La tabla II compila información técnica sobre el procedimiento de encuestas realizado en los distintos estudios. Esta y la siguientes tablas están ordenadas por fecha de publicación de los originales correspondientes a cada estudio. En todos los casos se empleó el método Delphi modificado en dos rondas que se desarrolló, de acuerdo a las previsiones temporales de cada estudio, en un periodo máximo de

tres meses (en dos tercios de los estudios, el proceso consiguió completarse en sólo dos meses naturales consecutivos).

La circulación de encuestas por vía electrónica (adjuntas a mensajes personales de email) aceleró el proceso y facilitó el seguimiento de quienes retrasaron su cumplimentación sobre los plazos previstos. Un seguimiento estrecho de esta cuestión por la oficina técnica de los estudios resultó determinante del éxito de todos los proyectos. En todos ellos, quienes aceptaron participar como expertos completaron las dos rondas de valoración solicitadas. El último proyecto (DOLOR 2013), que dispuso de medios para ello, permitió trabajar a los panelistas con mayor comodidad cumplimentando las rondas de encuesta sobre un formulario electrónico web diseñado ad-hoc para el proyecto que, además, automatizó parte del análisis descriptivo de los resultados entre rondas y la recogida de opiniones libres de los expertos.

Otro detalle técnico que facilita la rápida resolución del proceso de encuesta tiene que ver con el desenlace buscado de la primera ronda de estudio: se trata de obtener el refrendo o rechazo expreso del grupo respecto al contenido de cada ítem presentados por el Comité Científico. Por ello, salvo que algún panelista exprese un veto explícito que merezca la consideración del comité científico, las cuestiones analizadas que resultan consensuadas en la primera ronda de encuesta, se eliminan de la encuesta de la segunda vuelta. Esto reduce en gran medida el tamaño del segundo cuestionario, y aligera la tarea de revisión y reconsideración de las cuestiones sobre las que se mantuvieron posiciones discrepantes (polarización de opiniones contrarias) o se detectó una indeterminación de criterio (desagregación de posiciones entre distintos panelistas) en la ronda inicial.

Tabla II: Contenidos e implementación de las encuestas en los estudios presentados

Nombre breve	Dislipemia	Asma	Epilepsia	Pediatría	Depresión	Esquizofrenia	Inhaladores	Próstata	Dolor
Fecha de encuesta	Marzo-Abril 2008	Mayo-Junio 2009	Mayo-Julio 2009	Julio-Sept 2008	Abril-Junio 2008	Octubre-Nov 2009	Sept-Octubre 2010	Mayo-Junio 2011	Feb-Marzo 2012
Canal de encuesta	email	email	email	email	email	email	email	email	Formulario WEB
Nº de ítems <i>(criterios o recomendaciones clínicas evaluadas)</i>	46	56	194	70	54	66	40	64	61
Procedencia de ítems de encuesta	Revisión bibliográfica	Documento previo	Revisión bibliográfica	Documento previo y revisión bibliográfica	Revisión bibliográfica	Documento previo y revisión bibliográfica	Documento previo	Revisión bibliográfica	Revisión bibliográfica
Escala de acuerdo	Likert 5 puntos	Likert 5 puntos	Likert 5 puntos	Likert 5 puntos	Likert 5 puntos	Likert 9 puntos	Likert 9 puntos	Likert 9 puntos	Likert 9 puntos
Criterio de consenso	Media de puntuaciones e IC95%	Media de puntuaciones e IC _{95%}	Media de puntuaciones e IC95%	Media de puntuaciones e IC95%	Media de puntuaciones e IC95%	Mediana de puntuaciones y nivel concordancia	Mediana de puntuaciones y nivel concordancia	Mediana de puntuaciones y nivel concordancia	Mediana de puntuaciones y nivel concordancia

3.6. Criterios de análisis estadístico e interpretación de resultados de los consensos:

Disponer de un criterio estadístico expreso, conocido y aceptado a priori por los panelistas, sobre cuando se considerarán suficientemente unánimes las opiniones del panel para aceptarlas como aprobadas por consenso, allana la resolución del proceso, evita malentendidos y acelera su aprobación. En la tabla II se resume información al respecto de los nueve estudios.

Para analizar la opinión grupal y el tipo de consenso alcanzado sobre cada cuestión planteada, en los estudios que emplearon escalas Likert de 9 puntos, se empleó la posición de la mediana de puntuaciones del grupo y el “nivel de concordancia” alcanzado por los encuestados, según los criterios que a continuación se detallan (132,133).

Se consideró consensuado un ítem cuando existía “concordancia” de opinión en el panel. Es decir, cuando los expertos que puntuaron fuera de la región de tres puntos ([1-3], [4-6], [7-9]) que contenía la mediana, eran menos de la tercera parte de los encuestados. En tal caso, el valor de la mediana determinaba el consenso grupal alcanzado: “*desacuerdo*” mayoritario con el ítem, si la mediana era igual o inferior a 3, o “*acuerdo*” mayoritario con el ítem si la mediana era igual o superior a 7. Los casos en los que la mediana se encontraba en el margen 4-6, fueron considerados como ítems “*dudosos*” para una mayoría representativa del grupo.

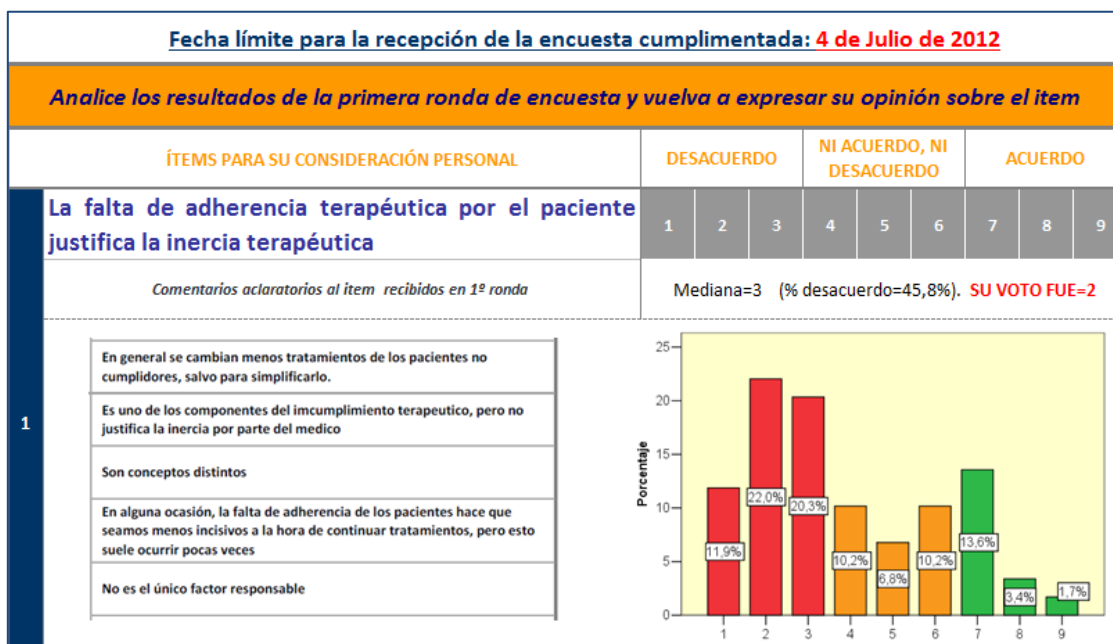
Por el contrario, se estableció que existía “discordancia” de criterio en el panel cuando las puntuaciones de un tercio o más de los panelistas estaban en la región [1-3], y de otro tercio o más en la región [7-9]. Los ítems restantes en los que no se observaba concordancia ni discordancia, se consideraron como nivel de consenso “indeterminado”.

Todos los ítems en los que el grupo no alcanzó un consenso manifiesto a favor o en contra de la cuestión planteada (los ítems dudosos, aquellos en los que se apreció discordancia y los que mostraron un nivel de consenso indeterminado) se propusieron a la reconsideración del panel en la segunda ronda Delphi. También se

sometieron a reevaluación los ítems en los que se apreció una alta dispersión de opiniones entre los encuestados, con un rango intercuartílico igual o superior a 4 puntos (rango de puntuaciones contenidas entre los valores p25 y p75 de la distribución).

Entre ambas rondas, los panelistas fueron informados de la distribución detallada de las respuestas del grupo en la primera encuesta (mediante gráficos de barras e indicadores de centralización y dispersión no paramétricos), facilitándoseles, en cada ítem, un recuerdo de su propia opinión, y los comentarios y aclaraciones aportadas por el resto de participantes sobre la cuestión. En la figura 8 se reproduce un ejemplo de esta información intermedia facilitada a los panelistas para su reflexión. Tras revisar esta información, se les solicitó la posible reconsideración de sus posiciones, y una nueva valoración de los ítems no consensuados en la primera ronda.

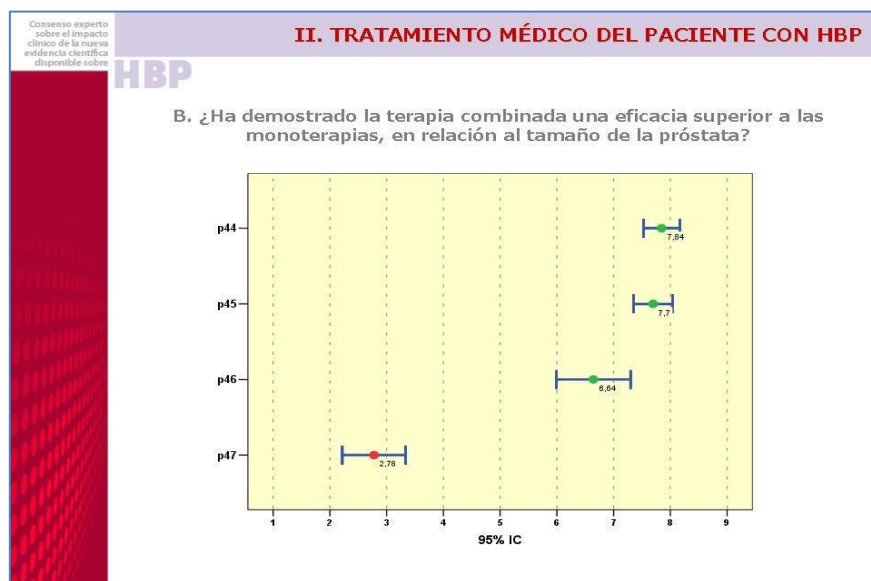
Figura 8: Ejemplo de retroalimentación informativa tras la 1ª ronda (gráficos e indicadores estadísticos, comentarios y aclaraciones aportadas por participantes) ofrecida cómo información intermedia a los panelistas para su reflexión.



Tras la segunda ronda de encuesta, se aplicaron criterios idénticos para discriminar los ítems definitivamente consensuados de aquéllos en los que no resultó posible unificar el criterio del panel. Los ítems en los que no se logró un consenso tras completar el proceso descrito se analizaron de forma descriptiva, para distinguir si tal situación era debida a que persistía discordancia de criterio, o al posicionamiento mayoritario del panel en la región de duda respecto al ítem, cuando la mayoría del grupo declaraba no tener un criterio definitivo (voto entre 4 y 6).

A efectos de comparación gráfica entre ítems, se calculó el promedio de puntuación de los panelistas en cada cuestión, con su intervalo de confianza al 95% (IC^{95%}). Cuanto más extremo resultó el promedio de puntuaciones de un ítem (más próximo a 1 o a 9), más manifiesto se consideró el consenso logrado, bien en el desacuerdo o en el acuerdo respectivamente, sobre la propuesta de cada ítem. En la figura X se reproduce este análisis gráfico sobre la encuesta de uno de los estudios, en el que se puede comparar el grado de respaldo de los expertos a algunos ítems de encuesta.

Figura 9: Ejemplo de representación comparativa de puntuaciones medias (e IC95%) entre distintos ítems de una encuesta



En el caso de los estudios que emplearon escalas Likert de cinco puntos la valoración del nivel de consenso alcanzado por cada ítem se analizó mediante el cálculo de los valores promedio de puntuaciones de cada ítem y su correspondiente intervalo de confianza al 95% (IC^{95%}). Se consideraron consensuados por el panel los ítems en los que el límite superior del IC^{95%} resultó inferior a 3 (acuerdo del panel con la afirmación), o aquellos en los que el límite inferior del IC^{95%} resultó superior a 3 (desacuerdo con la afirmación). Los ítems restantes que incluyeron el valor 3 en el IC^{95%} de la media se propusieron a la reconsideración de los miembros del panel en la segunda ronda Delphi.

Un criterio idéntico se aplicó en la segunda ronda para discriminar los ítems definitivamente consensuados de aquellos en los que no fue posible unificar el criterio del panel. A efectos comparativos, cuanto más extremo resulta el promedio de puntuaciones de un ítem (más próximo al valor 1 o al 5) más manifiesto se considera el consenso logrado, bien en el acuerdo o en el desacuerdo, respectivamente, sobre la propuesta expresada por cada ítem. Además, cuanto más estrecho sea el recorrido del intervalo de confianza, mayor unanimidad se constata entre las opiniones del grupo.

Aunque esta metodología de análisis está bien establecida y ha sido empleada en estudios previos similares (127, 129, 130,), se verificó el consenso resultante bajo criterios estadísticos alternativos más exigentes, usados por otros autores en estudios que emplearon las mismas escalas de evaluación (134, 135). Entre ellos, se comprobó la existencia de un coeficiente de variación menor de 0,3, una puntuación promedio inferior a 2,5 o superior a 3,5, la suma de respuestas mayoritarias superior al 70% del total de respuestas (1+2 o 4+5, respectivamente, para el acuerdo y el desacuerdo), y una mediana distinta al punto central = [3].

El resto de procedimientos del método no cambian en modo alguno por causa de la escala de Likert empleada (5 o 9 puntos).

3.7. Contexto clínico de los estudios presentados: tema, ámbito de investigación y fin estratégico del consenso pretendido

La tabla III recoge información comparativa sobre la temática y finalidad principal de los proyectos. En cuanto a esta última, los estudios se pueden clasificar en tres grupos:

1º grupo: Estudios que desarrollan, ex-novo, unas recomendaciones clínicas inéditas hasta la fecha, mediante un consenso Delphi:

Cuatro de los nueve estudios analizados (DISLIPEMIA 2009, EPILEPSIA 2010, DEPRESION 2011, PRÓSTATA 2012 y DOLOR 2013) desarrollaron el proceso completo de formulación de unas nuevas recomendaciones clínicas en sus temas de interés. En tal caso el desarrollo de los estudios incluyó todas las fases comentadas más atrás, con el fin de elaborar unas propuestas de actuación clínica de nuevo cuño: un comité científico de liderazgo en la materia realizó un proceso de búsqueda y lectura crítica de la bibliografía disponible sobre la cuestión al inicio del estudio; a partir de la información localizada (disponibilidad y calidad de las evidencias) el propio comité formuló una batería de recomendaciones clínicas; esta propuesta inicial se planteó a un panel de expertos, mediante proceso de encuesta a distancia en dos rondas, durante el que, ocasionalmente, se pudieron añadir, eliminar o modificar algunos ítems a sugerencia de los panelistas; completado el proceso de encuestas se determinó estadísticamente el grado de acuerdo/desacuerdo de los expertos con cada recomendación y la versión final del consenso apto para su publicación. En ninguno de los estudios antedichos, existía una versión previa de recomendaciones clínicas para la toma de decisiones en los ámbitos de incertidumbre clínica tratados por cada uno.

Grupo 2º: Estudios que actualizan unas recomendaciones clínicas preexistentes pero obsoletas, mediante el método Delphi.

A diferencia de los anteriores, el estudio PEDIATRIA 2010, planteó por primera vez en nuestro país el uso del método Delphi como una técnica dirigida a la revisión y puesta al día de unas recomendaciones clínicas preexistentes sobre la materia

(prevención de las infecciones graves por VRS en niños cardiopatas). Con este proyecto, se revisaron y pusieron al día, a fecha 2010, las recomendaciones previas de la Sociedad Española de Cardiología Pediátrica y Cardiopatías Congénitas difundidas en 2005 (136), que en 2010, y a la luz de recientes datos epidemiológicos y terapéuticos, eran objeto de controversia por considerarse insuficientes y desactualizadas. En cuanto a su metodología, el proceso fue prácticamente idéntico al descrito en el bloque anterior de estudios, por cuanto requirió, antes de elaborar el cuestionario de estudio, un procedimiento preliminar de revisión bibliográfica de las controversias latentes y de las pruebas científicas disponibles en la literatura sobre la materia de actualidad en el momento de su realización.

Grupo 3º: Estudios que validan unas recomendaciones clínicas actuales, mediante el refrendo experto de un panel Delphi.

Los restantes estudios (ASMA 2010, INHALADORES, 2012 y ESQUIZOFRENIA 2012) emplearon el método Delphi con otra finalidad estratégica específica: realizar la validación externa experta, por un grupo numeroso de clínicos de referencia, de las recomendaciones clínicas de una serie de guías preexistentes, de plena actualidad científica por su reciente publicación previa, con el fin principal de impulsar su limitada difusión y aplicación conseguida hasta el momento, a saber:

- “Guía Española para el Manejo del Asma. Grupo interprofesional GEMA, 2009”.
- “Terapia inhalada. Teoría y práctica. Grupo interprofesional Inhaladores, 2010”.
- “Consenso sobre la salud física del paciente con esquizofrenia de las Sociedades Españolas de Psiquiatría y de Psiquiatría Biológica, 2008”

En todos estos casos se buscaba que un amplio círculo multicéntrico y estatal de expertos de reconocido criterio en cada materia, analizaran y, en su caso, refrendasen las recomendaciones publicadas por un grupo profesional en el que no habían participado personalmente. Con tal iniciativa, los impulsores de los

proyectos pretendieron, por una parte, promover la más amplia difusión de calidad (entre los especialistas de referencia en la materia de interés) de las recomendaciones preexistentes y, por otra, ayudar a dar carta de naturaleza a dichos documentos como estándares a nivel estatal, precisamente por gozar del respaldo expreso de una amplia red de profesionales-clave (líderes de opinión en sus respectivas especialidades), procedentes todos los entornos asistenciales y geográficos del sistema sanitario español. La robustez de la metodología Delphi para seleccionar y para poner en juego sin sesgos a los panelistas participantes, fue el aval técnico del pretendido apoyo. En la discusión de la tesis se argumenta sobre la necesidad general de mejorar los procedimientos de difusión de las guías de práctica clínica y se analiza el porqué de la utilidad el método Delphi para tal fin.

La tabla III ofrece, además, otros detalles sobre la naturaleza específica del ámbito de controversia clínica sometido a debate en cada uno de los nueve estudios.

Al respecto, basta señalar la amplia variedad de contextos clínicos y especialidades que fueron analizados por los distintos proyectos, la diversidad de poblaciones-diana a la que se dirigen las recomendaciones producidas y el distinto grado de amplitud de foco de los estudios sobre el proceso asistencial de cada patología, desde una parte concreta de la actuación clínica sobre el problema estudiado (limitada, por ejemplo, al inicio del tratamiento antiepiléptico, en el estudio EPILEPSIA 2012, o a la prevención de las infecciones por VRS, en el estudio PEDIATRÍA 2012), hasta un barrido completo de recomendaciones sobre todas las áreas posibles de un proceso de atención integralmente concebido (prevención, diagnóstico, tratamiento, rehabilitación y pronóstico y cuidado paliativo). En la discusión se analiza el alcance y posible significado de esta variabilidad intrínseca de los estudios.

Tabla III: Objetivos, contenido temático y ámbito de los consensos

Nombre breve	Asma	Depresión	Epilepsia	Pediatría	Dislipemia	Esquizofrenia	Inhaladores	Próstata	Dolor
Tipo de consenso	Validación de recomendac. pre-existentes	Ex-novo	Ex-novo	Actualización de consenso previo	Ex-novo	Validación y ampliación de recomendac. pre-existentes	Validación de recomendac. pre-existentes	Ex-novo	Ex-novo
Tópico del consenso estudio	Manejo clínico del asma bronquial	Manejo de depresión en ancianos	Inicio de terapia antiepiléptica	Prevención infección VRS en cardiopatías	Uso clínico de ezetimiba	Salud física en la esquizofrenia	Terapia inhalada	Manejo clínico de la hiperplasia benigna próstata	Tratamiento tópico dolor neuropático periférico
Perfil demográfico de paciente diana	Adulto Infantil	Anciano	No especificado	Infantil	No especificado	No especificado	Adulto Infantil	No especificado	No especificado
Tipo de decisiones clínicas analizadas	Proceso asistencial completo	Proceso asistencial completo	Terapéutica	Prevención	Terapéutica	Proceso asistencial completo	Terapéutica	Proceso asistencial completo	Terapéutica

3.8. Selección de panel de expertos participantes en los consensos:

La selección del panel experto es un punto crítico en el desarrollo cualquier procedimiento de consenso profesional, al ser una tarea no exenta del riesgo de influencia por parte del Comité Científico que impulsa el proyecto. Este riesgo es particularmente delicado en la formulación tradicional del método Delphi modificado, dado el limitado tamaño del grupo (frecuentemente, entre 9 y 12 expertos), donde el reemplazo de un solo miembro por otro de distinto perfil podría resultar determinante de cambios en las conclusiones del consenso. Como se discute más adelante, se propone, a la luz del desarrollo y los resultados de los nuevos estudios presentados, una adaptación del método Delphi modificado que plantea el reclutamiento de un panel de expertos clínicos sustancialmente más numeroso, lo que reduce este riesgo.

Para garantizar la identificación y reclutamiento de aquellos expertos a quienes más mérito se les reconoce y más crédito se les concede dentro de su propia especialidad, la literatura recomienda buscar profesionales que reúnan el siguiente perfil ideal de atributos personales (137), como criterios de selección (a ponderar en cada estudio):

- Liderazgo clínico y científico reconocido.
- Amplitud de conocimiento e interés en el tema.
- Actitud y aptitud científica.
- Mérito docente o institucional.
- Diversidad geográfica.
- Diversidad de tipo de hospital y servicio.
- Capacidad y disponibilidad de trabajo en equipo, ausencia de visiones rígidas.
- Alto nivel de motivación intrínseca.
- Disponibilidad real de tiempo.
- Ausencia de conflicto de intereses, entendidos como aquellas situaciones, en la que puede encontrarse cualquier profesional, cuando su posición ante un problema puede verse influenciada por un interés secundario (tales como

obtención de un beneficio económico o similar) que le induzca a estar claramente a favor o en contra.

Para localizar tales perfiles, en los estudios presentados se han empleado, en general, dos procedimientos técnicos:

- Una estrategia de muestreo “en bola de nieve” (también llamado de arrastre, en cadena, o nominado), de acuerdo a la propuesta original de Goodman y Coleman (*snowball sample*) (138), iniciada a partir de los contactos personales de los miembros del Comité, quienes, a su vez han de proponer nuevos candidatos referentes en el tema, a quienes se solicita nominar a otros expertos de mérito indiscutible de su conocimiento, entorno profesional y ámbito geográfico. Tras identificar al menos dos generaciones consecutivas de posibles aspirantes, se eligen prioritariamente aquellos que resultan propuestos como referentes (bajo los criterios antedichos) desde las redes de contactos de al menos dos “antecesores”.
- En algunos estudios, el proceso anterior se completó con una búsqueda activa de autores nacionales de artículos originales relacionados con el tema de estudio en las bases de datos bibliográficas convencionales.

La tabla IV recoge información comparativa del tamaño, composición, procedencia y forma de selección del panel de expertos que abordó el procedimiento de consenso en cada estudio.

Dos tercios de los estudios se desarrollaron mediante el consenso de paneles conformados por expertos de la única especialidad responsable de la atención clínica a la patología a debate. Los tres estudios restantes emplearon paneles multidisciplinarios compuestos por profesionales de distintas especialidades médicas y ámbitos asistenciales, todas ellas implicadas en el manejo conjunto de la cuestión clínica tratada en el estudio. Así, en el panel del estudio ASMA 2010 participaron alergólogos, neumólogos, médicos de familia, alergólogos y neumólogos pediátricos, y especialistas en ORL. En el estudio INHALADORES 2012 colaboraron neumólogos, alergólogos, médicos de familia e internistas. Finalmente,

en el estudio DOLOR 2013 el panel se formó por anestesiólogos, neurólogos y traumatólogos. En conjunto, han participado en los distintos paneles expertos de los estudios presentados representantes de hasta 13 especialidades médico-quirúrgicas distintas.

Todos los estudios, por la amplia procedencia geográfica de los expertos reclutados, fueron proyectos multicéntricos de ámbito nacional. La vía electrónica facilitó en gran medida el contacto y coordinación de grupos de participantes tan dispersos. En cinco proyectos, por interés expreso de los promotores, se intentó distribuir la composición del panel entre expertos procedentes de todas las comunidades autónomas (peninsulares e insulares) de España. Los proyectos se consideraron de ámbito estatal cuando se logró garantizar una representación efectiva de expertos de cada comunidad. En ningún caso se registraron negativas expresas y completas a participar de los candidatos procedentes de una CA determinada que impidieran dicho objetivo. La finalidad de este procedimiento fue la de promover localmente la visibilidad del proyecto y fomentar su posterior aceptación generalizada, como se discutirá más adelante.

Para la identificación de los candidatos concretos a formar parte del panel experto de 8 de los 9 estudios, se desplegó la técnica de muestreo en bola de nieve descrita más atrás, iniciada desde la red de contactos profesionales del propio Comité Científico de cada estudio. Cada candidato localizado e invitado a participar estuvo, pues avalado por el respaldo expreso de al menos dos redes de referencia profesional directa de su misma especialidad. En los estudios en los que se intentó disponer de un panel de ámbito estatal, se desencadenó un proceso independiente de arrastre como el descrito en cada comunidad autónoma, para asegurar la identificación y selección de candidatos locales. El proceso se mostró rápido y eficaz en todos los casos.

El estudio ESQUIZOFRENIA 2010, dado el carácter innovador del tema y la relativa escasez de especialistas referentes en la materia, añadió al proceso de *snowball sample*, una búsqueda activa de autores nacionales de artículos originales relacionados con el tema de estudio (salud física del paciente con trastorno mental

grave), localizados mediante en una revisión bibliográfica de las bases de datos documentales PUBMED (Medline), EMBASE (Excerpta Médica) e IME (Índice Médico Español).

A diferencia de todos los anteriores, el comité científico del proyecto ASMA 2010, delegó en los órganos directivos colegiados de las 9 Sociedades Científicas que habían copatrocinado la guía GEMA, la identificación y propuesta del subgrupo de expertos que participaron en el panel en representación formal de sus respectivas especialidades y asociaciones profesionales. Esta distribución se hizo de forma proporcional a la composición del grupo redactor de la guía previa que fue objeto de análisis en el proyecto.

En general, los candidatos que rehusaron participar en los proyectos presentados declararon hacerlo por problemas de agenda profesional o circunstancias personales durante el periodo previsto para el trabajo de campo. Otros fueron rechazados por los respectivos comités científicos por conflictos de interés con el proyecto (por ejemplo, en los estudios ASMA 2010, INHALADORES, 2012 y ESQUIZOFRENIA 2012, se vetó a los quienes habían colaborado de forma directa o indirecta en la redacción de los documento-base objeto de análisis en cada proyecto, y a quienes declararon cualquier otro motivo de colisión de intereses con el desarrollo, la difusión o la implementación de dichas guías).

El tamaño de los paneles de los distintos estudios, en su composición final, osciló entre 15 y 87 participantes. Para conseguir la participación de estos profesionales, se contactó electrónicamente con una cantidad algo mayor de potenciales candidatos, a los que se les facilitó información preliminar del proyecto (entre otra, una breve justificación de sus objetivos, el tipo y duración estimada de la tarea propuesta, el perfil y método de identificación de los miembros del panel, y la identidad de los responsables científicos del proyecto), invitándoles a colaborar en el estudio. En este momento se garantizó plena confidencialidad sobre la decisión que adoptase cada especialista contactado y sobre su filiación. Cuatro de los estudios tuvieron una aceptación plena de los candidatos (100%). Los restantes estudios, salvo el proyecto PEDIATRIA 2010 que reclutó al 73,3% de los invitados,

alcanzaron muy altas tasas de participación (superior al 92% en todos los casos). En este último, la fecha elegida para el trabajo de campo (julio-septiembre de 2008), incluyendo el periodo estival, pueden resultar determinante del menor reclutamiento

En la experiencia de la oficina técnica conductora de los estudios, explicar a los invitados el sistema de identificación de los candidatos (basado en el reconocimiento profesional de sus compañeros de especialidad y/o su sociedad científica como experto de reconocido criterio en el tema a debate) y el hecho de que se tratase de una tarea a distancia de corta duración en tiempo, resultaron elementos determinantes de la aceptación general y de la alta tasa de reclutamiento logrados. Debe recordarse aquí que la participación solicitada fue, en todos los casos, voluntaria, anónima y no remunerada.

En general, cabe observar que se trata de paneles de expertos clínicos de un tamaño notable, en especial comparados con los grupos que suelen manejarse en estudios Delphi convencionales en otras áreas de conocimiento (en la mayoría de los proyectos con intención predictiva, por ejemplo, en el ámbito de la economía y la empresa, los grupos consultados suele ser más limitados en número, de 8-12 sujetos). Pero esta no es una cuestión sometida a restricciones metodológicas (66), sino un atributo condicionado por la disponibilidad de expertos en cada ámbito de estudio, y por la accesibilidad a los mismos (dependiente de ciertos usos profesionales diferenciales de las distintas disciplinas).

Como se discutirá más adelante, el concepto particular de quien puede ser considerado un experto clínico hace que los médicos con el nivel de pericia requerido para ello estén ubicuamente disponibles en el sistema sanitario. Además, el compromiso deontológico de la profesión médica con la mejora continuada de la asistencia, les hace sustancialmente más accesibles y participativos que los profesionales de otras materias técnicas.

Tabla IV: Características del panel de expertos de los estudios

Nombre breve	Asma	Depresión	Epilepsia	Pediatría	Dislipemia	Esquizofrenia	Inhaladores	Próstata	Dolor
Tipo de expertos	Multidisciplinar (6)	Psiquiatría Psicogeriatría	Neurología	Cardiología pediátrica	Atención Primaria	Psiquiatría	Multidisciplinar (4)	Urología	Multidisciplinar (3)
Nº panelistas expertos (invitados /aceptan)	80/74 (92,5%)	63/61 (96,8%)	60/60 (100%)	75/55 (73,3%)	91/87 (95,6%)	15/15 (100%)	60/59 (98,3%)	50/50 (100%)	52/52 (100%)
Ámbito geográfico de expertos	Estatal	Multicéntrico	Multicéntrico	Estatal	Multicéntrico	Multicéntrico	Estatal	Estatal	Estatal
Técnica selección	Expertos con representación institucional	Bola de Nieve	Bola de Nieve	Bola de Nieve	Bola de Nieve	Bola de Nieve y búsqueda bibliográfica	Bola de Nieve	Bola de Nieve	Bola de Nieve
Distribución panel	Proporcional por especialidad	Simple	Simple	Simple	Simple	Simple	Proporcional por CCAA	Proporcional por CCAA	Discrecional

Resultados

4

4. RESULTADOS

4.1. Referencia general de las publicaciones científicas derivadas de los estudios presentados en esta tesis:

La **tabla V** recoge información bibliográfica y bibliométrica de los estudios. Estas publicaciones son los resultados tangibles finales con los que se intentará responder a los objetivos de esta tesis. En ella se especifican también detalles de su autoría, auspicio y patrocinio.

Los estudios fueron publicados a lo largo del quinquenio 2009-2013 (116-124). La demora de los procesos de revisión por pares y de aprobación editorial de la versión definitiva de los originales produjeron un retraso medio aproximado de un año en la publicación de cada artículo respecto a las fechas de ejecución técnica de cada proyecto. Estos se desarrollaron entre marzo de 2008, el primero de ellos (DISLIPEMIA 2009), y marzo de 2012, el último (DOLOR 2013). Este detalle resulta de interés para considerar el estado del arte en cada tema de estudio en la fecha límite de revisión bibliográfica realizada durante las tareas preparatorias de cada uno de los estudios.

Todos los originales fueron publicados en revistas científicas indexadas en la base de datos MEDLINE (National Library of Medicine) que ofrecían, en la fecha de cada publicación, un índice de impacto variable entre 0,157 y 1.994 según el JCR Journal Impact Factor (Journal Citation Report de Thomson Reuters).

Pese a tratarse de publicaciones destinadas principalmente a una difusión nacional y a la aplicación local de sus recomendaciones, por la singularidad e interés de sus aportaciones, 2 de los 9 estudios fueron publicados en revistas en lengua inglesa y difusión internacional (DISLIPEMIA 2009 y EPILEPSIA 2010).

Seis de los nueve estudios fueron impulsados por sociedades científicas legalmente constituidas de diferentes especialidades, lo que acredita el interés y la pertinencia de los temas de controversia examinados en cada investigación, para el

colectivo profesional correspondiente. En general, estas sociedades se implicaron activamente en la supervisión de los proyectos y, de forma particular, en la tarea de identificación y selección de los expertos participantes en los paneles de debate de cada estudio. En tres ocasiones, (DISLIPEMIA 2009, INHALADORES 2012 y DOLOR, 2013) los comités científicos de los estudios los compusieron grupos de profesionales líderes de referencia en cada tema en cuestión, miembros de derecho de sus correspondientes asociaciones científicas, pero promotores a título personal de los estudios sin el auspicio corporativo de la institución.

Cuando el ámbito de incertidumbre abordado concernía transversalmente a distintas especialidades implicadas conjuntamente en la atención a un mismo problema clínico, los estudios se desarrollaron como proyectos colaborativos interprofesionales entre varias sociedades. Estos proyectos conllevan una complejidad añadida, derivada de la necesidad de ordenar las participaciones y de equilibrar los protagonismos de cada asociación científica, representantes, en ocasiones, de especialidades en litigio por la atención a la patología de interés.

El estudio ASMA 2010 resultó un paradigma de tal circunstancia por incluir entre sus promotores a 9 sociedades científicas relacionadas con la enfermedad, de las que procedieron sus 75 panelistas expertos: SEPAR (Sociedad Española de Neumología y Cirugía Torácica), SEAIC Sociedad Española de Alergología e Inmunología Clínica), semFYC (Sociedad Española de Medicina Familiar y Comunitaria), SEMERGEN (Sociedad Española de Médicos de Atención Primaria), SEMG (Sociedad Española de Médicos Generales y de Familia), GRAP (Sociedad de Respiratorio de Atención Primaria), SEICAP (Sociedad Española de Inmunología Clínica y Alergia Pediátrica), SENP (Sociedad Española de Neumología Pediátrica) y SEORL (Sociedad Española de Otorrinolaringología).

Salvo este último estudio, financiado directamente por el propio grupo impulsor de la guía GEMA ("Guía Española para el Manejo del Asma", edición 2009), cuyas recomendaciones fueron sometidas a reevaluación y refrendo en el consenso ASMA-2010, el resto de proyectos han contado con soporte financiero de distintos

representantes de la industria farmacéutica, a instancias de las propias sociedades científicas y/o grupos profesionales impulsores.

Como se expresa en las declaraciones de conflicto de interés de cada publicación, la labor de los laboratorios comerciales se limitó a garantizar el soporte financiero al comité científico de cada estudio para las tareas de desarrollo del trabajo de campo y de encuesta electrónica, que en todos los casos fueron asumidas por una consultora independiente de investigación por contrato. Los laboratorios promotores no participaron en modo alguno, ni supervisaron, el diseño de los estudios, el análisis de los datos, la redacción científica de sus conclusiones, ni la presentación de los originales finalmente publicados.

Los autores firmantes de cada estudio fueron los propios miembros del comité científico de cada proyecto (a título personal o en representación de su sociedad correspondiente), y el doctorando de esta tesis, como asesor metodológico y experto en la técnica, resultando grupos de entre 4 (DISLIPEMIA 2009) y 11 coautores (ASMA 2010). A fin de garantizar transparencia sobre la procedencia de los resultados de cada estudio, en todos los casos se referenció en un anexo del original la composición detallada del panel de expertos y la identidad concreta de cada miembro.

Tabla V: Referencia bibliográfica y bibliométrica, coautoría, auspicio y patrocinio de los estudios

Nombre breve	Asma	Depresión	Epilepsia	Pediatría	Dislipemia	Esquizofrenia	Inhaladores	Próstata	Dolor
Revista	<i>Arch Bronconeumol</i>	<i>Actas Esp Psiquiatr</i>	<i>Epilepsy & Behav</i>	<i>An Pediatr (Barc)</i>	<i>Av Diabetol</i>	<i>Actas Esp Psiquiatr</i>	<i>Arch Bronconeumol</i>	<i>Actas Urol Esp</i>	<i>Rev Soc Esp Dolor</i>
Año	2010	2011	2010	2010	2009	2012	2012	2012	2013
Idioma	Español	Español	Inglés	Español	Inglés	Español	Español	Español	Español
Factor de Impacto (*)	0.347 (**)	0.589	1.994	0.570	--	0.452	1.372	1.144	0.157
Auspicio Sociedades Científicas	9	1	1	1	No (grupo experto multidisciplinar)	2	No (grupo experto multidisciplinar)	1	No (grupo experto multidisciplinar)
Miembros Comité Científico	10	7	5	8	3	3	6	6	6
Patrocinio	Grupo GEMA	Lilly	UCB	Abbott	Schering-Plough / MSD	Lilly	Chiesi	GSK	Astellas

(*)Fuente: JCR Journal Impact Factor (Journal Citation Report, Thomson Reuters), excepto (**) SJR (Scimago Journal of Country Rank)

4.2. Breve resumen del marco teórico justificativo, finalidad y resultados principales de los estudios compilados:

Aunque para responder a los objetivos de esa tesis, ni el tema específico de cada estudio, ni sus resultados concretos (las recomendaciones clínicas consensuadas en cada caso) resulten asuntos de interés crítico, se presentan a continuación, con brevedad y en términos generales, algunos aspectos clave del contexto de controversia clínica que justificó cada proyecto.

Se pretende con ello familiarizar al lector con el entorno profesional en el que se desarrolló cada trabajo y aclarar la naturaleza de la incertidumbre clínica que se abordó, esto es, la materia que, cuando se inició cada estudio, era objeto de debate profesional o no tenía una respuesta expresa en pruebas científicas de alto nivel.

En la presentación de cada estudio, el comentario concluye con un juicio cualitativo sintético sobre el logro que representan los resultados de cada proyecto en su ámbito de interés. A modo de resumen introductorio, la tabla VI permite comparar en detalle los resultados cuantitativos concretos de cada estudio, en términos de las conclusiones finales alcanzadas sobre la batería de recomendaciones clínicas propuestas a consenso profesional.

Tabla VI: Resultados de los estudios. Grado de consenso final alcanzado sobre las recomendaciones analizadas

Nombre breve	Asma	Depresión	Epilepsia	Pediatría	Dislipemia	Esquizofrenia	Inhaladores	Próstata	Dolor
Nº de ítems (criterios o recomendaciones clínicas evaluadas)	56	54	194	70	46	66	40	64	61
Ítems consensuados en 1ª ronda	49 (87,5%)	39 (72,2%)	135 (69,6%)	44 (62,9%)	34 (73,9%)	46 (69,6%)	35 (87,5%)	50 (78,1%)	37 (60,6%)
Ítems consensuados final de estudio	53 (94,6%)	46 (85,2%)	148 (76,3%)	54 (77,1%)	41 (89,1%)	61 (92,4%)	30 (97,5%)	59 (92,2%)	46 (75,4%)
Criterio grupal sobre ítems A: acuerdo D: desacuerdo NC: no consenso	A: 53 (94,6%) D: 0 (0,0%) NC: 3 (5,4%)	A: 32 (59,3%) D: 14 (25,9%) NC: 8 (14,8%)	A: 85 (43,8%) D: 63 (32,5%) NC: 46 (23,7%)	A: 52 (74,3%) D: 2 (2,8%) NC: 16 (22,9%)	A: 37 (80,4%) D: 4 (8,7%) NC: 5 (10,9%)	A: 61(92,4%) D: 0 (0,0%) NC: 5 (7,6%)	A: 39 (97,5%) D: 0 (0,0%) NC: 1 (2,5%)	A: 58 (90,6%) D: 1 (1,6%) NC: 5 (7,8%)	A: 40 (65,6%) D: 6 (9,8.6%) NC: 15 (24,6%)
Meeting final de presentación	NO	NO	SI	NO	SI	SI	SI	SI	NO

1) Estudio DISLIPEMIA (Ampudia y cols, 2009)

El proyecto DISLIPEMIA 2009 analizó, apenas cinco años después de la autorización de ezetimiba en España como fármaco hipolipemiante, su rol en el control del riesgo cardiovascular asociado a la hipercolesterolemia. En su momento, se consideró una innovación terapéutica que atrajo poderosamente la atención de profesionales y pacientes.

El novedoso mecanismo de acción de este nuevo agente, sinérgico respecto al tratamiento de referencia entonces de la dislipemia (estatinas), produjo expectativas inmediatas sobre la posibilidad de optimizar el control de pacientes dislipémicos de alto riesgo, hasta la fecha expuestos a un insuficiente descenso de los niveles lipídicos plasmáticos, o a los efectos adversos de las altas dosis de estatinas necesarias para ello.

El estudio se focalizó en atención primaria, el entorno natural de detección y seguimiento de estos pacientes, dada las escasas pruebas disponibles sobre los resultados y riesgos del uso de la nueva molécula en la población general. La incertidumbre procedía de que, en ese momento, toda la información accesible derivaba de los grandes ensayos clínicos pivotaes de desarrollo de la molécula, realizados sobre pacientes muy seleccionados y de un perfil distinto al de la población atendida en atención primaria.

En los foros científicos sobre la cuestión se reiteraban controversias sobre el manejo (en monoterapia o en uso combinado) que los médicos de familia debían hacer de este nuevo agente terapéutico con sus pacientes, sin que se hubiera destacado una posición suficientemente mayoritaria. En particular arreciaba un debate entre los clínicos responsables de la atención directa a los pacientes y los servicios centrales responsables de la gestión de la prestación farmacéutica en el que se dirimían, con un argumentario necesariamente débil, cuestiones prácticas de coste/beneficio y oportunidad.

En este contexto, un comité multiprofesional de expertos en dislipemias (procedente de atención primaria, endocrinología y medicina interna),

promovieron el proyecto DISLIPEMIA 2009 en búsqueda de un consenso profesional que ayudase a clarificar la situación para los clínicos responsables de tomar decisiones y de la prescripción de hipolipemiantes.

Para ello, se convocó un panel multicéntrico estatal en el que colaboraron 87 médicos expertos del primer nivel asistencial particularmente interesados y capacitados en la valoración clínica y manejo de la dislipemia (pertenecientes a grupos de trabajo de sociedades científicas, referentes docentes o investigadores en su entorno asistencial).

Finalmente, el panel emitió por consenso 42 recomendaciones clínicas explícitas sobre aspectos tan concretos como las indicaciones, la eficacia y la tolerabilidad esperable de la molécula, consejos sobre su posología y uso en poblaciones especiales, o algunas consideraciones clínicas basadas en el coste/beneficio de la nueva opción, entre otras cuestiones.

2) Estudio ASMA (Caballero F y cols, 2010):

En el proyecto ASMA 2010 se empleó el método Delphi con el fin de validar externamente una batería de recomendaciones preexistentes sobre el diagnóstico y tratamiento del asma que habían sido publicadas el año previo, bajo el patrocinio conjunto de 9 sociedades científicas españolas de distintas especialidades implicadas en la atención a esta enfermedad (Guía GEMA: “Guía Española para el Manejo del Asma”, edición 2009) (139). Por su calidad metodológica, supervisada por el Centro Cochrane Iberoamericano, esta guía había sido incluida en el catálogo oficial de la iniciativa GuíaSalud (Guías de Práctica Clínica en el Sistema Nacional de Salud - Consejo Interterritorial SNS) (140). En la fecha del estudio, el comité ejecutivo de GEMA estaba ocupado en una ambiciosa estrategia de difusión e implementación de sus contenidos en todas las especialidades interesadas (un aspecto, en general, insuficientemente atendido en muchas otras guías de práctica clínica).

En este contexto, el proyecto ASMA 2010 intentó contrastar el grado de refrendo profesional que las recomendaciones de la guía habían alcanzado en su corta

trayectoria entre un amplio círculo de expertos (74 especialistas) que no había colaborado directamente en su redacción, pero que fueron identificados y propuestos por las distintas sociedades científicas como especialistas de referencia en la patología en los distintos servicios de salud regionales. Se pretendía con ello ratificar el trabajo de análisis y síntesis bibliográfica realizado por los redactores de GEMA y explorar en detalle el criterio de los expertos sobre aquellos aspectos de mayor controversia entre distintas fuentes o en los que la guía no había encontrado el respaldo de una sólida evidencia científica.

El principal logro del proyecto ASMA 2010 fue integrar el conocimiento científico disponible en la fecha con la experiencia clínica acumulada por una amplia red de expertos en asma procedentes de ámbitos especializados y geográficos muy diversos. Se solicitó a estos especialistas que incluyeran en el juicio a las recomendaciones GEMA su convicción personal sobre la aplicabilidad real de cada recomendación en su entorno socio-sanitario, tomando en consideración su experiencia clínica particular y el conocimiento directo de las expectativas de sus pacientes. De esta forma, se añadió a la síntesis de las mejores evidencias propuestas por GEMA el criterio procedente de la maestría individual de profesionales-clave que atesoran una valiosa experiencia clínica, de otra forma difícil de transferir y generalizar.

El consenso experto refrendó ampliamente la mayoría de las recomendaciones GEMA (acuerdo con el 94,6% de su corpus -53 de 56 ítems-) pero también detectó aspectos sobre los que la opinión de los expertos fue incapaz de resolver la controversia, identificándolos como áreas susceptibles de nuevo esfuerzo investigador futuro.

3) Estudio EPILEPSIA (Villanueva y cols, 2010)

El proyecto EPILEPSIA 2010 representa una aportación española de impacto internacional sobre un trastorno neurológico de alta prevalencia como la epilepsia, cuyo abordaje terapéutico estaba siendo discutido a la luz del cambio de paradigma clínico que la difusión, apenas 5 años antes, de una nueva definición de

los estados de epilepsia del International Bureau for Epilepsy (141) había inducido.

Desde entonces, el diagnóstico de epilepsia y su tratamiento no requiere necesariamente la presencia de crisis recurrentes, sino que puede establecerse tras una primera y única convulsión, si están presentes factores predisponentes para su reaparición. El balance beneficio-riesgo de los agentes anticonvulsionantes también había cambiado sustancialmente en los últimos años, con la aprobación de moléculas mejor toleradas y con menor riesgo de interacción medicamentosa que los complejos fármacos clásicos.

A consecuencia de todo ello existía una notable incertidumbre sobre cuando iniciar tratamiento ante una primera crisis y, en su caso, sobre la selección del primer agente en monoterapia con el que intentar su control. Estas decisiones se basan en numerosos factores (entre otros, el tipo de epilepsia, las características de cada fármaco, la edad, el sexo y la comorbilidad el paciente,...) con frecuencia difíciles de ponderar en la práctica neurológica.

Las guías de práctica disponibles en la fecha (142-145), basadas en datos procedentes de ensayos clínicos, no daban respuesta a varias controversias concretas en torno al inicio del tratamiento con antiepilépticos y al uso diferencial de dichos agentes. Consecuencia, en parte, de todo ello, era la llamativa variabilidad de prácticas observada (y criticada) entre los distintos neurólogos.

En este contexto, el estudio EPILEPSIA 2010, impulsado por un grupo independiente de especialistas de referencia, vinculados a un grupo de interés monográfico de la Sociedad Española de Neurología (SEN), consiguió que 60 expertos en epilepsia de ámbito nacional establecieran por consenso recomendaciones expresas para 148 de 194 escenarios expuestos a su consideración, que incluyen recomendaciones para la toma de decisión en una amplia variedad de situaciones clínicas.

Las conclusiones del estudio fueron presentadas en primicia, antes de su publicación internacional, en una sesión plenaria del congreso anual de SEN (Barcelona, 2009).

4) Estudio PEDIATRIA (Medrano y cols, 2010)

El proyecto PEDIATRIA 2010 se desarrolló a instancias de la Sociedad Española de Cardiología Pediátrica y Cardiopatías Congénitas (SECPC), tras la difusión previa (en 2009) de unas recomendaciones generales para la prevención en España de las infecciones respiratorias graves infantiles por el virus respiratorio sincitial (VRS), consensuadas por cuatro asociaciones científicas: Sociedad Española de Neonatología, Sociedad Española de Neumología Pediátrica, Sociedad Española de Hematología y Oncología Pediátrica y la propia SECPC (146).

Este agente causa una importante morbimortalidad evitable en la población pediátrica portadora de factores de riesgo (trastorno pulmonar, prematuridad, síndrome de Down, cardiopatía congénita, ...). Para su prevención está bien establecida la utilidad profiláctica de ciertas medidas higiénicas que limitan la transmisión del virus entre la población en riesgo, y la efectividad de una inmunoprofilaxis intramuscular con anticuerpos monoclonales humanizados (palivizumab) en los periodos epidémicos de la infección. Todo ello ha demostrado evitar infecciones agudas en niños de alto riesgo y reducir las hospitalizaciones pediátricas e ingresos en UCI.

Sin embargo, pocos meses después de la publicación del documento de recomendación interprofesional, se publicó el resultado de un estudio epidemiológico de calidad (estudio CIVIC 21) que evidenció el particular riesgo de hospitalización, entre otros, de niños con cardiopatías leves (147). Las recomendaciones previas no habían establecido una indicación taxativa de la inmunoprofilaxis con palimizumab en tal circunstancia. Además de ello, diferencias en la normativa de autorización y financiación pública del fármaco en las distintas comunidades autónomas habían favorecido una notable variabilidad de prácticas entre distintos centros y una inequidad obvia de acceso a la provisión del fármaco preventivo según el lugar de residencia de cada niño en riesgo.

En tal contexto, SECPCC impulsó la actualización de unas recomendaciones preventivas propias para la infección por VRS y el uso clínico de palivizumab, específicamente orientadas a pediatras y cardiólogos que atendían a niños portadores de cardiopatías congénitas (que, difundidas en 2005, se consideraron obsoletas) (136). La tradicional escasez de pruebas procedentes de investigación experimental en la terapéutica pediátrica, derivada de la rigurosa vigilancia ética de la investigación en niños y de la particular necesidad de prevenir el riesgo iatrogénico en la infancia, hacen especialmente pertinente el fin inspirador del proyecto PEDIATRIA 2010: ampliar el criterio procedente de la escasa evidencia científica para la toma de decisiones clínicas en este ámbito con otras perspectivas derivadas de la experiencia personal de reconocidos expertos en la materia.

Aspectos como la percepción del impacto socio-sanitario de la enfermedad, la viabilidad real de las recomendaciones preventivas en las poblaciones atendidas, el riesgo iatrogénico aceptable y asumible por las familias, o la prioridad que debe darse a las distintas recomendaciones (en términos de beneficio-riesgo, reducción del uso de los servicios sanitarios, o evitación de costes indirectos al sistema sanitario o a las familias), fueron algunos factores novedosos tomados en consideración a la hora de confeccionar las nuevas recomendaciones.

Al final del proceso de consenso, 75 expertos propuestos por SECPCC aprobaron 54 recomendaciones expresas (el 77% de los contenidos propuestos) que pretenden ayudar a reducir la variabilidad de prácticas en este ámbito.

5) Estudio DEPRESIÓN (Martín y cols, 2011)

El proyecto DEPRESIÓN 2011 fue impulsado por la Sociedad Española de Psicogeriatría (SEPG) con el fin de responder a un déficit asistencial evidente de nuestro sistema sanitario, como es la excesiva variabilidad de prácticas en la atención clínica a los pacientes ancianos y muy ancianos que padecen depresión (148). En buena medida, esta situación está condicionada por las limitaciones del conocimiento clínico-epidemiológico y de las evidencias de base experimental disponibles sobre la efectividad terapéutica en la depresión en esta población de edad avanzada.

Desde SEPG, asociación científica que agrupa a los clínicos subespecializados en la atención a los trastornos mentales de este grupo etario, se planteó la necesidad de unificar criterio profesional y optimizar los cuidados ofrecidos a una población particularmente frágil y frecuentemente comórbida.

La baja calidad de las pruebas científicas que sustentan una parte no menor de las decisiones que se toman en el ejercicio de la psiquiatría clínica es un problema ampliamente analizado y de causas ya comentadas. Además, en el caso concreto de las poblaciones geriátricas y la depresión, la exclusión específica de los ancianos en los ensayos clínicos con antidepresivos y la práctica ausencia de información referida a la subpoblación muy anciana, obligan al psiquiatra a trabajar en condiciones de alta incertidumbre, habitualmente extrapolando sus decisiones a partir de información originada en poblaciones no geriátricas.

En un intento de paliar esta situación, SEPG solicitó la participación de un panel multicéntrico de expertos de reconocido prestigio por su trayectoria previa en el campo de la depresión y la psicogeriatría, a fin de desarrollar por consenso unas recomendaciones clínicas sobre el manejo de la depresión geriátrica que aprovecharan la experiencia clínico-epidemiológica personal de los 61 expertos que finalmente aceptaron colaborar.

Tras un proceso Delphi reglado, el grupo emitió una batería de 46 recomendaciones clínicas expresas para optimizar la detección y tratamiento de los ancianos con depresión.

6) Estudio ESQUIZOFRENIA (Bobes J et al, 2012)

El proyecto ESQUIZOFRENIA 2012 fue promovido de formar multicéntrica por investigadores asociados colaborativamente al Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM).

Las Sociedades Españolas de Psiquiatría (SEP) y de Psiquiatría Biológica (SEPB) , a fin de paliar las deficiencias asistenciales existentes en la fecha y de garantizar la protección a la salud global de los pacientes con enfermedades mentales graves,

habían promovido en 2007 el desarrollo de un documento de consenso sobre la salud física del paciente con esquizofrenia (149).

Estos pacientes presentan una prevalencia de enfermedad física más alta y una mayor mortalidad por causas naturales que el resto de la población. Su mortalidad estandarizada debida a causas naturales es 2,3 veces mayor que la de la población general, lo que reduce su expectativa de vida en quince años, por causas mayoritariamente atribuibles a enfermedades físicas. Todo ello está relacionado con el hecho de que la comorbilidad física de los pacientes con esquizofrenia pasa frecuentemente inadvertida y no es tratada. Además, estos pacientes disfrutaban de un peor acceso a los cuidados de salud y, en general, muestran un cumplimiento terapéutico inferior al de la población general. Con frecuencia, la iatrogenia derivada del propio tratamiento antipsicótico, empeora el problema.

El documento SEP-SEPB 2007 recomendaba una serie de procedimientos diagnósticos y de intervenciones clínicas destinadas a controlar los factores de riesgo modificables que deterioran la esperanza y calidad de vida del paciente esquizofrénico, reclamando un mayor protagonismo del psiquiatra en el manejo de la repercusión biológica del trastorno. Aunque poco después de su publicación casi la mitad de una muestra estatal de psiquiatras clínicos declararon conocer sus recomendaciones, una auditoría real de sus historias clínicas constató que sólo una fracción testimonial de estos profesionales (inferior al 15%) había modificado sus hábitos de valoración e intervención sobre la salud física de sus pacientes esquizofrénicos (150).

Con el fin de mejorar la implementación del consenso SEP-SEPB 2007 entre el colectivo psiquiátrico español y de valorar el grado de aceptación general de sus recomendaciones, el proyecto de ESQUIZOFRENIA 2012 convocó a un nuevo panel de expertos de ámbito estatal que no habían participado directamente en la redacción del anterior consenso, a fin de que reevaluaran y, en su caso, refrendaran sus contenidos.

Quince psiquiatras de ámbito nacional, seleccionados por una demostrable trayectoria experta, asistencial y/o investigadora, en esquizofrenia participaron en

un proceso de consenso que validó 61 de las 66 recomendaciones anteriores (92,4% de lo planteado). El proyecto permitió, además, actualizar algunos aspectos del anterior consenso e incluir ciertos tópicos novedosos gracias a una nueva revisión crítica de la literatura localizada en una búsqueda bibliográfica que se completó en julio de 2009.

El proyecto esquizofrenia es una muestra del potencial de la técnica Delphi para la ratificación o la actualización de recomendaciones clínicas preexistentes.

7) Estudio INHALADORES (Plaza V et al, 2012)

El proyecto INHALADORES 2012 fue promovido por un grupo clínico independiente multidisciplinar de médicos sensibilizados con el problema de la mala utilización de la terapia inhalada para el tratamiento de las enfermedades respiratorias obstructivas.

La vía inhalatoria es la ruta de elección para la administración de fármacos que deben actuar en el árbol bronquial, en particular broncodilatadores y antiinflamatorios. Por ello en todas las guías de práctica clínica se recomienda su uso para el tratamiento de las enfermedades respiratorias obstructivas (151-153). Desafortunadamente una gran proporción de pacientes los emplean de forma incorrecta, resultando de ello una baja eficacia terapéutica y un insuficiente control de estas enfermedades.

Todas las guías de asma y de EPOC dedican una especial atención a la necesidad de educar y adiestrar a los pacientes en el correcto uso de los inhaladores (151-153). Pero el núcleo del problema está en que la mayoría de los médicos que prescriben inhaladores desconocen su uso correcto y no resultan agentes eficaces para la educación de sus pacientes (el 86% de 1.514 médicos españoles de diferentes especialidades entrevistados recientemente mostró un insuficiente nivel de conocimientos sobre la terapia inhalada) (154). La situación no parece haber mejorado recientemente a pesar de los esfuerzos docentes realizados por sociedades científicas y grupos de trabajo al respecto.

Para subsanar la situación un grupo clínico español multicéntrico, particularmente sensibilizado con el problema, impulsó una iniciativa docente para la mejora del conocimiento sobre la terapia inhalada y el uso de los inhaladores entre los profesionales sanitarios: el proyecto “Terapia Inhalada”. Entre otras acciones, el programa incluyó la elaboración de una monografía específica sobre el tema, en la que los autores propusieron una serie de recomendaciones clínicas prácticas sobre el manejo apropiado de la terapia inhalada (155).

El proyecto INHALADORES 2012 tuvo como objetivo validar mediante un procedimiento externo de consenso estatal, realizado con un grupo de expertos ajenos a la redacción de la monografía mencionada, las recomendaciones propuestas en el documento. El panel estuvo formado por 59 acreditados expertos en aparato respiratorio, vinculados a actividades docentes o de investigación sobre inhaladores y técnicas de inhalación.

Tras el proceso de consenso resultaron respaldadas 39 de las 40 recomendaciones analizadas, en base al criterio y experiencia personal de los expertos españoles. Este proceso ha mejorado sustancialmente la visibilidad y alcance de las recomendaciones del documento reseñado (155), otorgándole carta de naturaleza como referente en esta cuestión a nivel nacional.

8) Estudio PRÓSTATA (Cózar JM et al, 2012)

El proyecto PRÓSTATA 2012 fue promovido por un grupo especializado de trabajo vinculado a la Asociación Española de Urología (AEU).

La hiperplasia benigna prostática es una de las áreas de la urología que más cambios conceptuales y terapéuticos ha sufrido en los últimos 20 años, de la mano de un mejor conocimiento de los mecanismos patogénicos de la enfermedad y su historia natural, del mejor análisis de su repercusión en la calidad de vida de los pacientes y de los resultados logrados por nuevos agentes farmacológicos.

La identificación de pacientes en riesgo de progresión de la enfermedad, las terapias de combinación y la introducción de varias alternativas quirúrgicas frente

a la resección transuretral prostática convencional (vaporización prostática, enucleación con láser de Holmium, RTU bipolar, ...) han sido también novedades relevantes en la especialidad en los últimos años, en buena medida aún carentes de un análisis crítico riguroso.

Todo ello había fomentado, en la fecha de estudio, una enorme variabilidad de prácticas entre distintos centros y profesionales que era objeto de preocupación para el colectivo profesional. El proyecto PRÓSTATA 2012 se impulsó desde la propia AEU con la pretensión de ayudar a unificar criterio, facilitar la toma de decisiones clínicas y racionalizar el uso de las alternativas terapéuticas entonces disponibles.

El estudio se inició con amplia revisión sistemática de la literatura más reciente de calidad sobre esta materia. A partir de los hallazgos identificados se sometió a la consideración de un grupo estatal de 50 expertos en la patología un cuestionario con 64 recomendaciones clínicas concretas avaladas por distinto grado de evidencia. Las recomendaciones incluyeron nuevos enfoques sobre el diagnóstico y la estratificación de los pacientes con HBP, algunas novedades de alcance sobre su tratamiento médico y sugerencias sobre controversias referidas a las nuevas opciones quirúrgicas en HBP.

Finalmente se logró unificar criterio en 50 de las cuestiones analizadas, que han sido difundidas por la sociedad científica como recomendaciones consensuadas que deberían ser tomadas en consideración para tomar decisiones en la práctica clínica hasta la aparición de nuevos datos científicos que justifiquen su futura revisión.

9. Estudio DOLOR (Pérez y coles 2013)

El dolor neuropático (DN) es consecuencia directa de una lesión o enfermedad que afecta al sistema somatosensorial. Su prevalencia en España no ha sido estudiada con suficiente rigor, pero se estima que puede afectar al 3% de la población mundial. En nuestro país existen 2 estudios recientes sobre la cuestión: uno de ellos sugiere que la incidencia puede alcanzar al 25% de los pacientes que

consultan por dolor crónico en atención primaria (156) y el otro lo identifica como el motivo de consulta en el 52% de los atendidos en unidades del dolor (157). En cualquier caso, se trata de un problema que deteriora de forma significativa el funcionamiento cotidiano y la calidad de vida de los pacientes, tanto en aspectos físicos como psicológicos, y que induce un importante gasto sanitario.

Por incluir un grupo complejo de síndromes dolorosos, en el que se involucran distintas causas y mecanismos patogénicos, su abordaje terapéutico es un reto clínico particularmente difícil. Se dispone para ello de una amplia batería de tratamientos farmacológicos (analgésicos, antidepresivos, antiepilépticos, opiáceos, agentes tópicos, etc.) y no farmacológicos (neuromodulación, radiofrecuencia, bloqueos nerviosos, técnicas neuroquirúrgicas, etc.), a pesar de lo cual el 40-60 % de los afectados sólo consigue un alivio insuficiente o sufre efectos adversos de los tratamientos aplicados. Ello es debido, no solo al manejo inapropiado de estos pacientes, sino a la eficacia parcial de los tratamientos disponibles.

Parte de la variabilidad de prácticas en el manejo del DN en España es debida, además, a que los pacientes son atendidos por diversos especialistas de distintas áreas médico-quirúrgicas, con sus particulares enfoques y “estilos clínicos”, pese a que existen múltiples guías clínicas que han intentado unificar la actitud diagnóstica y terapéutica. La última de ellas es un documento interprofesional, promovido por 7 sociedades científicas, que revisó la evidencia disponible hasta 2009 (158).

Por su fecha de redacción, el citado documento no analizó ni incluyó consejos sobre 3 nuevos fármacos, aprobados aquel año para el DN periférico (2 tratamientos tópicos y un agente sistémico). Ninguno de ellos había sido incluido tampoco en las guías terapéuticas europeas (159) y su aportación al arsenal terapéutico para el tratamiento del DN fue objeto de una notable controversia profesional.

Con el proyecto DOLOR (2013) un grupo de especialistas vinculados a la Sociedad Española del Dolor intentaron recabar el apreciación personal al respecto de 52

expertos de reconocida trayectoria en dolor para valorar conjuntamente, a la luz de la experiencia personal de todos ellos, cuál debería ser el posicionamiento racional de los nuevos fármacos en el manejo del DN, hasta que nuevos datos científicos ayudasen a establecer un mejor criterio.

Completado el proceso se publicó el informe científico del proyecto que recogió 46 recomendaciones explícitas (el 75% de las propuestas planteadas por el comité científico del proyecto) para optimizar el manejo del DN. Entre sus conclusiones se propusieron acciones innovadoras como la inclusión de tratamientos tópicos en la primera línea analgésica para el DN periférico o consejos específicos para el uso tópico de capsaicina en parches de alta concentración (una significativa novedad del vademécum analgésico introducida aquel mismo año).

4.3. Separatas originales de las publicaciones en revistas científicas de los estudios presentados.

En el Anexo 3, en el orden cronológico en el que fueron aceptados para publicación, se incluye una copia facsímil del documento original de cada estudio, tal y como fue publicado en su revista correspondiente.

Discusión



5. DISCUSIÓN

5.1. Reflexiones sobre el origen de la hipótesis de estudio, sobre los fines de esta tesis y sobre la pertinencia de una tesis por compendio de publicaciones.

El método Delphi incluye, hoy, varios procedimientos de consenso profesional originalmente diseñados para la predicción (con el concurso de expertos con criterio y experiencia en la cuestión objeto de interés) de la evolución futura esperable en algún ámbito (social, económico, político,...) expuesto a incertidumbre (66). El método se ha desarrollado de forma continuada, a partir de su formulación original en los años sesenta (73), particularmente en ámbitos ajenos a la Medicina (90).

El trabajo clínico, como se ha expuesto ampliamente en la introducción de esta tesis, ha de ser con frecuencia asumido en condiciones inciertas, por no existir (o por no resultar accesible) una parte sustancial de la información clínica y científica que podría resolver las dudas que el médico se plantea (53, 56). Tomar decisiones en circunstancias de incertidumbre clínica conlleva un elemento de previsión insegura sobre el estado de salud concreto del paciente y sobre su devenir más probable en el caso de elegir cada una de las alternativas que puedan plantearse.

Existe, pues, un vínculo estrecho entre las necesidades de información del clínico durante el ejercicio y el objetivo de un experto en cualquier otra disciplina cuando intenta hacer previsiones sobre el devenir futuro de un determinado problema en un contexto incierto. En esta analogía conceptual tuvo su inspiración inicial esta tesis, que explora la posibilidad de trasladar y adaptar eficazmente el uso del método DEMO (de probada eficacia predictiva) al ámbito clínico.

Antes de iniciar el diseño del primero de la serie de estudios presentados en esta tesis, se procedió al estudio bibliográfico detallado (88-105) de las particularidades de uso del método DEMO y a la reflexión crítica sobre su adecuación y aplicabilidad para la medicina clínica. A priori, la pretensión parecía

razonable, por cuanto el procedimiento se había mostrado efectivo, en consumo de tiempo y esfuerzo negociador, para facilitar el debate y alcanzar consensos sobre asuntos controvertidos en otros entornos profesionales (predicción sociológica, planificación estratégica, análisis de mercados, gestión empresarial, etc.). El método DEMO ya había sido también empleado en la propia medicina, si bien con fines distintos a los pretendidos en esta tesis, especialmente en asuntos de administración sanitaria, calidad asistencial y gestión de la investigación (91,92). Además, desde hacía más de 20 años, otros autores han aplicado el método DEMO en el contexto del procedimiento RAND/UCLA o del "uso apropiado", un procedimiento estandarizado de evaluación de la adecuación de usos (*appropriateness*) de la tecnología sanitaria a través del consenso experto (197,198). En este procedimiento se busca la aprobación o el rechazo por expertos de las posibles indicaciones en las que podría aplicarse cierto procedimiento diagnóstico o terapéutico en evaluación. Las distintas indicaciones se formulan exhaustivamente atendiendo a todos los factores clínicos críticos que pueden condicionarla (factores de riesgo, síntomas y signos del paciente, resultados de pruebas diagnósticas, terapias y otras variables relevantes de su historia clínica previa). La técnica, a través de encuestas al panel experto, procesadas con un criterio estadístico estandarizado, ha sido muy resolutivo para elaborar informes de las Agencias de Evaluación de Tecnologías Sanitarias.

Tales precedentes permitían considerar plausible y oportuno el fin último de esta tesis: trasladar el método DEMO al ámbito asistencial y analizar su idoneidad para responder dudas clínicas y ayudar a unificar criterio de actuación en escenarios inciertos o controvertidos. Para decidir en tales situaciones, donde no existe el apoyo de pruebas científicas firmes que orienten adecuadamente el mejor quehacer (ni siquiera por analogía o inferencia de otros datos disponibles) resultaría de enorme ayuda poder acceder al consejo basado en el conocimiento y la experiencia de los mejores especialistas en ese problema. Y los métodos de consenso grupal ofrecen una vía probadamente efectiva para conseguirlo (70,71).

Sin embargo, el procedimiento descrito más arriba (RAND/UCLA) no tiene la versatilidad necesaria para poder ofrecer respuesta a los múltiples motivos

posibles de incertidumbre (dudas clínicas de todo orden) que los médicos pueden afrontar durante el proceso asistencial a un paciente. Además el propio método DEMO tiene ciertas dificultades de aplicación si se pretende trabajar con expertos clínicos, que cuestionarían su viabilidad:

- Requiere un trabajo preparatorio dilatado de los expertos hasta conseguir la versión definitiva de los ítems de encuesta. Este punto, señalado como el más complejo y comprometido del Delphi, requiere de los panelistas una particular motivación intrínseca y ciertas habilidades técnicas de la “medicina basada en pruebas” (criterio bibliográfico y de valoración crítica de la literatura) no siempre disponibles.
- Requiere una reunión presencial de los expertos entre las rondas de encuesta que conlleva un riesgo indiscutible de influencia incontrolada en el desarrollo del consenso, e impone unas exigencias organizativas y un gasto no despreciable si el número de panelistas crece.

Todo ello sugirió la conveniencia de diseñar, ensayar y evaluar una nueva adaptación de factura propia al método DEMO que simplificase el procedimiento, haciéndolo más viable y, probablemente, más eficiente para la toma de decisión en la atención clínica. Más adelante, en cada sección de este mismo apartado, se discuten en detalle los distintos aspectos innovadores respecto al DEMO, sus ventajas y sus restricciones de acuerdo a la experiencia reunida de aplicación en los estudios que sustentan esta tesis. Para facilitar su identificación, en lo sucesivo se empleará la abreviatura **DEMOCLI** en referencia a esta adaptación original de la técnica DEMO para la definición y consenso de recomendaciones clínicas.

Los nueve estudios presentados son proyectos multicéntricos de ámbito estatal que comparten haber empleado por primera vez el nuevo método DEMOCLI en su desarrollo. En ellos han participado activamente 513 expertos españoles de distintas especialidades que han alcanzado consensos satisfactorios en los plazos previstos y sin conflictos relevantes. Los estudios, promovidos bajo el auspicio de 14 sociedades científicas, han sido publicados en siete revistas científicas indexadas en bases de datos internacionales, cuyos revisores y comités editoriales han ratificado el interés y rigor técnico de los proyectos (116-124). Con carácter

previo a una futura posible validación instrumental de la nueva herramienta de consenso, consideramos todo ello un pilotaje suficiente de la propuesta y una constatación empírica de la aplicabilidad y rendimiento del nuevo método. Los resultados comprobados permiten posicionar privilegiadamente el procedimiento adaptado, para los fines pretendidos, frente a otras alternativas técnica de objetivo similar (grupos nominales, conferencias de consenso,...).

En términos curriculares, el compendio de 9 publicaciones científicas completadas a lo largo de 5 años de las que el doctorando es coautor avalan su experiencia investigadora y su producción científica continuada en una línea estable de investigación. La contribución a los estudios del doctorando incluyó la consultoría metodológica a los equipos investigadores (formados por líderes asistenciales, docentes o investigadores en sus respectivas especialidades) y las labores de conducción técnica, análisis de datos e interpretación de resultados. En el anexo 3 se han recogido declaraciones confirmatorias de este particular por el autor principal de las publicaciones en las que el doctorando no fue primer o último firmante (autor senior).

Nueve experiencias exitosas de uso real del DEMOCLI para el desarrollo de consensos clínicos expertos en especialidades distintas y sobre temas de interés clínico muy diversos, acreditan una experiencia de uso estable y versátil. La heterogeneidad de contextos y contenidos en los que el método se ha aplicado satisfactoriamente y la publicación de las conclusiones particulares de cada estudio en revistas indexadas de varias disciplinas (116-124), respaldarían también la pretensión de generalizar esta propuesta metodológica y propagarla a nuevos entornos clínicos.

5.2. Reflexiones sobre las aportaciones y la efectividad de la adaptación del método Delphi modificado (DEMOCLI) para el diseño apropiado de las encuestas.

Como se ha comentado, un determinante principal de la buena aceptabilidad al método DEMOCLI en los estudios presentados, puede estar relacionado con el procedimiento propuesto para la identificación, selección y redacción de los

contenidos de la encuesta de estudio (una batería de recomendaciones de actuación que intentar responder a dudas habituales de los clínicos).

Cada ítem de la encuesta, en general, es una breve recomendación clínica expresa para actuar en determinada situación incierta, frente a la cual se solicita al experto posicionarse (manifestarse en acuerdo/desacuerdo, según su criterio y experiencia personal). En el procedimiento Delphi tradicional, el procedimiento comienza por una ronda inicial con formato de encuesta abierta (texto libre) con el fin de explorar las opiniones individuales de cada miembro del panel experto y recoger sus propuestas iniciales independientes (sin constrarar opiniones con sus pares) sobre el tema a debate. Cada panelista propone libremente cual debería ser, a su criterio, el contenido del cuestionario Delphi, avalando sus opiniones con su experiencia personal y con la evidencia científica disponible. Un trabajo de síntesis de las propuestas recibidas, bajo la supervisión de los investigadores impulsores del proyecto lleva, tras sucesivas revisiones del material recopilado, a una versión final de cuestionario.

Esta etapa del diseño de un proyecto Delphi se ha identificado como una fase compleja y prolongada del método, y una de sus principales dificultades (99,100). Si, como en los proyectos de contenido clínico, el panel de expertos no se compone de 8-12 miembros sino de un grupo más numeroso, el proceso puede complicarse o demorarse hasta resultar ineficiente. Frente a este sistema, la adaptación DEMOCLI, propone una solución pragmática y posibilista: que sea el propio comité científico del proyecto (en general, un número limitado de profesionales líderes, expertos indiscutibles en la materia), quienes se encarguen del laborioso proceso de revisión de la evidencia bibliográfica disponible y de la redacción preliminar de la batería de ítems de encuesta. En el apartado de "Métodos", se describió detalladamente esta fase, para la que cuentan con asistencia técnica para garantizar la estructura y contenidos adecuados al método. El anexo 1 transcribe el documento de instrucciones explícitas que se ofrece al comité científico para abordar esta tarea coordinadamente.

A priori, dejar en manos de los investigadores los contenidos de las recomendaciones clínicas que se debatirán por los expertos implica un riesgo de que el enfoque del proyecto resulte parcial o sesgado, bien por la propia limitación de pericia del Comité, por tomar decisiones arbitrarias o por la intervención de intereses espurios. Pero al menos tres garantías protegen a este sistema alternativo de diseño de la encuesta del peligro de que comité científico pueda influir en la opinión de los panelistas expertos, o de cualquier otro intento de manipulación de las conclusiones del consenso:

- Cualquiera que sea la formulación de los ítems (redactados en forma asertiva o negativa, posicionados pro o contra de una determinada opción) cada panelista consultado dispondrá siempre de la última decisión, libremente adoptada y confidencialmente expresada, sobre si está o no de acuerdo con las propuestas del comité.
- Por si el trabajo del Comité Científico resultara incompleto, parcial o sesgado, se garantiza al panelista que podrá sugerir cualquier cambio que considere oportuno al cuestionario (incluida la impugnación, modificación o ampliación del ítems) a través de comentarios libres que se le solicitan, para cada ítem, durante el desarrollo de la 1ª ronda de encuesta.
- Para facilitar que los panelistas puedan abordar esta tarea con criterio y responsabilidad, al inicio del proyecto se pone a su disposición el dossier de referencias bibliográficas empleadas por el Comité Científico durante la tarea preparatoria de la encuesta, y el documento, si se dispuso, de síntesis de evidencias a fecha de inicio de estudio.

Aunque no intervengan en su formulación inicial, los contenidos de la encuesta resultan, por tanto, un material ampliamente supervisado y sometido al control de todos los expertos participantes en el proceso de consenso.

En ninguno de los nueve estudios presentados se detectaron problemas relevantes al respecto, aunque en todos ellos se recibieran y, en su caso, se atendieran, sugerencias de aclaración, modificación, anulación o adición de ítems a la propuesta inicial del comité. En tales casos, las decisiones finales de cambio se aprobaron colegiadamente por el comité científico responsable de cada estudio, en

atención a su pertinencia, al argumentario presentado a su favor y al grado de generalización de la opinión.

Sobre la aceptabilidad de estos mecanismos de control y la aptitud de la propuesta DEMOCLI para desarrollar una encuesta Delphi adecuada a sus fines, habla el hecho de que ninguno de los 513 participantes en los paneles expertos de los distintos estudios declinara su decisión de participar en el proyecto durante su desarrollo, ni vetara al final del proceso (sobre la revisión del informe final de resultados) ninguna de las conclusiones alcanzada por consenso.

Eximir a los expertos clínicos del protagonizar las tareas de esta fase, evita tener que seleccionar con criterio restrictivo al panel (solo a quienes tienen capacidad técnica demostrada para abordar un trabajo técnico de lectura crítica de la literatura y de redacción científica, para la que muchos clínicos no están suficientemente capacitados, por más que resulten expertos incuestionables por su pericia y larga trayectoria clínica). De alguna forma, el método reserva al panel experto para su principal cometido: aportar opinión, sin requerir a sus componentes un factor de motivación extra que ponga en riesgo su aceptación a participar. Lograr tasas de reclutamiento de expertos superiores al 92% de los invitados en la mayoría de los estudios (y del 100% en varios de ellos) parece relacionado con esta cautela técnica. Confiar que, sin tal precaución, más de 500 clínicos hubieran aceptado participar altruistamente en un proyecto no remunerado, hubiera sido una expectativa poco realista.

5.3. Reflexiones sobre los atributos identitarios del experto-clínico y sobre la adecuación del tamaño y composición de los paneles expertos empleados en los estudios.

La segunda característica particular de DEMOCLI se refiere al tamaño de los paneles de expertos empleados en los consensos presentados. En el apartado de “Métodos” se detalló (ver tabla 4) como los estudios presentados, en general, participaron grupos de expertos de un tamaño notable, en especial comparados con los paneles habituales en estudios DEMO en otras áreas de conocimiento (por ejemplo, en los frecuentes proyectos de intención predictiva en el ámbito de la

economía/empresa, es habitual consultar a grupos máximos de 8-12 sujetos). Pero esta no es una cuestión sometida a restricciones metodológicas, sino que se debe a la distinta disponibilidad de expertos en cada ámbito de estudio, a ciertos usos propios de cada profesión y, finalmente, a las posibilidades organizativas y recursos de quienes impulsan los proyectos .

En materias como las comentadas no resulta fácil implicar a muchos expertos en proyectos para los que se les pide opinión pero no protagonismo y que, en general, no se remuneran. Esta dificultad se comprende mejor si se considera que muchos de los potenciales expertos económicos que podrían invitarse a un proyecto Delphi (responsables institucionales públicos o privados, líderes en el mundo empresarial, académico o en escuelas de negocios) ejercen labores de consultoría estratégica para terceros como una legítima actividad lucrativa, o participan como miembros en Consejos de Administración de empresas, fundaciones e instituciones, donde se espera de ellos una función de asesoría sujeta a reserva frente a la competencia.

Por otra parte, la predicción sobre escenarios inciertos en el ámbito de economía/empresa es un asunto complejo, que requiere considerar simultáneamente múltiples moduladores interrelacionados (políticos, regulatorios, fiscales, contables, financieros, laborales y sindicales, comerciales, de estrategia empresarial y responsabilidad social, de ciclo económico, bursátiles, internacionales, sociológicos, etc.). La selección de expertos en tal contexto requiere asegurarse que cada candidato elegido no sólo dispone de pericia suficiente en especialidad sino, además, de un conocimiento profundo de la materia concreta objeto de análisis. En estas circunstancias, en general, los paneles Delphi se articulan laboriosamente intentando garantizar una representación mínima de expertos disponer del enfoque de cada una de las especialidades que influyen en el asunto a debatir. La identificación y selección de tales expertos resulta, en general, un proceso delicado y clave sobre el grado de acierto en las previsiones del consenso. Se trata por ello de un aspecto que los impulsores del consenso deben considerar muy detenidamente, personalizando la decisión sobre cada candidato (66).

Pero el perfil de profesionales-objetivo (clínicos expertos) requeridos para aplicar el método DEMOCLI a la medicina asistencial, con el fin de desarrollar recomendaciones clínicas por consenso experto, es muy distinto al descrito más arriba para la economía. Por una parte, los clínicos “expertos” representan una subpoblación relativamente numerosa del colectivo profesional de cualquier especialidad médica. Se discutirá a continuación sobre este asunto. Además, por las raíces éticas inherentes a la praxis médica, los médicos suelen ser profesionales particularmente receptivos a la reflexión sobre posibles déficits asistenciales y comprometidos con la participación en procesos de mejora. Todo ello promueve una mayor disponibilidad de expertos y explica, en alguna medida, las altas tasas de reclutamiento de participantes logradas en la mayoría de los estudios presentados.

Lo que se persigue en estos estudios es disponer del criterio fundado en la pericia (*expertise*) de los médicos que tengan una trayectoria clínica particularmente valiosa y singular para los fines del estudio. Bajo esta consideración, el juicio sobre si un médico es un “clínico-experto” es una apreciación obviamente subjetiva, pero posible de caracterizar. Se considera clínico experto aquel que, con un conocimiento profundo de su especialidad y tras un tiempo suficiente de ejercicio profesional, desarrolla una aptitud para abordar con excelencia todas las demandas posibles de actuación profesional que cierto perfil de paciente/patología/especialidad pueda requerir (111, 137). Es decir, un profesional resolutivo que tiene una alta capacidad para generar y perfeccionar hipótesis diagnósticas y para tomar las decisiones (diagnósticas o terapéuticas) más apropiadas y prudentes (de mejor beneficio/riesgo) en situaciones infrecuentes, inciertas o complejas.

Para lograr tal perfil, resulta imprescindible disponer del mejor conocimiento posible de la fisiopatología humana; de la prevalencia y características de las enfermedades y los síndromes (de la gama completa de sus manifestaciones probables); de su curso natural y pronóstico; de la utilidad, limitaciones y uso racional de las pruebas diagnósticas; y del estado vigente de la terapéutica más actual. Disponer de todos esos atributos, apela al compromiso ético ineludible del

médico para organizar con responsabilidad su *longlife learning* y garantizar la actualización de sus conocimientos.

Pero nada de lo anterior garantiza la pericia ni convierte en experto. De hecho, es común que el nivel del conocimientos teóricos sean máximo al inicio de la vida profesional, cuando el medico novel está recién egresado de una facultad. Pero nadie le consideraría experto por ello, ni confiaría a la pericia del recién graduado la atención de casos complejos o críticos. De hecho, bien al contrario, la ley española limita expresamente la capacidad autónoma de ejercicio del residente durante su primer año de formación, obligando a su supervisión directa y escrita por un especialista cualificado, de forma que se asuman responsabilidades de forma progresiva (160). Mantener un conocimiento actualizado es una condición necesaria pero no suficiente para que un médico adquiera pericia: su carencia imposibilita generar hipótesis razonables o tomar decisiones correctas, pero disponer de ellos no garantiza dichas capacidades, ya que los atributos clínicos de las enfermedades pueden resultar muy variables y estar influidos por muchos moduladores.

Para devenir en experto, el médico debe además transitar regularmente una vía formativa imprescindible y simultánea a la del mantenimiento actualizado de su saber, como es la de acumular experiencia clínica (161,162).

Conseguir la mayor cantidad de experiencias personales durante la atención a casos clínicos permite desarrollar la destreza clínica y acumular hallazgos procedentes de pacientes diversos. Cada nueva experiencia clínica facilita el reconocimiento futuro de patrones similares a los de otros casos ya vistos o estudiados y ayuda a desarrollar atajos heurísticos de razonamiento intuitivo, propios del experto. Este es un proceso en perfeccionamiento continuo que permite a configurar un modelo personal, cada vez más completo y sofisticado, de cada enfermedad o síndrome, “enriquecido” en la experiencia, que se empleará como patrón de comparación para afrontar con mayor eficacia nuevos casos.

Kassirer y cols (163) han estudiado detenidamente el proceso de razonamiento clínico de médicos expertos de distinta procedencia y especialidad, a través del

análisis detallado de sus decisiones en múltiples casos, y han identificado algunas claves de la pericia clínica.

Son muchos los elementos circunstanciales que pueden ocurrir durante el proceso de atención a un paciente concreto que resultan valiosos en la construcción de esos arquetipos personales “ampliados” de enfermedad: el seguimiento de cuadros atípicos o de situaciones variantes de la normalidad; la experiencia de atención a pacientes que presentan un debut, curso clínico o complicaciones inesperadas; la comprobación en la práctica de los fallos de especificidad y de sensibilidad inherentes a las pruebas diagnósticas; la identificación de efectos adversos o interacciones medicamentosas infrecuentes o paradójicas; la comorbilidad casi sistemática de la clínica humana; el encuentro directo con las expectativas, necesidades sentidas y preferencias del paciente y su familia; las dificultades para lograr la adherencia al consejo terapéutico; las interferencias económicas u otras del sistema sanitario sobre la práctica clínica, etc. De todo ello se puede extraer un nuevo conocimiento valioso e insustituible para el médico experto.

Otros autores han discutido sobre si este el lento proceso de adquisición de experiencia que requiere la exposición continuada a la tarea clínica podría acelerarse y asegurarse a través del análisis teórico de casos, de la práctica clínica en entornos simulados (mediante software o simulación robótica), mediante procedimientos de coach o de otras estrategias educacionales (164-166).

Lo que resulta indiscutible es que, después de un proceso más o menos prolongado (cifrado por algunos en un periodo de, al menos, 10 años de ejercicio autónomo de la especialidad) muchos médicos alcanzan un nivel de competencia que les permite ser considerados como expertos por sus compañeros de profesión (163). Este reconocimiento, en general tácito y no documentado, no es por ello débil ni incierto. Las asociaciones científicas tienen claramente identificados a sus interlocutores expertos en cada uno de los tópicos de interés en la especialidad. Los propios profesionales sanitarios en cada centro y servicio asistencial admiten quienes son sus compañeros de referencia en cada tema, en general, por una unanimidad no concertada. En otros países más frecuentemente, pero también en

España, las figuras de consultores, especialistas *seniors* o eméritos sancionan formalmente la pericia de algunos médicos.

Bajo esta acepción de que es y cómo se forma un experto clínico, cabe suponer que en todas las especialidades existe un buen número de médicos que han desarrollado con responsabilidad una carrera profesional suficientemente dilatada como para acumular la pericia requerida. Resulta razonable aventurar que en cualquier comunidad autónoma, en cualquier gran ciudad y en cualquier centro hospitalario general haya algún experto clínico, al menos en el manejo de problemas clínicos de alta prevalencia, como los tratados en los estudios presentados, cuyo abordaje no suele requerir la derivación a especializadas.

Así las cosas, los paneles de expertos empleados en los estudios presentados en esta tesis resultan más numerosos que en otros estudios Delphi desarrollados en disciplinas ajenas a la medicina clínica. En parte, esta decisión se justifica por el notable mayor tamaño (en general) del colectivo de profesionales que podrían, como clínicos expertos, ser considerados candidatos a participar en un proceso de consenso, respecto al marco de situación de otras disciplinas. Dado el mayor tamaño de la hipotética base muestral de expertos clínicos disponibles en España, aunque nunca se ha insistido en la conveniencia de realizar cálculos de tamaño muestral, parece razonable pensar que un grupo sobredimensionado (respecto al tamaño medio de 9 miembros) resulte más apropiado y representativo de una población teórica total de candidatos tan amplia como la comentada (los expertos clínicos que podrían opinar sobre la materia). Evidencias experimentales (96) han puesto de manifiesto la relación directa entre el tamaño del plantel de expertos y la precisión de la estimación grupal obtenida (el error de previsión tiende a disminuir exponencialmente a medida que se amplía el tamaño de la muestra).

Pero, además, reclutar un grupo de mayor tamaño aporta una ventaja técnica no despreciable en un proyecto de consenso profesional: permite emplear sistemas de estratificación de los expertos por criterios que puedan resultar de interés para los objetivos de cada estudio (su origen geográfico, su especialidad de procedencia en proyectos multiprofesionales, su ámbito de dedicación principal: asistencial, investigador, académico, gestor,...).

Otros criterios que se han tomado en consideración en la identificación de expertos clínicos, en particular en el caso de estudios que analizaron aspectos farmacológicos, es que el candidato disponga de una particular trayectoria investigadora en la materia, como sería haber participado en estudios preclínicos de desarrollo de los fármacos de interés, o en los ensayos clínicos pivotaes que permitieron su aprobación, o en estudios de farmacovigilancia postautorización que analizan el resultado de los tratamientos en el entorno clínico real y que permiten identificar los efectos adversos de baja incidencia. Todo ello, y cualquier otro tipo de experiencia investigadora básica, traslacional, clínica o epidemiológica puede representar una información privilegiada y de interés adicional para incluir a cierto profesional en el panel de expertos.

En la misma línea, en algunos estudios se ha puesto en valor el rango académico o la vinculación de un profesional a un centro de referencia en la materia de estudio, o su adscripción a un grupo de trabajo profesional sobre la cuestión, como potenciales elementos de interés. Sin duda, la experiencia en dichos entornos representa una oportunidad para estos profesionales de profundizar en el conocimiento y de ampliar su criterio personal.

5.4. Reflexiones sobre la idoneidad de las distintas escala de respuesta empleadas en las encuestas y sobre la validez del criterio estadístico de consenso empleado en los estudios presentados.

Los estudios analizados emplean una modalidad de encuesta que resulta muy sencilla de cumplimentar por usar una única escala de valoración ordinal, de tipo Likert, a lo largo de todo el cuestionario. Sin duda, este aspecto es un determinante del alto grado de cumplimiento de las encuestas.

Los cinco primeros estudios de la serie, cuyas rondas de encuesta se desarrollaron entre abril de 2008 y julio de 2009, replicaron exactamente el planteamiento original de De Beers, quien propuso en una serie de estudios hoy clásicos (127,129,130), el uso de una escala de valoración ordinal (tipo Likert) de cinco puntos. La escala tiene en sus extremos el pleno desacuerdo (1) y el pleno acuerdo (5) con el contenido del ítem, y permite al experto posicionarse dentro de ese

rango, incluyendo un punto central de indefinición (ni acuerdo ni desacuerdo = 3). Aunque se trata de una propuesta experimentada y efectiva, la opinión compartida por algunos panelistas que participaron en estos cinco estudios, al finalizar el proceso, era haber tenido la sensación de disponer de una limitada posibilidad de gradación de sus opiniones, y de la dificultad para expresar diferencias de matiz en su consideración sobre distintos ítems (aunque todos ellos estuvieran en la misma región de valoración).

En la misma línea, el comité científico de algún estudio se planteó la conveniencia de que el informe de conclusiones del consenso no se limitara sólo a clasificar las recomendaciones valoradas como “aceptadas” o “rechazadas” por consenso, sino que ofreciera información adicional sobre el grado diferencial de apoyo experto obtenido por cada recomendación, a fin de poder priorizar las más respaldadas, en caso de tener que seleccionar las más críticas, urgentes u oportunas.

Tales demandas se resolvieron, en los estudios realizados a partir de octubre de 2009, sustituyendo la anterior escala de valoración por una nueva escala ordinal más amplia (de 9 puntos de anclaje) subdividida en tres regiones de opinión: “desacuerdo” [puntos 1-3], “ni acuerdo ni desacuerdo” [puntos 4-6] y “acuerdo” [puntos 7-9]. Esta escala resulta más sensible y gradual a la hora de discriminar el nivel de acuerdo/desacuerdo entre distintos ítems. Por otra parte, requiere del respondedor un proceso de doble decisión entre tres elementos alternativos (en primer lugar, la región de opinión y después, el punto concreto dentro de la región elegida), lo que en opinión de algunos, resulta una toma de decisión mentalmente más satisfactoria que la elección dicotómica. Esta escala tiene un largo aval de empleo en el método de uso apropiado de tecnologías sanitarias RAND-UCLA (107-109, 132,133).

A diferencia de otras áreas de la medición clínica (por ejemplo, en la definición de los cuantificadores semánticos para escalas Likert de medición de frecuencia de eventos) (167) en las que se han establecido recomendaciones sobre cuál debe ser el número óptimo de puntos de anclaje para escalar la medición, no hemos encontrado un criterio técnico que avale preferentemente una u otra de las escalas empleadas en nuestros estudios (de 5 o 9 puntos). La decisión de cambio se tomó

en atención a las cuestiones comentadas y al mayor uso de la segunda opción en la literatura reciente de evaluación de tecnologías sanitarias. Además, el análisis estadístico de la escala de rango mayor muestra alguna ventaja adicional que se comenta a continuación.

En sentido estricto, las mediciones Likert son de escala ordinal (no interválica), por lo que no resulta posible determinar si la diferencia entre la puntuación 8 y 9 es necesariamente la misma que entre 7 y 8, y por tanto, deberían ser estadísticamente descritas mediante medidas de centralización y dispersión no paramétricas (mediana y cuartiles de distribución). Sin embargo, se puede defender que el cálculo de las medias aritméticas de las puntuaciones obtenidas con paneles tan amplios como los de estos estudios y con una escala de medición suficientemente discriminativa de las diferencias de opinión (9 puntos), sirva para la construcción de un ranking orientador de la prioridad relativa concedida por los expertos a cada ítem. Este argumento tiene respaldo en algún informe técnico (168) que, a pesar de no disponer de una teoría justificativa para la "escalabilidad" de las puntuaciones Likert, considera que, en la práctica, las mediciones de este instrumento pueden proporcionar la base para una primera ordenación del grado o nivel medio alcanzado por el grupo en la característica medida con la escala (en nuestro caso, el grado de apoyo a cada recomendación).

En la formulación estándar del método de "uso apropiado" para evaluar intervenciones sanitarias, se considera que el panel está de acuerdo en su valoración cuando al menos dos tercios de los panelistas expresan una misma valoración en una de las regiones extremas de opinión (inapropiado=[1-3] o apropiado=[7-9]). En expresión estadística, se requiere que no más de un tercio de los consultados se posicione en contra de la mayoría, eligiendo puntuaciones situadas fuera de la región de tres puntos en la que está contenida la mediana del grupo.

Este criterio, obviamente arbitrario, es análogo al empleado para cierta toma de decisiones sancionadas con criterio jurídico. Así, en los sistemas de decisión parlamentaria y, por extensión, en cualesquiera órganos colegiados públicos y privados, se entiende por mayoría cualificada que la votación sobre un asunto

sometido a consideración requiera, para su aprobación, un porcentaje importante de apoyo a su favor, siempre superior al de la mayoría simple. Se consideran, así, cualificadas las mayorías que superan el 50% de los votos.

Aunque para superar ese dintel bastaría una mayoría absoluta (*mitad más uno*), cuando se abordan decisiones muy sensibles que requieren su adopción por muy amplia mayoría, se exigen mayorías cualificadas superiores a las absolutas, denominadas mayorías reforzadas. La mayoría reforzada requerida habitualmente en tal situación (por ejemplo, para la reforma constitucional o para la aprobación de ciertas leyes orgánicas que afectan a derechos fundamentales) suele ser superior a los dos tercios de la cámara (169). Como puede comprobarse, se trata de un criterio conceptual y operativamente superponible a los empleados en el método Delphi.

En cualquier caso, definir arbitrariamente el punto a partir del cual las opiniones de un grupo se consideran consensuadas es una decisión técnica que no repugna los usos habituales de la estadística en otras materias (por ejemplo, al definir el umbral del valor p de error máximo asumible al rechazar una hipótesis sometida a un test de contraste, al interpretar el grado de correlación entre dos variables cuantitativas a través del valor concreto del coeficiente r de Pearson, al clasificar el grado de dispersión de una variable a través del valor particular de su coeficiente de variación, o al aceptar un nivel de confianza determinado en la construcción del intervalo de confianza de un estimador).

Conceptualmente, consensuar es alcanzar un criterio compartido por una fracción suficientemente alta, pero indeterminada, de un grupo que sujetos que se ha de posicionar en algún punto del trayecto entre la mayoría simple y la unanimidad plena. La primera opción sería un sistema débil e incierto para extraer conclusiones, pero exigir la unanimidad total sería dejar expuesto al veto paralizante de un solo miembro cualquier posibilidad de acuerdo. Aunque pueden defenderse otras opciones (170), establecer en dos tercios de los encuestados el umbral del consenso es una decisión sancionada por los usos aprobados en otras esferas de la vida social que requieren del acuerdo grupal.

En los cinco estudios iniciales que usaron escalas de Likert de 5 puntos, se respetaron los mismos criterios de análisis para determinar la existencia de consenso que otros autores han venido empleando sistemáticamente en una serie de prestigiados informes, publicados entre 1991 y 2003, sobre usos inapropiados de fármacos en ancianos (127,129,130). Estos estudios fueron evaluados y difundidos en una prestigiosa revista internacional de alto impacto (*Arch Intern Med*) pese a su heterodoxo empleo de estimadores paramétricos (media e intervalo de confianza del promedio al 95%) para el análisis estadístico de datos ordinales.

5.5. Reflexiones sobre el impacto de la sustitución de la reunión presencial de panelistas entre rondas por el proceso de interacción no presencial propuesto en el método DEMOCLI.

La organización y gestión de la reunión presencial que el método Delphi modificado convencional propone para la preparación de la segunda ronda de encuesta habría sido un serio problema en los estudios presentados, dada la convocatoria de paneles con gran número de expertos. Algunos inconvenientes operativos de tal requerimiento resultan obvios:

- La sesión de encuentro, manejable con una dinámica de conducción bien establecida para un grupo de debate de tamaño limitado, resultaría un reto metodológico para un grupo de trabajo tan numeroso como alguno de los propuestos (de hasta 90 expertos).
- Cubrir los gastos derivados de un encuentro presencial de 1-2 días en un lugar geográfico común y accesible a todos los panelistas conllevaría, si fuera posible de organizar, un coste inasumible para la mayoría de proyectos analizados.
- Se han comentado ampliamente los posibles sesgos de influencia que pueden operar en reuniones presenciales de trabajo en grupo, en especial si este es de gran tamaño y la sesión está poco estructurada.
- Dado el alto número de asistentes y el limitado tiempo de trabajo disponible en esa sesión, resultaría harto probable que las expectativas de participación de muchos panelistas no pudieran resultar adecuadamente satisfechas.

Ante todo ello, la propuestas de DEMOCLI para trabajar con grandes paneles suple esta sesión presencial de intercambio y matización de posturas por la entrega de una información detallada de los resultados de la primera ronda de encuesta (mediante informes gráficos y estadísticos apropiados a la naturaleza de escala de medida empleada, con la oportuna explicación interpretativa), junto al recuerdo de su propia opinión, y a un informe anonimizado que recoge todos los comentarios libres y aclaraciones aportadas por el resto de participantes sobre cada ítem y sobre su posicionamiento personal. En la sección “Métodos” se explica en detalle el procedimiento y en la figura 8 se reproduce una muestra de ejemplo. Este material de retroalimentación enriquece el *feedback* habitual ofrecido en la aplicación del método Delphi convencional (una simple distribución numérica de los votos del panel), en el que tampoco existía ningún tipo de sesión presencial.

Ese material, adecuadamente preparado, es una valiosa información intermedia facilitada a los panelistas para su reflexión. Tras revisar esta información, los panelistas pueden reconsiderar o no sus posiciones, antes de abordar la segunda y última valoración de los ítems no consensuados con anterioridad. Se garantiza, pues una interacción controlada del panel, se permite un espacio y tiempo para la reflexión confidencial, y se facilita, en su caso, el cambio de opinión, sin que esta circunstancia resulte evidente ni manifiesta al resto del panel. Gracias a ello, nuestra adaptación del método mantiene hasta el final de proceso el anonimato de los miembros del panel, dificultando el efecto de influencias espurias, o el establecimiento alianzas o rechazos concertados entre los panelistas.

Puede discutirse el alcance y posibilidades de interacción que este modelo de retroalimentación controlada produce (171), pero de nuevo cabe apelar al éxito en la conclusión de los proyectos y a la aceptación plena (por todos los panelistas) de las conclusiones de todos los estudios (respecto a todas las recomendaciones consensuadas).

Los cambios de criterio personal resultantes de este proceso, llevaron la aprobación media de ítems (promedio de todos los estudios) desde un 73,6% en la primera ronda hasta un 86,6% al finalizar la segunda. Se incrementó, por tanto, el porcentaje de consenso alcanzado tras la interacción descrita en el 13% de las

recomendaciones propuestas (según estudios, entre el 6,4% y el 33,8%). Al final del proceso, permanecieron sin consensuar el 13,4% de las recomendaciones globalmente consideradas (según estudios, entre el 2,5% y el 24,6%).

Debe destacarse el elevado nivel de acuerdo alcanzado con las recomendaciones propuestas por los comités científicos de cada estudio. La concordancia de criterio de los panelistas participantes en los proyectos es superior a la habitualmente observada en otros proyectos de metodología similar (130, 146). Aunque la legitimidad de tal comparación es discutible, ya que la capacidad de concitar unanimidad de opinión es propia de cada estudio (según la heterogeneidad del panel profesional consultado y el grado de controversia existente en el tema a debate), debe destacarse que los resultados de varios de estos estudios se han logrado con grupos multiprofesionales, con diferentes intereses y responsabilidades asistenciales, en ocasiones, competitivas. De hecho, un valor principal de esos estudios es la integración del conocimiento y la experiencia clínica de expertos procedentes de diversos ámbitos especializados y geográficos.

Debe aclararse, finalmente, que en algunos proyectos presentados se celebraron sesiones presenciales de presentación de los resultados donde el panel de desarrollo de cada estudio pudo conocer en primicia los resultados del consenso y donde se desveló la identidad de todos los componentes del panel. Ello se pudo hacer cuando la sociedad científica que auspiciaba y/o los promotores dispusieron de recursos para ello. Algunas de las comentarios libres de la sala en esas reuniones fueron aportaciones de valor para los investigadores de cada proyecto, a la hora de acometer la discusión del original en el que publicaron las conclusiones de sus estudios.

5.6. Reflexiones sobre el valor del uso alternativo de la adaptación clínica del método Delphi modificado (DEMOCLI) para la validación y refrendo de recomendaciones preexistentes (otras Guías).

En el apartado “Métodos” se explicó como tres estudios de la serie se desarrollaron sobre las recomendaciones de guías o documentos previos de reciente publicación (ASMA 2010, INHALADORES 2012 y ESQUIZOFRENIA2012), con el fin de

refrendarlos y ayudar a su implantación. No debe minusvalorarse la utilidad marginal de la adaptación del método DEMOCLI a grandes grupos para ayudar a conseguir una difusión amplia y selectiva de un estado de opinión entre un colectivo profesional.

Tal uso es de más interés si se considera que la divulgación e implementación de las guías de práctica clínica es un elemento frecuentemente mal atendido y una causa principal del bajo conocimiento y uso por los profesionales a quienes se destina. Conseguir que las recomendaciones de una guía clínica se apliquen en la atención sanitaria real es un proceso complejo y costoso. La mayoría de las acciones de implementación que han sido evaluadas con rigor se han mostrado poco efectivas y, en general, de efectos pasajeros. Casi siempre, los resultados resultan muy alejados de las expectativas teóricas iniciales de influencia de las guías (172).

Entre otras barreras bien conocidas que dificultan esta implementación (complejidad del formato, baja comprensibilidad del texto, carencias formativas profesionales sobre MBE, comorbilidad y otras restricciones de aplicación en los pacientes reales etc.), hay cuestiones que tienen que ver con el pobre respaldo a la guía por pares, superiores jerárquicos o líderes de opinión, así como la percepción entre los profesionales de poco realismo en los objetivos y tareas clínicas recomendados (173-177).

Ambos aspectos podría ser intervenidos a través del apoyo expreso y amplio a las recomendaciones de la guía por parte de los expertos que acumulan la más reconocida y larga experiencia en la materia. Esta circunstancia, adecuadamente explicada, es un valor añadido interesante del uso del método DEMOCLI con los fines descritos (validación y refrendo a documentos previos).

5.7. Reflexiones sobre posibles limitaciones de esta tesis, sugerencias sobre nuevas líneas de investigación sobre la materia y secuelas científicas de esta tesis.

En este último apartado se discuten algunas de las posibles limitaciones y mejoras que caben plantear a este trabajo y que, en otras circunstancias y con mejores recursos, podrían justificar futuros proyectos en esta misma línea de investigación.

La tesis ha demostrado el rendimiento de un nuevo método para desarrollar recomendaciones expertas de ayuda a la toma de decisión clínica en incertidumbre. Pero el doctorando omitió deliberadamente entre sus objetivos la posibilidad de realizar una validación instrumental convencional del nuevo método, esto es, una comprobación del grado en el que la herramienta alcanza con excelencia sus cometidos (que las recomendaciones clínicas desarrolladas resulten de calidad objetiva y representen el mejor criterio clínico que el panel experto pueda alcanzar en común).

Tal cuestión no se ha planteado por un determinante insuperable: la necesidad de recursos que exceden las posibilidades del doctorando, en particular para comparar los resultados del consenso respecto a otras alternativas técnicas de consenso profesional. Pero, además, por la convicción de que tal objetivo resulta de cuestionable, al menos tal y como ha sido planteado con anterioridad, en condiciones similares, por otros autores. Un estudio previo (178) abordó una validación instrumental, *sensu stricto*, del método DEMO para el desarrollo de recomendaciones expertas (en concreto, sobre hipertensión arterial). Para ello se empleó un enfoque similar al usado en los estudios de validación de pruebas diagnósticas. Los autores concluyeron que el método DEMO mostró una buena fiabilidad entre rondas y una satisfactoria validez, tanto predictiva (respecto a las recomendaciones de guías clínicas basadas en la evidencia), como de contenido (en qué medida las recomendaciones desarrolladas se relacionaban con el perfil de competencias de los médicos destinatarios) y de constructo (comparación de los resultados obtenidos por el método Delphi con los de otros métodos). Resulta discutible si tiene sentido comprobar la “fiabilidad entre rondas” de un método

que pretende mejorar el nivel de acuerdo del grupo, y cuyo éxito está, por tanto, relacionado con el máximo cambio posible de resultados entre ambas encuestas. De igual forma, cabe cuestionar el intento de validar predictivamente a un método de consenso experto respecto a los contenidos de guías basadas en la evidencia de referencia, cuando el objetivo principal del método es, precisamente, desarrollar recomendaciones de ayuda a la decisión en situaciones clínicas no clarificadas en las guías clínicas, por ausencia de pruebas científicas suficientes.

En el caso de la adaptación DEMOCLI, ya se ha discutido anteriormente como la permanencia del 100% de los panelistas en la dinámica de trabajo propuesta por el nuevo método y la aceptación formal por todos ellos de la totalidad de las conclusiones derivadas de los estudios como procedentes de un consenso legítimo, son indicios sólidos de la calidad y la utilidad clínica del producto final de estos consensos. Pero constatar científicamente tales conclusiones requeriría desarrollar alguno o varios de los siguientes estudios que se proponen como posibles líneas de investigación futura:

- Estudiar el conocimiento y opiniones de una muestra representativa de clínicos del colectivo profesional al que se destinan las conclusiones de los estudios, como “usuarios finales”, sobre el grado de difusión alcanzado por las recomendaciones expertas, sobre su aplicabilidad y sobre su utilidad clínica. Un estudio observacional basado en un número suficiente y bien distribuido de encuestas podría describir fiablemente dichas opiniones. El doctorando tiene alguna experiencia previa en este sentido (179-181).
- Confirmar la potencial reducción de la variabilidad de prácticas lograda en dichos colectivos al usar las recomendaciones expertas para manejar la patología de estudio. Como en otros estudios (150, 182), esta comprobación podría hacerse mediante procesos de auditoria directa de registros clínicos producidos en la práctica asistencial real.
- Verificar la mejora objetiva en resultados de salud (intermedios o finales) resultantes de la implementación de las recomendaciones clínicas en la práctica, con un diseño idealmente experimental.

Todos ellos y otros posibles serían proyectos de innegable interés, pero inaccesible para las posibilidades y medios del doctorando.

Por otra parte, se debe reconocer el carácter parcial (aunque diverso) de los ámbitos clínicos en los que el nuevo método de consenso ha sido ensayado. Pese a aplicarse en diversas especialidades médicas y quirúrgicas, y tratar temas que incluyen distintos aspectos del proceso clínico, no puede asumirse que los resultados presentados avalen sin discusión la utilidad del método en cualquier otro contexto clínico. En otros entornos asistenciales, la disponibilidad de expertos y la naturaleza de la incertidumbre clínica podrían resultar distintas a los comprobados en nuestros estudios, y todo ello suponer algún tipo de obstáculo insalvable, hasta el momento no detectado, en la aplicación del nuevo método.

Aunque la experiencia investigadora presentada, a lo largo de cinco años, parece abundar en que la convicción de que el método resulta universalizable, con este fin serían bienvenidos nuevos estudios confirmatorios difundidos en publicaciones de prestigio. Resultaría, además, conveniente que algunos de estos nuevos estudios procediesen de grupos investigadores ajenos al equipo responsable de la nueva adaptación del método, como prueba de su difusión efectiva y de su viabilidad en otras manos.

En este sentido, una búsqueda centinela sobre esta cuestión realizada a la finalización de la tesis, ha permitido localizar al menos cuatro publicaciones recientes que siguen nuestra estela y son las primeras réplicas exactas de nuestra propuesta (método DEMOCLI) que otros grupos finalizan con éxito:

- Un gran consenso de ámbito estatal sobre el manejo de la dislipemia aterogénica, impulsado por un grupo multiprofesional procedente de cinco grandes hospitales universitarios y un prestigioso instituto de investigación público (Institut de Recerca Vall d'Hebron, Universitat Autònoma de Barcelona), publicado en la revista **Clínica e Investigación en Arterioesclerosis** en junio de 2013 (183). Estos resultados han sido objeto de una doble publicación, se supone que autorizada, en **Revista Española de Cardiología** en enero de 2014 (184).

- Un consenso monográfico (dirigido por urólogos) que pretende extender recomendaciones clínicas expertas sobre un aspecto concreto de actualidad y controversia como es el uso de la toxina botulínica en la vejiga hiperactiva publicado en **Actas Urológicas Españolas** en mayo de 2014 (185).
- Un consenso monográfico (dirigido por nefrólogos) con recomendaciones clínicas para optimizar el manejo de las alteraciones del metabolismo óseo-mineral en la enfermedad renal crónica, publicado en la revista **Nefrología** en marzo de 2014 (186).

Los autores de estos estudios, pese a replicar con precisión el método DEMOCLI y reproducir literalmente amplios fragmentos de nuestros originales en la sección de “Material y Métodos” de sus publicaciones, no referencian, como es norma, el origen de su propuesta metodológica. No obstante, las fechas de remisión, aceptación y publicación de los artículos analizados en esta tesis, que anteceden hasta en 4 años a los nuevos estudios, clarifican el origen del método y nos permite reivindicar sin ninguna duda la autoría intelectual sobre este procedimiento.

Conclusiones



6. CONCLUSIONES

- La tesis presenta, pilota y evalúa, sobre la experiencia objetiva de una amplia serie de estudios ya publicados (9), los resultados de una innovadora adaptación técnica original del método Delphi Modificado para la clínica (**DEMOCLI**), como procedimiento de consenso no presencial para elaborar recomendaciones expertas que orienten la actuación de los médicos en situaciones de incertidumbre clínica y/o alta variabilidad de prácticas.
- **DEMOCLI** es una adaptación del método Delphi modificado en dos rondas, específicamente introducida para su empleo en el ámbito clínico, que incluye tres innovaciones técnicas: a) la responsabilidad exclusiva de los investigadores que lideran cada proyecto (su comité científico) en la confección de los cuestionarios Delphi, b) el reclutamiento de un panel de expertos clínicos de mayor tamaño que lo previamente convencional, y c) la sustitución de la preceptiva reunión presencial previa a la encuesta final, por un procedimiento a distancia para la interacción controlada entre los expertos, antes de que completen la encuesta de cierre del proyecto.
- En todos los proyectos en los que se empleó **DEMOCLI** se ha mostrado como una técnica de negociación y consenso profesional operativa para que grupos heterogéneos de expertos clínicos confronten su criterio personal, superen discrepancias iniciales y desarrollen por consenso amplios listados de recomendaciones expertas para unificar prácticas en situaciones de incertidumbre clínica.
- El alto grado de consenso experto alcanzado (superior al 85% de las proposiciones analizadas en todos los estudios) sobre las recomendaciones preliminarmente propuestas por el comité científico de cada proyecto, o modificadas a sugerencia de sus expertos durante la ronda inicial de encuesta, avala la efectividad del proceso adaptado de confección de cuestionarios Delphi

que **DEMOCLI** introduce y confirma la funcionalidad del procedimiento de participación voluntaria de los panelistas previsto para esta tarea.

- El éxito en el reclutamiento de candidatos invitados a participar como expertos en los paneles Delphi de todos los estudios (en promedio, superior al 95%) apoya que la pretensión de **DEMOCLI** de configurar paneles de expertos más numerosos de lo convencional resulta viable en el entorno clínico (en promedio n=57) y confirma el alto grado de aceptación y participación logrado por este procedimiento de consenso profesional para elaborar recomendaciones clínicas en escenarios de incertidumbre.
- En todos los estudios, las conclusiones de los consensos (el listado final de recomendaciones clínicas a difundir entre los especialistas destinatarios de cada proyecto) fueron aprobadas con la conformidad final del 100% de panelistas participantes, sin registrarse ningún veto, abandono ni otros conflictos. Los resultados avalan que el procedimiento de interacción no presencial de los expertos ofrecida por **DEMOCLI** y los criterios estadísticos de interpretación empleados resultan aceptables y suficientes para que los expertos respalden los resultados finales de la investigación como procedentes de un proceso legítimo de consenso grupal.
- **DEMOCLI** ha mostrado su versatilidad en distintos tipos de ámbitos asistenciales y/o de especialidades médicas (hasta 13), resultando ser igualmente aplicable y eficiente para tratar asuntos de controversia clínica en todas ellas. El aval científico de los proyectos por diversas sociedades científicas médico-quirúrgicas y otros grupos especializados de referencia estatal en muy diversas disciplinas, sugiere la generabilidad de este procedimiento técnico y respalda la pretensión de propagar la propuesta metodológica a nuevos entornos clínicos. A la conclusión de esta tesis, al menos tres grupos profesionales ajenos al investigador habían empleado ya el método y publicado nuevos consensos en otras especialidades.

- Las innovaciones técnicas introducidas por **DEMOCLI** han superado con éxito la valoración de pertinencia, adecuación, rigor metodológico y calidad de la investigación exigidos por los comités editoriales de diversas revistas científicas que siguen los procedimientos normalizados de revisión y aprobación del Comité Internacional de Editores de Revistas Médicas (ICMJE) y que están internacionalmente indexadas. Todo ello, confirmado repetidamente en 9 estudios y a lo largo de 5 años, avala la robustez de la adaptación técnica propuesta como una metodología apta para desarrollar estudios con los fines descritos en el ámbito clínico, que aspiren a su difusión en revistas científicas de impacto.

Epílogo

Epílogo

Los primeros pasos de esta tesis nacieron entre serias dudas sobre la oportunidad y la pertinencia de dedicar el esfuerzo de un proyecto doctoral en Medicina al estudio de un procedimiento más propio del territorio de las “ciencias inexactas”, poco conocido entre los clínicos y de posicionamiento discutido entre los métodos convencionales de investigación (187). Hacerlo, además, para confrontar las debilidades de un paradigma tan establecido y prestigioso como la medicina basada en pruebas, y para reivindicar el mérito debido al criterio del experto en la medicina del siglo XXI, indujo muy poco entusiasmo entre los más cercanos a quienes presenté la idea.

A algunos, a quienes más respeto por sensatos, el proyecto les pareció más apto para el debate epistemológico que para ser defendido científicamente ante un tribunal universitario de Medicina. Los más prudentes pensaron que apenas era una provocación intelectual, y que acabaría aparcándolo *motu proprio*. Escribiendo estas últimas líneas, sospecho que ni la sensatez ni la prudencia son rasgos relevantes de mi carácter. Pero desde la convicción de hoy, alcanzada paso a paso, sobre proyectos reales y publicaciones tangibles, intuyo también que las carencias a veces ayudan.

Esta tesis hubiera sido imposible sin la oportunidad de articularla poco a poco, mediante el compendio de publicaciones, y sin la acogida incondicional de mi Director a tan lento proceso. La confirmación de cada nueva publicación ha sido como hacer escala en un rumbo de 5 años, durante los que he colaborado con gente que forman ya parte del resto de mi camino profesional, mientras este dure. Cada uno de ellos ha aportado algo concreto y palpable a esta tesis. Mi agradecimiento y admiración para todos ellos.

La pretensión compartida de ayudar a clínicos que afrontan cada día la incertidumbre clínica, a veces con medios precarios, sin caer en el desánimo ni dejar de ayudar a sus pacientes, ha sido una pieza de persuasión infalible para mantener el rumbo en algunas *tormentas*.

Acabar la redacción de estas páginas constatando que otros han tomado el relevo y emprenden su propio rumbo con nuestra hoja de ruta es más de lo que podía imaginar al escribir en una hoja de papel un primer título tentativo para esta tesis, hace hoy más de cinco años. Concluyo con la satisfacción de comprobar como las expectativas de impacto profesional y de difusión de la propuesta metodológica en otras especialidades y en otros campos de interés, han sido precozmente cubiertas.

En Madrid, miércoles 17 de septiembre de 2014.

*Festividad de Santa Hildegardis de Bingen, Doctora de la Iglesia, reconocida **experta** medieval y prolífica autora de libros médicos: "Liber Simplicis Medicinae", "Liber Compositae Medicinae", "Hildegardis Curae et Causae".*

Bibliografía

BIBLIOGRAFIA

1. Wennberg JE. Dealing with medical practice variations: a proposal for action. *Health Affairs*. 1984;3(2):6-32
2. Mercuri M, Gafni A. Medical practice variations: what the literature tells us (or does not) about what are warranted and unwarranted variations. *J Eval Clin Pract*. 2011;17(4):671-7
3. Nathens AB, Rivara FP, Mack CD, Rubenfeld GD, Wang J, Jurkovich GJ, et al. Variations in rates of tracheostomy in the critically ill trauma patient. *Critical Care Medicine*. 2006;34(12):2919-2924.
4. Landrigan CP, Conway PH, Stucky ER, Chiang VW, Ottolini MC. Variation in pediatric hospitalists' use of proven and unproven therapies: A study from the pediatric research in inpatient settings (PRIS) network. *Journal of Hospital Medicine*. 2008;3(4):292-298.
5. Schneiderman R, Kirkby S, Turenne W, Greenspan J. Incubator weaning in preterm infants and associated practice variation. *Journal of Perinatology*. 2009;29(8):570-574.
6. Cromwell DA, Mahmood TA, Templeton A, Van Der Meulen JH. Surgery for menorrhagia within English regions: Variation in rates of endometrial ablation and hysterectomy. *BJOG: An International Journal of Obstetrics and Gynaecology*. 2009;116(10):1373-1379.
7. Chen Y, Briesacher BA, Field TS, Tjia J, Lau DT, Gurwitz JH. Unexplained variation across US nursing homes in antipsychotic prescribing rates. *Archives of Internal Medicine*. 2010;170(1):89-95.
8. Fernandez-de-Maya J, Richard-Martinez M. Variations in clinical practice. Current status and challenges for nursing. *Enfermería Clínica*. 2010;20(2):114-118.
9. Dossett LA, Adams RC, Cotton BA. Unwarranted national variation in the use of prophylactic inferior vena cava filters after trauma: An analysis of the national trauma databank. *Journal of Trauma Injury, Infection and Critical Care*. 2011;70(5):1066-1070.
10. Elferink MAG, Siesling S, Lemmens VEPP, Visser O, Rutten HJ, Van Krieken JHJM, et al. Variation in lymph node evaluation in rectal cancer: A Dutch nationwide population-based study. *Annals of Surgical Oncology*. 2011;18(2):386-395.
11. Wennberg JH. *Tracking Medicine. A Researcher's Quest to Understand Health Care*. New York: Oxford University Press; 2010.
12. Pearce-Smith N. *Unwarranted Variation: a Reading List produced by QIPP Right Care*. Right Care Essential Reading, September 2011.
13. Gómez de la Cámara A, Bravo Toledo R. Medicina basada en la Evidencia. En: Fernández-Cruz A. *Manual de habilidades para la Práctica Clínica*. Madrid: Grupo MSD; 1999.
14. Eddy DM. The Challenge. *JAMA*. 1990; 263:2420-2425.
15. de Jong JD. *Explaining medical practice variation. Social organization and institutional mechanisms*. Utrecht: Nivel Ed.; 2008: 9-10.

16. Anderson TF, Mooney G. Medical practice variations: where are we? In: Anderson TF, Mooney G (eds). The challenges of medical practice variations. London: The Macmillan Press; 1990:1-15
17. Feinstein AR, Horwitz RI. Problems in the "evidence" of "evidence-based medicine". *Am J Med.* 1997;103(6):529-35.
18. Evidence-based medicine working group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA.*1992; 268(17):2420-2425.
19. Guyatt GH. Evidence-based medicine. *ACP Journal Club.* 1991;114: A16.
20. Sackett DL, Rosenberg WM, Muir Gray JA, Haynes RB, Richardson WS.: Evidence based medicine: what it is and isn't. *British Medical Journal.* 1996;312:71-72.
21. Gol-Freixa JM, Del Llano Señaris JE. El fenómeno de la "medicina basada en la evidencia". *Med Clin.* 1999;112(Supl 1):3-8.
22. McKibbon A, Eady A, Marks S. PDQ Evidence-based principles and practice. Hamilton (Canada): B.C. Decker Inc.; 1999.
23. Evidence-based medicine in its place (editorial). *Lancet.* 1995;346:785
24. Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem solving. *BMJ.* 1995; 310:1122-6.]
25. Smith BH, Taylor RJ. Medicine, a healing or a dying art? *Br J Gen Pract.* 1996; 46:249-51.
26. Glass RM. The Patient-Physician relationship. *JAMA.* 1996;275:147-8.
27. Crawley L. Evidence-Based Medicine: A New Paradigm for the Patient. *JAMA.* 1993;269(10):1253.
28. Sackett DL, Richardson WS, Rosenberg W, Haynes RB: Medicina Basada en la Evidencia. Cómo ejercer y enseñar la MBE. Madrid: Churchill Livingstone; 1997.
29. Pocock SJ. The historical development of clinical trials. Cap. 2. En: *Clinical Trials.* Toronto: John Wiley & Sons;1984: 14-15.
30. Antman EM, Lau J, Kuperlink B, Mosteller F, Chalmers T. A comparison of results of meta analyses of randomized controlled trials and recommendations of clinical experts: treatments for myocardial infarction. *JAMA.* 1992;268:240-8.
31. Sacket D. Applying overviews and meta-analyses at the bedside. *J Clin Epidemiol.* 1995;48:61-66.
32. Feinstein AR, Horowitz RI. Problems in the "evidence" of "evidence based medicine". *Am J Med.* 1997;103:529-35.
33. Koessler M. Les faux-amis des vocabulaires anglais et americains. Paris: Vuivert Ed; 1975.
34. Real Academia Española. Diccionario Panhispánico de Dudas [Internet]. (Acceso26/08/2014). Disponible en: <http://lema.rae.es/dpd/?key=EVIDENCIA>
35. Bravo Toledo R, Campos Asensio C. Medicina basada en pruebas (Evidence-based medicine). *JANO.* 1997; LIII(1218):71-72.
36. Puigdollers Colós JM. Precisiones terminológicas a la medicina basada en la evidencia. *Med Clin (Barc).*1999;113:79.
37. Sanz Rubiales A, Del Valle Rivero ML, Garavís Vicente M. Medicina basada en la evidencia: ¿no hay una traducción mejor?. *Med Clin (Barc).*1999;112:119.
38. Cuenca Boy R, Cuenca Boy MI: Medicina basada en la evidencia: un caso de pereza lingüística. *Med Clin (Barc).* 1999;113(13):518.
39. Pozo-Rodríguez F. La medicina basada en la evidencia. Una perspectiva desde la clínica. *Med Clin (Barc).* 1999;112:12-16.

40. White KL. Archie Cochrane's legacy: an American perspective. En: Maynard A, Chalmers I, Eds. Non-random reflections on Health Services research: on the 25th anniversary of Archie Cochrane's Effectiveness and Efficiency. London: BMJ Publishing Group, 1997. [Internet] (acceso 3 de agosto de 2013). Disponible en: http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/non-random_reflections_on_health_service_research.pdf
41. Booth A, Djulbegovic B, Guthrie B, Perleth M, Sackett D, Endersly S, Jenkins D, et al. What proportion of healthcare is evidence based. Resource Guide. [Internet]. (acceso 3 de agosto de 2013). Disponible en: www.shef.ac.uk/scharr/ir/percent.html.
42. Brook RH. Health, Health Insurance, and the Uninsured. *Journal of the American Medical Association*. 1991; 265(22):2998-3002.
43. Stewart M, Brown JB, Weston WW, McWhinney IR, McWilliam CL, Freeman TR. Patient-centered medicine: transforming the clinical method. 2 ed. Oxford (UK): Radcliffe Medical Press Ltd.;2003.
44. Miles A. On the interface between science, medicine, faith and values in the individualisation of clinical practice: a review and analysis of *Medicine of the Person*. *Journal of Evaluation in Clinical Practice*. 2009;15:1000-1024.
45. Miles A., Mezzich JE. The care of the patient and the soul of the clinic: person-centered medicine as an emergent model of modern clinical practice. *International Journal of Person Centered Medicine*. 2011;1(2):207-222.
46. Ely JW, Burch RJ, Vinson DC. The information needs of family physicians: case-specific clinical questions. *J Fam Pract*. 1992;35(3):265-269.
47. Gorman PN, Ash J, Wykoff L. Can primary care physicians' questions be answered using the medical journal literature?. *Bull Med Libr Assoc*. 1994; 82(2):140-146.
48. Gorman PN, Helfand M. Information seeking in primary care: how physicians choose which clinical questions to pursue and which to leave unanswered. *Med Decis Making*. 1995;15(2):113-119.
49. Smith R. What clinical information do doctors need?. *BMJ*. 1996;313:1062-1068.
50. Chambliss ML, Conley J. Answering clinical questions. *J Fam Pract*. 1996;43(2):140-144.
51. Ely JW, Osheroff JA, Ebell MH, Bergus GR; Levy BT, Chambliss ML; Evans ER. Analysis of questions asked by family doctors regarding patient care. *BMJ*. 1999; 319:358-361.
52. Ely JW, Osheroff JA, Ebell MH, et al. Obstacles to answering doctors' questions about patient care with evidence: qualitative study. *BMJ*. 2002;324:710-3.
53. Ely JW, Osheroff JA, Chambliss ML, Ebell MH, Rosenbaum ME. Answering physicians' clinical questions: obstacles and potential solutions. *J Am Med Inform Assoc*. 2005;12(2):217-224.
54. Alper BS, White DS, Ge B. Physicians answer more clinical questions and change clinical decisions more often with synthesized evidence: a randomized trial in primary care. *Ann Fam Med*. 2005;3(6):507-513.
55. Coumou HC, Meijman FJ. How do primary care physicians seek answers to clinical questions?. A literature review. *J Med Libr Assoc*. 2006;94(1):55-60.
56. Covell DG, Uman GC, Manning PR. Information needs in office practice: are they being met?. *Ann Intern Med*. 1985;103(4):596-9.

57. Richardson WS, Wilson MC, Nishikawa J, Hayward RS: The well-built clinical question: a key to evidence-based decisions. *ACP J Club*. 1995;123(3):A12-3.
58. González-González AI, Dawes M, Sánchez-Mateos J, Riesgo-Fuertes R, Escortell-Mayor E, Sanz-Cuesta T, Hernández-Fernández T. Information needs and information-seeking behavior of primary care physicians. *Ann Fam Med*. 2007; 5(4):345-52.
59. Pellegrino ED. *The virtues in medical practice*. N. York: Oxford University Press US.; 1993.
60. Letelier LM, Moore P. La medicina basada en evidencia: Visión después de una década. *Rev Med Chile*. 2003;131:939-946
61. Martí-Carvajal A. ¿Es más útil la revisión narrativa que la sistemática en la toma de decisión en salud?. *Gac Méd Caracas*. [Internet]. 2010;118(1):37-40. [acceso 28 de agosto de 2014]. Disponible en:http://www.scielo.org.ve/scielo.php?script=sci_arttext&pid=S0367-47622010000100005&lng=es.
62. Martí-Carvajal A, Salanti G, Cardona AF. Human recombinant activated protein C for severe sepsis. *Cochrane Database Syst Rev*. 2007 Jul 18;(3):CD004388.
63. Mackenzie I, Lever A. Management of sepsis. *BMJ*. 2007;335:929-932.
64. Gudiol F. "Ojo clínico" y evidencia científica. *Educ Med*. 2006; 9(1):21-23.
65. Asua Batarrita J. Entre el consenso y la evidencia. *Gaceta Sanitaria*. 2005;19(1):65-70.
66. Landeta J. *El método Delphi. Una técnica de previsión del futuro*. Barcelona: Editorial Ariel; 2002.
67. Berra A, Marin I, Alvarez R. *Metodología de Expertos. Consenso en Medicina. Serie Monografías nº 14*. Granada: Escuela Andaluza de Salud Pública; 1996.
68. Yañez R, Cuadra R. La técnica Delphi y la investigación en los servicios de salud. *Ciencia y Enfermería*. 2008;14(1):9-15.
69. Peiró S, Portella E. No todo es acuerdo en el consenso. Limitaciones de los métodos de consenso en los servicios de salud. *Gaceta Sanitaria*. 1993;7:294-300.
70. Lorge I, Fox D, Davitz J, Brenner M. A survey of studies contrasting the quality of group performance and individual performance (1920-1957). *Psychological Bulletin*. 1958;55:337-372.
71. Miner FC, Jr. Group versus individual decision making: A investigation of performance measures, decision strategies, and process losses/gains. *Organizational Behaviour and Human Performance*. 1984;33:112-24.
72. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CFB, Askham J, et al. Consensus development methods, and their use in clinical guideline development. *Health Technology Assessment*. 1998; 2(3).
73. Dalkey NC, Helmer O. An experimental application of the Delphi method to the use of experts. *Management Science*. 1963;9(3):458-467.
74. Pill J. The Delphi Method: substance, context, a critique and an annotated bibliography. *Soc Econ Plan Sci*. 1971;3:37-71.
75. Delbecq AL, Van de Ven AH. A group process model for problem identification and program planning. *Journal of Applied Behavioral Science*. 1971;7(4):466-492.
76. U.S. Department of Health & Human Services. National Institutes of Health. NIH Consensus Development Program [Internet]. Acceso 4 de agosto de 2013. Disponible en: <http://consensus.nih.gov/>.

77. Pérez-Andrés C. ¿Deben estar las técnicas de consenso incluidas entre las técnicas de investigación cualitativa?. *Revista Española de Salud Pública* 2000; 74:319-321.
78. Isaacs D, Fitzgerald D. Seven alternatives to evidence based medicine. *BMJ* 1999;319:1618.
79. Tudor R. Técnicas en grupo basadas en los principios y la práctica del brainstorming. En: *La creatividad. Análisis y solución de problemas empresariales*. Bilbao: Deusto;1977.
80. Briones E, Marín I, Álvarez R, Reyes A. Fundamentos de consenso en el ámbito de las ciencias de la salud. En: *Consenso en Medicina. Metodología de Expertos*. Granada: Monografías EASP; 1996.
81. Eddy D. Guidelines for policy statements: the explicit approach. *JAMA*. 1990; 263:2239-43.
82. Eddy D. Resolving conflicts in clinical policies. *JAMA*. 1990;263:3077-84.
83. Eddy D. Resolving conflicts in clinical policies. *JAMA*. 1990;264:389-91.
84. Granados A. Las conferencias de consenso. ¿Un método para la evaluación de las tecnologías sanitarias?. *Gac Sanit*. 1991;5:201-2.
85. Peiró S, Portella E, El grupo nominal en el entorno sanitario. *Quaderns de Salut Publica i Administració de Serveis de Salut 1*. Valencia: Escola Valenciana d'Estudis per a la Salut; 2003.
86. Scheele DS. Reality construction as a product of Delphi interaction. En: Linstone HA, Turoff M. *The Delphi method: Techniques and applications* Massachusetts: Addison-Wesley Eds.; 1975.
87. Grimal P. *Diccionario de mitología griega y romana*. Madrid: Ediciones Paidós Ibérica, S.A.; 1992.
88. Dalkey NC, Helmer O. The use of experts for estimation of bombing requirements. A project Delphi experimental. Santa Monica (California): The Rand Corporation 1951. Publication RM-727-PR.
89. Dalkey NC. The Delphi Method: an experimental study of group opinion. Santa Monica (California): The Rand Corporation 1969. Publication RM-59999 PR.
90. Linstone HA, Turoff M. *The Delphi method: Techniques and applications* Massachusetts: Addison-Wesley Eds.; 1975.
91. Boulkedid R, Abdoul H, Loustau M, Sibony O, Alverti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One*. 2011;6(6):e20476.
92. Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. *PLoS Med*. 2011;8(1):e1000-393.
93. Hsu CC, Sandford BA. The Delphi Technique: Making Sense Of Consensus. *Practical Assessment, Research & Evaluation*, 2007; 12(10):1-8. [Internet]. (Acceso 28 de agosto de 2014). Disponible en: <http://pareonline.net/getvn.asp?v=12&n=10>.
94. Yousuf, MI. Using Experts' Opinions through Delphi Technique. *Practical Assessment, Research & Evaluation* 2007; 12(4). [Internet]. (Acceso 28 de agosto de 2014). Disponible en: <http://pareonline.net/getvn.asp?v=12&n=4>.
95. Galane T; Miku J. The choice of an optimum group of experts. *Technological Forecasting and Social Change*. 1986;30:245-50.

96. Dalkey N, Brown B, Cochran S. The Delphi Method, III: Use of Self Ratings to Improve Group Estimates. Santa Monica (California): The Rand Corporation 1969. Publication RM-6115-PR.
97. Spetzler C, Von Holstein CAS. Probability encoding in decision analysis. *Management Science*. 1975;22:340-358.
98. Dajani JS, SIncoff MZ, Talley WK. Delphi stability and agreement criteria. *Technological forecasting and Social Change*. 1980;16:67-73.
99. Williams PL, Webb C. The Delphi technique: a methodological discussion. *J Adv Nurs*. 1994;19:180-186.
100. Powell C. The Delphi tecniqe: myths and realities. *J Adv Nurs*. 2003; 41(4):376-382.
101. Nelms KR, Porter AL. EFTE: An Interactive Delphi Method. *Technological Forecasting and Social Change*. 1985;28(1):43-61.
102. Murray JW, Hammons JO. Delphi: A versatile methodology for conducting qualitative research, *The Review of Higher Education* 1995;18(4),423-436.
103. Custer RL, Scarcella JA, Stewart BR. The modified Delphi technique: A rotational modification. *Journal of Vocational and Technical Education*. 1999;15(2),1-10.
104. Timbie JW, Damberg CL, Schneider EC, Bell DS. A conceptual framework and protocol for defining clinical decision support objectives applicable to medical specialties. *BMC Med Inform Decis Mak*. 2012;12(93):1-12.
105. Stewart J, O'Halloran C, Harrigan P, Spencer JA, Barton JR, Singleton SJ et al. Identifying appropriate tasks for the preregistration year: modified Delphi technique. *BMJ*. 1999; 319:224-229.
106. Canadian Partnership Against Cancer's Capacity Enhancement Program (CEP). Mc Master University. Modified Delphi Consensus summary 2010. [Internet]. Acceso 7 de septimebr de 2014. Disponible en: <http://fhswedge.csu.mcmaster.ca/cepftp/qasite/ModifiedDelphi.htm>.
107. Martínez-Sahuquillo ME, Echevarría MC. Métodos de consenso. Uso adecuado de la evidencia en la toma de decisiones. «Método RAND/UCLA». *Rehabilitación (Madr)*. 2001;35(6):388-392.
108. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, Lacalle JR, Lázaro P, van het Loo M, McDonnell J, Vader JP, Kahan JP. The RAND/UCLA Appropriateness Method User's Manual. Santa Monica (California): RAND publication MR-1269; 2001.
109. Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care*. 1986;2:53-63.
110. García Valdés M, Suárez Marín M. El método Delphi para la consulta a expertos en la investigación científica. *Revista Cubana de Salud Pública*. 2013; 39(2). [Internet]. Acceso 6 de septiembre de 2014. Disponible en: http://bvs.sld.cu/revistas/spu/vol39_2_13/spu07213.htm
111. Baker J, Lovell K, Harris N. How expert are the experts? An exploration of the concept of 'expert' within Delphi panel techniques. *Nurse Res*. 2006;14(1):59-70.
112. Hutchings A, Raine R. A systematic review of factors affecting the judgment produced by formal consensus development methods in health care. *Journal of Health Services Research & Policy*. 2006;11:172-179.

113. Hutchings A, Raine R, Sanderson C, Black N. A comparison of formal consensus methods used for developing clinical practice. *Journal of Health Services Research & Policy*. 2006; 11:218-224.
114. Goodman C Baratz S. *Improving Consensus Development for Health Technology Assessment: an international perspective*. Washington DC: Council on Health Care technology. Institute of Medicine; 1990.
115. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ*. 1995; 311:376-80.
116. Ampudia Blasco FJ, Baos Vicente V, Caballero Martínez F, Álvarez-Sala Walther LA. Criterios profesionales sobre el papel de ezetimiba en el manejo clínico del paciente dislipémico en el ámbito de atención primaria. *Av Diabetol*. 2009;25(5):399-407.
117. Caballero Martínez F, Plaza Moral V, Quirce Gancedo S, Fernández Benítez M, Gómez Ruiz F, López Viña A, Molina París J, Quintano Jiménez JA, Soler Vilarrasa R, Villa Asensi JR, Balmes Estrada S. Valoración externa de las recomendaciones de la GEMA 2009 por un panel multiprofesional de expertos en asma. *Arch Bronconeumol*. 2010;46(8):411-419.
118. Villanueva Haba V, Sánchez Álvarez JC, Peña Mayor P, Salas Puig J, Caballero Martínez F, Gil-Nagel Rein A. Treatment initiation in epilepsy. An expert consensus in Spain. *Epilepsy and Behaviour*. 2010;19(3):332-42.
119. Medrano López C, García-Guereta L, Fernández Pineda L, Malo Concepción P, Maroto Álvaro E, Santos de Soto J, Lirio Casero J, Suárez Cabrera P, Caballero Martínez F. Consenso clínico sobre la profilaxis de la infección por virus respiratorio sincitial (VRS) y el uso del palivizumab en cardiología pediátrica. *An Pediatr (Barc)*. 2010; 72(6): 432.e1-432.e13.
120. Martín Carrasco M, Agüera Ortiz L, Caballero Martínez L, Cervilla Ballesteros J, Menchón Magriñá JM, Montejo González AL, Moríñigo Domínguez A, Caballero Martínez F. Consenso clínico sobre la depresión en el paciente anciano. *Actas Esp Psiquiatr*. 2011;39(1):20-31.
121. Bobes García J, Saiz Ruiz J, Bernardo Arroyo M, Caballero Martínez F, Gilaberte Asín I, Ciudad Herrera A, en representación del Grupo de Trabajo sobre la Salud Física del Paciente con Trastorno Mental Grave. Consenso Delphi sobre la salud física del paciente con esquizofrenia: valoración por un panel de expertos de las recomendaciones de las Sociedades Españolas de Psiquiatría y de Psiquiatría Biológica. *Actas Esp Psiquiatr*. 2012;40(3):114-28.
122. Plaza Vicente, Calle M, Molina J, Quirce S, Sanchis J, Viejo JL, Caballero F. Validación, mediante el método Delphi, del Consenso sobre Terapia Inhalada para el tratamiento de las enfermedades respiratorias obstructivas. *Arch Bronconeumol*. 2012;48(6):189-196.
123. Cózar Olmo JM, Hernández Fernández C, Miñana López B, Amón Sesmero JH, Montlleó González M, Rodríguez Antolín A, Caballero Martínez F. Consenso experto sobre el impacto clínico de la nueva evidencia científica disponible sobre Hiperplasia Benigna Prostática. *Actas Urol Esp*. 2012;36(5):265-275.
124. Pérez C, Rodríguez MJ, Guerrero A, Margarit C, Martín-Estefanía C, Oteo-Álvaro A Caballero-Martínez F. Consenso experto sobre el uso clínico de los tratamientos por vía tópica en el manejo del dolor neuropático periférico. *Rev Soc Esp Dolor*. 2013; 20(6):308-323.

125. Polit DF, Hungler BP. Diseños de muestreo. En: Polit DF, Hungler BP. Investigación científica en ciencias de la salud. 6ª ed. México: McGraw-Hill Interamericana; 2000:285-289.
126. Holey EA, Feeley JL, Dixon J, Whittaker VJ. An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. *BMC Med Res Methodol.* 2007;7:52.
127. Beers MH, Ouslander JG, Rollinger J, Reuben DB, Beck JC. Explicit criteria for determining inappropriate medication use in nursing home residents. *Arch Intern Med.* 1991;151(9):1825-32.
128. Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A et al (OCEBM Levels of Evidence Working Group). The Oxford 2011 Levels of Evidence. Oxford Centre for Evidence- Based Medicine. [Internet]. Acceso 23 de agosto de 2014. Disponible en: <http://www.cebm.net/index.aspx?o=5653>.
129. Beers MH. Explicit criteria for determining potentially inappropriate medication use by the elderly. *Arch Intern Med.* 1997;157:1531-6.
130. Fick DM, Cooper JW, Wade WE, Waller JL, Maclean JR, Beers MH. Updating the Beers Criteria for Potentially Inappropriate Medication Use in Older Adults. Results of a US Consensus Panel of Experts. *Arch Intern Med.* 2003;163:2716-24.
131. Isaacs David, Fitzgerald Dominic. Seven alternatives to evidence based medicine *BMJ.* 1999;319:1618.
132. Lacalle JR, Pastor L, Reyes A, Pérez MJ, Álvarez R. Metodología Delphi aplicada a la evaluación de procedimientos diagnósticos y terapéuticos. En: Berra de Unamuno A, Marín León I, Álvarez Gil R, editores. Metodología de expertos. Consenso en medicina. Granada: Escuela Andaluza de Salud Pública; 1996: 53-69.
133. Brook RH, Chassin MR, Fink A, Solomon DH, Kosecoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care.* 1986;2:53-63.
134. Almansa C, Rey E, Bolaño E, Palma M, Álvarez A, Díaz-Rubio M. Opinión de los médicos españoles sobre el síndrome de intestino irritable: resultados de un estudio utilizando el método Delphi. *Rev Esp Enferm Dig.* 2007;99:210-7.
135. Dago A, Arcos P, Álvarez de Toledo F, Baena MI, Martínez J, Gorostiza I. Indicadores de riesgo de morbilidad prevenible causada por medicamentos. *Gac Sanit.* 2007;21:29-36.
136. Suárez Cabrera P, Malo Concepción P, Maroto E, Santos de Soto J. Recomendaciones para la prevención de la infección por virus respiratorio sincitial en pacientes con cardiopatía congénita. Sociedad Española de Cardiología Pediátrica y Cardiopatías Congénitas 2005. [Internet]. Acceso 1 de agosto de 2014. Disponible en: http://www.secardioped.org/Descargas/PyB/PubSEC_004.pdf.
137. Pastor L, Lacalle JR, de Villar E, Pérez I, Gómez E. Sistemática de la composición de los paneles de expertos. En: Berra de Unamuno A, Marín León I, Álvarez Gil R, editores. Metodología de expertos. Consenso en medicina. Granada: Escuela Andaluza de Salud Pública; 1996: 71-80.
138. Goodman LA. Snowball Sampling. *The Annals of Mathematical Statistics.* 1961;32:148-70.

139. Plaza V, Alvarez FJ, Casan P, Cobos N, Lopez Viña A, Llauger MA, et al. en calidad de Comité Ejecutivo de la GEMA y en representación del grupo de redactores. Guía Española para el Manejo del Asma (GEMA). Arch Bronconeumol. 2003; 39(Supl 5):1-42.
140. GEMA 2009: Guía española para el manejo del asma. Guía Salud: Guías de Práctica Clínica en el Sistema Nacional de Salud. Consejo Interterritorial. Sistema Nacional de Salud. [Internet]. Acceso 24 de noviembre de 2009. Disponible en: <http://www.guiasalud.es/viewGPC.asp?idGuia=304>.
141. Fisher RS, van Emde Boas W, Blume W, et al. Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). Epilepsia. 2005;46:470-2.
142. French JA, Kanner AM, Bautista J, et al. Efficacy and tolerability of the new antiepileptic drugs. I. Treatment of new onset epilepsy: report of the Therapeutics and Technology Assessment Subcommittee and Quality Standards Subcommittee of the American Academy of Neurology and the American Epilepsy Society. Neurology. 2004;62:1252-60.
143. Glauser T, Ben-Menachem E, Bourgeois B, et al. ILAE treatment guidelines: evidence-based analysis of antiepileptic drug efficacy and effectiveness as initial monotherapy for epileptic seizures and syndromes. Epilepsia. 2006;47:1094-120.
144. Spanish Neurology Society Epilepsy Study Group. Official guide for the diagnosis and treatment of epilepsy 2008. Barcelona: Prous Science; 2008.
145. Andalucian Epilepsy Society. Andalucian epilepsy guide 2009. Madrid: Editorial EMISA; 2009.
146. Figueras J, López J, Medrano C, Bermúdez-Cañete R, Fernández L, Bonillo PA, et al. Consenso multidisciplinar español sobre la profilaxis de la infección respiratoria pediátrica por virus respiratorio sincitial. An Pediatr (Barc). 2008;69:63-71.
147. Medrano C, García-Guereta L, Lirio J, García J. Infecciones respiratorias, síndrome de Down y cardiopatías congénitas. Estudio CIVIC 21. Anales de Pediatría. 2009; 71(1):38-46.
148. Pincus HA, Tanielian TL, Marcus SC, Olfson M, Zarin DA, Thompson J, Magno Zito J. Prescribing trends in psychotropic medications: primary care, psychiatry, and other medical specialties. JAMA 1998;280:132-3.
149. Sáiz Ruiz J, Bobes García J, Vallejo Ruiloba J, Giner Ubago J, García-Portilla González MP; Grupo de Trabajo sobre la Salud Física del Paciente con Esquizofrenia. Consenso sobre la salud física del paciente con esquizofrenia de las Sociedades Españolas de Psiquiatría y de Psiquiatría Biológica. Actas Esp Psiquiatr. 2008;36:251-64.
150. Saiz-Ruiz J, Saiz-González MD, Alegría AA, Mena E, Luque J, Bobes J. Impacto del Consenso Español sobre la Salud Física del Paciente con Esquizofrenia. Rev Psiquiatr Salud Ment (Barc.). 2010;3:119-27.
151. GINA. Global Initiative for Asthma. Global Strategy for Asthma Management and Prevention NHLBI/WHO Workshop Report. 2009. [Internet]. Acceso 23 de agosto de 2014. Disponible en: <http://www.ginasthma.com>.
152. GOLD. NHLBI/WHO. Global Initiative for Chronic Obstructive Lung Disease. NHLBI/WHO workshop report. National Institutes of Health, National Heart,

- Lung and Blood Institute. Publication Number 2701. 2010. [Internet]. Acceso 23 de agosto de 2014. Disponible en: <http://www.goldcopd.com>
153. GEMA. Guía Española para el Manejo del Asma. Arch Bronconeumol 2009; 45 Supl. 7:2-35. [Internet]. Acceso 23 de agosto de 2014. Disponible en: <http://www.gemasma.com>.
 154. Plaza V, Sanchis J, Roura P, Molina J, Calle M, Quirce S, et al. Physician's knowledge of inhaler devices and inhalation techniques remain poor in Spain. J Aerosol Med Pulm Drug Deliver. 2012;25:16-22.
 155. Plaza V, Calle M, Molina J, Quirce S, Sanchis J, Viejo JL. Terapia inhalada. Teoría y práctica. Madrid: Luzán 5; 2010.
 156. Pérez C, Saldaña MT, Navarro A, Vilardaga I, Rejas J. Prevalence and characterization of neuropathic pain in a primary-care setting in Spain: A cross-sectional, multicentre, observational study. Clin Drug Investig. 2009;29:441-50.
 157. Pérez C, Ribera MV, Gálvez R, Micó JA, Barutell C, Failde I, et al. High prevalence of confirmed, but also of potential and believed, neuropathic pain in pain clinics. Eur J Pain. 2013;17(3):347-56.
 158. Martínez-Salió A, Gómez de la Cámara A, Ribera MV, Montero J, Blanco E, Collado A, et al. Conferencia de Consenso: Diagnóstico y tratamiento del dolor neuropático. Med Clin (Barc). 2009;133(16):629-36.
 159. Busquets C, Faulí A. Novedades en el tratamiento del dolor neuropático. Semin Fund Esp Reumatol. 2012;13(03):103-9.
 160. RD 183/2008, de 8 de febrero, por el que se determinan y clasifican las especialidades en Ciencias de la Salud y se desarrollan determinados aspectos de formación sanitaria especializada. Boletín Oficial del Estado. Núm. 45. Jueves 21 febrero 2008.
 161. Ericsson KA, Charness N. Expert performance: its structure and acquisition. Am Psychol. 1994;49:725-747.
 162. Norman G. Building on experience. The development of clinical reasoning. N Engl J Med. 2006;355:2251-2252.
 163. Kassirer JP, Wong JB, Kopelman RI. Learning Clinical Reasoning. Philadelphia: Wolters Kluwer, Lippincott Williams and Wilkins; 2009.
 164. Kassirer JP. Teaching Clinical Reasoning: Case-Based and Coached. Acad Med. 2010;85:1118-1124.
 165. Eva KW. What every teacher needs to know about clinical reasoning. Med Educ. 2004;39:98-106.
 166. Bowen JL. Educational strategies to promote clinical diagnostic reasoning. N Engl J Med. 2006;355:2217-2225.
 167. Cañadas Osinski I, Sánchez Bruno A. Categorías de respuesta en escalas tipo Likert. Psicothema. 1998;10(3):623-631.
 168. Fernández de Pinedo I. NTP 15: Construcción de una escala de actitudes tipo Likert. Barcelona: Centro de Investigación y Asistencia Técnica. Ministerio de Trabajo y Asuntos Sociales;1982.
 169. Constitución Española de 1978. Título X, de la Reforma Constitucional: arts. 167,168.
 170. Crespo MG, Almenar L, Alonso-Pulpón L; Campreciós M, Cuenca JJ, De la Fuente L. Conferencia de Consenso de los Grupos Españoles de Trasplante Cardíaco. Rev Esp Cardiol Supl. 2007;7:4-54.

171. Campbell SM, Hann M, Roland MO, Quayle JA, Shekelle PG. The effect of panel membership and feedback on ratings in a two-round Delphi survey: results on randomized controlled trial. *Med Care*. 1999;37:964-8.
172. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA*. 1999;282:1458-1465.
173. Francke AL, Smit MC, de Veer AJE, Mistiaen P. Factors influencing the implementation of clinical guidelines for health care professionals: A systematic meta-review. *BMC Medical Informatics and Decision Making*. 2008;8:38. [Internet]. Disponible en: <http://www.biomedcentral.com/1472-6947/8/38>.
174. Grol R, Buchan H. Clinical guidelines: what can we do to increase their use?. *MJA*. 2006;185(6):301-302.
175. Grol R, Wensing M, Eccles M. Improving patient care: the implementation of change in clinical practice. Oxford: Elsevier;2004.
176. Grimshaw JM, Thomas RE, MacLennan G, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess*. 2004;8(6):1-72.
177. Grol R, Grimshaw J: From best evidence to best practice: effective implementation of change in patients' care. *Lancet*. 2003;362:1225-1230.
178. Tomasik T. Reliability and validity of the Delphi method in guideline development for family physicians. *Qual Prim Care*. 2010;18(5):317-26.
179. Plaza V, Sanchís J, Roura P, Molina J, Calle M, Quirce S, et al. Physician's knowledge of inhaler devices and inhalation techniques remain poor in Spain. *J Aerosol Med Pulm Drug Deliv*. 2012;25(1):16-22.
180. Egocheaga MI, Alcázar R, Lobos JM, Górriz JL, Martínez-Castelao A, Pastor A, Martínez I, Caballero F, Sánchez-Celaya M. Conocimiento y aplicación en la práctica real del Documento de consenso S.E.N.-semFYC sobre la enfermedad renal crónica. *Nefrología*. 2012;32(6):797-808.
181. Escobar Y, Rodriguez C, Caballero F, Recuero V, Camps C. Professional survey on knowledge and clinical patterns of pain management in Spanish medical oncology. *Clin Transl Oncol* 2010; 12: 819-824.
182. Caballero-Martínez F, León-Vázquez F, Paya-Pardo A, Díaz-Holgado A. Uso de recursos sanitarios y pérdida de productividad en pacientes con depresión atendidos en atención primaria: Estudio InterDep. *Actas Españolas de Psiquiatría*, 2014 (en prensa, ref. 1857).
183. Foro Dislipidemia Aterogénica. Consenso multidisciplinar sobre dislipidemia aterogénica. *Clin Inves Arterioscl*. 2013;25(2):83-91.
184. Millán Núñez-Cortés J, Pedro-Botet J, Brea-Hernando Á, Díaz-Rodríguez Á, González-Santos P, Hernández-Mijares A, Mantilla-Morató T, Pintó-Sala X, Simó R. Use of expert consensus to improve atherogenic dyslipidemia management. *Rev Esp Cardiol (Engl Ed)*. 2014; 67(1):36-44.
185. Esteban M, Salinas J, Arlandis S, Díez J, Jiménez M, Rebassa M, Angulo JC. Consenso experto sobre la evidencia científica disponible acerca del uso de toxina botulínica en vejiga hiperactiva. *Actas Urol Esp*. 2014;38(4):209-16.

186. Cannata-Andía J, Vicente Torregrosa J. Los nefrólogos españoles y el manejo de las alteraciones del metabolismo óseo-mineral en la enfermedad renal crónica. *Nefrología* 2014;34(2):175-88.
187. Pérez C. ¿Deben estar las técnicas de consenso incluidas entre las técnicas de investigación cualitativa?. *Rev Esp Salud Pública* 2000;74: 319-321.

Anexos

ANEXOS

ANEXO 1: Instrucciones al comité científico para la elaboración del cuestionario Delphi.

ANEXO 2: Niveles de evidencia. Oxford Centre for Evidence-Based Medicine.

ANEXO 3: Declaración aclaratoria del autor principal de los estudios en los que el doctorando no es primer o último firmante (autor senior).

ANEXO 4. Originales en facsímil de las publicaciones compendiadas en esta tesis, en orden cronológico de publicación (2009-1013).

ANEXO 1. INSTRUCCIONES AL COMITÉ CIENTÍFICO PARA LA ELABORACIÓN DEL CUESTIONARIO DELPHI

1. DEFINICIÓN BLOQUES TEMÁTICOS

Inicie este proceso tratando de identificar de la forma más exhaustiva posible los bloques de contenidos generales que deberían ser incluidos en el consenso, a modo de índice temático de la encuesta a desarrollar (por ejemplo: criterios diagnósticos, uso de pruebas complementarias, tratamiento médico, etc.).

2. DEFINICIÓN DE PREGUNTAS CLÍNICAS

Tras ello debe formular preguntas abiertas que revisen exhaustivamente los posibles contenidos clínicos relevantes de cada bloque temático, esto es, las dudas habituales o las preocupaciones razonables que pueda plantearse un médico en su ejercicio clínico (por estar sometidas a incertidumbre clínica o controversia profesional).

Esta tarea inicial no requiere ningún trabajo bibliográfico previo, sino que puede ser abordada a partir del conocimiento especializado y la experiencia clínica personal de cada miembro del Comité.

No limite a priori las posibles “preguntas clínicas” sobre las que se articulará argumentalmente este proyecto. Lo que sí habrá que acotar, en un momento posterior, será el número de “ítems” que conformarán el cuestionario, esto es, el número de recomendaciones o criterios clínicos con los que se intentará dar respuesta a las preguntas clínicas que ahora les pedimos identificar.

3. ELABORACIÓN ÍTEMS

Cuando disponga de las preguntas clínicas, puede proceder a la redacción de los ítems del cuestionario Delphi definitivo para cada bloque temático.

Los criterios y recomendaciones que se someterán finalmente a la valoración experta del panel Delphi, como ítems de encuesta, intentarán responder expresamente a las preguntas clínicas que han definido previamente. Es decir, trate de posicionarse explícitamente, según su opinión, definiendo cuáles serían las prácticas o los criterios profesionales que considera más acertados respecto a cada pregunta clínica. Para cada pregunta formulada pueden redactar uno o varios *ítems*.

Normas básicas para redactar los ítems:

- Se deben redactar de forma que resulten autoexplicativos y que puedan ser respondidos mediante la siguiente escala de opinión única prevista para todas las respuestas:

DESACUERDO			NI ACUERDO, NI DESACUERDO			ACUERDO		
1	2	3	4	5	6	7	8	9
↑						↑		
PLENO DESACUERDO						PLENO ACUERDO		

- Cada ítem debe recoger, en forma de aseveración, un criterio profesional o una recomendación clínica concreta sobre cualquier aspecto que, a su juicio, pueda ser motivo de controversia o insuficiente conocimiento. Al elegir los contenidos recuerde que (en caso de lograrse un consenso), los resultados del consenso serán difundidos como la opinión experta de los colectivos implicados sobre el asunto valorado.
- No son apropiados ítems que exploren “como un examen” el conocimiento de los expertos sobre datos bien establecidos (p. ej., tasas de prevalencia poblacional de un trastorno, eficacia de un fármaco demostrada en determinado estudio, etc.), o que propongan consensuar recomendaciones sobre aspectos que no admiten debate o controversia profesional alguna (p. ej., contraindicaciones formales de los medicamentos según ficha técnica, que son aceptadas unánimemente por el colectivo).
- Otro error común es incluir en un mismo ítem varios aspectos conceptuales independientes sobre los que el encuestado puede tener opiniones diversas, y ante lo que no podría posicionarse con una única opción acuerdo/desacuerdo. En tal caso la cuestión debe subdividirse en varias preguntas.
- Es preferible que los ítems se redacten de forma afirmativa para evitar errores de interpretación de los panelistas por doble negación (“**No** estoy de acuerdo con que **no** está indicado...”). A veces esto no será posible.
- Los ítems pueden recoger tanto la opinión personal de quien los redacta, como, por el contrario, un posicionamiento con el que no se está de acuerdo. El proceso de encuesta libre se encargará de constatar el acuerdo/desacuerdo real del panel experto con dichas afirmaciones.

Ejemplos de preguntas clínicas (rojo) y redactado final de ítems del cuestionario (*cursiva*)

(Nota: El cuestionario final solo incluiría los ítems en cursiva)

- **¿Qué papel tiene la hiperglucemia en el daño renal del diabético?**
La hiperglucemia es tanto un factor desencadenante como un factor que contribuye a la progresión en el daño renal del diabético.
- **¿Cuántas determinaciones son necesarias para establecer el diagnóstico de nefropatía en la DM?**
Para establecer el diagnóstico de nefropatía en la DM se precisan 2 de 3 determinaciones positivas en un plazo de 3-6 meses.
- **¿Deben ser distintos los objetivos de HbA1c en el paciente con diabetes y enfermedad renal avanzada?**
Los objetivos de HbA1c en el paciente con DM y enfermedad renal avanzada han de ser los mismos que en el resto de pacientes diabéticos.
- **¿Puede revertir la microalbuminuria con la mejora del control metabólico en el paciente diabético?**
El control metabólico del paciente con DM puede revertir la progresión de microalbuminuria a nefropatía.

ANEXO 2. NIVELES DE EVIDENCIA. Oxford Centre for Evidence-Based Medicine.

(Reproducido de: Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A et al (OCEBM Levels of Evidence Working Group). The Oxford 2011 Levels of Evidence. Oxford Centre for Evidence- Based Medicine.).

Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence

Question	Step 1 (Level 1*)	Step 2 (Level 2*)	Step 3 (Level 3*)	Step 4 (Level 4*)	Step 5 (Level 5)
How common is the problem?	Local and current random sample surveys (or censuses)	Systematic review of surveys that allow matching to local circumstances**	Local non-random sample**	Case-series**	n/a
Is this diagnostic or monitoring test accurate? (Diagnosis)	Systematic review of cross sectional studies with consistently applied reference standard and blinding	Individual cross sectional studies with consistently applied reference standard and blinding	Non-sequential studies, or studies without consistently applied reference standards**	Case-control studies, or poor or non-independent reference standard**	Mechanism-based reasoning
What will happen if we do not add a therapy? (Prognosis)	Systematic review of inception cohort studies	Inception cohort studies	Cohort study or control arm of randomized trial*	Case-series or case-control studies, or poor quality prognostic cohort study**	n/a
Does this intervention help? (Treatment Benefits)	Systematic review of randomized trials or n-of-1 trials	Randomized trial or observational study with dramatic effect	Non-randomized controlled cohort/follow-up study**	Case-series, case-control studies, or historically controlled studies**	Mechanism-based reasoning
What are the COMMON harms? (Treatment Harms)	Systematic review of randomized trials; systematic review of nested case-control studies; n-of-1 trial with the patient you are raising the question about, or observational study with dramatic effect	Individual randomized trial or (exceptionally) observational study with dramatic effect	Non-randomized controlled cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.)**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning
What are the RARE harms? (Treatment Harms)	Systematic review of randomized trials or n-of-1 trial	Randomized trial or (exceptionally) observational study with dramatic effect			
Is this (early detection) test worthwhile? (Screening)	Systematic review of randomized trials	Randomized trial	Non-randomized controlled cohort/follow-up study**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning

* Level may be graded down on the basis of study quality, imprecision, indirectness (study PICO does not match questions PICO), because of inconsistency between studies, or because the absolute effect size is very small; Level may be graded up if there is a large or very large effect size.

** As always, a systematic review is generally better than an individual study.

How to cite the Levels of Evidence Table

OCEBM Levels of Evidence Working Group*. "The Oxford 2011 Levels of Evidence".

Oxford Centre for Evidence-Based Medicine. <http://www.cebm.net/index.aspx?oc=5653>

* OCEBM Table of Evidence Working Group = Jeremy Howick, Iain Chalmers (James Lind Library), Paul Glasziou, Trish Greenhalgh, Carl Heneghan, Alessandro Liberati, Ivan Moschetti, Bob Phillips, Hazel Thornton, Olive Goddard and Mary Hodgkinson

ANEXO 3. DECLARACIÓN ACLARATORIA DEL AUTOR PRINCIPAL DE LOS ESTUDIOS EN LOS QUE EL DOCTORANDO NO ES PRIMER O ULTIMO FIRMANTE (AUTOR SENIOR).

1. Ampudia Blasco FJ, Baos Vicente V, **Caballero Martínez F**, Álvarez-Sala Walther LA. Criterios profesionales sobre el papel de ezetimiba en el manejo clínico del paciente dislipémico en el ámbito de atención primaria. *Av Diabetol* 2009; 25(5): 399-407.
2. Villanueva Haba V, Sánchez Álvarez JC, Peña Mayor P, Salas Puig J, **Caballero Martínez F**, Gil-Nagel Rein A. Treatment initiation in epilepsy. An expert consensus in Spain. *Epilepsy and Behaviour* 2010; 19(3): 332-42.
3. Bobes García J, Saiz Ruiz J, Bernardo Arroyo M, **Caballero Martínez F**, Gilaberte Asín I, Ciudad Herrera A, en representación del Grupo de Trabajo sobre la Salud Física del Paciente con Trastorno Mental Grave. Consenso Delphi sobre la salud física del paciente con esquizofrenia: valoración por un panel de expertos de las recomendaciones de las Sociedades Españolas de Psiquiatría y de Psiquiatría Biológica. *Actas Esp Psiquiatr* 2012; 40(3): 114-28.

DECLARACIÓN ACLARATORIA DE AUTORIA

JULIO BOBES GARCIA, como autor principal y primer firmante de la publicación:

Bobes García J, Saiz Ruiz J, Bernardo Arroyo M, Caballero Martínez F, Gilaberte Asín I, Ciudad Herrera A, en representación del Grupo de Trabajo sobre la Salud Física del Paciente con Trastorno Mental Grave. Consenso Delphi sobre la salud física del paciente con esquizofrenia: valoración por un panel de expertos de las recomendaciones de las Sociedades Españolas de Psiquiatría y de Psiquiatría Biológica. Actas Esp Psiquiatr 2012; 40(3): 114-28

Declaro que **FERNANDO CABALLERO MARTÍNEZ**, coautor de dicho estudio, se responsabilizó de las tareas de apoyo metodológico al Comité Científico del proyecto, durante el diseño y desarrollo del proceso de consenso. Asimismo, coordinó el trabajo de procesamiento y análisis de datos, y asistió al Comité en la interpretación de los resultados del consenso.

Los resultados publicados, hasta donde conozco, no han sido empleados como material principal de ninguna otra tesis doctoral desde su publicación.

Para que a si conste, y a petición de la Universidad Autónoma de Madrid, firmo la presente en Oviedo, a día 9 de septiembre de 2014.



Fdo. Julio Bobes García

DECLARACIÓN ACLARATORIA DE AUTORIA

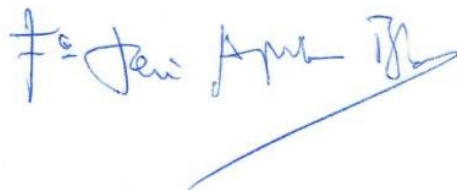
FRANCISCO JAVIER AMPUDIA BLASCO, como autor principal y primer firmante de la publicación:

Ampudia Blasco FJ, Baos Vicente V, Caballero Martínez F, Álvarez-Sala Walther LA. Criterios profesionales sobre el papel de ezetimiba en el manejo clínico del paciente dislipémico en el ámbito de atención primaria. Av Diabetol 2009; 25(5): 399-407.

Declaro que **FERNANDO CABALLERO MARTÍNEZ**, coautor de dicho estudio, se responsabilizó de las tareas de apoyo metodológico al Comité Científico del proyecto, durante el diseño y desarrollo del proceso de consenso. Asimismo, coordinó el trabajo de procesamiento y análisis de datos, y asistió al Comité en la interpretación de los resultados del consenso.

Los resultados publicados, hasta donde conozco, no han sido empleados como material principal de ninguna otra tesis doctoral desde su publicación.

Para que a sí conste, y a petición de la Universidad Autónoma de Madrid, firmo la presente en Valencia, a día 5 de septiembre de 2014.



Fdo. Francisco Javier Ampudia Blasco

DECLARACIÓN ACLARATORIA DE AUTORIA

VICENTE VILLANUEVA HABA, como autor principal y primer firmante de la publicación:

Villanueva Haba V, Sánchez Álvarez JC, Peña Mayor P, Salas Puig J, Caballero Martínez F, Gil-Nagel Rein A. Treatment initiation in epilepsy. An expert consensus in Spain. Epilepsy and Behaviour 2010; 19(3): 332-42.

Declaro que **FERNANDO CABALLERO MARTÍNEZ**, coautor de dicho estudio, se responsabilizó de las tareas de apoyo metodológico al Comité Científico del proyecto, durante el diseño y desarrollo del proceso de consenso. Asimismo, coordino el trabajo de procesamiento y análisis de datos, y asistió al Comité en la interpretación de los resultados del consenso.

Los resultados publicados, hasta donde conozco, no han sido empleados como material principal de ninguna otra tesis doctoral desde su publicación.

Para que a si conste, y a petición de la Universidad Autónoma de Madrid, firmo la presente en Valencia, a día 8 de septiembre de 2014.



Fdo. Vicente Villanueva Haba

ANEXO 4. ORIGINAL EN FACSIMIL DE LAS PUBLICACIONES COMPENDIADAS EN ESTA TESIS, EN ORDEN CRONOLÓGICO DE PUBLICACIÓN (2009-1013).

1. Ampudia Blasco FJ, Baos Vicente V, **Caballero Martínez F**, Álvarez-Sala Walther LA. Criterios profesionales sobre el papel de ezetimiba en el manejo clínico del paciente dislipémico en el ámbito de atención primaria. *Av Diabetol* 2009; 25(5): 399-407.
2. **Caballero Martínez F**, Plaza Moral V, Quirce Gancedo S, Fernández Benítez M, Gómez Ruiz F, López Viña A, Molina París J, Quintano Jiménez JA, Soler Vilarrasa R, Villa Asensi JR, Balmes Estrada S. Valoración externa de las recomendaciones de la GEMA 2009 por un panel multiprofesional de expertos en asma. *Arch Bronconeumol* 2010; 46(8): 411-419.
3. Villanueva Haba V, Sánchez Álvarez JC, Peña Mayor P, Salas Puig J, **Caballero Martínez F**, Gil-Nagel Rein A. Treatment initiation in epilepsy. An expert consensus in Spain. *Epilepsy and Behaviour* 2010; 19(3): 332-42.
4. Medrano López C, García-Guereta L, Fernández Pineda L, Malo Concepción P, Maroto Álvaro E, Santos de Soto J, Lirio Casero J, Suárez Cabrera P, **Caballero Martínez F**. Consenso clínico sobre la profilaxis de la infección por virus respiratorio sincitial (VRS) y el uso del palivizumab en cardiología pediátrica. *An Pediatr (Barc)*. 2010; 72(6): 432.e1-432.e13.
5. Martín Carrasco M, Agüera Ortiz L, Caballero Martínez L, Cervilla Ballesteros J, Menchón Magriñá JM, Montejo González AL, Moríñigo Domínguez A, **Caballero Martínez F**. Consenso clínico sobre la depresión en el paciente anciano. *Actas Esp Psiquiatr* 2011; 39(1): 20-31.
6. Bobes García J, Saiz Ruiz J, Bernardo Arroyo M, **Caballero Martínez F**, Gilaberte Asín I, Ciudad Herrera A, en representación del Grupo de Trabajo

sobre la Salud Física del Paciente con Trastorno Mental Grave. Consenso Delphi sobre la salud física del paciente con esquizofrenia: valoración por un panel de expertos de las recomendaciones de las Sociedades Españolas de Psiquiatría y de Psiquiatría Biológica. *Actas Esp Psiquiatr* 2012; 40(3): 114-28.

7. Plaza Vicente, Calle M, Molina J, Quirce S, Sanchis J, Viejo JL, **Caballero F**. Validación, mediante el método Delphi, del Consenso sobre Terapia Inhalada para el tratamiento de las enfermedades respiratorias obstructivas. *Arch Bronconeumol* 2012; 48(6): 189-196.
8. Cózar Olmo JM, Hernández Fernández C, Miñana López B, Amón Sesmero JH, Montleó González M, Rodríguez Antolín A, **Caballero Martínez F**. Consenso experto sobre el impacto clínico de la nueva evidencia científica disponible sobre Hiperplasia Benigna Prostática. *Actas Urol Esp* 2012; 36(5): 265-275.
9. Pérez C, Rodríguez MJ, Guerrero A, Margarit C, Martín-Estefanía C, Oteo-Álvaro A **Caballero-Martínez F**. Consenso experto sobre el uso clínico de los tratamientos por vía tópica en el manejo del dolor neuropático periférico. *Rev. Soc. Esp. Dolor*. 2013; 20(6): 308-323.