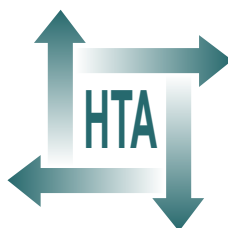


Efectividad y eficiencia de las estrategias de difusión e implementación de directrices

J.M. Grimshaw, R.E. Thomas, G. MacLennan, C. Fraser,
C.R. Ramsay, L. Vale, P. Whitty, M.P. Eccles, L. Matowe, L. Shirran,
M. Wensing, R. Dijkstra y C. Donaldson



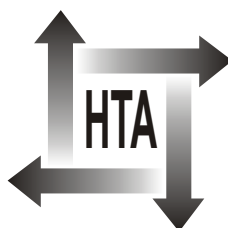
Health Technology Assessment

SOCIEDAD ESPAÑOLA DE FARMACÉUTICOS DE ATENCIÓN PRIMARIA



Efectividad y eficiencia de las estrategias de difusión e implementación de directrices

J.M. Grimshaw, R.E. Thomas, G. MacLennan, C. Fraser,
C.R. Ramsay, L. Vale, P. Whitty, M.P. Eccles, L. Matowe, L. Shirran,
M. Wensing, R. Dijkstra y C. Donaldson



Health Technology Assessment

SOCIEDAD ESPAÑOLA DE FARMACÉUTICOS DE ATENCIÓN PRIMARIA





Efectividad y eficiencia de las estrategias de difusión e implementación de directrices

J.M. Grimshaw, R.E. Thomas, G. MacLennan, C. Fraser,
C.R. Ramsay, L. Vale, P. Whitty, M.P. Eccles, L. Matowe, L. Shirran,
M. Wensing, R. Dijkstra y C. Donaldson

Reservados todos los derechos. Ninguna parte de esta publicación puede ser reproducida, transmitida en ninguna forma o medio alguno, electrónico o mecánico, incluyendo las fotocopias, grabaciones o cualquier sistema de almacenamiento de información sin el permiso escrito de los titulares del Copyright.

De la edición original:

© Crown copyright material is reproduced with the permission of the Controller of HMSO.

Este material es la traducción de:

Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004;8(6).

De la edición en español:

© 2005, Prous Science, S.A.

Provenza 388, 08025 Barcelona

Dep. legal: ????????

Impreso en España

Junta directiva de la SEFAP

Presidente

MÓNICA AUSEJO SEGURA
Sociedad Española de Farmacéuticos de
Atención Primaria
C/ Villanueva, 11 7ª planta
28001 Madrid
Tfno: 914312560 Fax: 915763905
mausejos@sefap.org

Tesorero

ÁNGEL MATAIX SANJUÁN
Sociedad Española de Farmacéuticos de
Atención Primaria
C/ Villanueva, 11 7ª planta
28001 Madrid
Tfno: 914312560 Fax: 915763905
amataix@sefap.org

Secretario

ENCARNACIÓN CRUZ MARTOS
Sociedad Española de Farmacéuticos de
Atención Primaria
C/ Villanueva, 11 7ª planta
28001 Madrid
Tfno: 914312560 Fax: 915763905
ecruz1@sefap.org

Vocales

Andalucía Occidental
JUAN CARLOS MORALES SERNA
Jmoraless1@sefap.org

Andalucía Oriental
ESTER MARTOS LÓPEZ
emartosl@sefap.org

Aragón – Cantabria
MARÍA CONCEPCIÓN CELAYA LECEA
mcelayal@sefap.org

Asturias
BELÉN CONCEJO ALFARO
bconcejoa1@sefap.org

Canarias
VICENTE OLMO QUINTANA
volmoq@sefap.org

Castilla la Mancha
ÁNGELES LLORET CALLEJO
allorete1@sefap.org

Castilla y León
Mª MONTSERRAT PAJARES FERNÁNDEZ
mpajaresf@sefap.org

Cataluña
ARANTXA CATALÁN RAMOS
acatalanr@sefap.org

Extremadura
VENANCIO JESÚS PÉREZ GARCÍA
vperezg@sefap.org

Madrid
MARÍA JOSÉ CALVO ALCÁNTARA
mcalvoa1@sefap.org

Navarra
CRISTINA AGUDO PASCUAL
cagudop1@sefap.org

País Vasco
SUSANA FERNÁNDEZ GALLASTEGUI
sfernandezga@sefap.org

Valencia
MARÍA SOLEDAD GALEOTE MAYOR
mgaleotem1@sefap.org

Galicia
CARMEN DURÁN PARRONDO
cduranp1@sefap.org

Carta de presentación

La colección **Series Divulgativas SEFAP** recoge en este nuevo número la traducción de la monografía *Effectiveness and efficiency of guideline dissemination and implementation strategies*, recientemente publicada en *Health Technology Assessment 2004*; vol. 8: No. 6.

Se trata de una revisión sistemática de las estrategias de difusión e implementación de directrices más extensa (abarca estudios incluidos en el periodo 1966-1998), metodológicamente sofisticada y actualizada de todas las revisiones disponibles hasta la fecha.

Esta investigación, realizada por autores de máximo prestigio, incluye, además, la primera revisión sistemática de las evaluaciones económicas y los análisis de costes realizados en evaluaciones de este tipo de estrategias. De esta revisión destaca la escasez de evaluaciones económicas de alta calidad, lo que limita la capacidad de extraer conclusiones firmes acerca de los recursos que se necesitan y la eficiencia de las distintas estrategias de difusión e implementación de directrices empleadas.

Una pequeña encuesta efectuada entre informantes clave del sistema sanitario británico que explora la factibilidad y los recursos que se necesitarían para desarrollar diversas intervenciones en este entorno sanitario completa este documento.

De la lectura de este documento se desprende que existen pocas evidencias científicas en las que fundamentar las decisiones sobre cuáles son las estrategias de difusión e implantación de directrices que tienen probabilidad de ser más eficientes en diferentes circunstancias. Es obvio que dicha decisión debe orientarse a utilizar de la mejor manera posible los escasos recursos disponibles maximizando los beneficios para la población.

En cualquier caso, espero que el análisis de los conocimientos disponibles en este campo que nos proporciona esta nueva publicación sea de tu interés y que contribuya a la mejora del ejercicio de nuestra profesión.

Arantxa Catalán
Coordinadora Series Divulgativas
Vocal de Cataluña y Baleares
Sociedad Española de Farmacéuticos de Atención Primaria

Efectividad y eficiencia de las estrategias de difusión e implementación de directrices

JM Grimshaw,^{1*} RE Thomas,¹ G MacLennan,¹ C Fraser,¹ CR Ramsay,¹ L Vale,^{1,2}
P Whitty,³ MP Eccles,⁴ L Matowe,^{1†} L Shirran,¹ M Wensing,⁵ R Dijkstra⁵ y C Donaldson^{6‡}

¹Unidad de Investigación de Servicios Sanitarios, Universidad de Aberdeen, Reino Unido

²Unidad de Investigación de Economía Sanitaria, Universidad de Aberdeen, Reino Unido

³Departamento de Epidemiología y Salud Pública, Universidad de Newcastle-upon-Tyne, Reino Unido

⁴Centro de Investigación de Servicios Sanitarios, Universidad de Newcastle-upon-Tyne, Reino Unido

⁵Centro de Investigación en la Calidad Asistencial, Universidad de Nijmegen, Holanda

⁶Departamento de Ciencias de la Salud Comunitaria, Universidad de Calgary, Canadá

*Autor para correspondencia. Afiliación actual: Programa de Epidemiología Clínica, Instituto de Investigación de la Salud de Ottawa y Centro de Mejores Prácticas, Instituto de Salud Poblacional, Universidad de Ottawa, Canadá

†Afiliación actual: Departamento de Práctica Farmacéutica, Facultad de Farmacia, Universidad de Kuwait, Kuwait

‡Afiliación actual: Centro de Investigación de Servicios Sanitarios, Universidad de Newcastle-upon-Tyne, Reino Unido

Resumen

Efectividad y eficiencia de las estrategias de difusión e implementación de directrices

JM Grimshaw,^{1*} RE Thomas,¹ G MacLennan,¹ C Fraser,¹ CR Ramsay,¹ L Vale,^{1,2} P Whitty,³ MP Eccles,⁴ L Matowe,^{1†} L Shirran,¹ M Wensing,⁵ R Dijkstra⁵ y C Donaldson^{6‡}

¹ Unidad de Investigación de Servicios Sanitarios, Universidad de Aberdeen, Reino Unido

² Unidad de Investigación de Economía Sanitaria, Universidad de Aberdeen, Reino Unido

³ Departamento de Epidemiología y Salud Pública, Universidad de Newcastle-upon-Tyne, Reino Unido

⁴ Centro de Investigación de Servicios Sanitarios, Universidad de Newcastle-upon-Tyne, Reino Unido

⁵ Centro de Investigación en la Calidad Asistencial, Universidad de Nijmegen, Holanda

⁶ Departamento de Ciencias de la Salud Comunitaria, Universidad de Calgary, Canadá

* Autor para correspondencia. Afiliación actual: Programa de Epidemiología Clínica, Instituto de Investigación de la Salud de Ottawa y Centro de Mejores Prácticas, Instituto de Salud Poblacional, Universidad de Ottawa, Canadá

† Afiliación actual: Departamento de Práctica Farmacéutica, Facultad de Farmacia, Universidad de Kuwait, Kuwait

‡ Afiliación actual: Centro de Investigación de Servicios Sanitarios, Universidad de Newcastle-upon-Tyne, Reino Unido

Objetivos: Llevar a cabo una revisión sistemática de la efectividad y los costes de diferentes estrategias de desarrollo, difusión e implementación de directrices. Calcular los recursos que necesitan estas estrategias. Elaborar un marco de referencia para decidir cuándo es eficiente desarrollar e introducir unas directrices clínicas.

Fuentes de datos: MEDLINE, Healthstar, Cochrane Controlled Trial Register, EMBASE, SIGLE y el registro especializado del grupo Cochrane Effective Practice and Organisation of Care (EPOC).

Métodos de revisión: De cada comparación de estudios, se derivaron estimaciones individuales de variables dicotómicas de proceso basadas en la variable principal de valoración o en la mediana de diversas variables de valoración. Se hicieron análisis separados de las comparaciones de distintos tipos de intervención. El estudio también exploró si los efectos de intervenciones múltiples aumentaban con el número de componentes de la intervención. También se evaluaron los estudios que aportaban datos económicos. Se llevó a cabo una encuesta para valorar si las estrategias de difusión e implementación de directrices eran factibles en el Reino Unido, y qué recursos serían necesarios con más probabilidad; para ello se contó con informantes clave procedentes de la atención primaria y de la atención especializada.

Resultados: En total, 235 estudios con un total de 309 comparaciones cumplieron los criterios de inclusión; de éstos, el 73% de las comparaciones evaluaban intervenciones múltiples, aunque el número máximo de replicaciones de una intervención múltiple específica fue de 11 comparaciones. Globalmente, la mayoría de las comparaciones que aportaban datos dicotómicos de proceso observaron mejoras en la asistencia; sin embargo, hubo una variación considerable en los efectos observados tanto dentro de una misma intervención como entre diferentes intervenciones. Las intervenciones individuales evaluadas con mayor frecuencia fueron las de recordatorios, difusión de material educativo y auditoría y retroalimentación. Hubo 23 comparaciones de intervenciones múltiples que incluían reuniones educativas. La mayoría de las intervenciones observaron mejoras entre modestas y moderadas de la atención asistencial. No se encontró ninguna relación entre el número de componentes de las intervenciones y los efectos de las intervenciones múltiples. Sólo el 29,4% de las comparaciones aportaron datos económicos. La mayoría de los estudios sólo informaron sobre los costes del tratamiento; únicamente 25 estudios facilitaron datos sobre los costes del desarrollo de unas directrices o de su difusión e implementación. La mayoría de los estudios utilizaron medidas del proceso como variable principal de valoración, a pesar del hecho de que sólo tres directrices se basaron explícitamente en pruebas (y pueden no

haber sido eficientes). En pocas ocasiones los informantes clave encuestados identificaron presupuestos para financiar las estrategias de difusión e implementación de directrices. En general, opinaron que con los recursos existentes sólo eran factibles la difusión de material educativo y los encuentros educativos de corta duración (durante el almuerzo).

Conclusiones: Existe una base imperfecta de pruebas en la que fundamentar las decisiones respecto a las estrategias de difusión e implementación de directrices que tienen más probabilidades de ser eficientes en determinadas circunstancias. A la hora de tomar la decisión, para maximizar los beneficios para la población, se debe tener un gran criterio a la hora de hacer el mejor uso de los limitados recursos de que se dispo-

ne para la gestión clínica y las actividades relacionadas. Se debe pensar en las áreas susceptibles de ser objeto de actividades de efectividad clínica, en los costes y los beneficios probables asociados a la introducción de directrices y a los asociados a cualquier cambio en la conducta de los proveedores. Se necesita más investigación para: desarrollar y validar un marco teórico coherente de comportamiento del profesional sanitario y organizativo y del cambio conductual para tomar decisiones con más información a la hora de elegir las intervenciones en entornos de investigación y servicios, así como para estimar la eficiencia de las estrategias de difusión e implementación cuando se den distintas barreras y distintos efectos moduladores.

Índice

Lista de abreviaturas	XV
Resumen ejecutivo	XVII
1. Introducción	1
Antecedentes	1
Objetivos del estudio	2
Estructura de este informe	2
2. Revisión sistemática de las estrategias de difusión e implementación de directrices: objetivos y métodos	3
Objetivos	3
Métodos	3
3. Revisión sistemática de las estrategias de difusión e implementación de directrices: información sobre los estudios incluidos	11
Resultado de la búsqueda de bibliografía	11
Calidad metodológica de los estudios incluidos	11
Otros aspectos metodológicos	13
Datos de los participantes	14
Datos de las intervenciones	15
4. Revisión sistemática de las estrategias de difusión e implementación de directrices: resultados	17
Evaluación de intervenciones únicas frente a grupos de control sin intervención	17
Evaluación de intervenciones únicas frente a grupos de control con intervención	23
Evaluación de intervenciones múltiples	23
5. Revisión sistemática de las estrategias de difusión e implementación de directrices: evaluaciones económicas y análisis de costes	43
Estructura para la evaluación económica de la implementación de directrices	43
Métodos	44
Resultados	45
Discusión	47
Conclusiones	51
6. Estimación de la factibilidad y de los recursos que son necesarios para las estrategias de difusión e implementación de directrices en el marco asistencial británico	53
Métodos	53
Resultados	54
Uso relativo de recursos en las distintas estrategias	59
Discusión	59
7. Discusión	61
Principales resultados del estudio	61
Puntos fuertes y puntos débiles	62
Interés del estudio: posibles mecanismos y repercusiones para médicos y para diseñadores de políticas	63
Preguntas sin contestar y futuras investigaciones	65
Bibliografía	69

Lista de abreviaturas

ANOVA	Análisis de la varianza	LS	L. Shirran
AR	Auditoría y retroalimentación	Mateduc	Material educativo
ARIMA	Proceso autorregresivo integrado de media móvil	ME	M.P. Eccles
C-ECA	Ensayo clínico controlado con asignación aleatoria de conglomerados (cluster) (ámbito profesional)	Med.pac.	Intervención mediada por el paciente
C-ECC	Ensayo clínico controlado con asignación de conglomerados (cluster) (ámbito profesional)	MRC	Medical Research Council
CF	C. Fraser	MW	M. Wensing
DME	Diferencia media estandarizada	NC	No claro
ECA	Ensayo clínico controlado aleatorizado	NP	No procede
ECC	Ensayo clínico controlado	NS	No significativo
Enc.educ.	Encuentros educativos	Org.	Intervención organizativa
EPOC	Cochrane Effective Practice and Organisation of Care (grupo)	PCL	Procesos de consenso locales
Estr.	Intervención estructural	P-ECA	Ensayo clínico controlado con asignación aleatoria individual en el ámbito del paciente
EUA	Error de la unidad de análisis	P-ECC	Ensayo clínico controlado con asignación individual en el ámbito del paciente
Fin.	Intervención financiera	PW	P. Whitty
GM	G. MacLennan	RD	R. Dijkstra
Int.prof.	Otra intervención profesional	Rec.	Recordatorios
IQR	Rango intercuartil	Reg.	Intervención reguladora
JG	J.M. Grimshaw	RT	R.E. Thomas
LM	L. Matowe	S	Significativo
		Vis.educ.	Visitas educativas

Todas las abreviaturas utilizadas en este informe se incluyen aquí, a no ser que la abreviatura sea muy conocida, o se haya utilizado sólo una vez, o sea una abreviatura no estándar utilizada sólo en figuras/tablas/anexos, en cuyo caso la abreviatura se define en la leyenda de la figura o al final de la tabla.

Resumen ejecutivo

ANTECEDENTES

Las directrices o guías clínicas se están convirtiendo en un elemento cada vez más común en la asistencia sanitaria de todo el mundo. Estas directrices tienen el potencial de mejorar la atención al paciente, ya que favorecen las intervenciones con un efecto beneficioso comprobado, y facilitan que no se lleven a cabo intervenciones probablemente ineficaces. Sin embargo, desarrollar e introducir unas directrices tiene un coste y, en algunas circunstancias, éste puede ser superior a los beneficios que puede aportar. En algunas ocasiones puede resultar más eficiente adoptar estrategias de difusión e implementación menos costosas aunque resulten menos efectivas. Las organizaciones sanitarias locales disponen de relativamente pocos recursos para actividades centradas en la efectividad clínica así que, a la hora de diseñar una política, se debe plantear cómo usar mejor estos recursos para maximizar los beneficios.

OBJETIVOS

Los objetivos del estudio fueron:

- hacer una revisión sistemática de la efectividad y los costes de distintas estrategias de desarrollo, difusión e implementación de directrices
- estimar los recursos que requieren las distintas estrategias de desarrollo, difusión e implementación de directrices
- desarrollar un marco de referencia para decidir cuándo es eficiente elaborar e implantar unas directrices clínicas, en función de los costes y beneficios potenciales de la actividad clínica seleccionada, así como de la efectividad y los costes del desarrollo e implantación de esas directrices.

MÉTODOS

Revisión sistemática de la efectividad y la eficiencia de diferentes estrategias de difusión e implementación de directrices

Fuentes de datos

Se hicieron búsquedas en MEDLINE (1966–1998), Healthstar (1975–1998), Cochrane Controlled Trial Register (4ª edición

1998), EMBASE (1980–1998), SIGLE (1980–1988) y en el registro especializado del grupo Cochrane Effective Practice and Organisation of Care (EPOC), utilizando una estrategia de búsqueda estandarizada desarrollada a partir de búsquedas manuales en revistas clave. La estrategia de búsqueda tenía una sensibilidad del 93% y una especificidad del 18%.

Selección de estudios (criterios de inclusión)

- *tipos de diseño*: ensayos aleatorizados controlados, ensayos clínicos controlados, estudios controlados pre-post y series temporales interrumpidas
- *tipos de participantes*: profesionales sanitarios licenciados en medicina
- *tipos de intervenciones*: estrategias de difusión e implementación de directrices
- *tipos de resultados*: medidas objetivas del cumplimiento del profesional y/o del resultado para el paciente.

Recogida de datos (y valoración de su validez)

Dos revisores recogieron datos de forma independiente sobre la calidad metodológica de los estudios (utilizando los criterios de calidad metodológica del grupo Cochrane EPOC), las características del marco de estudio, los participantes, los comportamientos estudiados y las características de las intervenciones. Asimismo, los estudios que aportaban evaluaciones económicas y análisis de costes fueron valorados de acuerdo con las pautas del *British Medical Journal* para los revisores de evaluaciones económicas.

Síntesis de datos

De cada comparación basada en un criterio principal de valoración (definido por los autores del estudio) o en la mediana de varios criterios de valoración, se obtuvieron estimaciones individuales de variables de proceso dicotómicas (por ejemplo, la proporción de pacientes que recibían el tratamiento apropiado). Se intentó volver a analizar estudios que tenían algún defecto metodológico habitual. Se llevaron a cabo análisis

separados para comparaciones de intervenciones únicas frente a controles sin intervención, para intervenciones únicas frente a controles con intervención, para intervenciones múltiples frente a controles sin intervención y para intervenciones múltiples frente a controles con intervención. El estudio también exploró si los efectos de las intervenciones múltiples aumentaban con un mayor número de componentes. Para cada intervención, se comunicaron el número de comparaciones que mostraban una dirección positiva del efecto, la magnitud del efecto promedio entre todas las comparaciones, la magnitud del efecto promedio entre comparaciones sin errores de la unidad de análisis y el número de comparaciones que mostraban resultados estadísticamente significativos. No se pudo llevar a cabo el análisis de metarregresión que se había programado inicialmente debido al elevado número de distintas combinaciones de intervenciones múltiples.

Encuesta para estimar los recursos necesarios y la factibilidad de las diferentes estrategias de difusión e implementación de directrices en el Reino Unido

Se llevaron a cabo entrevistas telefónicas con informantes clave, tanto de la atención primaria como de la especializada.

RESULTADOS (HALLAZGOS DE LA INVESTIGACIÓN)

Revisión sistemática de la efectividad y la eficiencia de las estrategias de difusión e implementación de directrices

En total, cumplieron los criterios de inclusión 235 estudios que informaban de 309 comparaciones. La calidad global de los estudios era deficiente. El 73% de las comparaciones evaluaron intervenciones múltiples, aunque el número máximo de replicaciones de una intervención múltiple específica fue de 11 comparaciones. En conjunto, la mayoría de las comparaciones que presentaban datos de proceso dicotómicos (86,6%) observaron mejoras de la asistencia; sin embargo, hubo una variación considerable en los efectos observados tanto dentro de la misma intervención como entre diferentes intervenciones. Las intervenciones únicas evaluadas fueron fundamentalmente de recordatorios (38 comparaciones), difusión de material educativo (18 comparaciones) y auditoría y retroalimentación (12 comparaciones). Hubo 23 comparaciones de intervenciones múltiples que incluyeron visitas educativas. La mayoría de las intervenciones observaron mejoras asistenciales entre moderadas y modestas. Por ejemplo, la mediana de la mejora absoluta del cumplimiento en todas las intervenciones fue del 14,1% en 14 comparaciones aleatorizadas de conglomerados (*cluster*) sobre los recordatorios, del 8,1% en 4 comparaciones aleatorizadas de conglomerados

sobre la difusión de material educativo, del 7,0% en 5 comparaciones aleatorizadas de conglomerados de la auditoría y la retroalimentación, y del 6,0% en 13 comparaciones aleatorizadas de conglomerados de intervenciones múltiples que incluían visitas educativas. No se encontró ninguna relación entre el número de componentes de las intervenciones y los efectos de las intervenciones múltiples.

Sólo el 29,4% de las comparaciones aportaron datos económicos. En 11 casos se analizó la relación entre coste y efectividad, en 38 se analizó la relación entre costes y resultados (analizaron la relación entre las diferencias de costes y los diversos grados de efectividad) y, finalmente, 14 analizaron los costes (el análisis se centró en algún aspecto del coste pero no se relacionó con los beneficios). La mayoría de los estudios sólo aportaron los costes del tratamiento; sólo 25 estudios aportaron datos sobre los costes del desarrollo de directrices o de su difusión y su implementación. La mayoría de los estudios utilizaron medidas de proceso como criterio principal de valoración, a pesar del hecho de que sólo tres directrices se basaron explícitamente en pruebas (y pueden no haber sido eficientes). En conjunto, los métodos utilizados para las evaluaciones económicas y los análisis de costes fueron deficientes. Sólo en 10 estudios se expuso el punto de vista que se había adoptado en las evaluaciones económicas. Los métodos para calcular los costes fueron completos en aproximadamente la mitad de los estudios, y pocos estudios detallaron el uso de recursos. Debido a la mala calidad de presentación de la evaluación económica, en la mayoría de estudios no se disponía de datos sobre el uso de recursos y el coste del desarrollo, de la difusión y la implementación de las directrices; sólo 4 estudios aportaron datos suficientemente robustos como para poder recogerlos para la revisión.

Encuesta para estimar los recursos que necesitan y la factibilidad de las diferentes estrategias de difusión e implementación de directrices en el Reino Unido

En muy pocas ocasiones los encuestados identificaron presupuestos destinados a apoyar estrategias de difusión e implementación de directrices, e hicieron comentarios frecuentes sobre el uso de fondos de fuentes privadas o recursos para iniciativas específicas para costear estas actividades. En general, se opinó que, con los recursos disponibles, sólo eran factibles la difusión de material educativo y los encuentros educativos de corta duración (durante el almuerzo).

CONCLUSIONES: REPERCURSIONES EN LA ATENCIÓN SANITARIA Y RECOMENDACIONES PARA LA INVESTIGACIÓN

Existe una base imperfecta de pruebas en las que fundamentar las decisiones acerca de las estrategias de difusión e imple-

mentación de las directrices que tienen más probabilidades de ser eficientes en distintas circunstancias. A la hora de decidir, se debe tener un gran criterio para darle el mejor uso a los limitados recursos que hay para ese fin, de forma que se maximicen los beneficios para la población. Se debe pensar en aquellas áreas susceptibles de ser objeto de actividades de efectividad clínica, en los costes y los probables beneficios asociados a la introducción de directrices, así como los que se asocian a un cambio en la conducta de los profesionales. Se

necesita más investigación para: desarrollar y validar un marco teórico coherente de comportamiento profesional y organizativo y de cambio conductual, de forma que se pueda tener una mayor información a la hora de elegir las intervenciones en entornos de investigación y servicios, y para estimar la eficiencia de las diferentes estrategias de difusión e implementación cuando se den determinadas barreras y modificadores del efecto.

ANTECEDENTES

Las directrices o guías clínicas se están convirtiendo en un elemento cada vez más común en la atención clínica en todo el mundo. Los sistemas sanitarios están invirtiendo recursos considerables en el desarrollo y la introducción de directrices, en la creencia de que tienen una buena relación coste-eficacia y de que se favorece una mayor eficiencia en la atención sanitaria al incorporarlas en la práctica clínica. En el Reino Unido se han creado dos instituciones, el National Institute for Clinical Effectiveness (NICE) y la Scottish Intercollegiate Guidelines Network (SIGN), para desarrollar directrices de ámbito nacional basadas en pruebas.

Estas directrices tienen el potencial de mejorar la atención recibida por los pacientes, ya que favorecen las intervenciones con un efecto beneficioso comprobado, y facilitan que no se lleven a cabo intervenciones con probabilidad de ser ineficaces¹. Sin embargo, desarrollar e introducir unas directrices tiene un coste y, en algunas circunstancias, éste puede ser superior a los beneficios que pueden aportar las directrices. En otras ocasiones puede resultar más eficiente adoptar estrategias de difusión e implementación menos costosas aunque menos efectivas.

Las organizaciones sanitarias locales disponen de relativamente pocos recursos para las actividades centradas en la efectividad clínica, así que los gestores sanitarios deberían plantearse cómo usarlos mejor para maximizar los beneficios. Deberían considerar una serie de factores para que sus decisiones partan de una buena información. En primer lugar, ¿cuáles son las áreas clínicas en las que se podría mejorar la efectividad? Las actividades encaminadas a su mejora pueden reflejar prioridades nacionales o locales. Se tiene que tener en cuenta la prevalencia clínica, así como si las intervenciones sanitarias son efectivas y eficientes para obtener el mejor resultado para el paciente, y si hay evidencias de que la práctica actual es subóptima. En segundo lugar, ¿qué costes y qué beneficios supone la introducción de las directrices? Los gestores sanitarios deben considerar la efectividad probable de distintas estrategias de difusión e implementación, así como los recursos que necesitan esos procesos en el ámbito correspondiente. En tercer lugar, ¿qué costes y qué beneficios cabe esperar de un cambio en el comportamiento de los profesionales?

A pesar del actual interés por las directrices, aún persisten las dudas sobre la efectividad y los recursos que requieren las distintas estrategias para su difusión y su implantación. Un número reciente del boletín *Effective Health Care*² identificó siete revisiones sistemáticas de distintas estrategias de difusión e implementación de directrices³⁻⁹. En general, las revisiones sistemáticas llegaron a la conclusión de que las directrices podían producir cambios en la conducta del profesional sanitario. Por otro lado, también infieren que las estrategias tienen más probabilidades de ser efectivas si tienen en cuenta las circunstancias locales, se difunden mediante una intervención educativa activa y se aplican mediante recordatorios específicos para los pacientes. Oxman y sus colaboradores⁵ concluyeron que no existen “soluciones milagrosas” para cambiar el comportamiento de los profesionales; un conjunto de intervenciones puede posibilitar ese cambio pero ninguna intervención por sí sola es siempre eficaz para conseguirlo. Wensing y sus colaboradores⁸ concluyeron que las intervenciones múltiples que combinaban más de una intervención tendían a ser más efectivas, aunque también pueden resultar más caras. Sin embargo, muchos de estos análisis no tuvieron en cuenta los puntos débiles de la metodología de los estudios primarios (como errores potenciales de la unidad de análisis) y tendían a usar técnicas de recuento de votos (véase el Anexo 1 para una exposición más detallada de estos aspectos). El resultado es que aportan poca información sobre los probables tamaños de efecto de las distintas intervenciones, así como sobre cómo pueden variar éstos, por ejemplo, en función del ámbito, del grupo al que vaya destinado y de la actividad clínica. Ninguno de los estudios consideró los recursos necesarios para implantar la intervención o el grado probable de eficiencia de las intervenciones. Como resultado, compradores y proveedores tienen pocos datos en los que basarse a la hora de determinar el momento adecuado para dedicar los recursos que necesitan unas directrices para unas problemas clínicos específicos para desarrollarse e implantarse.

El estudio se propuso desarrollar un marco de referencia para decidir cuándo vale la pena desarrollar e implantar unas directrices clínicas, tomando como base los costes y los beneficios potenciales en la actividad clínica seleccionada, los niveles de actividad del momento (y, por tanto, el beneficio potencial para la salud), así como la probabilidad de que las directrices clínicas reporten los costes y los beneficios planteados.

OBJETIVOS DEL ESTUDIO

Los objetivos del estudio fueron:

- hacer una revisión sistemática de la efectividad y de los costes de distintas estrategias de desarrollo, difusión e implementación de directrices
- estimar los recursos que necesitan las distintas estrategias de desarrollo, difusión e implementación
- desarrollar un marco para decidir en qué circunstancias es eficiente desarrollar e introducir las directrices, en función de los costes y los beneficios potenciales de la actividad en cuestión, así como la efectividad y los costes del desarrollo y la implantación de las directrices.

ESTRUCTURA DE ESTE INFORME

El Capítulo 2 expone los métodos de la revisión sistemática de estrategias de difusión e implementación de directrices. Los

Anexos 1-3 aportan más información sobre los métodos estadísticos que se han usado, el desarrollo de estrategias de búsqueda y una lista de verificación de datos. El Capítulo 3 describe las características de los estudios incluidos. El Capítulo 4 describe los resultados de la revisión sistemática. Los Anexos 4-6 aportan información bibliográfica de los estudios incluidos, tablas de las características de los estudios y resultados. El Capítulo 5 describe los métodos y los resultados de una revisión de las evaluaciones económicas y los análisis de costes de las estrategias de implementación de directrices aportadas por los estudios incluidos. El Capítulo 6 describe los resultados de las entrevistas telefónicas realizadas con informantes clave respecto a la factibilidad de usar distintas estrategias de difusión e implementación, así como los factores que pueden influir en los recursos que se necesitan para hacerlas realidad. El Capítulo 7 expone los resultados del estudio, sus puntos fuertes y sus limitaciones, y presenta las principales conclusiones y recomendaciones del estudio.

Revisión sistemática de las estrategias de difusión e implementación de directrices: objetivos y métodos

OBJETIVOS

Los objetivos de la revisión fueron:

- Identificar las evaluaciones rigurosas sobre la implantación de directrices en la práctica médica
- estimar la efectividad de las estrategias de difusión e implementación de las directrices a la hora de promover mejoras en la práctica profesional

MÉTODOS

El estudio siguió los métodos propuestos por el grupo Cochrane Effective Practice and Organisation of Care (EPOC)¹⁰.

Criterios de selección de los estudios

Los criterios de inclusión de los estudios fueron los siguientes:

Diseño de los estudios

- ensayos controlados aleatorizados (ECA), con aleatorización individual en el ámbito del paciente (P-ECA), o aleatorización de conglomerados (*cluster*), en el ámbito del profesional, el consultorio o la organización sanitaria (C-ECA)
- ensayos clínicos controlados (ECC), con asignación individual en el ámbito del paciente (P-ECC) o bien con asignación de conglomerados (*cluster*) en el ámbito del profesional, el consultorio o la organización sanitaria (C-ECC)
- estudios controlados pre-post
- diseños de series temporales interrumpidas

En el anexo sobre estadística (Anexo 1) se puede encontrar más información sobre estos diseños de estudio y las razones por las que se incluyeron en el análisis.

Participantes

- Profesionales sanitarios con titulación médica

Si las directrices estaban dirigidas a grupos de profesiones distintas o bien a otros profesionales de la salud, sólo se incluyeron si más del 50% de la población seleccionada eran licenciados en medicina o si los datos procedentes de médicos se ofrecían por separado. Se excluyeron los estudios que evaluaban la implantación de directrices dirigidas a estudiantes de medicina, a pacientes o al público general.

Intervenciones

A los efectos de esta revisión, se adoptó la definición amplia de directrices o guía clínica como “enunciados desarrollados sistemáticamente para facilitar las decisiones de los facultativos respecto a la atención sanitaria idónea en circunstancias clínicas específicas”¹¹. Se incluyeron en la revisión los estudios que evaluaban una intervención que, en opinión de los revisores, cumplía los criterios anteriores, independientemente de que mencionara explícitamente el término ‘directrices’ o ‘guía clínica’.

Resultados

- Medidas objetivas del comportamiento del proveedor y/o resultado para el paciente.

Se excluyeron los estudios que usaban datos agregados autorreferidos sobre el comportamiento del profesional médi-

co, por ejemplo: '¿generalmente mide la presión arterial de los pacientes durante la visita?'. Se excluyeron asimismo los estudios que evaluaban sólo el nivel de conocimientos y/o la satisfacción de los profesionales médicos.

Estrategia de búsqueda

Búsquedas electrónicas

Se llevaron a cabo búsquedas electrónicas en las bases de datos siguientes:

- MEDLINE (1966–1998)
- Healthstar (1975–1998)
- Cochrane Controlled Trial Register (4ª edición 1998)
- EMBASE (1980–1998)
- SIGLE (1980–1988)
- el registro especializado del grupo EPOC.

En el Anexo 2 se ofrece más información sobre las estrategias de búsqueda y su desarrollo.

Bibliografía de revisiones sistemáticas anteriores

Por otro lado, los revisores comprobaron las listas de referencias bibliográficas de 51 revisiones sistemáticas sobre las estrategias para cambiar el comportamiento de profesionales aparecidas en 'Llevar las pruebas a la práctica médica' del boletín *Effective Health Care*¹².

No hubo ninguna restricción respecto al idioma de publicación.

Examen de los resultados de la búsqueda

RT y CF examinaron los resultados de la búsqueda para identificar estudios potencialmente válidos, de los cuales se obtuvieron copias impresas. RT y LS valoraron si cumplían los criterios de inclusión de la revisión. Las discrepancias fueron resueltas por consenso con JG.

Recogida de datos y valoración de la calidad (véase el Anexo 3 para la lista de verificación de la recogida de datos)

Dos revisores (RT y RD, ME, JG, LM, GM, LS, MW o PW) recogieron datos de forma independiente sobre:

- diseño de los estudios
- calidad metodológica de los estudios (según los criterios de calidad metodológica del EPOC; véase el Cuadro 1)
- participantes (características de los profesionales, con datos como: especialidad clínica, nivel de formación, edad, años transcurridos desde la licenciatura y proporción de profesionales que cumplen los criterios que participan en

el estudio; características de los pacientes, con datos como: problema clínico, edad, género y etnia; número de pacientes, de profesionales y de centros (consultorios, hospitales o comunidades) que participan en el estudio)

- marco del estudio (sistema de reembolso, marco asistencial, categoría académica, país)
- comportamientos examinados
- características de las intervenciones (características de las directrices clínicas, si se había llevado a cabo un análisis de necesidades (análisis GAP)¹³ y tipo de intervención (véase abajo para más información)
- resultados de los estudios (véase abajo para más información).

Clasificación de las intervenciones

Las estrategias de difusión e implementación se clasificaron según una taxonomía elaborada por el grupo EPOC, en función de si las intervenciones estaban orientadas a profesionales, o bien eran de tipo organizativas, financieras o reguladoras (véase el Cuadro 2 para las categorías de intervenciones profesionales). La mayoría de las intervenciones estaban dirigidas a profesionales. También hubo algunas intervenciones dirigidas a los pacientes (por ejemplo, intervenciones mediadas por el paciente, o con incentivos económicos para el paciente) aunque, dado el reducido número de ellas, se agruparon en una sola categoría (intervenciones dirigidas al paciente). De modo similar, todas las intervenciones organizativas y financieras profesionales se agruparon en categorías individuales (organizativas y financieras, respectivamente).

Derivación del tamaño del efecto del estudio (véase el Anexo 1 para más información)

Un estudio podía informar sobre una o todas las variables de valoración siguientes: variable dicotómica del proceso asistencial (ejemplo: proporción de pacientes que recibieron un tratamiento apropiado); variable continua del proceso asistencial (ejemplo: número de recetas expedidas por los profesionales sanitarios); variable dicotómica del resultado asistencial (ejemplo: proporción de pacientes que han dejado de fumar); y variable continua del resultado asistencial (ejemplo: puntuación media de los síntomas). Se recogieron datos sobre cada tipo de variable de valoración. Cuando los estudios informaron sobre más de una medida de cada variable de valoración, se recogió la medida principal (tal como la habían definido los autores del estudio) o la medida mediana. Por ejemplo, si la comparación informaba sobre cinco variables y no se identificaba ninguna como principal, se clasificaron los tamaños de efecto de las cinco variables y se obtuvo el valor mediano.

La dirección esperable del efecto variaba en función del estudio; algunos esperaban un aumento del resultado y otros una disminución. En todos los casos se estandarizó el tamaño del efecto, de modo que una diferencia positiva entre los

Cuadro 1. Criterios de calidad metodológica del Cochrane EPOC.

1. Criterios de calidad de los ensayos controlados aleatorizados (ECA) y los ensayos clínicos controlados (ECC)

(a) Ocultación de la asignación

HECHO	La unidad de asignación era una institución, un equipo o un profesional, y se realizó con cualquier método aleatorio descrito explícitamente, como el uso de tablas de números aleatorios. O la unidad de asignación era un paciente o un episodio asistencial, con algún tipo de plan de aleatorización centralizado, a través de un sistema informático en el propio centro o mediante sobres opacos sellados
NO CLARO	Procedimiento de asignación no descrito explícitamente. O la unidad de asignación era un paciente o un episodio asistencial y se refirió el uso de 'lista' o 'tabla', 'sobres' o 'sobres sellados' para la asignación
NO HECHO	Uso de alternancia, como la referencia a número de historia, fecha de nacimiento, día de la semana o cualquier otro método similar. O la unidad de asignación era un paciente o un episodio asistencial y se refirió el uso de cualquier proceso de asignación totalmente transparente desde el principio, como una lista abierta de números o asignaciones aleatorias. O la asignación fue modificada por los investigadores, los profesionales o los pacientes.

(b) Seguimiento de profesionales (protección contra el sesgo de exclusión)

HECHO	Medidas del resultado para $\geq 80\%$ de los profesionales distribuidos al azar (no supone un seguimiento del 100% a no ser que se indique explícitamente)
NO CLARO	No especificado
NO HECHO	Medidas del resultado para $< 80\%$ de los profesionales distribuidos al azar.

(c) Seguimiento de pacientes o episodios asistenciales

HECHO	Medidas del resultado para $\geq 80\%$ de los pacientes distribuidos al azar o de los pacientes que habían entrado en el ensayo (no supone un seguimiento del 100% a no ser que se indique explícitamente)
NO CLARO	No especificado
NO HECHO	Medidas del resultado para $< 80\%$ de los pacientes distribuidos al azar o de los pacientes que habían entrado en el ensayo.

(d) Valoración a ciegas de las variables principales del resultado^a (protección contra el sesgo de detección)

HECHO	Se indicó explícitamente que las variables principales del resultado fueron valoradas a ciegas. O las variables del resultado son objetivas, como la duración de la estancia hospitalaria, los niveles de fármaco determinados mediante una prueba estandarizada
NO CLARO	No especificado
NO HECHO	Resultados no valorados a ciegas.

(e) Medición basal

HECHO	Cumplimiento o resultados del paciente medidos antes de la intervención, sin que haya diferencias sustanciales entre grupos de estudio
NO CLARO	No se comunican las medidas basales o no está claro si las medidas basales son distintas entre diferentes grupos de estudio
NO HECHO	Las diferencias basales en las principales medidas del resultado probablemente afectarían negativamente a las diferencias que se encuentran después de la intervención; ejemplo: las diferencias entre distintos grupos antes de la intervención son similares a las halladas después de ésta.

(f) Medida(s) fiable(s) de las variables principales del resultado^b

HECHO	Dos o más evaluadores con una concordancia $\geq 90\%$ o kappa $\geq 0,8$. O la valoración del resultado es objetiva, como la duración de la estancia hospitalaria, o los niveles de fármaco determinados mediante una prueba estandarizada
NO CLARO	No se comunica la fiabilidad de las medidas del resultado mediante extracción a partir de un gráfico o recogidas por un individuo
NO HECHO	Dos o más evaluadores con una concordancia $< 90\%$ o kappa $< 0,8$.

(g) Protección contra la contaminación

HECHO	Asignación según comunidad, institución o consultorio y es poco probable que el grupo de control recibiera la intervención
NO CLARO	La asignación es de los profesionales que pertenecen a un centro (clínica o consultorio) y es posible que haya habido comunicación entre profesionales del grupo experimental y el grupo de control
NO HECHO	Es probable que el grupo de control recibiera la intervención, como en ensayos cruzados o si la aleatorización era sobre pacientes y no sobre profesionales.

Continúa

Cuadro 1. Criterios de calidad metodológica del Cochrane EPOC (Cont.).

2. Criterios de calidad para diseños controlados pre-post**(a) Medida basal**

HECHO	Cumplimiento o resultados del paciente valorados antes de la intervención, sin que haya diferencias sustanciales entre grupos de estudio
NO CLARO	Medidas basales no comunicadas o no está claro si las medidas basales son distintas entre grupos de estudio
NO HECHO	Las diferencias basales en las principales medidas del resultado probablemente afectarían negativamente a las diferencias que se encuentran después de la intervención; ejemplo: las diferencias entre grupos antes de la intervención son similares a las halladas después de ésta.

(b) Características de los grupos de estudio y de control

HECHO	Se especificaron las características de los grupos de estudio y de control, y eran similares
NO CLARO	No queda claro. Por ejemplo, se mencionan las características de ambos grupos, pero no se aportan datos.
NO HECHO	No se hace referencia a las características de los grupos (ni en el texto ni en una tabla). O se mencionan las características iniciales y se aprecian diferencias entre los grupos de estudio y el de control.

(c) Valoración a ciegas de las variables principales del resultado^a (protección contra el sesgo de detección)

HECHO	Se indicó explícitamente que se había hecho una valoración a ciegas de las variables principales del resultado. O las variables del resultado son objetivas; como en: duración de la estancia hospitalaria, niveles de fármaco determinados mediante una prueba estandarizada
NO CLARO	No se especifica
NO HECHO	No se hace una valoración a ciegas de los resultados.

(d) Protección contra la contaminación

HECHO	Asignación según comunidad, institución o consultorio, y es poco probable que el grupo de control recibiera la intervención
NO CLARO	La asignación es de los profesionales que pertenecen a un centro (clínica o consultorio) y es posible que haya habido comunicación entre profesionales del grupo experimental y el grupo de control
NO HECHO	Es probable que el grupo de control recibiera la intervención, como en ensayos cruzados o si la aleatorización era sobre pacientes y no sobre profesionales.

(e) Medida(s) fiable(s) de las variables principales del resultado^b

HECHO	Dos o más evaluadores con una concordancia $\geq 90\%$ o $\kappa \geq 0,8$. O la valoración del resultado es objetiva, como la duración de la estancia hospitalaria, o los niveles de fármaco determinados mediante una prueba estandarizada
NO CLARO	No se comunica la fiabilidad de las medidas del resultado mediante extracción a partir de un gráfico o recogidas por un individuo
NO HECHO	Dos o más evaluadores con una concordancia $< 90\%$ o $\kappa < 0,8$.

(f) Seguimiento de profesionales (protección contra el sesgo de exclusión)

HECHO	Medidas del resultado para $\geq 80\%$ de los profesionales distribuidos al azar (no supone un seguimiento del 100% a no ser que se indique explícitamente)
NO CLARO	No especificado
NO HECHO	Medidas del resultado para $< 80\%$ de los profesionales distribuidos al azar.

(g) Seguimiento de pacientes

HECHO	Medidas del resultado del $\geq 80\%$ de los pacientes distribuidos al azar o de los pacientes que habían entrado en el ensayo (no supone un seguimiento del 100% a no ser que se indique explícitamente)
NO CLARO	No especificado
NO HECHO	Medidas del resultado de $< 80\%$ de los pacientes distribuidos al azar o de los pacientes que habían entrado en el ensayo.

3. Criterios de calidad de los estudios de series temporales interrumpidas*Protección contra cambios seculares***(a) La intervención es independiente de otros cambios**

HECHO	La intervención se produjo de forma independiente de otros cambios que se dieron en el transcurso del estudio
NO CLARO	No se especifica (se tratará como NO HECHO si no se puede obtener información de los autores)
NO HECHO	La intervención no fue independiente de otros cambios que se dieron en el transcurso del estudio.

Continúa

Cuadro 1. Criterios de calidad metodológica del Cochrane EPOC (Cont.).

(b) Hay suficientes datos indicativos como para inferir una estadística fiable

HECHO	Hay al menos 20 puntos documentados antes de la intervención Y los autores han hecho un análisis tradicional de series temporales (modelo ARIMA) (o puede realizarse un análisis <i>post hoc</i>) O hay al menos 3 puntos documentados antes y después de la intervención Y los autores han realizado un análisis de las medidas repetidas (o puede realizarse un análisis <i>post hoc</i>) O hay al menos 3 puntos documentados antes y después de la intervención Y los autores han realizado un análisis de varianza (ANOVA) o una prueba de la t múltiple (o puede realizarse un análisis <i>post hoc</i>) Y hay al menos 30 observaciones por punto de datos
NO CLARO	No está especificado, como en el caso de que no se mencione el número específico de puntos de datos en el texto o en las tablas (se tratará como NO HECHO si no se puede obtener información de los autores)
NO HECHO	No se cumple ninguna de las condiciones anteriores.

(c) Prueba formal de tendencia (esta sección se debe rellenar si los autores han usado el modelo de análisis de varianza -ANOVA-)

HECHO	Se informa sobre una prueba formal de cambio en la tendencia mediante un método apropiado (por ejemplo, puede verse el de Cook y Campbell) (o puede repetirse)
NO CLARO	No está especificado (se tratará como NO HECHO si no se puede obtener información de los autores)
NO HECHO	No se ha realizado una prueba formal de cambio en la tendencia.

Protección contra el sesgo de detección

(d) La intervención tiene pocas probabilidades de afectar a la recogida de datos

HECHO	Se informa de que la intervención tiene pocas probabilidades de afectar a la recogida de datos, como en el caso de que las fuentes y los métodos de recogida de datos sean los mismos antes y después de la intervención
NO CLARO	No especificado (se tratará como NO HECHO si no se puede obtener información de los autores)
NO HECHO	Es probable que la intervención haya afectado a la recogida de datos, como en el caso de que se informe de cualquier cambio en la fuente o en el método de recogida de datos.

 (e) Valoración a ciegas de las variables primarias del resultado^a

HECHO	Se indicó explícitamente que las variables primarias del resultado fueron valoradas a ciegas. O las variables del resultado son objetivas, como la duración de la estancia hospitalaria, los niveles de fármaco determinados mediante una prueba estandarizada
NO CLARO	No especificado
NO HECHO	Resultados no valorados a ciegas.

(f) Integridad del conjunto de datos

HECHO	El conjunto de datos cubre el 80%-100% del número total de participantes o episodios asistenciales en el estudio
NO CLARO	No especificado (se tratará como NO HECHO si no se puede obtener información de los autores)
NO HECHO	El conjunto de datos cubre menos del 80% del número total de participantes o episodios asistenciales en el estudio.

 (g) Medida(s) fiable(s) de las variables primarias del resultado^b

HECHO	Dos o más evaluadores con una concordancia $\geq 90\%$ o kappa $\geq 0,8$. O la valoración del resultado es objetiva, como la duración de la estancia hospitalaria, o los niveles de fármaco determinados mediante una prueba estandarizada
NO CLARO	No se comunica la fiabilidad de los valores del resultado obtenidos mediante un gráfico o recogidos por un individuo
NO HECHO	Dos o más evaluadores con una concordancia $< 90\%$ o kappa $< 0,8$.

^aLas variables primarias del resultado son aquellas que corresponden a la hipótesis o la pregunta primaria definida por los autores. En el caso de que algunas variables primarias del resultado sean valoradas a ciegas y otras no, se puntúa cada una por separado en el dorso del formulario y se etiqueta claramente cada variable del resultado.

^bEn el caso de que algunas variables primarias del resultado sean valoradas de forma fiable y otras de forma no fiable, se puntúa cada una por separado en el dorso del formulario y se etiqueta claramente cada variable del resultado.

ARIMA: modelo autorregresivo integrado de media móvil.

Cuadro 2. Clasificación de las intervenciones profesionales a partir de la taxonomía del EPOC.

- **Distribución de material educativo:** distribución de recomendaciones publicadas o impresas para la atención médica, incluyendo directrices clínicas, material audiovisual y publicaciones electrónicas. El material puede haber sido entregado personalmente o a través de correo.
- **Encuentros educativos:** profesionales sanitarios que han participado en congresos, conferencias, talleres o periodos de prácticas.
- **Proceso de consenso local:** inclusión de profesionales participantes que aportaron su opinión para convenir la elección del problema clínico por su importancia y el enfoque más adecuado para resolver el problema.
- **Visitas educativas:** una persona formada acudió a los centros médicos para facilitar información a los profesionales con la intención de cambiar sus actuaciones clínicas. La información facilitada pudo incluir retroalimentación sobre las acciones del profesional/es.
- **Líderes locales de opinión:** utilización de profesionales nominados por sus colegas como de "opinión influyente". Los investigadores deben haber indicado explícitamente que sus colegas identificaron a los líderes de opinión.
- **Intervenciones mediadas por el paciente:** información clínica nueva (no disponible anteriormente) recogida directamente a través de los pacientes y facilitada al profesional, por ejemplo, puntuación en una escala de depresión.
- **Auditoría y retroalimentación:** un resumen del comportamiento clínico de la atención sanitaria durante un período determinado. El resumen puede haberse obtenido mediante las historias clínicas, las bases de datos informáticas o las observaciones de los pacientes. Se excluyen las intervenciones siguientes:
 - información clínica nueva que no refleja directamente la actuación del profesional y que fue recogida a través de los pacientes, por ejemplo, puntuaciones en una escala de depresión, resultados anormales en pruebas... Estas intervenciones deben describirse como mediadas por el paciente
 - retroalimentación en un formato alternativo (por ejemplo, en soporte informático) de la información incluida en la historia clínica de algunos pacientes individuales. Estas intervenciones deben describirse como organizativas.
- **Recordatorios:** información específica facilitada verbalmente, sobre papel o a través del ordenador, cuya finalidad o intención es ayudar a un profesional sanitario a recordar información, especialmente para realizar o evitar determinadas acciones en el tratamiento de determinados pacientes. Esta información normalmente se obtendría a través de su formación general, de archivos médicos o de otros colegas. Esta categoría incluiría la ayuda por ordenador para determinar la dosificación de fármacos o el apoyo en las decisiones.
- **Marketing:** uso de entrevistas personales, discusión en grupo (grupos focales) o una encuesta a profesionales seleccionados para identificar las resistencias al cambio y diseñar una intervención para superarlas.
- **Medios de comunicación:** (1) mediante medios de comunicación de amplia difusión, como radio, televisión, periódicos, carteles, folletos o cuadernos, solos o combinados con otras intervenciones; (2) dirigidos a la población general.
- **Otras:** otras categorías que se pueden convenir tras consultar con el equipo editorial del grupo EPOC.

porcentajes o medias después de la intervención constituía un buen resultado.

El tipo de datos obtenidos variaba en función del diseño del estudio. Para los ensayos controlados aleatorizados con conglomerados, los C-ECA, los P-ECA, los ECC y los estudios controlados pre-post, se recogieron datos sobre el cumplimiento basal, el obtenido después de la intervención, la diferencia absoluta en el cumplimiento después de la intervención y el nivel de significación para la diferencia absoluta en el período posterior a la intervención (y, sólo para datos continuos, la diferencia relativa en el cumplimiento después de la intervención y la diferencia media estandarizada). Para las series temporales interrumpidas, se recogieron datos sobre el cambio en la pendiente y el nivel. Los revisores intentaron volver a analizar los estudios de series temporales interrumpidas que no habían sido analizados adecuadamente o que no

ofrecían resultados comparables (véase abajo y el Anexo 1 para más información).

Cuando no era posible recoger datos comparables, generalmente por la complejidad del diseño o del análisis, se trasladaron los resultados a la tabla de resultados (Anexo 6) y se resumieron por separado en el texto (Capítulo 4).

Reanálisis de los estudios (véase el Anexo 1 para más información)

Comparaciones pre-post dentro de un mismo grupo

En los C-ECA, los P-ECA, los ECC y los estudios controlados pre-post, se intentó volver a analizar la comparación después de la intervención entre grupos si los estudios sólo aportaban comparaciones pre-post dentro de un mismo grupo. Si no fue

posible, se comunicaron las estimaciones puntuales para la comparación después de la intervención entre grupos sin una estimación de la significación estadística.

Errores potenciales de la unidad de análisis

En muchos ensayos C-ECA la aleatorización se había hecho sobre el facultativo, pero durante los análisis estadísticos, los datos individuales de los pacientes se habían analizado como si no hubiera ninguna conglomeración (*clustering*) por facultativos. Los C-ECA que no tienen en cuenta la conglomeración durante el análisis tienen ‘errores de la unidad de análisis’¹⁴. Aunque es poco probable que la estimación puntual del efecto sea sesgada, los valores de *p* pueden ser artificialmente extremos y los intervalos de confianza excesivamente estrechos, con lo que es más probable que se den resultados falsamente significativos y conclusiones engañosas¹⁵. Si los C-ECA tenían errores potenciales de la unidad de análisis, los revisores intentaron volver a analizar el ensayo utilizando datos de conglomerado. Si no era posible hacerlo, se comunicaron las estimaciones puntuales para la comparación entre grupos después de la intervención sin una estimación de la significación estadística.

Estudios de series temporales interrumpidas

Para los estudios de series temporales interrumpidas, se volvió a analizar cada estudio (si era posible) utilizando métodos de regresión de series temporales (véase el Anexo 1). Los datos se derivaron de tablas de resultados presentados en los estudios originales. Cuando sólo se comunicaba la información sobre valores individuales en formato gráfico, los datos se derivaron mediante medición manual de cada punto de datos. La imagen digital del gráfico se obtuvo mediante un escáner; a continuación, se importaron las imágenes en un paquete gráfico, se aplicaron retículas y se leyeron los valores correspondientes. En algunos estudios no fue posible derivar los puntos de datos a partir de los gráficos debido a las escalas que se habían utilizado. En ese caso, se llevaron a cabo regresiones de series temporales para estimar dos tamaños de efecto para cada comparación de series temporales interrumpidas. Primero, se estimó el cambio en el nivel de resultado inmediatamente posterior a la implantación de la intervención. Para ello, se extrapoló la recta de regresión antes de la intervención hasta el primer punto después de la intervención. El cambio en la estimación del nivel se obtuvo a partir de la diferencia entre este punto extrapolado y la estimación de la regresión después de la intervención para el mismo punto. Los autores pueden facilitar más detalles sobre los métodos matemáticos. En segundo lugar, se estimó el cambio en las pendientes de las rectas de regresión (calculado como la pendiente después de la intervención menos la pendiente antes de la intervención). Fueron necesarios ambos tamaños del efecto para interpretar los resultados de cada estudio. Por ejemplo, aunque el nivel no haya cambiado inmediatamente después de la intervención, puede haberse producido un cambio significativo en la pendiente. Sería engañoso comunicar sólo el cambio de la estimación de nivel en este ejemplo. Se estanda-

rizó la dirección del efecto, de modo que una estimación positiva del nivel o la pendiente se consideró un buen resultado; y una estimación negativa, un mal resultado.

Tratamiento de los datos

A todos los estudios incluidos se les asignó un número identificador exclusivo. Los datos se introdujeron en una base de datos de Access. Cualquier discrepancia se resolvió mediante consenso.

Síntesis y presentación de los datos

Dada la extrema heterogeneidad que se esperaba dentro de la revisión y el número de estudios con errores potenciales de la unidad de análisis, no estaba previsto llevar a cabo un metanálisis formal. Para los comentarios sobre este tema, se puede consultar la sección sobre el marco analítico en el Anexo 1. Las revisiones sistemáticas cualitativas anteriores sobre estrategias de implementación han utilizado mayormente el método de recuento de votos, de forma que suman el número de estudios positivos y negativos y, en función del resultado, concluyen si las intervenciones fueron eficaces^{2,16}. Sin embargo, este método no incorpora la precisión de la estimación a partir de estudios primarios, no aporta una estimación del tamaño del efecto de la intervención y tiene un bajo poder estadístico. En esta revisión, se utilizó un marco analítico más explícito.

Se llevaron a cabo análisis separados para las comparaciones de intervenciones únicas frente a controles ‘sin intervención’ (es decir, grupos de tratamiento estándar o grupos de control que no recibieron ninguna intervención), intervenciones únicas frente a controles ‘con intervención’ (es decir, grupos de control que sí recibieron una intervención), intervenciones múltiples frente a controles ‘sin intervención’ e intervenciones múltiples frente a controles ‘con intervención’. Para las intervenciones múltiples, se llevaron a cabo análisis separados para las combinaciones que incluían visitas educativas y las combinaciones con más de cuatro comparaciones. Los revisores también investigaron si los efectos de las intervenciones múltiples aumentaban paralelamente al número de componentes. Aunque, según el plan original, debía emprenderse un análisis de metarregresión para estimar los efectos de distintas intervenciones, el número de combinaciones distintas de intervenciones múltiples evaluadas resultó ser problemático (véase el Anexo 1 para más información).

La bibliografía de los estudios incluidos se facilita en el Anexo 4, información sobre las características de los estudios (que incluye una valoración de la calidad) se exponen en el Anexo 5, y los resultados de todas las comparaciones se indican en el Anexo 6. Como tamaño del efecto primario para cada comparación, se utilizaron medidas dicotómicas del proceso asistencial por dos motivos de índole pragmática: primero, se comunicaron con una frecuencia mucho mayor en los estudios; y, segundo, las medidas continuas del proceso asistencial fueron menos estables. Por ejemplo, un cambio porcentual relativo en una medida continua depende de la esca-

la que se esté utilizando: una comparación que cambia de una media de 1 a 2 mostrará la misma mejora relativa que una comparación que cambia de 25 a 50. Para contrarrestar este efecto, se calcularon diferencias medias estandarizadas siempre que fue posible. Sin embargo, por regla general, los estudios no presentaban datos suficientes para poder hacer este cálculo. Para asegurar un análisis completo, también se comunicaron datos dicotómicos y continuos del resultado.

Para las comparaciones C-ECA, P-ECA, ECC y de los estudios controlados pre-post, se comunicaron (por separado para cada diseño de estudio):

- el número de comparaciones que mostraban una dirección positiva del efecto
- el tamaño del efecto mediano de todas las comparaciones
- el tamaño del efecto mediano de las comparaciones sin errores de la unidad de análisis
- y el número de comparaciones que mostraban efectos estadísticamente significativos

Esto permite al lector valorar la consistencia de los efectos entre distintos diseños de estudio y entre comparaciones cuando se sabe la significación estadística.

Este trabajo también presenta diagramas de barras apiladas de las estimaciones de medidas dicotómicas del proceso (para todos los diseños) frente al tamaño del efecto observado, diferenciando entre los estudios que comunicaban efectos significativos, los de efectos no significativos y los que tenían errores de la unidad de análisis. De este modo, resulta más fácil para el lector valorar la mediana y el rango de los tamaños del efecto de todos los estudios.

Para describir el tamaño del efecto (diferencia absoluta entre medidas postintervención) para las medidas dicotómicas del proceso, se utilizaron los términos siguientes:

- 'pequeño' para describir tamaños de efecto $\leq 5\%$
- 'modesto' para describir tamaños de efecto entre $> 5\%$ y $\leq 10\%$
- 'moderado' para describir tamaños de efecto entre $> 10\%$ y $\leq 20\%$
- 'grande' para describir tamaños de efecto $> 20\%$.

Para las comparaciones de series temporales interrumpidas, se aporta la significación de los cambios de nivel y pendiente.

Revisión sistemática de estrategias de difusión e implementación de directrices: información sobre los estudios incluidos

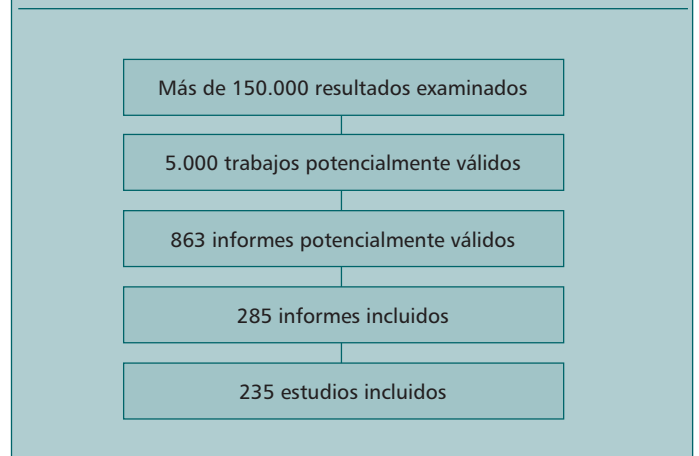
RESULTADOS DE LA BÚSQUEDA BIBLIOGRÁFICA

La estrategia de búsqueda produjo más de 150.000 resultados. Se examinaron los títulos y los resúmenes de los estudios, y se identificaron inicialmente unos 5000 como potencialmente válidos. Se recuperó el texto completo de 863 informes de estudios potencialmente válidos y se valoraron de cara a su inclusión en esta revisión. En total, se incluyeron en esta revisión 285 informes de 235 estudios que dieron 309 comparaciones separadas (Figura 1). La bibliografía de los estudios incluidos (incluyendo los propios trabajos) se facilitan en el Anexo 4. Asimismo, las tablas del Anexo 5 ofrecen datos sobre todos los estudios incluidos. Los autores pueden facilitar información sobre los estudios excluidos.

CALIDAD METODOLÓGICA DE LOS ESTUDIOS INCLUIDOS

Se incluyeron 110 (46,8%) C-ECA, 29 (12,3%) P-ECA, 7 (3,0%) C-ECC, 10 (4,3%) P-ECC, 40 estudios controlados pre-post (17,0%) y 39 (16,6%) diseños de series temporales interrumpidas.

Figura 1. Diagrama de flujo de los estudios incluidos.



Calidad metodológica

La calidad metodológica se valoró según los criterios del grupo Cochrane EPOC (véase el Anexo 3 para más información). En conjunto, fue difícil determinar la calidad de los estudios por deficiencias en la comunicación de datos. En las Tablas 1-3 se resume la calidad de los estudios en función del diseño y el nivel de asignación.

Ensayos controlados aleatorizados (ECA) y ensayos controlados (ECC) (Tabla 1)

La ocultación de asignación fue 'no clara' o 'no hecha' para la mayoría de los ECA y los ECC. La protección contra la contaminación fue 'hecha' en poco más de la mitad (54%) de los ensayos controlados aleatorizados con conglomerados (es decir, era poco probable que el grupo de control recibiera la intervención), pero fue 'no clara' en el 41% de ellos. En contraste, la protección contra la contaminación fue 'no hecha' en la mayoría de los P-ECA y P-ECC.

No estuvo claro si el resultado había sido valorado a ciegas en más del 70% de los C-ECA y los ECC, y en el 59% de los P-ECA. La fiabilidad de la medición del resultado y el seguimiento del $\geq 80\%$ de los profesionales implicados fueron 'no claros' a partir de los informes de más del 70% de los ensayos clínicos controlados y aleatorizados. El seguimiento de los

pacientes fue 'no claro' en el 71% de los C-ECA. La medición basal del resultado y que los grupos no presentaran diferencias fue 'hecha' en el 42% de los C-ECA y en el 14% de los C-ECC.

Estudios controlados pre-post (Tabla 2)

El criterio 'características de los centros de estudio y de control comunicadas y similares' fue 'no claro' o 'no hecho' en el 77,5% de los estudios controlados pre-post, y 'resultado medido en la valoración basal y ausencia de diferencias' fue 'no claro' o 'no hecho' en el 57,5% de los estudios controlados pre-post. La protección contra la contaminación fue 'hecha' en el 62,5% de los estudios.

La valoración del resultado a ciegas, la fiabilidad de la medición del resultado y el seguimiento de $\geq 80\%$ de los

Tabla 1. Criterios de calidad para los ensayos controlados aleatorizados y los ensayos controlados.

Criterio		Hecho n (%)	No claro n (%)	No hecho n (%)
Ocultación de la asignación	C-ECA	29 (26,4)	80 (72,7)	1 (0,9)
	P-ECA	4 (13,8)	24 (82,8)	1 (3,4)
	C-ECC	0	2	5 (71,4)
	P-ECC	0	0	10 (100)
Protección contra la contaminación	C-ECA	60 (54,5)	45 (41)	5 (4,5)
	P-ECA	0	5 (17,2)	24 (82,8)
	C-ECC	0	5 (71,4)	2 (28,6)
	P-ECC	0	0	10 (100)
Valoración del resultado a ciegas	C-ECA	21 (19,1)	85 (77,3)	4 (3,6)
	P-ECA	11 (37,9)	17 (58,6)	1 (3,4)
	C-ECC	2 (28,6)	5 (71,4)	0
	P-ECC	2 (20,0)	8 (80,0)	0
Medida fiable del resultado	C-ECA	27 (24,5)	80 (72,7)	3 (2,7)
	P-ECA	3 (10,3)	26 (89,7)	0
	C-ECC	1 (14,3)	5 (71,5)	1 (14,3)
	P-ECC	0	10 (100)	0
Resultados medidos en la valoración basal con ausencia de diferencias	C-ECA	46 (41,8)	58 (52,7)	6 (5,5)
	P-ECA	9 (31,0)	20 (69,0)	0
	C-ECC	1 (14,3)	6 (85,7)	0
	P-ECC	0	10 (100)	0
Seguimiento de profesionales	C-ECA	28 (25,5)	78 (70,9)	4 (3,6)
	P-ECA	0	29 (100)	0
	C-ECC	0	7 (100)	0
	P-ECC	1 (10)	9 (90)	0
Seguimiento de pacientes	C-ECA	29 (26,4)	78 (70,9)	3 (2,7)
	P-ECA	16 (55,2)	10 (34,5)	3 (10,3)
	C-ECC	4 (57,1)	3 (42,9)	0
	P-ECC	3 (30)	7 (70)	0

C-ECA: ensayo controlado aleatorizado con conglomerados; P-ECA: ensayo controlado aleatorizado con pacientes; C-ECC: ensayo clínico controlado con asignación de conglomerados; P-ECC: ensayo clínico controlado con asignación de pacientes.

Tabla 2. Criterios de calidad para los estudios controlados pre-post.

Criterio	Hecho n (%)	No claro n (%)	No hecho n (%)
Características de los grupos de estudio y control comunicadas y similares	9 (22,5)	17 (42,5)	14 (35,0)
Protección contra la contaminación	25 (62,5)	14 (35)	1 (2,5)
Medición basal y ausencia de diferencias	17 (42,5)	16 (40,0)	7 (17,5)
Valoración del resultado a ciegas	6 (15)	31 (77,5)	3 (7,5)
Medida fiable del resultado	6 (15)	34 (85)	0
Seguimiento de profesionales	12 (30)	27 (67,5)	1 (2,5)
Seguimiento de pacientes	4 (10)	35 (87,5)	1 (2,5)

pacientes fueron 'no claros' en más del 75% de los estudios. El seguimiento de $\geq 80\%$ de los profesionales fue 'no claro' en el 67,5% de los estudios.

Estudios de series temporales interrumpidas (Tabla 3)

La intervención se consideró independiente de otros cambios en sólo el 36% de los estudios de series temporales interrumpidas, y se consideró que no afectaba a la recogida de datos en el 92% de los estudios. Sin embargo, la integridad del conjunto de datos fue 'no clara' en el 90% de los estudios. Dieciséis (41%) de los 39 estudios de series temporales interrumpidas fueron analizados adecuadamente. La valoración a ciegas y la medida fiable del resultado fueron 'no claras' en 14/39 (36%) de los estudios de series temporales interrumpidas y 'hechas' en el resto.

OTROS ASPECTOS METODOLÓGICOS

Cálculo del tamaño de la muestra

La mayoría de los estudios ECA, ECC y de los estudios controlados pre-post no comunicaron un cálculo del tamaño de la muestra (135/196, 69%). En 44 (22%) estudios, estaba claro que sí se había calculado y se había comunicado adecuadamente, pero en 15 (34%) de estos estudios, parecía ser un cálculo retrospectivo. En los 17 estudios restantes no estaba claro si se había calculado el tamaño de la muestra.

Errores potenciales de la unidad de análisis

Había errores potenciales en la unidad de análisis en los principales análisis del 53% de los C-ECA, en el 86% de los C-ECC y en el 83% de los estudios controlados pre-post.

Tabla 3. Criterios de calidad para los diseños de series temporales interrumpidas.

Criterio	Hecho n (%)	No claro n (%)	No hecho n (%)
Intervención independiente de otros cambios	14 (36)	24 (61)	1 (3)
Es poco probable que la intervención afecte a la recogida de datos	36 (92)	3 (8)	0
Valoración del resultado a ciegas	25 (64)	14 (36)	0
Medida fiable del resultado	25 (64)	14 (36)	0
Conjunto de datos completo	3 (7)	35 (90)	1 (3)
Estudio analizado adecuadamente	16 (41)	11 (28)	12 (31)

Puntos de datos e intervalos: diseños de series temporales interrumpidas

Los números medianos de puntos de datos antes y después de la intervención para los estudios de series temporales interrumpidas fueron 10 [rango intercuartil (IQR) 5–17] y 12 (IQR 7–24), respectivamente. El intervalo entre puntos de datos fue de 1 mes en el 64% de los estudios. El intervalo de tiempo mínimo fue de 5 días y el intervalo máximo fue de 1 año.

Datos aportados

Se midió el proceso asistencial en el 95% de los estudios; el 67% de los estudios aportaron medidas dicotómicas del proceso; el 39%, medidas continuas del proceso. El resultado asistencial se comunicó en el 22% de los estudios; el 14% aportaron medidas dicotómicas del resultado; y el 14%, medidas continuas del resultado.

El 29% de los estudios aportaron datos económicos (véase el Capítulo 5 para más información).

Detalle de las comparaciones

En el 79% de los estudios hubo una sola comparación entre un grupo de intervención frente a un grupo de control; en el 13% hubo dos comparaciones (por ejemplo, un C-ECA de tres brazos); y en el 8% se hicieron tres comparaciones o más.

En el 69% de los ECA, los ECC y los estudios controlados pre-post, el grupo de control no recibió ninguna intervención; en el 22%, el grupo de control recibió una sola intervención; y en el 10%, el grupo de control recibió más de una intervención.

DATOS DE LOS PARTICIPANTES

Número de unidades de asignación: ECA, ECC y ensayos controlados pre-post

El número de unidades de asignación en cada grupo de estudio se comunicó en el 77% de los estudios. El resto tendía a comunicar un número global que participaba en el estudio o el número de otras unidades por grupo de estudio; por ejemplo, se asignan consultorios a grupos de estudio pero se expone el número de médicos de cabecera por grupo. El número mediano de unidades de asignación por brazo de estudio fue de 18 (IQR 2–65). Los estudios que asignaron al brazo de estudio por conglomerado (*cluster*) (C-ECA, C-ECC, estudios controlados pre-post) presentaron un número mediano de 7 unidades por brazo (IQR 1–24). El número mediano de los estudios sin conglomerado (P-ECA, P-ECC) fue de 113 (IQR 61–423).

Proporción de proveedores o unidades de asignación elegibles que participaron

El porcentaje de proveedores o unidades de asignación elegibles en el marco del muestreo seleccionado fue 'no claro' en 186 (79%) estudios. Cuando se comunicó, el valor mediano fue del 80% (IQR 45–95%).

Marco asistencial y profesionales

Los estudios se llevaron a cabo en 14 países distintos: el 71% se efectuaron en Estados Unidos, el 11% en el Reino Unido, el 6% en Canadá, el 3% en Australia y otro 3% en Holanda. Los estudios restantes se llevaron a cabo en Dinamarca, Francia, Alemania, Israel, Méjico, Nueva Zelanda, Noruega, Suecia y Tailandia.

El marco asistencial más común fue la asistencia primaria (39%), seguido de atención hospitalaria (19%) y atención ambulatoria (19%) (aunque estos últimos estudios se hicieron todos en Estados Unidos, y el marco puede equipararse a la asistencia primaria). Treinta y seis (15%) estudios se basaron en marcos mixtos (atención hospitalaria y ambulatoria), en la interfaz entre marcos o en una combinación de consultas externas y atención hospitalaria. Los estudios restantes se llevaron a cabo en residencias para enfermos crónicos (3%), en servicios de urgencias (2%), en atención especializada ambulatoria (1%) y en una clínica militar (0,4%). El marco asistencial no queda claro en tres (1%) estudios.

En 174 (74%) estudios, la intervención se dirigía sólo a médicos; el resto de estudios se dirigían a médicos y a otros profesionales de la organización sanitaria, como enfermeras, farmacéuticos, dietistas, auxiliares médicos o personal administrativo.

La mayoría de los estudios (57%) se dirigieron sólo a una especialidad médica. La especialidad más común fue la de medicina general o de familia (24%).

El nivel de formación de los proveedores no se explicita en los informes del 51% de los estudios. El 26% incluyeron tanto profesionales 'plenamente formados', como médicos en prácticas, el 14% incluyeron sólo profesionales en vías de formación, y el 8% incluyeron sólo profesionales plenamente formados.

DATOS DE LAS INTERVENCIONES

Características de las directrices clínicas

La fuente de las directrices fue, en el 35% de los estudios, un organismo nacional de profesionales expertos o un organismo gubernamental de ámbito nacional; médicos locales, en el 30%; y alguna otra fuente, en el 10%. En el 25% de los estudios no se explicitó la fuente de las directrices; y en el 81% no se expresa la composición del grupo de desarrollo de las directrices. La base de pruebas de las directrices recomendadas no estuvo clara en el 94% de los estudios. Parecían basarse en pruebas sólidas sólo en un 3% de los estudios.

La finalidad de las recomendaciones fue administrar un tratamiento correcto en el 81% de los estudios; la reducción de costes en el 4%; y ambas cosas (administrar un tratamiento correcto y reducir costes) en el 14%. No se explicitó la finalidad o era otra distinta en 3 (1%) estudios.

El cambio que se perseguía era, en el 40% de los estudios, conseguir un incremento del tratamiento establecido; en el 25%, una modificación del tratamiento (es decir, aumentar la frecuencia de una actividad y disminuirla en otra); y una reducción del tratamiento establecido en el 15% de los estudios. La naturaleza del cambio que se perseguía no estuvo clara en el 11% de los estudios, y se perseguían combinaciones de cambios en el 9%.

Comportamientos estudiados

El 47% de los ensayos estudiaron sólo un tipo de comportamiento del profesional. De éstos, el comportamiento que se estudió con mayor frecuencia fue el enfoque global del problema, con un 19% de los estudios, seguido de la expedición de recetas en el 14% de los estudios y de la petición de pruebas en el 8%. Los comportamientos combinados estudiados con mayor frecuencia fueron: expedición de recetas (el 34% de los estudios), prevención (32%), educación y asesoramiento al paciente (31%), petición de pruebas (23%) y enfoque global (23%). Otros comportamientos investigados fueron: diagnóstico, planificación del alta, gestión financiera, intervenciones, comunicación entre profesional y paciente, archivo y derivación a otros profesionales.

Estrategias de intervención

En 84 de las 309 comparaciones (27%) hubo un grupo de estudio que recibió una sola estrategia de intervención para la implementación de directrices frente a un grupo de control 'sin intervención' o con tratamiento habitual o estándar; y en el 1% de las comparaciones hubo grupos de estudio con una sola intervención frente a un grupo de control que recibió también una única intervención. En 136 (44%), se comparó un grupo que recibió una intervención múltiple (>1 intervención) con un grupo de control sin intervención o de tratamiento estándar. Los grupos de intervención múltiple se compararon con un grupo de control que recibió una intervención (≥ 1 intervención) en 85 (27%) de las comparaciones.

La intervención única evaluada con mayor frecuencia frente a un control sin intervención fue la de los recordatorios en el 13% de todas las comparaciones; seguida de difusión de material educativo, en el 6% de las comparaciones; auditoría y retroalimentación, en el 4%; e intervenciones dirigidas al paciente, en el 3% (Tabla 4).

La estrategia de intervención utilizada con mayor frecuencia como parte de intervenciones múltiples fue la de material educativo (evaluada en el 48% de todas las comparaciones) (Tabla 5), seguida de encuentros educativos (41%), recordatorios (31%), y auditoría y retroalimentación (24%).

Ciento diecisiete estudios (que incluían 136 comparaciones) evaluaron un total de 68 combinaciones distintas de

intervenciones frente a un grupo de control 'sin intervención', y 61 estudios (que incluían 85 comparaciones) evaluaron 58 combinaciones distintas de intervenciones frente a un grupo de control que también recibió una o varias intervenciones.

Tabla 4. Intervenciones únicas con grupo de control 'sin intervención'.

Intervención	Nº de comparaciones (% del total, 309)
Material educativo	18 (6)
Encuentros educativos	3 (1)
Procesos de consenso	0 (0)
Visitas educativas	1 (0,3)
Líderes de opinión	0 (0)
Intervenciones dirigidas al paciente	8 (3)
Auditoría y retroalimentación	12 (4)
Recordatorios	38 (13)
Otras intervenciones dirigidas a profesionales (incluyendo medios de comunicación y <i>marketing</i>)	2 (1)
Intervenciones financieras	0 (0)
Intervenciones organizativas	2 (1)
Intervenciones estructurales	0 (0)
Intervenciones reguladoras	0 (0)

Tabla 5. Tipos de intervenciones múltiples utilizadas

Intervención	Nº de comparaciones (% del total, 309)
Material educativo	147 (48)
Encuentros educativos	126 (41)
Procesos de consenso	16 (5)
Visitas educativas	35 (11)
Líderes de opinión	6 (2)
Intervenciones dirigidas al paciente	56 (18)
Auditoría y retroalimentación	73 (24)
Recordatorios	95 (31)
Otras intervenciones dirigidas a profesionales (incluyendo medios de comunicación y <i>marketing</i>)	19 (6)
Intervenciones financieras	12 (4)
Intervenciones organizativas	45 (15)
Intervenciones estructurales	37 (12)
Intervenciones reguladoras	3 (1)

Revisión sistemática de estrategias de difusión e implementación de directrices: resultados

Este capítulo presenta los resultados de la revisión sistemática de las diferentes estrategias de difusión e implementación de directrices. Los resultados se sintetizaron y se comunicaron en el siguiente formato estructurado (véase también 'Marco analítico usado en esta revisión', en el Anexo 1):

- Evaluación de intervenciones únicas frente a grupos de control 'sin intervención'.
- Evaluación de intervenciones únicas frente a grupos de control 'con intervención'.
- Evaluación de intervenciones múltiples frente a grupos de control 'sin intervención':
 - todas las comparaciones
 - intervenciones múltiples que incluían visitas educativas
 - combinaciones de material educativo y encuentros educativos
 - combinaciones de material educativo, y auditoría y retroalimentación
 - combinaciones de recordatorios e intervenciones mediadas por el paciente
 - combinaciones de material educativo, encuentros educativos, y auditoría y retroalimentación
 - combinaciones de material educativo, encuentros educativos e intervenciones organizativas
- Evaluación de intervenciones múltiples frente a grupos de control 'con intervención':
 - todas las comparaciones
 - intervenciones múltiples que incluían visitas educativas
 - material educativo y recordatorios comparados con material educativo

- material educativo, encuentros educativos y recordatorios comparados con material educativo y encuentros educativos

Para cada una de las intervenciones indicadas, se comunican las características demográficas del estudio, las medidas dicotómicas del proceso, las medidas continuas del proceso, las medidas dicotómicas del resultado, las medidas continuas del resultado, los resultados de las series temporales interrumpidas y el resumen. Las referencias bibliográficas de los estudios identificados con un número más la letra A se facilitan en el Anexo 4. Los resultados de todas las comparaciones se presentan en el Anexo 6.

EVALUACIÓN DE INTERVENCIONES ÚNICAS FRENTE A GRUPOS DE CONTROL SIN INTERVENCIÓN

Ochenta y un estudios (que incluían 84 comparaciones) evaluaron intervenciones únicas comparadas con un control sin intervención. En ninguna comparación se evaluaron como intervención única las visitas educativas, los procesos de consenso locales, los líderes de opinión locales, ni las intervenciones financieras, estructurales o reguladoras.

Material educativo

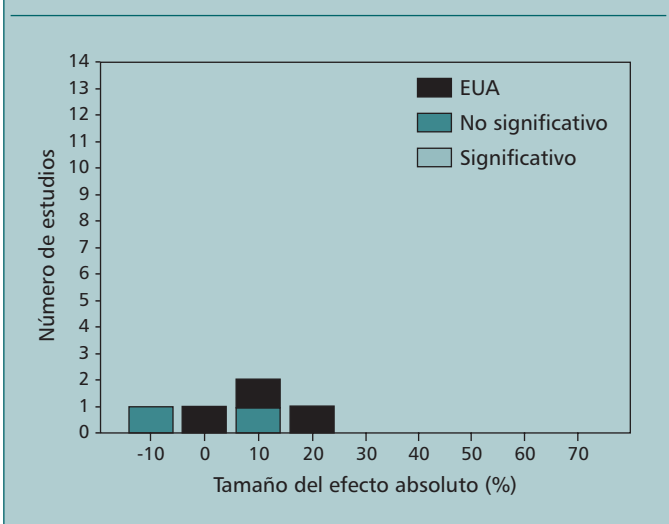
Dieciocho estudios (que incluían 18 comparaciones) evaluaron los efectos de la difusión de material educativo, que incluían 7 C-ECA (A15, A44, A60, A86, A140, A151, A168), 2 P-ECA (A76, A91), 2 estudios controlados pre-post (A167, A198) y 7 estudios de series temporales interrumpidas (A19, A121, A167, A184,

A188, A206, A234). Un C-ECA (A140) también presentó datos sobre variables de valoración continuas del proceso basadas en una comparación pre-post. Ocho estudios se llevaron a cabo en Estados Unidos, 4 en el Reino Unido, 3 en Australia, 2 en Canadá y 1 en Holanda. La mayoría de los estudios se realizaron en marcos de asistencia primaria ($n = 10$). Los comportamientos estudiados fueron el enfoque global de un problema clínico en 7 estudios, la expedición de recetas en 3 estudios, los servicios de prevención en 3 estudios, la petición de pruebas en 3 estudios y los procedimientos en 2 estudios.

Resultados de las comparaciones de los ECA, los ECC y los ensayos controlados pre-post

- *Medidas dicotómicas del proceso (Figura 2).* Cinco comparaciones aportaron datos dicotómicos del proceso: 4 comparaciones C-ECA (A15, A60, A151, A168) y 1 comparación P-ECA (A76). Todas las comparaciones C-ECA observaron mejoras en la asistencia; el efecto mediano fue una mejora absoluta del cumplimiento del +8,1% (límites: +3,6 a +17%). Dos comparaciones tenían errores potenciales de la unidad de análisis (A15, A60) y no se pudo determinar la significación de 1 comparación (A151). La otra comparación, sin error potencial de la unidad de análisis, observó un efecto del +6% (no significativo) (A168). La comparación P-ECA observó un deterioro absoluto de la asistencia del -8,3% (no significativo).
- *Medidas continuas del proceso.* Cinco comparaciones comunicaron datos continuos del proceso: 3 comparaciones C-ECA (A44, A86, A151) y 2 comparaciones controladas pre-post (A140, A198). Una comparación C-ECA (A151) informó de una mejora relativa del cumplimiento del +34,7% y una diferencia media estandarizada de +0,25 (no significativa).

Figura 2. Tamaños del efecto observados en datos dicotómicos del proceso en comparaciones controladas aleatorizadas, controladas y controladas pre-post del material educativo frente a un control sin intervención. (EUA: error de la unidad de análisis).



No fue posible extraer datos comparables en 2 comparaciones C-ECA (A44, A86). Ambos informaron de efectos no significativos, pero 1 comparación tenía un error potencial de la unidad de análisis (A86). Las 2 comparaciones controladas pre-post observaron mejoras relativas del cumplimiento del +77,3% (A140) y +100,3% (A198). Sólo fue posible calcular una diferencia media estandarizada en 1 comparación: +0,53 (A140). Ambas comparaciones tenían un error potencial de la unidad de análisis.

- *Medidas dicotómicas del resultado.* Dos comparaciones comunicaron datos dicotómicos del resultado: 1 comparación P-ECA (A91) y 1 controlada pre-post (A167). La comparación P-ECA comunicó un efecto mediano de un deterioro absoluto del resultado del -4,6% (no significativo). No fue posible extraer datos comparables de la comparación controlada pre-post (A167), que comunicó un efecto no significativo.

- *Medidas continuas del resultado.* Una comparación C-ECA (A140) señaló que el efecto mediano de una mejora relativa del resultado era del +17,1%, diferencia media estandarizada +0,86 ($p < 0,05$).

Resultados de las comparaciones de los ensayos de series temporales interrumpidas

En 6 casos se volvieron a analizar las comparaciones mediante regresión de series temporales. La comparación restante no aportaba datos suficientes para hacer un nuevo análisis, pero había realizado una regresión de series temporales del cambio de pendiente (A121). Dos comparaciones comunicaron mejoras significativas del nivel (A19, A170); sin embargo, ambas también comunicaron un deterioro de la pendiente (estadísticamente significativo para A170), lo que sugiere un efecto de decrecimiento. Una comparación comunicó un deterioro significativo tanto del nivel como de la pendiente (A234). La comparación que no volvió a ser analizada (A121) observó una mejora significativa de la pendiente (no se comunicaron los datos sobre el cambio de nivel).

RESUMEN

La mayoría de los estudios que evaluaron la difusión de material educativo observaron mejoras en el proceso asistencial. Los efectos fueron modestos (mejora absoluta en 4 comparaciones C-ECA del +8,1%, límites: +3,6 a +17%, con una mejora relativa en una de +0,25 de diferencia media estandarizada). Una comparación P-ECA observó un deterioro de la asistencia. Pudo determinarse la significación estadística de 2 comparaciones dicotómicas del proceso y de 1 comparación continua del proceso; ninguna era significativa. Tres de 6 comparaciones de series temporales interrumpidas observaron mejoras significativas del cumplimiento (aunque 2 también observaron posibles efectos de decrecimiento) y 1 de las comparaciones de series temporales interrumpidas observó un deterioro significativo del cumplimiento. Sólo se llevaron a cabo 4 estudios en marcos asistenciales en el Reino Unido. Estos resultados sugieren que el material educativo puede tener un efecto modesto sobre la

implementación de directrices y que puede ser de corta duración. Sin embargo, la base de pruebas es escasa y de poca calidad.

Encuentros educativos

Tres C-ECA (A86, A191, A202), que incluían 3 comparaciones, evaluaron los efectos de encuentros educativos frente a un control sin intervención. Un estudio se llevó a cabo en Estados Unidos, 1 en el Reino Unido y el otro en Holanda. Dos estudios se realizaron en un marco de asistencia primaria y 1 en un marco de atención hospitalaria. En 2 casos, el comportamiento estudiado fue el enfoque global, y la derivación de pacientes en el otro.

Resultados de las comparaciones ECA, ECC y de los ensayos controlados pre-post

- **Medidas dicotómicas del proceso.** Una comparación C-ECA aportó resultados dicotómicos del proceso asistencial (A191); el tamaño del efecto mediano fue una mejora absoluta del cumplimiento del +1%; sin embargo, había un error potencial de la unidad de análisis.
- **Medidas continuas del proceso.** Una comparación C-ECA facilitó medidas continuas del proceso (A202); el efecto mediano fue una mejora relativa del cumplimiento del +27%. Los datos que se aportaban eran insuficientes para calcular una diferencia media estandarizada, y también tenía un error potencial de la unidad de análisis. No fue posible extraer datos comparables en 1 C-ECA (A86), que comunicó efectos no significativos pero tenía un error potencial de la unidad de análisis.
- **Medidas continuas del resultado.** Una comparación C-ECA aportó medidas continuas del resultado (A191) y observó un tamaño del efecto mediano de un cambio relativo del cumplimiento del -3,6%. Se presentaron datos insuficientes para calcular una diferencia media estandarizada y la comparación también tenía un error potencial de la unidad de análisis.

RESUMEN

Hay relativamente pocas evaluaciones de encuentros educativos frente a un grupo de control sin intervención. Los resultados sugieren que los efectos, si los hay, probablemente serán escasos.

Auditoría y retroalimentación

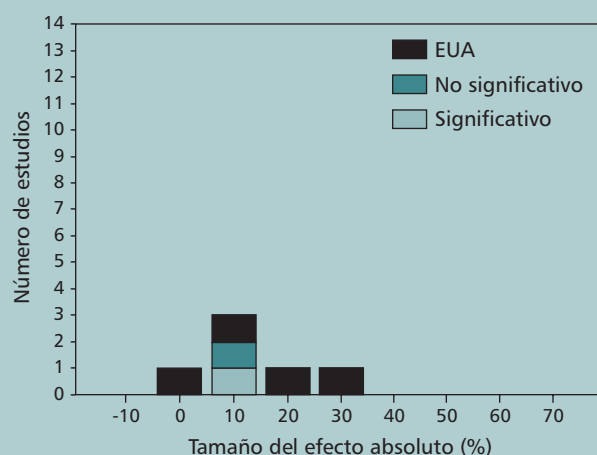
Diez estudios (que incluían 12 comparaciones) evaluaron los efectos de auditoría y retroalimentación: 7 C-ECA (A86, A125, A129, A130, A194, A209, A228), 1 P-ECA (A174 – 3 comparaciones), 1 estudio controlado pre-post (A147) y 1 estudio de series temporales interrumpidas (A75). Ocho estudios se llevaron a cabo en Estados Unidos y 2 en el Reino Unido. Cuatro estudios se realizaron en un marco ambulatorio, 3 en un marco de atención hospitalaria, 2 en centros de asistencia primaria y

en 1 no se explicita el marco asistencial. El comportamiento estudiado fue el enfoque global en 3 estudios, la prevención en 3 estudios, la petición de pruebas en 3 estudios y la planificación del alta en un estudio.

Resultados de las comparaciones ECA, ECC y de los ensayos controlados pre-post

- **Medidas dicotómicas del proceso (Figura 3).** Seis comparaciones aportaron datos dicotómicos del proceso: 5 comparaciones C-ECA (A129, A130, A194, A209, A228) y 1 comparación controlada pre-post (A147). Las 5 comparaciones C-ECA observaron mejoras en la asistencia. Para todas las comparaciones, el efecto mediano fue una mejora absoluta del cumplimiento del +7,0% (límites: +1,3 a +16,0%). Tres comparaciones tenían errores potenciales de la unidad de análisis (A129, A194, A209). Las 2 comparaciones restantes observaron efectos del +5,2% (no significativo) (A228) y +13% ($p < 0,05$) (A130). La comparación controlada pre-post (A147) observó una mejora absoluta del cumplimiento del +32,6%; este estudio tenía un error potencial de la unidad de análisis.
- **Medidas continuas del proceso.** Seis comparaciones aportaron medidas continuas del proceso: 3 comparaciones C-ECA (A86, A125, A129) y 1 comparación P-ECA (A174 – con 3 comparaciones). Una comparación C-ECA (A129) observó una mejora relativa del cumplimiento del +8,5%, diferencia media estandarizada +0,2; sin embargo, el estudio tenía un error potencial de la unidad de análisis. En 2 C-ECA (A86, A125), no fue posible extraer datos comparables. Los autores de ambos estudios indicaron que los efectos de la auditoría y la retroalimentación

Figura 3. Tamaños del efecto observados en datos dicotómicos del proceso en comparaciones ECA, ECC y controladas pre-post de la auditoría y la retroalimentación frente a un grupo de control sin intervención. (EUA: error de la unidad de análisis).



fueron no significativos; sin embargo, había un error potencial de la unidad de análisis en ambas comparaciones.

Dos de 3 comparaciones P-ECA (A174) observaron mejoras en la asistencia; para todas las comparaciones, el efecto mediano fue una mejora relativa del cumplimiento del +15,4% (límites: 0 a +20,3%). Los datos que se aportaron fueron insuficientes para calcular una diferencia media estandarizada y no se pudo determinar la significación de estas comparaciones.

Resultados de las comparaciones de los ensayos de series temporales interrumpidas

Se volvió a analizar 1 comparación (A75) mediante regresión de series temporales, y se observó un cambio significativo del nivel pero no de la pendiente.

RESUMEN

Todos los estudios que evaluaban la intervención de la auditoría y la retroalimentación observaron mejoras en la asistencia. Los efectos fueron modestos, con una mejora absoluta en 5 comparaciones C-ECA del +7,0% (límites: +1,3 a +16,0%) y una mejora relativa de una diferencia media estandarizada de +0,2 en un estudio. Se pudo determinar la significación estadística de 2 comparaciones dicotómicas del proceso; 1 era estadísticamente significativa. Una comparación controlada pre-post observó mejoras absolutas importantes del cumplimiento (+32,6%), pero tenía un error de la unidad de análisis. Una comparación de series temporales interrumpidas observó un cambio significativo del nivel. Sólo 2 estudios se realizaron en el Reino Unido y ambos investigaron la petición de pruebas. Los resultados sugieren que la intervención de la auditoría y la retroalimentación puede tener un efecto modesto sobre la implementación de directrices.

Intervenciones dirigidas al paciente

Siete estudios (que incluían 8 comparaciones) evaluaron los efectos de intervenciones dirigidas al paciente: 4 C-ECA (A24, A111, A179, con 2 comparaciones, A219) y 3 P-ECA (A18, A34, A76). Seis estudios se llevaron a cabo en Estados Unidos y 1 en Canadá. Cuatro estudios se realizaron en centros de asistencia primaria; y 3, en marcos de atención ambulatoria de medicina general. El comportamiento objeto de estudio fue el de los servicios de prevención en 5 estudios, y el tratamiento global del problema en los 2 restantes.

Resultados de las comparaciones ECA, ECC y de los ensayos controlados pre-post

- *Medidas dicotómicas del proceso (Figura 4).* Seis comparaciones comunicaron datos dicotómicos del proceso: 3 comparaciones C-ECA (A111, A179, con 2 comparaciones) y 3 comparaciones P-ECA (A18, A34, A76). Las 3 comparaciones

C-ECA observaron mejoras de la asistencia; el tamaño del efecto mediano fue una mejora absoluta del cumplimiento del +20,8% (límites: +10,0 a +25,4%) (sin embargo, las 3 comparaciones tenían errores potenciales de la unidad de análisis). Las 3 comparaciones P-ECA que comunicaron variables dicotómicas del proceso asistencial observaron mejoras de la asistencia. El tamaño del efecto mediano fue una mejora absoluta del cumplimiento del +1,0% (límites: 0,8 a +9,0%); no pudo determinarse la significación estadística de 1 comparación (A18) y los resultados de las otras 2 comparaciones fueron no significativos (A34, A76).

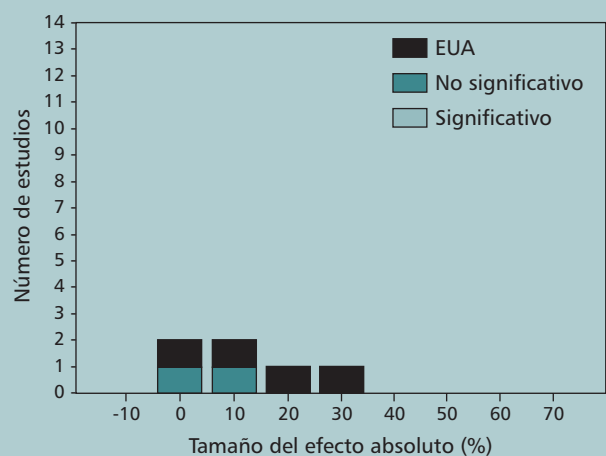
- *Medidas continuas del proceso.* Una comparación cC-ECA (A24) observó un efecto mediano de un deterioro relativo del cumplimiento del -9,1% y una diferencia media estandarizada de -0,67; sin embargo, el estudio tenía un error potencial de la unidad de análisis.

- *Medidas continuas del resultado.* Dos comparaciones C-ECA (A24, A219) observaron efectos medianos de una mejora relativa del cumplimiento del +9,10% y +5,1% y una diferencia media estandarizada de +6,00 y +0,09; sin embargo, ambas comparaciones tenían errores potenciales de la unidad de análisis.

RESUMEN

Todos los estudios observaron mejoras de la asistencia. Los efectos fueron entre moderados y grandes. Los efectos medianos absolutos fueron del +20,8% (límites: +10,0 a +25,4%) en 3 comparaciones C-ECA y de +1,0% (+0,8 a +9,0%) en 3 comparaciones P-ECA. Los tamaños del efecto de las comparaciones P-ECA pueden subestimar los efectos de la interven-

Figura 4. Tamaños del efecto observados en datos dicotómicos del proceso en comparaciones ECA, ECC y controladas pre-post de las intervenciones dirigidas al paciente frente a un grupo de control sin intervención. (EUA: error de la unidad de análisis).



ción si hubo un efecto de contaminación (véase el Anexo 1). Por consiguiente, se debería dar más peso a los C-ECA. Por desgracia, éstos tenían errores de la unidad de análisis. Ninguno de los estudios se llevó a cabo en el Reino Unido, y la mayoría de los estudios se centraron en los servicios de prevención. Estos resultados sugieren que las intervenciones mediadas por el paciente pueden producir mejoras de moderadas a importantes del cumplimiento, sobre todo cuando están centradas en los servicios de prevención.

Recordatorios

Treinta y ocho estudios (que incluían 38 comparaciones) evaluaron los efectos de los recordatorios frente a un control sin intervención/con tratamiento habitual: 16 C-ECA (A10, A38, A40, A47, A81, A111, A119, A128, A134, A135, A146, A156, A179, A180, A182, A209), 9 P-ECA (A13, A16, A34, A37, A42, A98, A133, A204, A207), 2 C-ECC (A46, A189), 8 P-ECC (A90, A112, A142, A222, A223, A224, A225, A226), 2 C-ECC (A101, A187) y 1 estudio de series temporales interrumpidas (A12).

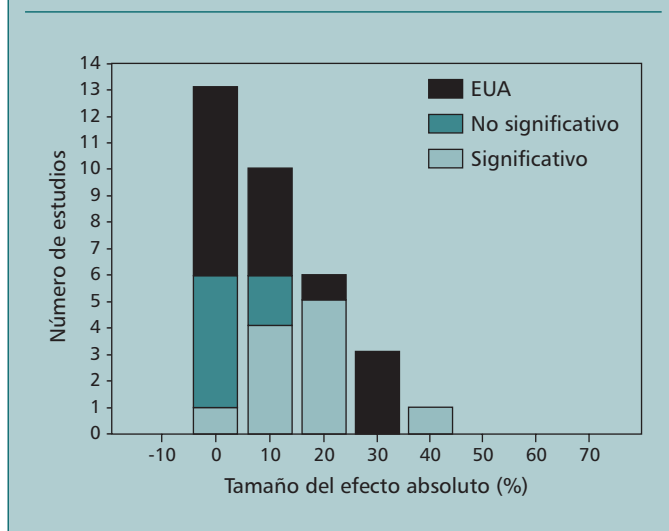
Se llevaron a cabo 34 estudios en Estados Unidos, 2 en Israel, 1 en Canadá y 1 en Tailandia. Quince estudios (40%) se realizaron en marcos de atención ambulatoria de medicina general, 11 (29%) en un marco de asistencia primaria, 8 en uno de atención hospitalaria, 2 en marcos asistenciales mixtos, 1 en un centro ambulatorio especializado y 1 en un centro médico militar. El comportamiento investigado fueron las actividades enfocadas a la prevención en 19 estudios, el enfoque global en 13 estudios, la expedición de recetas en 3 estudios y la planificación del alta, la gestión financiera y los procedimientos en los 3 estudios restantes. Cinco estudios se llevaron a cabo en el Regenstrief Institute (A133, A134, A135, A156, A209) y otros 5 en el Ceders Sinai Medical Center (A222, A223, A224, A225, A226).

Resultados de las comparaciones ECA, ECC y de los ensayos controlados pre-post

- *Medidas dicotómicas del proceso (Figura 5).* Treinta y tres comparaciones comunicaron datos dicotómicos del proceso: 15 comparaciones C-ECA (A10, A38, A40, A47, A81, A111, A119, A128, A134, A135, A156, A179, A180, A182, A209), 8 comparaciones P-ECA (A13, A16, A34, A37, A98, A133, A204, A207), 8 comparaciones ECC (A46, A90, A112, A142, A222, A224, A225, A226) y 2 comparaciones controladas pre-post (A101, A187).

Doce de las 14 comparaciones C-ECA informaron de mejoras en la asistencia; el efecto mediano fue una mejora absoluta del cumplimiento del +14,1% (límites: -1,0 a +34,0%). Once comparaciones tenían errores potenciales de la unidad de análisis; las 3 comparaciones restantes (A119, A134, A135) observaron un efecto mediano del +20,0% (límites: +13 a +20%); todas fueron significativas. No se pudieron extraer datos comparables en 1 ensayo C-ECA (A10), que no comu-

Figura 5. Tamaños del efecto observados en datos dicotómicos del proceso en comparaciones ECA, ECC y en comparaciones controladas pre-post de los recordatorios frente a un grupo de control sin intervención. (EUA: error de la unidad de análisis).



nicó ningún cambio significativo en el cumplimiento global y tenía un error potencial de la unidad de análisis.

Siete de 8 comparaciones P-ECA comunicaron mejoras en la asistencia; el efecto mediano fue del +5,4% (límites: -1 a +25,7%). Tres comparaciones fueron estadísticamente significativas (A13, A37, A207) y 1 estudio tenía un error potencial de la unidad de análisis (A133).

Un C-ECC (A46) comunicó una mejora absoluta de la asistencia del +4,3%, pero tenía un error potencial de la unidad de análisis. Seis de las 7 comparaciones P-ECC comunicaron mejoras en la asistencia; el efecto mediano fue una mejora absoluta en la asistencia del +10,0% (límites: 0 a +40,0%). Cuatro comparaciones fueron estadísticamente significativas (A90, A142, A222, A224).

Dos comparaciones controladas pre-post (A101, A187) observaron efectos consistentes en mejoras absolutas del cumplimiento del +3,6% y +10%; ambas tenían errores potenciales de la unidad de análisis.

- *Medidas continuas del proceso.* Diez comparaciones aportaron datos continuos del proceso: 2 comparaciones C-ECA (A146, A182), 1 comparación P-ECA (A207), 1 comparación C-ECC (A189) y 6 comparaciones P-ECC (A90, A112, A223, A224, A225, A226). Una comparación C-ECA (A182) observó una mejora relativa del cumplimiento del +16,7%; sin embargo, no se pudo calcular la diferencia media estandarizada y había un error potencial de la unidad de análisis. No se pudieron extraer datos comparables para la otra comparación C-ECA (A146). La comparación P-ECC observó un deterioro relativo del cumplimiento del -3,3% y una diferencia media estandarizada de -0,28; sin embargo, no se pudo determinar la significación de la comparación.

El C-ECC (A189) observó una mejora relativa del cumplimiento del +32,0%, y una diferencia media estandarizada de +0,15; sin embargo, no fue significativo. Cinco de 6 comparaciones P-ECC observaron mejoras del cumplimiento; el efecto mediano fue una mejora relativa del cumplimiento del +5,7% (límites: -41,8 a +36,0%). Se pudo calcular la diferencia media estandarizada para 6 comparaciones. Se observaron mejoras en 5 de ellas, y el efecto mediano fue de +0,11 (límites: -0,81 a +0,22). Tres comparaciones fueron estadísticamente significativas. Una de ellas (A225) observó un deterioro relativo del cumplimiento del -41,8% (diferencia media estandarizada -0,81), y las otras 2 observaron mejoras relativas del cumplimiento del +25,7% (diferencia media estandarizada +0,22) (A224) y +36% (diferencia media estandarizada no calculable) (A223).

- *Medidas dicotómicas del resultado.* Cuatro comparaciones comunicaron medidas dicotómicas del resultado. Un C-ECA (A182) observó una mejora absoluta del 3% pero tenía un error potencial de la unidad de análisis. Tres comparaciones ECC que aportaron medidas dicotómicas del resultado (A90, A225, A224) observaron mejoras; el efecto mediano fue de una mejora absoluta del cumplimiento del +1,9% (límites: +1,0 a +6,8%). Ninguno de los efectos fue estadísticamente significativo.

- *Medidas continuas del resultado.* Cuatro comparaciones aportaron datos continuos del resultado: 1 comparación P-ECA (A42) y 3 comparaciones ECC (A90, A225, A224). La comparación P-ECA fue diseñada para comprobar la equivalencia y sugirió que los resultados eran similares entre pacientes atendidos por ordenador o por el médico; sin embargo, había un error potencial de la unidad de análisis. No fue posible extraer datos de la comparación acerca del tamaño del efecto. Dos de las 3 comparaciones ECC observaron mejoras; el efecto mediano fue una mejora relativa del cumplimiento del +1,9% (límites: -2,0 a +7,2%). La diferencia media estandarizada mediana fue de +0,11 (límites: -0,07 a +0,11). Ninguno de los estudios fue estadísticamente significativo.

Resultados de las comparaciones de los ensayos de series temporales interrumpidas

Se volvió a analizar 1 comparación (A12) mediante regresión de series temporales; en ella se observan cambios positivos y estadísticamente significativos tanto en el nivel como en la pendiente.

RESUMEN

Los recordatorios fueron la intervención única evaluada con más frecuencia. Los resultados fueron moderados. Se observaron mejoras en las medidas dicotómicas del proceso en 28 de las 33 comparaciones de todos los diseños. En todos los

C-ECA, el tamaño del efecto mediano fue del +14,1% (límites: -1,0 a +34,0%). Se pudo determinar la significación estadística de 3 comparaciones dicotómicas del proceso mediante C-ECA; todas las comparaciones fueron significativas. Entre todos los P-ECA, el efecto mediano fue del +5,4% (límites: -1 a +25,7%); 3 de las 8 comparaciones fueron estadísticamente significativas. De nuevo, la diferencia en tamaños del efecto entre los C-ECA y los P-ECA sugiere que puede haber un efecto de contaminación en los P-ECA (y en los P-ECC). Un C-ECC observó una mejora del +4,3%, pero tenía un error potencial de la unidad de análisis. En todos los P-ECC, el efecto mediano fue del +10,0% (límites: +0 a +40%); 4 de 7 comparaciones fueron estadísticamente significativas. Dos comparaciones controladas pre-post observaron efectos del +3,6% y del +10%; ambas tenían errores de la unidad de análisis. Se han comparado los recordatorios en una amplia gama de comportamientos estudiados y en una amplia gama de marcos asistenciales (aunque ninguno de los estudios se realizó en el Reino Unido). Estos resultados sugieren que los recordatorios pueden tener un efecto moderado sobre la implementación de directrices.

Otras intervenciones profesionales

Otros 3 estudios (que incluían 3 comparaciones) evaluaron una intervención profesional frente a un grupo de control sin intervención o bien con un tratamiento estándar: 2 C-ECA (A227, A86) y 1 P-ECA (A80). Un estudio se llevó a cabo en Estados Unidos, 1 en Canadá y 1 en el Reino Unido. Dos se realizaron en centros de asistencia primaria, el otro en un marco hospitalario. El enfoque global fue el comportamiento estudiado en 2 comparaciones, y la derivación de pacientes del tercer estudio.

Una comparación C-ECA (A227) evaluó la entrega a los médicos de chicles de nicotina como sustitutos, y observó una mejora absoluta del cumplimiento del +39,1% (porcentaje de pacientes que refieren que los médicos mencionaron el hábito tabáquico) y una mejora absoluta del resultado del +0,5% (porcentaje medio de pacientes por consultorio que dejaron de fumar, prevalencia a 1 año). Había un error potencial de la unidad de análisis en el análisis dicotómico del proceso, y no se pudo determinar la significación del análisis dicotómico del resultado.

La otra comparación C-ECA (A86) evaluó los efectos de entrevistar a los médicos de cabecera respecto a la derivación de pacientes a unidades ambulatorias. No se pudieron extraer datos comparables. El autor afirma que la intervención no tuvo ningún efecto, pero había un error potencial de la unidad de análisis.

La comparación P-ECA (A80) evaluó los efectos de un protocolo para poder descartar rápidamente algunas posibles causas del dolor torácico, y observó una mejora relativa del cumplimiento del +71,8% y una diferencia media estandarizada de +0,31 (estadísticamente significativa).

Intervenciones organizativas

Dos estudios evaluaron los efectos de intervenciones organizativas: 1 C-ECA (A181) y 1 estudio controlado pre-post (A93). Uno de los estudios se llevó a cabo en Estados Unidos, el otro en Holanda. Uno se realizó en un marco de asistencia primaria, el otro en uno de atención hospitalaria. El enfoque global fue el comportamiento estudiado en 1 de las comparaciones, y la expedición de recetas en la otra.

La comparación C-ECA (A181) evaluó los efectos de la continuidad de la atención en la diabetes en la asistencia primaria. Observó una mejora relativa del nivel de hemoglobina glucosilada del +2,1%; no se pudo calcular la diferencia media estandarizada, y el estudio tenía un error potencial de la unidad de análisis.

La comparación controlada pre-post (A93) evaluó los efectos de los servicios de farmacia clínica (revisión de la función profesional) sobre la expedición de recetas en una unidad ortopédica. Observaron un deterioro relativo del cumplimiento del -16,2% y una diferencia media estandarizada de -0,21; sin embargo, el estudio tenía un error potencial de la unidad de análisis.

EVALUACIÓN DE INTERVENCIONES ÚNICAS FRENTE A GRUPOS DE CONTROL CON INTERVENCIÓN

Tres estudios, con 3 comparaciones, evaluaron intervenciones únicas frente a grupos de control que recibieron intervenciones: 1 comparación C-ECA (A118) y 2 comparaciones controladas pre-post (A78, A210). Dos estudios se llevaron a cabo en Estados Unidos y 1 en Canadá. Dos se realizaron en un marco ambulatorio y 1 en un marco de atención hospitalaria. Dos estudiaron el enfoque global del problema y el tercero la expedición de recetas.

La comparación C-ECA (A118) evaluó si requerir a los médicos que respondieran a recordatorios generados por ordenador aumentaba el cumplimiento comparado con los recordatorios solos. Se observó una mejora absoluta del 7% en el grupo de estudio; sin embargo, hubo un error potencial de la unidad de análisis.

Una comparación controlada pre-post (A78) comparó material educativo (grupo de estudio) y recordatorios (grupo de control). Hubo un deterioro absoluto del -67% de la asistencia en el grupo de estudio; el estudio tenía un error potencial de la unidad de análisis.

La otra comparación del estudio controlado pre-post (A210) comparó recordatorios (grupo de estudio) e intervención dirigida al paciente (grupo de control). Hubo una mejora absoluta del +5,6% de la asistencia en el grupo de estudio; el estudio tenía un error potencial de la unidad de análisis.

EVALUACIÓN DE INTERVENCIONES MÚLTIPLES

Un total de 178 estudios (que incluían 222 comparaciones) evaluaron intervenciones múltiples: 137 comparaciones (44,3%) evaluaron 68 combinaciones distintas comparadas con un grupo de control que no recibió ninguna intervención; y 85 (27,5%) comparaciones evaluaron 55 combinaciones distintas comparadas con un grupo de control que también recibió una intervención. El número máximo de comparaciones de la misma combinación de intervenciones frente a un grupo de control sin intervención fue de 11. El número máximo de comparaciones de intervenciones frente a un grupo de control que recibió una intervención fue de 6. Este hecho planteó dificultades importantes para la síntesis y la interpretación de estos estudios. Originalmente, se planeaba llevar a cabo un análisis de metarregresión para estimar los efectos de distintas intervenciones, pero el número de estudios de intervenciones múltiples resultó ser problemático. No fue posible realizar una metarregresión que tuviera en cuenta los efectos de interacción entre intervenciones debido al gran número de variables adicionales necesarias y porque las combinaciones de algunas intervenciones estaban altamente correlacionadas (véase el Anexo 1 para más información). Esta sección describe los resultados de comparaciones de intervenciones múltiples frente a un grupo de control que no recibió ninguna intervención (grupo de control sin intervención) y frente a un grupo de control que sí recibió intervenciones, y resume los resultados para distintas combinaciones de intervenciones con más de 5 comparaciones. Por último, el estudio investigó si la efectividad de las intervenciones múltiples aumentaba con el número de intervenciones.

Intervenciones múltiples frente a un grupo de control sin intervención

Un total de 117 estudios (que incluían 136 comparaciones) evaluaron 68 combinaciones distintas de intervenciones: 26 combinaciones de 2 intervenciones, 19 combinaciones de 3 intervenciones, 16 combinaciones de 4 intervenciones, y 7 intervenciones de 5 intervenciones o más (Tabla 6). Cuarenta y seis eran C-ECA (A1, A9, A17, A22, A23, A24, A28, A31, A32, A39, A48, A51, A55, A56, A59, A64, A65, A85, A95, A96, A100, A102, A106, A111, A116, A117, A129, A137, A138, A141, A144, A149, A154, A158, A168, A183, A186, A192, A194, A196, A202, A208, A209, A219, A227, A235), 13 eran P-ECA (A3, A16, A18, A21, A34, A53, A76, A103, A109, A123, A145, A163, A177), 1 era un C-ECC (A97), hubo 27 estudios controlados pre-post (A5, A27, A49, A50, A54, A66, A69, A71, A74, A77, A83, A101, A110, A115, A127, A147, A160, A161, A167, A171, A172, A173, A175, A176, A178, A213, A216) y 30 estudios de series temporales interrumpidas (A8, A11, A25, A26, A29, A41, A52, A58, A63, A68, A70, A73, A82, A87, A88, A107, A113, A124, A143, A150, A159, A164, A190, A195, A200, A201, A203, A217, A218, A231).

Ochenta y dos estudios se llevaron a cabo en Estados Unidos, 16 en el Reino Unido, 6 en Canadá, 5 en Australia y los 8 restantes en Francia, Méjico, Nueva Zelanda, Holanda, Noruega, Suecia y Tailandia. Cuarenta y cuatro estudios se realizaron en centros de asistencia primaria, 25 en marcos mixtos, 20 en un marco de atención hospitalaria, 14 en un marco de consultas externas de medicina general, y 12 en residencias o instalaciones para enfermos crónicos, servicios de urgencias, consultorios de especialistas y un centro médico militar. El marco no se especificaba en 2 estudios. El comportamiento estudiado fue el enfoque global de un problema en 45 estudios, la prevención en 30 estudios, la expedición de recetas en 21 estudios, la petición de pruebas en 12 estudios, y procedimientos, gestión financiera y derivación de pacientes en los 9 estudios restantes.

Resultados de las comparaciones ECA, ECC y de ensayos controlados pre-post

- *Medidas dicotómicas del proceso.* Setenta y ocho comparaciones (derivadas de 62 estudios), que evaluaron 49 combinaciones distintas de intervenciones, aportaron datos dicotómicos del proceso: 41 comparaciones C-ECA (A1 – 2 comparaciones, A9, A17 – 2 comparaciones, A23, A28 – 2 comparaciones, A31, A32 – 2 comparaciones, A102, A55 – 3 comparaciones, A48, A51, A56, A64, A65, A95, A111, A116, A117, A129 – 2 comparaciones, A137, A141, A144, A154 – 2 comparaciones, A168, A183, A186, A194, A196, A209, A227, A235, A208, A39), 10 comparaciones P-ECA (A3, A16, A18 – 2 comparaciones, A34, A76, A103, A123, A145, A177), 1 comparación ECC (A97), 26 comparaciones controladas pre-post (A5, A49, A50, A54, A66, A69 – 4 comparaciones, A71, A101, A110, A115, A147, A160, A172, A171, A173, A175, A176 – 4 comparaciones, A213 – 2 comparaciones, A216). La mayoría de las comparaciones evaluaron combinaciones de 2 ($n = 36$), 3 ($n = 25$) o 4 ($n = 10$) intervenciones. Hubo pocas repeticiones de evaluaciones para combinaciones distintas de intervenciones.

La Tabla 7 resume el tamaño del efecto mediano de la mejora absoluta del cumplimiento entre estudios para cada combinación de intervenciones.

No se pudieron extraer datos comparables en 1 C-ECA (A208) y en 1 estudio controlado pre-post (A110). La comparación C-ECA evaluó una combinación de material educativo, encuentros educativos e intervenciones organizativas, y observó mejoras significativas de la asistencia. La comparación controlada pre-post evaluó una combinación de material educativo, encuentros educativos, auditoría y retroalimentación y visitas educativas; y observó una mejora significativa en el uso de la profilaxis antibiótica en cirugía.

- *Medidas continuas del proceso.* Treinta comparaciones (derivadas de 27 estudios) aportaron datos continuos del proceso: 18 comparaciones C-ECA (A9, A22, A24, A56, A59, A85, A96, A100, A116, A129 – 2 comparaciones, A138, A149, A158, A186, A192, A202, A208), 3 comparaciones P-ECA (A109, A145, A177), 1 comparación ECC (conglomerado) (A97) y 8 comparaciones de estudios controlados pre-post (A27, A74, A77, A83 – 3 comparaciones, A161, A178). Catorce comparaciones evaluaron combinaciones de 2 intervenciones, 5 comparaciones evaluaron combinaciones de 3 intervenciones, 4 comparaciones evaluaron combinaciones de 4 intervenciones y 1 comparación evaluó una combinación de 5 intervenciones. La Tabla 8 resume el tamaño del efecto mediano de la mejora relativa del cumplimiento y la diferencia media estandarizada entre comparaciones para cada combinación de intervenciones.

No se pudieron extraer datos comparables en 6 comparaciones (A83 – 3 comparaciones, A85, A96, A208). Dos comparaciones C-ECA (A85, A96) no observaron efectos significativos. El tercer C-ECA (A208) observó una reducción significativa del tiempo de espera después de la derivación del paciente. Un estudio controlado pre-post comunicó 3 comparaciones (A83). Dos comparaciones que evaluaban la combinación de material educativo y encuentros educativos observaron incrementos significativos de la monitorización del flujo máximo. La tercera comparación, que evaluaba una combinación de material educativo, encuentros educativos y recordatorios,

Tabla 6. Resumen de comparaciones de intervenciones múltiples frente a un grupo de control sin intervención.

Nº de intervenciones en intervenciones múltiples	Nº de combinaciones distintas de intervenciones múltiples	Nº de comparaciones de intervenciones múltiples	Nº máximo de comparaciones relacionadas con una combinación específica de intervenciones
2	26	62	11
3	19	44	8
4	16	21	3
5	4	6	3
6	2	2	1
7	1	1	
Total	68	136	

observó un incremento significativo del uso de antagonistas B2 orales.

- *Medidas dicotómicas del resultado.* Quince comparaciones aportaron datos dicotómicos del resultado: 8 comparaciones C-ECA (A9, A59, A64, A117, A141, A144, A183, A227), 4 comparaciones P-ECA (A21, A109, A163, A177), 1 comparación ECC (A97) y 2 comparaciones de estudios controlados pre-post (A66, A127). Seis estudios evaluaron 5 combinaciones de 2 intervenciones, 4 estudios evaluaron 4 combinaciones de 3 intervenciones, 3 estudios evaluaron 3 combinaciones de cuatro intervenciones, y 1 estudio evaluó una combinación de 6 intervenciones. La Tabla 9 resume el tamaño del efecto mediano de la mejora absoluta del cumplimiento entre comparaciones para cada combinación de intervenciones. No se disponía de datos comparables en un C-ECA (A117), que observó mejoras del resultado.

- *Medidas continuas del resultado.* Dieciséis comparaciones (derivadas de 15 estudios) aportaron datos continuos del resultado: 11 comparaciones C-ECA (A24, A59, A64, A85, A106, A116, A141, A219 – 2 comparaciones, A235, A208), 4 comparaciones P-ECA (A53, A145, A163, A177) y 1 comparación controlada pre-post (A173). Cinco comparaciones evaluaron 4 combinaciones de 2 intervenciones, 1 comparación evaluó una combinación de 3 intervenciones, 5 comparaciones evaluaron 4 combinaciones de 4 intervenciones y 2 comparaciones evaluaron 2 combinaciones con 5 intervenciones o más. La Tabla 10 resume el tamaño del efecto mediano de la mejora relativa del resultado y la diferencia media estandarizada entre comparaciones para cada combinación de intervenciones.

No se pudieron extraer datos comparables de 3 comparaciones C-ECA (A85, A106, A208). Ninguna de las 3 comparaciones observó efectos significativos de la intervención.

Resultados de las comparaciones de los ensayos de series temporales interrumpidas

Treinta comparaciones evaluaron los efectos de intervenciones múltiples (A8, A11, A25, A26, A29, A41, A52, A58, A63, A70, A68, A73, A82, A87, A88, A107, A113, A124, A143, A150, A159, A164, A190, A195, A200, A201, A203, A217, A218, A231). Se volvieron a analizar 19 comparaciones mediante regresión de series temporales (A8, A11, A25, A41, A52, A70, A73, A82, A88, A113, A124, A143, A150, A164, A190, A201, A203, A218, A231), y se disponía de datos comparables de 3 comparaciones (A29, A195, A217). Diecinueve de 21 comparaciones observaron mejoras en los cambios de nivel, 9 comparaciones fueron estadísticamente significativas (A11, A25, A29, A70, A82, A88, A150, A217, A231) (Tabla 11). Catorce de las 21 comparaciones observaron mejoras en los cambios de pendiente; 4 comparaciones fueron significativas (A25, A29, A124, A190). Una de 7 comparaciones que observaron un deterioro del cambio de pendiente fue estadísticamente significativa (A195).

De 8 estudios no se pudieron extraer datos comparables ni volver a hacer un análisis (A26, A58, A63, A68, A87, A107, A159, A200). Un estudio (A200) presentó datos en gráficos de control e informó de mejoras significativas de la asistencia, pero no se facilitó ninguna cuantificación.

Análisis adicionales

Varias comparaciones evaluaron combinaciones discretas de intervenciones (por ejemplo, 11 comparaciones evaluaron combinaciones de material educativo y encuentros educativos). En esta sección se resumen los resultados de combinaciones de intervenciones que tuvieron más de 4 comparaciones. En 23 de las comparaciones que evaluaron intervenciones múltiples, una de las intervenciones era la de las visitas educativas. Seis combinaciones adicionales de intervenciones múltiples tuvieron más de 4 comparaciones, entre las que encontramos: material educativo con encuentros educativos; material educativo con auditoría y retroalimentación, recordatorios e intervenciones dirigidas al paciente; material educativo, encuentros educativos y auditoría y retroalimentación; material educativo, encuentros educativos y auditoría y retroalimentación; y material educativo, encuentros educativos e intervenciones organizativas.

Intervenciones múltiples que incorporaban visitas educativas

Veintidós estudios (que incluían 23 comparaciones) evaluaron intervenciones múltiples que incluían visitas educativas frente a un grupo de control sin intervención: 14 C-ECA (A9, A28, A48, A55 – 2 comparaciones, A56, A59, A64, A65, A138, A141, A144, A154, A186, A195), 6 estudios controlados pre-post (A49, A110, A161, A171, A172, A173) y 2 estudios de series temporales interrumpidas (A195, A200).

Se evaluaron 11 intervenciones múltiples distintas que incorporaban visitas educativas: 4 combinaciones de 2 intervenciones, 1 combinación de 3 intervenciones, 5 combinaciones de 4 intervenciones y 1 combinación de 6 intervenciones. Las combinaciones más comunes fueron sobre: material educativo y visitas educativas (8 comparaciones); material educativo, encuentros educativos y visitas educativas (4 comparaciones); auditoría y retroalimentación y visitas educativas (2 comparaciones); y material educativo, encuentros educativos, auditoría y retroalimentación y visitas educativas (2 comparaciones).

Doce estudios se llevaron a cabo en Estados Unidos, 4 en el Reino Unido, 4 en Australia y 2 en Suecia. Doce estudios se realizaron en centros de asistencia primaria y 5 en residencias para enfermos crónicos. El resto de estudios se realizaron en un marco asistencial mixto (2) o en un marco de atención hospitalaria (2), y en 1 el marco no estuvo claro. Diez estudios examinaron la expedición de recetas, 9 estudiaron el enfoque global de un problema clínico, 2 la prevención y 1 los procedimientos.

Tabla 7. Resumen de comparaciones entre intervenciones múltiples y grupos de control sin intervención con datos dicotómicos del proceso.

Combinación de intervenciones	Nº de comparaciones	% mediano de la diferencia absoluta entre estudios	Intervalo % de diferencia absoluta entre estudios	Código del estudio
<i>Combinación de dos intervenciones</i>				
AR, Fin	1	6,7	NP	A95
AR, PCL	1	-9	NP	A194
AR, Vis.educ.	2	13,7	+10 a +17,4	A55, A64
AR, Int.prof.	1	68	NP	A123
Mateduc, AR	2	7,4	+7,0 a +7,8	A69, A176
Mateduc, Enc.educ.	5	3,0	-3,0 a +10	A28, A55, A97 A129, A154
Mateduc, Org	1	12,1	NP	A103
Mateduc, Vis.educ.	6	1,3	-5,6 a +13,1	A48, A49, A56, A154, A172, A171
Mateduc, Med.pac.	1	7,8	NP	A76
Mateduc, Int.prof.	1	12	NP	A168
Mateduc, Rec.	1	1,1	NP	A69
Mateduc, Org	1	5,0	NP	A51
Enc.educ., Vis.educ.	1	7,0	NP	A144
Enc.educ., Rec.	1	15,0	NP	A175
Org, Estr.	1	45,0	NP	A18
Vis.educ., Org	1	2,2	NP	A186
Rec., AR	2	9,2	+2,7 a +15,7	A32, A209
Rec., Med.pac.	5	17,0	+1,3 a +25,1	A16, A31, A34, A111, A147
Rec., Estr.	2	12,5	+8,0 a + 17,0	A177, A235
<i>Combinación de tres intervenciones</i>				
Mateduc, AR, PCL	1	18,0	NP	A176
Mateduc, Enc.educ., AR	3	43,0	+2,6 a +9,0	A23, A39, A129
Mateduc, Enc.educ., Org	3	1,0	+0,4 a +6,3	A115, A213, A213
Mateduc, Enc.educ., Vis.educ.	3	11,0	+8,4 a +16,4	A65, A173, A196
Mateduc, Enc.educ., Int.prof.	1	60,0	NP	A227
Mateduc, Org, Estr.	1	2,2	NP	A69
Mateduc, Rec., AR	2	26,0	+25,0 a + 27,0	A176, A176
Mateduc, Rec., Med.pac.	2	9,5	+5,6 a +13,4	A117, A137
Mateduc, Rec., Estr.	1	11,1	NP	A69
Enc.educ., AR, PCL	2	24,0	+16,8 a +33,2	A102, A160
Enc.educ., Rec., AR	1	-2,0	NP	A18
Enc.educ., Rec., Med.pac.	2	23,5	+20,0 a +27,0	A17, A17
Rec., AR, Med.pac.	1	16,9	NP	A32
Rec., AR, Estr.	1	1,5	NP	A183
<i>Combinación de cuatro intervenciones</i>				
Mateduc, Enc.educ., AR, Org	1	2,0	NP	A116
Mateduc, Enc.educ., AR, Vis.educ.	1	6,0	NP	A55
Mateduc, Enc.educ., AR, Estr.	1	0,3	NP	A216
Mateduc, Enc.educ., Int.prof., Org	1	10,0	NP	A50
Mateduc, Enc.educ., PCL, Int.prof.	1	6,0	NP	A71
Mateduc, Enc.educ., Vis.educ., Med.pac.	1	-4,0	NP	A28
Mateduc, Enc.educ., Vis.educ., Int.prof.	1	2,2	NP	A9
Mateduc, Enc.educ., Rec., Org	1	4,0	NP	A1
Mateduc, Rec., Vis.educ., Med.pac.	1	-2,0	NP	A141
Mateduc, Rec., Med.pac., Org, Estr.	1	24,0	NP	A3
Rec., Fin, Org, Estr.	1	54,0	NP	A145
Rec., Med.pac., Org, Estr.	1	8,9	NP	A54

Continúa

Tabla 7. Resumen de comparaciones entre intervenciones múltiples y grupos de control sin intervención con datos dicotómicos del proceso (Cont.).

Combinación de intervenciones	Nº de comparaciones	% mediano de la diferencia absoluta entre estudios	Intervalo % de diferencia absoluta entre estudios	Código del estudio
<i>Combinación de cinco intervenciones</i>				
Mateduc, Enc.educ., Med.pac., Org, Estr.	1	24,3	NP	A101
Mateduc, Enc.educ., Rec., AR, Org	1	-2,0	NP	A1
Mateduc, Enc.educ., Rec., Org, Estr.	1	19,6	NP	A5
<i>Combinación de seis intervenciones</i>				
Mateduc, Enc.educ., AR, Med.pac., Int.prof., Fin	1	15,0	NP	A66

(AR: auditoría y retroalimentación; Fin: financiera; PCL: procesos de consenso locales; Vis.educ.: visitas educativas; Int.prof.: otra intervención profesional; Mateduc: material educativo; Enc.educ.: encuentros educativos; Org: organizativa; Med.pac.: intervención mediada por el paciente; Rec.: recordatorios; Estr.: estructural; NP: no procede.)

Tabla 8. Resumen de comparaciones entre intervenciones múltiples y grupos de control sin intervención con datos continuos del proceso.

Combinación de intervenciones	Nº de comparaciones	Diferencia relativa % (intervalo)	Diferencia media estandarizada (intervalo)	Código del estudio
<i>Combinación de dos intervenciones</i>				
AR, Int.prof.	1	-17,2	NC	A149
Mateduc, AR	2	+9,1 (0 a +18,1)	0,23	A178, A192
Mateduc, Enc.educ.	4	+18,75 (+1,2 a +40,2)	0,09 (+0,03 a +0,15)	A22, A97, A109, A129
Mateduc, PCL	1	-38,6	NC	A100
Mateduc, Vis.educ.	2	+15,7 (+11,3 a +20,0)	NC	A56, A161
Enc.educ., Org	1	+245,0	NC	A202
Org, Vis.educ.	1	+5,5	NC	A186
Rec., Estr.	2	+130,9 (+81 a +180,7)	NC	A27, A177
<i>Combinación de tres intervenciones</i>				
Mateduc, Enc.educ., AR	3	+19,6 (-13,1 a +400,0)	0,12 (-2,1 a +0,45)	A74, A129, A158
Mateduc, Enc.educ., Vis.educ.	1	+24,0	NC	A138
<i>Combinación de cuatro intervenciones</i>				
Mateduc, Enc.educ., AR, Org	1	+32,2	NC	A116
Mateduc, Enc.educ., Vis.educ., Profoth	1	+15,0	NC	A9
Mateduc, Enc.educ., Rec., Med.pac.	1	+23,1	1,5	A24
Enc.educ., Vis.educ., Lid.Op., Int.prof.	1	+1,7	NC	A59
Rec., Fin, Org, Estr.	1	+7,6	4	A145
<i>Combinación de cinco intervenciones</i>				
Mateduc, Enc.educ., Rec., Med.pac., Org	1	NC	2,5	A77

(AR: auditoría y retroalimentación; Fin: financiera; PCL: procesos de consenso locales; Vis.educ.: visitas educativas; Int.prof.: otra intervención profesional; Mateduc: material educativo; Enc.educ.: encuentros educativos; Org: organizativa; Med.pac.: intervención mediada por el paciente; Rec.: recordatorios; Estr.: estructural; Lid.Op.: líderes de opinión; NC: no claro.)

Tabla 9. Resumen de comparaciones entre intervenciones múltiples y grupos de control sin intervención con datos dicotómicos del resultado.

Combinación de intervenciones	Nº de comparaciones	% mediano de la diferencia absoluta entre estudios	Intervalo % de diferencia absoluta entre estudios	Código del estudio
<i>Combinación de dos intervenciones</i>				
AR, Vis.educ.	1	3	NP	A64
Mateduc, Enc.educ.	2	11,5	10,0 a 13,0	A97, A109
Enc.educ., Vis.educ.	1	-1	NP	A144
Rec., Org	1	21	NP	A21
Rec., Estr.	1	8	NP	A177
<i>Combinación de tres intervenciones</i>				
Mateduc, Enc.educ., AR	1	0,8	NP	A127
Mateduc, Enc.educ., Int.prof.	1	3,8	NP	A227
Enc.educ., Org, Int.prof.	1	-8,3	NP	A163
Rec., AR, Estr.	1	-2,2	NP	A183
<i>Combinación de cuatro intervenciones</i>				
Mateduc, Enc.educ., Vis.educ., Int.prof.	1	-7	NP	A9
Mateduc, Rec., Vis.educ., Med.pac.	1	-7	NP	A141
Enc.educ., Vis.educ., Líd.Op., Int.prof.	1	0,1	NP	A59
<i>Combinación de seis intervenciones</i>				
Mateduc, Enc.educ., AR, Med.pac., Int.prof., Fin.	1	3	NP	A66

(AR: auditoría y retroalimentación; Fin: financiera; PCL: procesos de consenso locales; Vis.educ.: visitas educativas; Int.prof.: otra intervención profesional; Mateduc: material educativo; Enc.educ.: encuentros educativos; Org: organizativa; Med.pac.: intervención mediada por el paciente; Rec.: recordatorios; Estr.: estructural; NP: no procede.)

Tabla 10. Resumen de comparaciones entre intervenciones múltiples y grupos de control sin intervención con datos continuos del resultado.

Combinación de intervenciones	Nº de comparaciones	Diferencia relativa % (intervalo)	Diferencia media estandarizada (intervalo)	Código del estudio
<i>Combinación de dos intervenciones</i>				
AR, Vis.educ.	1	0	0	A64
Rec., Estr.	2	-0,4 (-3,0 a +2,2)	0,115 (-0,13 a +0,36)	A235, A177
Med.pac., Int.prof.	1	32		A53
Enc.educ., Org	1	19,4		A163
<i>Combinación de tres intervenciones</i>				
Mateduc, Enc.educ., Vis.educ.	1	13,9	2,375	A173
<i>Combinación de cuatro intervenciones</i>				
Enc.educ., Vis.educ., Líd.Op., Int.prof.	1	2,7		A59
Mateduc, Enc.educ., Rec., Med.pac.	1	9,3	6,0	A24
Rec., Fin, Org, Estr.	1	2		A145
Mateduc, Rec., Vis.educ., Med.pac.	1	-1,4		A141
Mateduc, Enc.educ., AR, Org	1	0,1		A116
<i>Combinación de cinco intervenciones o más</i>				
Mateduc, Enc.educ., Rec., AR, Org	1	0,9	0,003	A219
Mateduc, Enc.educ., Rec., AR, Med.pac., Int.prof., Org	1	3	0,1	A219

(AR: auditoría y retroalimentación; Fin: financiera; PCL: procesos de consenso locales; Vis.educ.: visitas educativas; Int.prof.: otra intervención profesional; Mateduc: material educativo; Enc.educ.: encuentros educativos; Org: organizativa; Med.pac.: intervención mediada por el paciente; Rec.: recordatorios; Estr.: estructural).

Tabla 11. Resumen de los resultados de las comparaciones de series temporales interrumpidas de intervenciones múltiples.

Código de estudio	Combinación de intervenciones	Estudio vuelto a analizar	Cambio de nivel	Significación	Cambio de pendiente	Significación
<i>Combinación de dos intervenciones</i>						
A164	AR, Org	Sí	+	NS	+	NS
A8	Mateduc, Enc.educ.	Sí	+	NS	+	NS
A41	Mateduc, Enc.educ.	Sí	+	NS	-	NS
A203	Mateduc, Org	Sí	+	NS	+	NS
A150	Mateduc, Org	Sí	+	S	-	NS
A195	Mateduc, Vis.educ.	No			-	S
A201	Mateduc, Reg	Sí	+	NS	+	NS
A190	Legislación, Fin	Sí	-	NS	+	S
A218	Rec., Med.pac.	Sí	+	NS	-	NS
A88	Rec., Org	Sí	+	S	+	NS
<i>Combinación de tres intervenciones</i>						
A29	Mateduc, AR, Org	No	+	S	+	S
A11	Mateduc, Enc.educ., AR S	Sí	+	S	+	NS
A70	Mateduc, Enc.educ., AR	Sí	+	S	-	NS
A217	Mateduc, Fin, Org	No	+	S		
A25	Mateduc, Fin, Org	Sí	+	S	+	S
A124	Mateduc, Medios Com, Org	Sí	+	NS	+	S
A122	Mateduc, Rec., Med.pac.	Sí	+	NS	+	NS
A82	Enc.educ., AR, PCL	Sí	+	S	+	NS
A73	Enc.educ., AR, Org	Sí	+	NS	+	NS
<i>Combinación de cuatro intervenciones</i>						
A231	Mateduc, Enc.educ., Rec., Org	Sí	+	S	-	NS
A143	Enc.educ., Rec., AR, Org	Sí	-	NS	+	NS
<i>Combinación de cuatro intervenciones</i>						
A52	Mateduc, Enc.educ., Rec., AR, Org	Sí	+	NS	-	NS

Sólo se disponía de datos comparables en 22 comparaciones. (AR: auditoría y retroalimentación; Fin: financiera; PCL: procesos de consenso locales; Vis.educ.: visitas educativas; Int.prof.: otra intervención profesional; Mateduc: material educativo; Enc.educ.: encuentros educativos; Org: organizativa; Med.pac.: intervención mediada por el paciente; Rec.: recordatorios; Estr.: estructural; Reg, reguladora; NS: no significativo; S: significativo)

Resultados de las comparaciones ECA, ECC y de ensayos controlados pre-post

- **Medidas dicotómicas del proceso.** Dieciocho comparaciones comunicaron datos dicotómicos del proceso: 13 comparaciones C-ECA (A9, A28, A55 – 2 comparaciones, A48, A56, A64, A65, A141, A144, A154, A186, A196) y 5 comparaciones de estudios controlados pre-post (A49, A110, A171, A172, A173).

Once de las comparaciones C-ECA observaron mejoras del cumplimiento; el efecto mediano fue una mejora absoluta del cumplimiento del +6,0% (límites: -4 a +17,4%). Se pudo determinar la significación estadística de 5 comparaciones (A28, A55, A56, A64); el tamaño del efecto mediano para estos estudios fue una mejora absoluta del cumplimiento del +10,0% (límites: -4 a +17,4%) (sólo un estudio que observó una mejora absoluta del cumplimiento del +17,4% fue estadísticamente significativo: A64). Siete estudios tenían

errores potenciales de la unidad de análisis (A48, A65, A141, A144, A154, A186, A196) y no se pudo determinar la significación de la comparación después de la intervención en un estudio (A9).

Dos de las 4 comparaciones de estudios controlados pre-post (A49, A171, A172, A173) que aportaron resultados dicotómicos del proceso asistencial observaron mejoras positivas del cumplimiento. Para todos los estudios, el efecto mediano fue una mejora absoluta del cumplimiento del +7,3% (límites: -5,6 a + 16,4%). Todos estos estudios tenían errores potenciales de la unidad de análisis. No se pudieron extraer datos comparables en un estudio (A110), que comunicó mejoras significativas en el uso de la profilaxis antibiótica en cirugía.

- **Medidas continuas del proceso.** Seis comparaciones aportaron datos continuos del proceso: 5 comparaciones C-ECA (A9, A56, A59, A138, A186) y 1 comparación controlada pre-

post (A161). Todas las comparaciones C-ECA observaron mejoras del cumplimiento; el efecto mediano fue una mejora relativa del cumplimiento del +15,0% (límites: +1,7 a +24,0%). Ningún estudio aportó datos suficientes para calcular la diferencia media estandarizada. Un estudio tenía un error potencial de la unidad de análisis (A186), y no se pudo determinar la comparación después de la intervención en dos estudios (A9, A56). Los dos estudios restantes no fueron estadísticamente significativos.

Un estudio controlado pre-post (A161) observó una mejora relativa del cumplimiento del 11,3%. No se pudo calcular la diferencia media estandarizada y el estudio tenía un error potencial de la unidad de análisis.

- *Medidas dicotómicas del resultado.* Cinco comparaciones C-ECA comunicaron datos dicotómicos del resultado (A9, A59, A64, A141, A144); el efecto mediano fue una mejora absoluta del resultado del -1,0% (límites: -7,0 a +3,0%). Sólo 2 comparaciones observaron mejoras del resultado. Tres estudios tenían errores potenciales de la unidad de análisis (A9, A64, A141), y no se pudo determinar la significación de una comparación (A144). La otra comparación fue no significativa (A59).

- *Medidas continuas del resultado.* Cuatro comparaciones comunicaron datos continuos del resultado: 3 comparaciones C-ECA (A59, A64, A141) y 1 comparación controlada pre-post (A173). El efecto mediano de las comparaciones C-ECA fue del 0% (límites: -1,4 a 2,7%). Se pudo calcular la diferencia media estandarizada para 1 comparación, que fue de 0 (A64). Dos comparaciones tenían errores potenciales de la unidad de análisis (A64, A141) y la tercera comparación fue no significativa (A59). Una comparación controlada pre-post (A173) observó una mejora relativa del resultado del +13,9% y una diferencia media estandarizada de +2,8; sin embargo, el estudio tenía un error potencial de la unidad de análisis.

Resultados de las comparaciones de ensayos de series temporales interrumpidas

Dos estudios comunicaron datos continuos del proceso (A195, A200). Un estudio comunicó mejoras estadísticamente significativas del cumplimiento (A195), pero no pudo volver a analizarse. El segundo estudio (A200) mostró los resultados en forma de gráficos de control; se comunicó que había cambios significativos, pero no se facilitó ninguna cuantificación.

Análisis adicionales

Se llevaron a cabo análisis exploratorios de subgrupos de las comparaciones que evaluaron combinaciones de material educativo y visitas educativas (8 comparaciones), y material educativo, encuentros educativos y visitas educativas (4 comparaciones).

Material educativo y visitas educativas

Ocho estudios evaluaron los efectos de material educativo y visitas educativas frente a un grupo de control sin intervención: 3 C-ECA (A48, A56, A154), 4 estudios controlados pre-post (A49, A161, A171, A172) y 1 estudio de series temporales interrumpidas (A195). Seis comparaciones comunicaron datos dicotómicos del proceso: 3 comparaciones C-ECA (A48, A56, A154) y 3 comparaciones controladas pre-post (A49, A171, A172). Se observaron mejoras de la asistencia en 4 comparaciones; el tamaño del efecto mediano fue una mejora absoluta del cumplimiento del +1,2% (límites: -5,6 a +13,1%). Cinco comparaciones tenían errores potenciales de la unidad de análisis, y el estudio restante fue no significativo (A56). Los 2 estudios (A56, A161) que aportaron datos continuos del resultado observaron una mejora de la asistencia, pero ninguno de ellos aportó datos suficientes para calcular un tamaño del efecto medio estandarizado, y la significación de las comparaciones fue incierta. Una comparación de series temporales interrumpidas refirió mejoras significativas del cumplimiento pero no pudo volver a analizarse (A195).

Material educativo, encuentros educativos y visitas educativas

Cuatro estudios evaluaron los efectos de material educativo, encuentros educativos y visitas educativas frente a un grupo de control sin intervención: 3 C-ECA (A65, A138, A196) y 1 estudio controlado pre-post (A173). Tres comparaciones comunicaron datos dicotómicos del proceso: 2 comparaciones C-ECA (A65, A196) y 1 comparación controlada pre-post (A173). Se observaron mejoras de la asistencia en todas las comparaciones, y el tamaño del efecto mediano fue una mejora absoluta del cumplimiento del +11,0% (límites: +8,4 a +16,4%). Todas las comparaciones tenían errores potenciales de la unidad de análisis. Una comparación C-ECA aportó datos continuos del proceso y observó una mejora relativa no significativa del cumplimiento del +24,0%; se aportaron datos insuficientes para calcular una diferencia media estandarizada (A138). Una comparación controlada pre-post aportó datos continuos del resultado: observaron una mejora relativa del 13,9% y una diferencia media estandarizada de +2,38; la comparación tenía un error potencial de la unidad de análisis (A173).

Relación entre número de intervenciones y tamaño del efecto

Se llevaron a cabo análisis exploratorios, a través de datos dicotómicos del proceso, para determinar si el número de intervenciones influía en la efectividad de intervenciones múltiples —que incluían visitas educativas— comparado con un grupo de control sin intervención. El efecto mediano de la combinación de 2 intervenciones fue del +4,6%, de 3 intervenciones fue del +11,0%, y de 4 intervenciones fue del +0,1% (Tabla 12).

RESUMEN

La mayoría de los estudios que evaluaron visitas educativas múltiples frente a un grupo de control sin intervención observaron mejoras absolutas del cumplimiento de la asistencia. Sin embargo, los efectos fueron modestos, con mejoras absolutas del +6,0% (límites: -4 a +17,4%) en 11 C-ECA y del +7,3% (límites: -5,6 a +17,4%) en 4 estudios controlados pre-post. Se pudo calcular la significación estadística de 5 estudios y sólo 1 fue significativo. Ninguno de los estudios con medidas continuas del proceso aportó información suficiente para calcular la diferencia media estandarizada. El análisis exploratorio de los subgrupos sugirió que:

- Las combinaciones de material educativo y visitas educativas pueden ser relativamente ineficaces
- Las combinaciones de material educativo, encuentros educativos y visitas educativas pueden tener efectos de modestos a moderados
- No existía una relación clara entre el número de intervenciones y el tamaño del efecto.

Cuatro estudios se llevaron a cabo en el Reino Unido. Diez estudios investigaron el comportamiento en la expedición de recetas. Estos resultados sugieren que las estrategias de las intervenciones múltiples que incluyen visitas educativas pueden tener un efecto modesto sobre la implementación de directrices, sobre todo cuando se centran en el comportamiento de la expedición de recetas.

Combinación de material educativo y encuentros educativos

Diez estudios (que incluían 11 comparaciones) evaluaron la combinación de material educativo y encuentros educativos: 5 C-ECA (A22, A28, A55, A129, A154), 1 P-ECA (A109), 1 C-ECC (A97), 1 estudio controlado pre-post (A83 – 2 comparaciones) y 2 estudios de series temporales interrumpidas (A8, A41).

Siete estudios (8 comparaciones) se llevaron a cabo en Estados Unidos, 2 en el Reino Unido y 1 en Francia. Cinco estudios se realizaron en centros de asistencia primaria. El comportamiento estudiado fue el enfoque global de un problema en 7 estudios (8 comparaciones).

Resultados de las comparaciones ECA, ECC y de los ensayos controlados pre-post

- *Medidas dicotómicas del proceso.* Siete comparaciones aportaron datos dicotómicos del proceso: 4 comparaciones C-ECA (A28, A55, A129, A154), 1 comparación C-ECC (A97) y 2 comparaciones de estudios controlados pre-post (A83). Tres comparaciones C-ECA observaron mejoras de la asistencia; el efecto mediano fue una mejora absoluta de la asistencia del +1,9% (límites, -3,0 a +5,0%). Dos comparaciones tenían errores potenciales de la unidad de análisis (A129, A154), y las otras dos comparaciones fueron no significativas. La compa-

ración C-ECC observó una mejora absoluta del +10%; sin embargo, el estudio tenía un error potencial de la unidad de análisis. No se pudieron extraer datos comparables de los estudios controlados pre-post; ambos refirieron aumentos significativos de la monitorización del flujo máximo (A83).

- *Datos continuos del proceso.* Cuatro comparaciones aportaron datos continuos del proceso: 2 comparaciones C-ECA (A22, A129), 1 comparación P-ECA (A109) y 1 comparación C-ECC (A97). Las 2 comparaciones C-ECA observaron mejoras relativas del cumplimiento del +1,2% (A129) y del 33,8% (A22). Sólo se pudo calcular una diferencia media estandarizada en una comparación, +0,03 (A129). Una comparación (A129) tenía un error potencial de la unidad de análisis, y la significación posterior a la intervención de la otra comparación no se pudo determinar. La comparación P-ECA observó una mejora no significativa del 3,7%, pero no se pudo calcular la diferencia media estandarizada. La comparación C-ECC (A97) observó una mejora relativa del +40,2%, con una diferencia media estandarizada de +0,15; sin embargo, el estudio tenía un error potencial de la unidad de análisis.

- *Datos dicotómicos del resultado.* Una comparación C-ECC (A97) observó un incremento del +13,0% pero tenía un error potencial de la unidad de análisis.

Resultados de las comparaciones de los ensayos de series temporales interrumpidas

Dos estudios de series temporales interrumpidas (A8, A41) observaron mejoras no significativas de la asistencia.

RESUMEN

Todas las comparaciones observaron mejoras entre pequeñas y modestas del proceso asistencial. Los efectos fueron

Tabla 12. Efectividad de intervenciones múltiples que incluían visitas educativas por número de intervenciones

Nº de intervenciones	Datos dicotómicos del proceso (% mejora absoluta del cumplimiento)
<i>Dos</i>	
Nº de estudios	10
Efecto mediano	+4,6
Intervalo	-5,6 a +17,4
<i>Tres</i>	
Nº de estudios	3
Efecto mediano	+11,0
Intervalo	+8,4 a +16,4
<i>Cuatro</i>	
Nº de estudios	4
Efecto mediano	+0,1
Intervalo	-4,0 a +6,0

pequeños (mejora absoluta en 4 comparaciones C-ECA +1,9%, límites: -3,0 a +5,0%) o modestos (1 comparación C-ECC +10%). Dos comparaciones C-ECA dicotómicas del proceso fueron no significativas. Dos comparaciones de estudios controlados pre-post dicotómicas del proceso aportaron resultados estadísticamente significativos. Una comparación P-ECA dicotómica del proceso fue no significativa. Dos comparaciones de series temporales interrumpidas fueron no significativas. Sólo 2 estudios se llevaron a cabo en centros del Reino Unido. Los resultados sugieren que la combinación de material educativo y encuentros educativos puede tener, en el mejor de los casos, un efecto pequeño sobre la implementación de directrices. Sin embargo, la base de pruebas es escasa y de poca calidad.

Combinación de material educativo y de auditoría y retroalimentación

Cuatro estudios evaluaron los efectos de una combinación de material educativo y de auditoría y retroalimentación: 1 comparación C-ECA (A192) y 3 comparaciones de estudios controlados pre-post (A69, A176, A178). Dos estudios se llevaron a cabo en el Reino Unido, 1 en Noruega y 1 en Estados Unidos. Dos estudios se realizaron en un marco de atención hospitalaria, 1 en un consultorio de medicina general o de familia y 1 en un marco mixto. El enfoque global fue el comportamiento examinado en 2 estudios, mientras que la petición de pruebas y la expedición de recetas fueron los comportamientos estudiados en los otros 2.

Resultados de las comparaciones ECA, ECC y de ensayos controlados pre-post

- *Datos dicotómicos del proceso.* Dos comparaciones de estudios controlados pre-post comunicaron datos dicotómicos del proceso, y observaron mejoras del +7% (A176) y del +7,8% (A69); ambos estudios tenían errores potenciales de la unidad de análisis.
- *Datos continuos del proceso.* Una comparación C-ECA (A192) observó una mejora relativa del +18,1% y una diferencia media estandarizada de 0,23. Una comparación controlada pre-post no observó ningún cambio (0%) en la eficacia; no se pudo calcular la diferencia media estandarizada. Ambos estudios tenían errores potenciales de la unidad de análisis.

RESUMEN

Hubo relativamente pocas comparaciones de la combinación de material educativo y auditoría y retroalimentación. Tres de las 4 comparaciones observaron mejoras de la asistencia. Los efectos fueron modestos (mejora absoluta mediana del 7,4% en 2 comparaciones C-ECA). Estos resultados sugieren que los efectos probablemente serán modestos.

Combinación de recordatorios e intervenciones dirigidas al paciente

Seis estudios evaluaron combinaciones de recordatorios e intervenciones dirigidas al paciente: 2 comparaciones C-ECA (A31, A111), 2 comparaciones P-ECA (A16, A34), 1 comparación controlada pre-post (A147) y 1 comparación de series temporales interrumpidas (A218). Todos los estudios se llevaron a cabo en Estados Unidos. Tres estudios se realizaron en un marco de asistencia primaria y 3 en un marco de atención ambulatoria. En todos ellos se estudió el comportamiento de enfoque global.

Resultados de las comparaciones ECA, ECC y de ensayos controlados pre-post

- *Medidas dicotómicas del proceso.* Cinco comparaciones comunicaron medidas dicotómicas del proceso: 2 comparaciones C-ECA (A31, A111), 2 comparaciones P-ECA (A16, A34) y 1 comparación controlada pre-post (A147). Las 2 comparaciones C-ECA observaron mejoras absolutas del +17,0% (A31) y +20,0% (A111), pero ambas tenían errores potenciales de la unidad de análisis. Las 2 comparaciones P-ECA observaron mejoras absolutas del +1,3% (A34) y +6,0% (A16); ninguna fue estadísticamente significativa. La comparación controlada pre-post observó una mejora absoluta del +25,1% pero tenía un error potencial de la unidad de análisis.

Resultados de las comparaciones de ensayos de series temporales interrumpidas

Una comparación de series temporales interrumpidas observó la ausencia de mejoras significativas del cumplimiento (A218).

RESUMEN

Todas las comparaciones observaron mejoras del cumplimiento. Los tamaños del efecto derivados de 2 comparaciones C-ECC 1 comparación controlada pre-post sugieren efectos de moderados a grandes (todas las comparaciones tenían errores potenciales de la unidad de análisis). Los tamaños del efecto derivados de 2 P-ECA fueron mucho menores y no significativos, probablemente debido a un efecto de contaminación. Los resultados sugieren que la combinación de recordatorios e intervenciones dirigidas al paciente pueden tener efectos moderados.

Combinación de material educativo, encuentros educativos y auditoría y retroalimentación

Ocho estudios evaluaron combinaciones de material educativo, encuentros educativos y auditoría y retroalimentación: 4 C-ECA (A23, A158, A129, A39), 2 estudios controlados pre-post (A74, A127) y 2 estudios de series temporales interrumpidas.

pidas (A11, A70). Seis estudios se llevaron a cabo en Estados Unidos y 2 en el Reino Unido. Tres estudios se realizaron en un marco de asistencia primaria, 2 en un marco de atención ambulatoria, 2 en un marco de atención hospitalaria y 1 en un marco mixto. El comportamiento estudiado fue el enfoque global en 4 estudios, y la petición de pruebas en otros 4.

Resultados de las comparaciones ECA, ECC y de ensayos controlados pre-post

- *Medidas dicotómicas del proceso.* Tres comparaciones C-ECA (A23, A129, A39) aportaron datos dicotómicos del proceso; el efecto mediano fue una mejora absoluta de la asistencia del +3,0% (límites: +2,6 a +9,0%). Los 3 estudios tenían errores potenciales de la unidad de análisis.
- *Medidas continuas del proceso.* Tres comparaciones aportaron datos continuos del proceso. Las 2 comparaciones C-ECA observaron una mejora relativa del cumplimiento del +19,6% (A129) y del +400% (A158). Fue posible calcular una diferencia media estandarizada de +0,45 en 1 estudio (A129). Un estudio tenía un error potencial de la unidad de análisis (A129) y no fue posible determinar la significación estadística en el otro estudio (A158). Una comparación de estudios controlados pre-post (A74) observó un deterioro relativo del cumplimiento del -13,1% y tenía una diferencia media estandarizada de -0,21. El estudio tenía un error potencial de la unidad de análisis.
- *Medidas dicotómicas del resultado.* Una comparación de estudios controlados pre-post (A127) observó una mejora del +1,0%, pero el estudio tenía un error potencial de la unidad de análisis.

Resultados de las comparaciones de ensayos de series temporales interrumpidas

Dos comparaciones (A11, A70) observaron mejoras significativas del nivel después de la intervención.

RESUMEN

Los efectos observados fueron pequeños (mejora absoluta mediana en 3 comparaciones C-ECA del +3,0%, mejora relativa en 1 comparación C-ECA de +0,45 de diferencia media estandarizada). Estos resultados sugieren que la combinación de material educativo, encuentros educativos y auditoría y retroalimentación puede tener, en el mejor de los casos, un efecto pequeño sobre la implementación de directrices.

Combinación de material educativo, encuentros educativos e intervenciones organizativas

Seis estudios (con 7 comparaciones) evaluaron la combinación de material educativo, encuentros educativos e intervenciones

organizativas: 1 C-ECA (A208), 2 estudios controlados pre-post (A115, A213 – 2 comparaciones) y 3 estudios de series temporales interrumpidas (A63, A68, A87). Cuatro estudios se llevaron a cabo en Estados Unidos y 2 en el Reino Unido. Cuatro estudios se realizaron en centros de asistencia primaria. El comportamiento estudiado fue el enfoque global en 3 estudios, la expedición de recetas en dos estudios y la petición de pruebas en el otro estudio.

Resultados de las comparaciones ECA, ECC y de ensayos controlados pre-post

- *Medidas dicotómicas del proceso.* Tres comparaciones de los estudios controlados pre-post (A115, A213 – 2 comparaciones) aportaron datos dicotómicos del proceso; el efecto mediano fue una mejora absoluta del cumplimiento de +1,0% (límites: +0,4 a +6,3%). Los 3 estudios tenían errores potenciales de la unidad de análisis.
- *Medidas continuas del proceso.* Una comparación C-ECA (A208) comunicó que se había producido una reducción significativa de los tiempos de espera.
- *Medidas continuas del resultado.* Una comparación C-ECA (A208) refirió una ausencia de diferencias significativas en el resultado para el paciente.

Resultados de las comparaciones de ensayos de series temporales interrumpidas

Se identificaron 3 comparaciones de series temporales interrumpidas (A63, A68, A87); sin embargo, no habían sido analizadas correctamente y no fue posible volver a analizarlas.

RESUMEN

Los efectos observados fueron pequeños (mejora absoluta mediana en 3 comparaciones de los estudios controlados pre-post del +1,0%). Estos resultados sugieren que la combinación de material educativo, encuentros educativos e intervenciones organizativas puede tener, en el mejor de los casos, un efecto pequeño sobre la implementación de directrices.

INTERVENCIONES MÚLTIPLES FRENTE A UN GRUPO DE CONTROL CON INTERVENCIÓN

Sesenta y un estudios (con 85 comparaciones) evaluaron 58 combinaciones distintas de intervenciones; había 14 combinaciones de 2 intervenciones, 19 combinaciones de 3 intervenciones, 14 combinaciones de 4 intervenciones y 11 combina-

ciones de 5 intervenciones o más. El número máximo de estudios que evaluaron una misma combinación específica de intervenciones de estudio frente a una combinación específica de intervenciones control fue de 6 (Tabla 13).

De estos estudios, 56 eran comparaciones C-ECA (A2, A6, A7, A20 – 3 comparaciones, A35, A36, A43, A45 – 3 comparaciones, A57, A61 – 3 comparaciones, A62, A67, A72, A79 – 2 comparaciones, A84 – 2 comparaciones, A92 – 2 comparaciones, A94 – 2 comparaciones, A114, A120, A122 – 2 comparaciones, A126, A132, A136 – 2 comparaciones, A139, A148, A152, A153 – 2 comparaciones, A155 – 3 comparaciones, A157, A166, A169, A197, A212, A215, A219, A220, A221 – 2 comparaciones, A230 – 2 comparaciones, A232), 8 comparaciones P-ECA (A30, A33, A104, A105, A193 – 2 comparaciones, A204, A214), 5 comparaciones C-ECC (A199 – 2 comparaciones, A205, A211, A229), 2 comparaciones P-ECC (A108, A165) y 14 comparaciones de los estudios controlados pre-post (A4 – 3 comparaciones, A14 – 2 comparaciones, A89, A99, A131 – 3 comparaciones, A162, A185, A210, A233).

Cuarenta y seis estudios se llevaron a cabo en Estados Unidos, 6 en Canadá, 4 en el Reino Unido, 3 en Holanda, 1 en Francia y 1 en Alemania. Veintiocho estudios se realizaron en un marco de asistencia primaria, 16 en un marco de atención ambulatoria de medicina general, 8 en un marco de atención hospitalaria, 7 en marcos mixtos y 2 en un servicio de urgencias. El comportamiento estudiado fue el enfoque global en 27 estudios, la prevención en 27 estudios, la expedición de recetas en 4 estudios, la petición de pruebas en 2 estudios y la planificación del alta en un estudio.

Resultados de las comparaciones ECA, ECC y de los ensayos controlados pre-post

- *Medidas dicotómicas del proceso.* Setenta y nueve comparaciones (derivadas de 55 estudios), que evaluaron 49 combinaciones distintas de intervenciones, aportaron datos dicotómicos del proceso: 52 comparaciones C-ECA (A2, A6, A7, A20 – 3 comparaciones, A36, A43, A45 – 3 comparaciones, A57, A61 – 3 comparaciones, A62, A67, A72, A79 – 2 comparaciones, A84 – 2 comparaciones, A92 – 2 comparaciones, A94 – 2 comparaciones, A114, A120, A122 – 2 comparaciones, A126, A132, A136 – 2 comparaciones, A148, A153 – 2 comparaciones, A155 – 3 comparaciones, A157, A169, A197, A212, A215, A219, A220, A221 – 2 comparaciones, A230 – 2 comparaciones, A232), 7 comparaciones P-ECA (A30, A33, A104, A105, A193 – 2 comparaciones, A204), 5 comparaciones C-ECC (A199 – 2 comparaciones, A205, A211, A229), 1 comparación ECC (A165) y 14 comparaciones de estudios controlados pre-post (A4 – 3 comparaciones, A14 – 2 comparaciones, A89, A99, A131 – 3 comparaciones, A162, A185, A210, A233).

La Tabla 14 resume el tamaño del efecto mediano de la mejora absoluta del cumplimiento entre estudios para cada

combinación de intervenciones de estudio e intervenciones de control. No se pudieron extraer datos comparables en 2 comparaciones de los estudios controlados pre-post (A14 – 2 comparaciones), que evaluaron los efectos de una combinación de material educativo, encuentros educativos y auditoría y retroalimentación frente a una combinación de material educativo y encuentros educativos. Se comunicó que se había dado un aumento mediano del +3,9%, pero no estaba clara la significación estadística de este aumento.

- *Medidas continuas del proceso.* Trece comparaciones (derivadas de 10 estudios) comunicaron datos continuos del proceso: 10 comparaciones C-ECA (A45 – 3 comparaciones, A62, A152, A153 – 2 comparaciones, A157, A166, A232), 2 comparaciones P-ECA (A104, A105) y 1 comparación ECC (A108). La Tabla 15 resume el tamaño del efecto mediano de la mejora relativa del cumplimiento en todos los estudios para cada combinación de intervenciones de estudio y de intervenciones de control.

No se pudieron extraer datos comparables en 1 comparación C-ECA (A166). El estudio comparaba una combinación de material educativo, auditoría y retroalimentación y visitas educativas con una combinación de auditoría y retroalimentación y visitas educativas. Los resultados fueron mayormente no significativos, pero el efecto de la participación en la valoración del tratamiento del paciente fue significativo.

- *Medidas dicotómicas del resultado.* Diecisiete comparaciones (derivadas de 11 estudios) comunicaron medidas dicotómicas del resultado: 13 comparaciones C-ECA (A7, A45 – 3 comparaciones, A61 – 3 comparaciones, A79 – 2 comparaciones, A132, A139, A230 – 2 comparaciones), 3 comparaciones P-ECA (A105, A104, A214) y 1 comparación controlada pre-post (A233). La Tabla 16 resume el tamaño del efecto mediano de la mejora absoluta del resultado en todos los estudios para cada combinación de intervenciones de estudio y de intervenciones de control.

- *Medidas continuas del resultado.* Nueve comparaciones (de 8 estudios) comunicaron medidas continuas del resultado: 7 comparaciones C-ECA (A6, A35, A36, A79 – 2 comparaciones, A132, A232), 1 comparación P-ECA (A214) y 1 comparación ECC (A108). La Tabla 17 resume el tamaño del efecto mediano de la mejora relativa del resultado entre estudios para cada combinación de intervenciones de estudio frente a intervenciones de control.

Análisis adicionales

Esta sección resume los resultados de las intervenciones múltiples que incluían visitas educativas comparadas con otras intervenciones; y de 3 combinaciones de intervenciones de estudio comparadas con combinaciones de intervenciones de control que tenían más de 4 comparaciones: material educativo y recordatorios comparados con material educativo; encuentros educativos y recordatorios comparados con encuentros educativos; y material educativo, encuentros educativos y recordatorios comparados con material educativo y encuentros educativos.

Tabla 13. Resumen de comparaciones entre intervenciones múltiples y grupos de control con intervención

Nº de intervenciones en intervenciones múltiples	Nº de intervenciones en el grupo de control	Nº de distintas combinaciones de intervenciones múltiples	Nº de comparaciones	Nº máximo de comparaciones de combinaciones específicas de intervenciones de estudio y de control
2	1	13	30	6
	2	1	1	1
3	1	9	9	1
	2	8	13	4
	3	2	2	1
4	1	5	5	1
	2	5	7	2
	3	4	6	2
5	1	3	4	1
	2	3	3	1
	3	1	1	1
6	1	1	1	1
7	1	3	3	1
Total		58	85	

Intervenciones múltiples que incluían visitas educativas comparadas con otras intervenciones

Diez estudios (que incluían 11 comparaciones) evaluaron intervenciones múltiples que incluían visitas educativas frente a una intervención de control. Hubo 9 comparaciones C-ECA (A62, A114, A126, A153 – 2 comparaciones, A166, A169, A215, A221) y 2 comparaciones de estudios controlados pre-post (A99, A233). Seis estudios se llevaron a cabo en Estados Unidos, 2 en Holanda, 1 en Canadá y 1 en el Reino Unido. La mayoría de los estudios (n = 7) se realizaron en centros de asistencia primaria. El comportamiento investigado fue la prevención en 5 estudios, el enfoque global de un problema en 3 estudios y la expedición de recetas en 2 estudios. Todas las comparaciones excepto 3 tenían una única intervención de control. La intervención de control más común fue la difusión de material educativo (n = 5).

Resultados de las comparaciones ECA, ECC y de ensayos controlados pre-post

- *Medidas dicotómicas del proceso.* Ocho comparaciones C-ECA (A62, A114, A126, A153 – 2 comparaciones, A169, A215, A221) y 2 comparaciones de los estudios controlados pre-post (A233, A99) aportaron medidas dicotómicas del proceso. El efecto mediano en los 8 C-ECA fue una mejora absoluta del cumplimiento del +4,5% (límites: -5,2 a +24,3%); 5 comparaciones observaron mejoras de la asistencia. Tres comparaciones tenían errores potenciales de la unidad de análisis (A114, A153 – 2 comparaciones) y no se pudo determinar la

significación estadística de una comparación (A114). Sólo 1 de las comparaciones restantes fue estadísticamente significativa (A62). Las 2 comparaciones de los estudios controlados pre-post observaron un deterioro absoluto de la asistencia del -1,2% (A233) y una mejora absoluta de la asistencia del +6,0% (A99), respectivamente; ambos estudios tenían errores potenciales de la unidad de análisis.

Cinco comparaciones C-ECA (A114, A126, A153 – 2 comparaciones, A215) evaluaron intervenciones múltiples que incluían visitas educativas comparadas con un grupo de control con material educativo. Tres de los 5 estudios observaron mejoras de la asistencia; en todos los estudios, el tamaño del efecto mediano fue del +8,1% (límites: 0 a +24,3%). Tres comparaciones tenían errores potenciales de la unidad de análisis (A126, A153 – 2 comparaciones), una no tenía datos suficientes para determinar la significación (A114) y la otra comparación fue no significativa (A215).

Una comparación C-ECA (A221) y 1 comparación controlada pre-post (A99) evaluaron intervenciones múltiples que incluían visitas educativas comparadas con un control con auditoría y retroalimentación. Ambas observaron mejoras de la asistencia del +0,9% (no significativo) (A221) y del +6,0% (error potencial de la unidad de análisis) (A99), respectivamente.

- *Medidas continuas del proceso.* Cuatro comparaciones C-ECA comunicaron medidas continuas del proceso (A62, A153 – 2 comparaciones, A166). Para 3 comparaciones (A62, A153 – 2 comparaciones), el efecto continuo mediano del proceso fue una mejora relativa del cumplimiento del +34,9% (límites: -1 a +149,0%). En 2 comparaciones se aportaron datos insuficientes para calcular la diferencia media estandarizada y no estaba clara la significación estadística de estas

Tabla 14. Resumen de las comparaciones, con datos dicotómicos del proceso, de intervenciones múltiples frente a grupos de control que recibieron intervenciones

Intervenciones de estudio	Intervenciones de control	Nº de estudios	Tamaño del efecto mediano, (%) diferencia absoluta entre estudios	Intervalo (%) diferencia absoluta entre estudios	Código del estudio
<i>Dos intervenciones de estudio comparadas con una intervención de control</i>					
AR, PCL	Mateduc	1	-2,7	NP	A122
AR, Vis.educ.	AR	1	6	NP	A99
AR, Med.pac.	Med.pac.	1	9	NP	A136
Mateduc, AR	Mateduc	2	3,8	-17 a 24,6	A57, A185
Mateduc, AR	AR	1	0,9	NP	A221
Mateduc, Enc.educ.	Enc.educ.	1	-1	NP	A61
Mateduc, Enc.educ.	Mateduc	2	2	-5, 9	A230, A230
Mateduc, Rec.	Mateduc	2	20,2	19 a 21,4	A157, A220
Enc.educ., AR	AR	1	5,8	NP	A89
Enc.educ., Org	Mateduc	1	8,7	NP	A232
Enc.educ., Med.pac.	Enc.educ.	1	10,3	NP	A193
Enc.educ., Rec.	Enc.educ.	5	8,8	3,9 a 32	A43, A92, A92, A148, A165
Rec., AR	Rec.	1	29,2	NP	A120
Rec., Med.pac.	Med.pac.	2	12,3	3,6 a 21	A136, A210
Rec., Med.pac.	Enc.educ.	1	11	NP	A61
Rec., Med.pac.	Rec.	1	9,9	NP	A211
Rec., Estr.	Rec.	1	2,8	NP	A205
Rec., Estr.	Med.pac.	1	2	NP	A212
<i>Dos intervenciones de estudio comparadas con dos intervenciones de control</i>					
Mateduc, Vis.educ.	Mateduc, Vis.educ.	1	-5,2	NP	A169
<i>Tres intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, AR, Vis.educ.	AR	1	1,5	NP	A221
Mateduc, Enc.educ., Lid.Op.	Mateduc	1	10,8	NP	A122
Mateduc, Enc.educ., Vis.educ.	Mateduc	1	24,3	NP	A153
Mateduc, Enc.educ., Rec.	Int.prof.	1	23	NP	A7
Mateduc, Rec., Profoth	Int.prof.	1	2,6	NP	A30
Enc.educ., AR, Med.pac.	Mateduc	1	-0,3	NP	A132
Enc.educ., Med.pac., Org	Enc.educ.	1	30,7	NP	A104
Enc.educ., Rec., Med.pac.	Enc.educ.	1	13,7	NP	A193
Rec., AR, PCL	PCL	1	0,5	NP	A229
<i>Tres intervenciones de estudio comparadas con dos intervenciones de control</i>					
Mateduc, AR, PCL	Mateduc, AR	1	5,3	NP	A2
Mateduc, Enc.educ., AR	Mateduc, Enc.educ.	1	3,1	NP	A20
Mateduc, Enc.educ., Med.pac.	Mateduc, Enc.educ.	1	2,8	NP	A94
Mateduc, Enc.educ., Rec.	Mateduc, Enc.educ.	4	8,7	-2,7 a 34	A20, A45, A84, A131
Mateduc, Med.pac., Estr.	Mateduc, Med.pac.	1	25	NP	A162
Rec., Med.pac., Estr.	Rec., Med.pac.	1	10	NP	A72
Rec., Med.pac., Estr.	Dir. pac., Estr.	1	2	NP	A204
<i>Tres intervenciones de estudio comparadas con tres intervenciones de control</i>					
Mateduc, Rec., Org	Mateduc, Rec., Org	1	13	NP	A4
Enc.educ., Rec., AR	Enc.educ., Rec., AR	1	4,4	NP	A155

Continúa

Tabla 14. Resumen de las comparaciones, con datos dicotómicos del proceso, de intervenciones múltiples frente a grupos de control que recibieron intervenciones (Cont.).

Intervenciones de estudio	Intervenciones de control	Nº de estudios	Tamaño del efecto mediano, (%) diferencia absoluta entre estudios	Intervalo (%) diferencia absoluta entre estudios	Código del estudio
<i>Cuatro intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., AR, Org	Mateduc	1	0,9	NP	A79
Mateduc, Enc.educ., AR, Vis.educ.	Mateduc	1	0	NP	A215
Mateduc, Enc.educ., Líd.Op., Int.prof.	Mateduc	1	2	NP	A197
Mateduc, Enc.educ., Med. pac., Estr.	Estr.	1	9	NP	A67
Mateduc, Enc.educ., Rec., Med.pac.	Enc.educ.	1	33	NP	A61
<i>Cuatro intervenciones de estudio comparadas con dos intervenciones de control</i>					
Mateduc, Enc.educ., Med.pac., Org	Mateduc, Enc.educ.	2	13,7	9,1 a 18,2	A94, A105
Mateduc, Enc.educ., Rec., AR	Mateduc, Enc.educ.	1	-9,6	NP	A20
Mateduc, Enc.educ., Rec., Med.pac.	Mateduc, Enc.educ.	2	43	34 a 56	A45, A45
Mateduc, Enc.educ., Rec., Estr.	Mateduc, Enc.educ.	1	6	NP	A131
<i>Cuatro intervenciones de estudio comparadas con tres intervenciones de control</i>					
Mateduc, Rec., Líd.Op., Org	Mateduc, Rec., Org	1	2	NP	A4
Enc.educ., Rec., AR, Med. pac.	Enc.educ., Rec., AR	2	4,5	2,8 a 6,2	A155, A155
Enc. educ., Rec., Med.pac., Org	Rec., Med.pac., Org	1	-5,7	NP	A199
Enc. educ., Rec., Med.pac., Org	Enc.educ., Med.pac., Org	1	1,5	NP	A199
Enc.educ., Rec., Med.pac., Estr.	Enc.educ., Med.pac., Estr.	1	12	NP	A33
Mateduc, Enc.educ., AR,	Mateduc, Líd.Op., Org	1	2,5	NP	A6
<i>Cinco intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., Rec., Vis.educ., Med.pac.	Int.prof.	1	-1,2	NP	A233
Mateduc, Enc.educ., Rec., Vis.educ., Med.pac.	Mateduc	1	0	NP	A153
Mateduc, Enc.educ., Rec., Med.pac., Org	Enc.educ.	1	-1	NP	A36
<i>Cinco intervenciones de estudio comparadas con dos intervenciones de control o más</i>					
Mateduc, Enc.educ., Vis.educ., Fin, Org	Mateduc, Fin	1	14	NP	A62
Mateduc, Enc.educ., Rec., AR, Fin	Mateduc, Enc.educ.	1	6,2	NP	A84
Mateduc, Enc.educ., Rec., Org, Estr.	Mateduc, Enc.educ.	1	8	NP	A131
Mateduc, Rec., PCL, Org, Estr.	Mateduc, Rec., Org	1	7	NP	A4
<i>Seis intervenciones de estudio o más comparadas con una intervención de control</i>					
Mateduc, Enc.educ., AR, PCL, Org, Estr.	Mateduc	1	2,2	NP	A79
Mateduc, Enc.educ., Rec., AR, Vis.educ., Líd.Op., PCL	Mateduc	1	12	NP	A114
Mateduc, Enc.educ., Rec., AR, Vis.educ., Med.pac., Estr.	Mateduc	1	8,1	NP	A126
Mateduc, Enc.educ., Rec., AR, Med.pac., Consulta, Teléfono de emergencias	Med.pac.	1	5	NP	A219

(AR: auditoría y retroalimentación; Fin: financiera; PCL: procesos de consenso locales; Vis.educ.: visitas educativas; Int.prof.: otra intervención profesional; Mateduc: material educativo; Enc.educ.: encuentros educativos; Org: organizativa; Med.pac.: intervención mediada por el paciente; Rec.: recordatorios; Estr.: estructural; NP: no procede.)

Tabla 15. Resumen de las comparaciones, con datos continuos del proceso, de intervenciones múltiples frente a grupos de control que recibieron intervenciones.

Intervenciones de estudio	Intervenciones de control	Nº de estudios	% diferencia relativa (intervalo)	diferencia media estandarizada (intervalo)	Código del estudio
<i>Dos intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Rec.	Mateduc	2	9,6	NC	A108, A157
Enc.educ., Rec.	Enc.educ.	1	-1,2	NC	A152
Enc.educ., Org	Mateduc	1	9,5	0,27	A232
<i>Tres intervenciones de estudio comparadas con una intervención de control</i>					
Enc.educ., Med.pac., Org	Enc.educ.	1	21,6	0,33	A104
Mateduc, Enc.educ., Vis.educ.	Mateduc	1	34,9	NC	A153
<i>Tres intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., Rec.	Mateduc, Enc.educ.	1	157	0,95	A45
<i>Cuatro intervenciones de estudio comparadas con dos intervenciones de control</i>					
Mateduc, Enc.educ., Rec., Med.pac.	Mateduc, Enc.educ.	2	257,1	1,57	A45, A45
Mateduc, Enc.educ., Med.pac., Org	Mateduc, Enc.educ.	1	12,2	0,25	A105
<i>Cinco intervenciones de estudio comparadas con dos intervenciones de control</i>					
Mateduc, Enc.educ., Rec., Vis.educ., Med.pac.	Mateduc	1	-1	NC	A153
<i>Cinco intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., Vis.educ., Fin, Org	Mateduc, Fin	1	149	4,9	A62

(AR: auditoría y retroalimentación; Fin: financiera; PCL: procesos de consenso locales; Vis.educ.: visitas educativas; Int.prof.: otra intervención profesional; Mateduc: material educativo; Enc.educ.: encuentros educativos; Org: organizativa; Med.pac.: intervención mediada por el paciente; Rec.: recordatorios; Estr.: estructural; NC: no claro)

comparaciones (A153). La otra comparación (A62) observó una mejora del +149,0% y una diferencia media estandarizada de 4,9 ($p < 0,001$). No se pudieron extraer datos comparables de la otra comparación C-ECA (A166), que comunicó mayormente resultados no significativos, aunque el efecto sobre la participación en la valoración del cuidado del paciente fue significativo.

Dos comparaciones C-ECA (A153 – 2 comparaciones) evaluaron intervenciones múltiples que incluían visitas educativas comparadas con un grupo de control con material educativo, y observaron cambios relativos del -1,0% y del +34,9%. Se aportaron datos insuficientes para calcular la diferencia media estandarizada, y no estaba clara su significación estadística.

- *Medidas dicotómicas del resultado.* Una comparación controlada pre-post (A233) observó un deterioro absoluto del -3,6% en una medida dicotómica del resultado; sin embargo, el estudio tenía un error potencial de la unidad de análisis.

RESUMEN

La mayoría de los estudios que evaluaron múltiples visitas educativas frente a una intervención de control observaron mejoras absolutas del cumplimiento de la asistencia. Los efectos fueron modestos e inferiores a los observados en comparaciones de intervenciones múltiples que incluían visitas educativas con un grupo de control sin intervención. Las visitas educativas parecían ser más efectivas que el material educativo en 3 de las 5 comparaciones, y más efectivas que la auditoría y la retroalimentación en los 2 estudios que comparaban estas intervenciones.

Material educativo y recordatorios comparados con material educativo

Cuatro estudios evaluaron encuentros educativos y recordatorios comparados con encuentros educativos: 2 comparaciones C-ECA (A157, A220), 1 comparación P-ECA (A214) y 1 compa-

Tabla 16. Resumen de las comparaciones, con datos dicotómicos del resultado, de intervenciones múltiples frente a grupos de control que recibieron intervenciones.

Intervenciones de estudio	Intervenciones de control	Nº de comparaciones	Mediana (%) diferencia absoluta entre estudios	Intervalo (%) diferencia absoluta entre estudios	Código del estudio
<i>Dos intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ.	Enc.educ.	2	-6,5	-1 a +14	A61, A139
Mateduc, Enc.educ.	Mateduc	2	-2	-3 a -1	A230, A230
Mateduc, Rec.	Mateduc	1	7	NP	A214
Rec., Med.pac.	Enc.educ.	1	20	NP	A61
<i>Tres intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., Rec.	Int.prof.	1	-2	NP	A7
Enc.educ., Med.pac., Org	Enc.educ.	1	20,4	NP	A104
Enc.educ., AR, Med.pac.	Mateduc	1	1,4	NP	A132
<i>Tres intervenciones de estudio comparadas con dos intervenciones de control</i>					
Mateduc, Enc.educ., Rec.	Mateduc, Enc.educ.	1	6,4	NP	A45
<i>Cuatro intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., AR, Org	Mateduc	1	-0,9	NP	A79
Mateduc, Enc.educ., Rec., Med.pac.	Enc.educ.	1	24	NP	A61
<i>Cuatro intervenciones de estudio comparadas con dos intervenciones de control</i>					
Mateduc, Enc.educ., Rec., Med.pac.	Mateduc, Enc.educ.	2	3,5	+3,2 a +3,7	A45, A45
Mateduc, Enc.educ., Med.pac., Org	Mateduc, Enc.educ.	1	21,7	NP	A105
<i>Cinco intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., Rec., Vis.educ., Med.pac.	Int.prof.	1	-3,6	NP	A233
<i>Seis intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., AR, PCL, Org, Estr.	Mateduc	1	2,2	NP	A79

(AR: auditoría y retroalimentación; Fin: financiera; PCL: procesos de consenso locales; Vis.educ.: visitas educativas; Int.prof.: otra intervención profesional; Mateduc: material educativo; Enc.educ.: encuentros educativos; Org: organizativa; Med.pac.: intervención mediada por el paciente; Rec.: recordatorios; Estr.: estructural; NP: no procede.)

ración ECC (A108). Dos estudios se llevaron a cabo en Estados Unidos, 1 en Holanda y otro en el Reino Unido. Dos estudios se realizaron en un marco de atención hospitalaria, 1 en un marco hospitalario mixto y 1 en un servicio de urgencias. El comportamiento investigado fue el enfoque global en 2 estudios y la petición de pruebas y la planificación del alta en los otros 2 estudios.

- *Medidas dicotómicas del proceso.* Dos comparaciones C-ECA aportaron datos dicotómicos del proceso, y observaron mejoras absolutas del proceso asistencial del +19% (error potencial de la unidad de análisis) (A220) y del 21,4% ($p < 0,05$) (A157).

- *Medidas continuas del proceso.* Una comparación C-ECA (A157) y 1 comparación ECC (A108) aportaron medidas continuas del proceso, y observaron mejoras relativas del proceso

asistencial del +6,2% (error potencial de la unidad de análisis) (A157) y del +14,0% (no significativo) (A108). Ninguno de los 2 estudios aportó datos suficientes para calcular la diferencia media estandarizada.

- *Medidas dicotómicas del resultado.* Una comparación P-ECA (A214) observó una mejora absoluta del +7,0% del resultado para el paciente, pero la significación de la comparación fue incierta.

- *Medidas continuas del resultado.* Una comparación P-ECA (A214) observó un cambio relativo del 0% y una diferencia media estandarizada de 0,0 (no significativo). Una comparación ECC observó un deterioro relativo del resultado para el paciente del -16,3% y una diferencia media estandarizada de -0,32 (no significativa).

Tabla 17. Resumen de las comparaciones, con datos continuos del resultado, de intervenciones múltiples frente a intervenciones de control.

Intervenciones de estudio	Intervenciones de control	Nº de comparaciones	Mediana (%) diferencia relativa entre estudios	Diferencia media estandarizada	Código del estudio
<i>Dos intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Rec.	Mateduc	2	-8,2 (-16,3 a 0)	-0,32	A108, A214
Enc.educ., Org	Mateduc	1	10	0,28	A232
<i>Tres intervenciones de estudio comparadas con una intervención de control</i>					
Enc.educ., AR, Med.pac.	Mateduc	1	6,4	NC	A132
<i>Tres intervenciones de estudio comparadas con dos intervenciones de control</i>					
Enc.educ., Med.pac., Org	Mateduc, Enc.educ.	1	31,3	NC	A35
<i>Cuatro intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., AR, Org	Mateduc	1		NP	A79
<i>Cinco intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., AR, Líd.Op., Org	Mateduc	1	6,1	NC	A6
Mateduc, Enc.educ., Rec., Med.pac., Org	Enc.educ.	1	-20	NC	A36
<i>Seis intervenciones de estudio comparadas con una intervención de control</i>					
Mateduc, Enc.educ., AR, PCL, Org, Estr.	Mateduc	1	5,7	NP	A79

(AR: auditoría y retroalimentación; Fin: financiera; PCL: procesos de consenso locales; Vis.educ.: visitas educativas; Int.prof.: otra intervención profesional; Mateduc: material educativo; Enc.educ.: encuentros educativos; Org: organizativa; Med.pac.: intervención mediada por el paciente; Rec.: recordatorios; Estr.: estructural; NP: no procede.)

RESUMEN

La combinación de material educativo y recordatorios parece ser más efectiva que el material educativo solo.

Encuentros educativos y recordatorios comparados con encuentros educativos

Cinco estudios (que incluían 6 comparaciones) evaluaron los encuentros educativos y los recordatorios comparados con los encuentros educativos: 5 comparaciones C-ECA (A43, A92 – 2 comparaciones, A148, A152) y 1 comparación ECC (A165). Todos los estudios se llevaron a cabo en Estados Unidos, en un marco de atención ambulatoria y estudiaron el enfoque global de un problema.

- *Medidas dicotómicas del proceso.* Cinco comparaciones, incluyendo 4 comparaciones C-ECA (A43, A92 – 2 comparaciones, A148) y 1 comparación ECC (A165), comunicaron datos dicotómicos del proceso. Todas las comparaciones C-ECA observaron mejoras del cumplimiento; el tamaño del efecto mediano fue del +7,9% (límites: +3,9 a +32,0%) (3 comparaciones tenían errores potenciales de la unidad de

análisis, 1 tenía datos insuficientes para determinar la significación). El ensayo clínico controlado observó una mejora del cumplimiento del +24% (no significativa).

- *Medidas continuas del proceso.* Una comparación C-ECA (A152) observó un deterioro relativo del cumplimiento del -1,2%. Se aportaron datos insuficientes para calcular la diferencia media estandarizada y el estudio tenía un error potencial de la unidad de análisis.

RESUMEN

La combinación de encuentros educativos y recordatorios parece ser más efectiva que los encuentros educativos solos.

Material educativo, encuentros educativos y recordatorios comparados con material educativo y encuentros educativos

Cuatro estudios evaluaron material educativo, encuentros educativos y recordatorios comparados con material educati-

Tabla 18. Resumen de los tamaños del efecto medianos de intervenciones múltiples por número de intervenciones en los grupos de estudio y de control

Nº de componentes en el grupo de estudio	Nº de componentes en el grupo de control	Tamaño del efecto absoluto mediano (nº de estudios)			
		0	1	2	3
1		10,2 (n = 52)	5,3 (n = 4)	-	-
2		7,4 (n = 36)	8,8 (n = 26)	-5,2 (n = 1)	-
3		11,0 (n = 25)	10,8 (n = 9)	4,7 (n = 10)	8,7 (n = 2)
4		4,0 (n = 11)	2,0 (n = 5)	13,7 (n = 6)	2,4 (n = 6)
5		21,8 (n = 4)	-0,5 (n = 4)	8,0 (n = 3)	8,0 (n = 1)
6		15,0 (n = 1)	2,2 (n = 1)	-	-
7		-	10,1 (n = 1)	-	-

vo y encuentros educativos: 3 comparaciones C-ECA (A20, A45, A84) y 1 comparación controlada pre-post (A131). Todos los estudios se llevaron a cabo en Estados Unidos y estudiaron el enfoque global. Dos estudios se realizaron en un marco de atención ambulatoria, 1 en un marco de atención hospitalaria y 1 en un marco de asistencia primaria.

- *Medidas dicotómicas del proceso.* Las 4 comparaciones comunicaron medidas dicotómicas del proceso. Dos de 3 comparaciones C-ECA observaron mejoras del cumplimiento; el efecto mediano fue una mejora absoluta del cumplimiento del +13,3% (límites: 2,7 a +34%). Dos comparaciones tenían errores potenciales de la unidad de análisis; la tercera comparación fue estadísticamente significativa (A84). La comparación del estudio controlado pre-post observó una mejora del

+4%; sin embargo, había un error potencial de la unidad de análisis.

- *Medidas continuas del proceso.* Una comparación C-ECA (A45) observó una mejora relativa del +157,0% y una diferencia media estandarizada de +0,95 ($p < 0,05$).

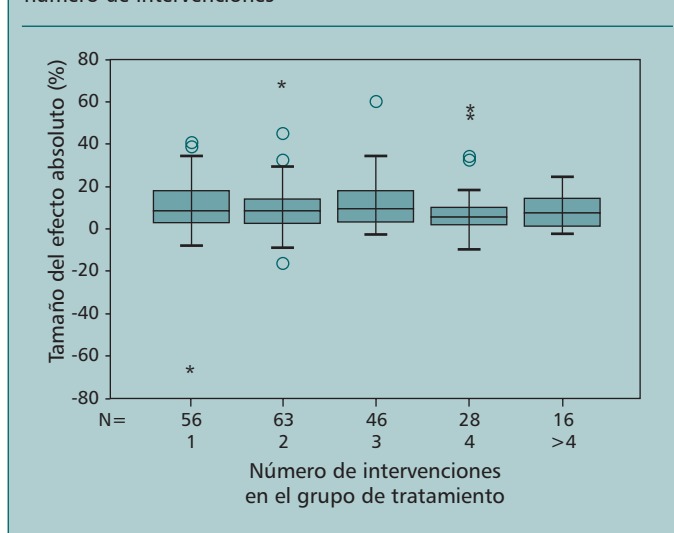
- *Medidas dicotómicas del resultado.* Una comparación C-ECA (A45) observó una mejora absoluta del +6,4%, pero no se pudo determinar la significación estadística de esta comparación.

RESUMEN

La combinación de material educativo, encuentros educativos y recordatorios parece ser más eficaz que el material educativo y los encuentros educativos solos.

¿Aumenta la efectividad de las intervenciones múltiples con el número de intervenciones?

La Tabla 18 muestra los tamaños del efecto medianos de las intervenciones múltiples por número de intervenciones en grupos de estudio y grupos de control. La Figura 6 usa diagramas de caja (*boxplots*) para ilustrar la distribución de los tamaños del efecto en un número creciente de intervenciones en el grupo de estudio. Visualmente, no parecía existir relación alguna entre tamaño del efecto y número de intervenciones. Para los estudios con grupos de control sin intervención, no hubo pruebas estadísticas de una relación entre el número de intervenciones en el grupo de estudio y el tamaño del efecto (prueba de Kruskal Wallis, $p = 0,18$). Tampoco hubo pruebas de una diferencia entre estudios que usaban grupos de control con intervenciones múltiples y estudios que usaban grupos de estudio con intervenciones múltiples (prueba de Kruskal Wallis, $p = 0,69$).

Figura 6. Tamaños del efecto de intervenciones múltiples por número de intervenciones


Revisión sistemática de estrategias de difusión e implementación de directrices: evaluaciones económicas y análisis de costes

Este capítulo revisa el uso de la valoración económica en las evaluaciones de estrategias de difusión e implementación de directrices. La finalidad es determinar la frecuencia y los métodos de valoración económica utilizados y resumir las pruebas existentes sobre la eficiencia relativa de las estrategias de difusión e implementación de directrices, y sobre los costes de desarrollo, difusión e implementación de directrices. A partir de aquí, se extraerán conclusiones acerca de los puntos fuertes y los débiles de los análisis económicos de la implementación de directrices, y se ofrecerán recomendaciones respecto a cómo mejorar los análisis.

La siguiente sección de este capítulo explica brevemente cómo una evaluación económica de las estrategias de implementación de directrices puede diferir de una evaluación económica de intervenciones clínicas, y señala por qué un analista puede optar legítimamente por no incluir todas las fases de difusión e implementación de directrices en una evaluación económica. Después, se presentan los métodos y los resultados de la revisión de valoraciones económicas y, en la última sección, los puntos fuertes y los débiles de los análisis económicos y de costes de la implementación de directrices.

ESTRUCTURA PARA LA EVALUACIÓN ECONÓMICA DE LA IMPLEMENTACIÓN DE DIRECTRICES

La evaluación económica de las estrategias de implementación de directrices debe basarse en los mismos principios básicos que una evaluación de una tecnología médica estándar (un nuevo fármaco o un nuevo tipo de cirugía)¹⁷. La valoración

de una tecnología médica estándar se limita a considerar los costes y los beneficios de proporcionar un tratamiento (por ejemplo, administrar el fármaco A o el fármaco B) y las consecuencias de ese tratamiento. La evaluación de las estrategias de implementación de directrices es distinta en el sentido de que la amplitud de los costes y de los beneficios que se podrían considerar es mayor. Para determinar si vale la pena implementar una pauta, hay que determinar si la propia pauta representa un uso eficiente de recursos, así como la forma más eficiente de ayudar a los médicos a adoptar la pauta.

Hay tres fases diferenciadas en el desarrollo y la implantación de directrices que podrían considerarse en una evaluación económica¹⁸:

1. desarrollo de la pauta
2. difusión e implementación de la pauta
3. efectos y costes del tratamiento como consecuencia del cambio de actuación.

Aunque la estructura de una evaluación económica de las estrategias de implementación de directrices podría incluir los costes y los beneficios de cada una de las tres fases, a veces puede estar justificado diseñar una evaluación económica de alcance más limitado. Que estas limitaciones sean apropiadas depende de cuál sea la justificación. En todo caso, ésta debe ser explícita y debe estar fundamentada en pruebas apropiadas.

Una razón legítima para limitar la evaluación de los costes y los beneficios de cambiar de práctica puede ser el hecho de que la pauta ya haya demostrado representar una práctica eficiente¹⁹. Por lo tanto, los costes y las consecuencias de adoptar las recomendaciones de la pauta no necesariamente tienen que valorarse en un estudio primario si la práctica propugnada por la pauta ha demostrado ser eficiente en el marco sanitario

en que tendrá lugar la implementación y a la escala de implementación deseada.

Otra razón para limitar legítimamente el alcance de la evaluación económica tiene que ver con la perspectiva adoptada por la evaluación. Si se adopta una perspectiva social, deben tenerse en cuenta todos los costes (incluyendo los que recaen en el paciente) y todos los beneficios, como los obtenidos por los profesionales de la salud en conocimientos o satisfacción laboral durante el desarrollo, la difusión y la implementación de la pauta.

Una última razón para no medir algunos costes y beneficios es cuando se puede suponer que su inclusión no cambiará la política elegida. Por ejemplo, si se cree, a priori, que una directriz reducirá los costes del tratamiento a la vez que mantiene o mejora el resultado para el paciente, puede considerarse justificado excluir los costes de desarrollo, difusión e implementación, ya que de ningún modo podrían invalidar el ahorro en los costes del tratamiento. Sin embargo, los resultados sólo serían transferibles a la aplicación de la misma estrategia de implementación de unas directrices que resolvieran los mismos problemas en un marco similar.

Se valora como parte de la revisión el grado en que las evaluaciones económicas y los análisis de costes incluidos hayan considerado todos los costes y los beneficios de cada etapa. En los casos en que se han excluido costes y beneficios aplicables, se tendrá en cuenta si su exclusión está justificada por la perspectiva adoptada, por la irrelevancia para las conclusiones finales o por la existencia de un trabajo anterior que demuestre que sería eficiente adoptar las directrices recomendadas.

MÉTODOS

Criterios de búsqueda y selección

Se identificaron las evaluaciones económicas y los análisis de costes incluidos en esta revisión como parte de la revisión sistemática detallada en los Capítulos 2–4.

Criterios de selección para la revisión de los análisis de costes y económicos

Dos revisores identificaron todos los estudios que aportaban información sobre costes. A continuación, un tercer revisor (L. Vale) determinó si el estudio era un análisis de costes o una evaluación económica detallada. Los estudios se consideraron como análisis de costes si no relacionaban costes con eficacia/beneficios. Se utilizó una definición general de evaluación económica: los estudios debían aportar datos sobre costes y, al menos, criterios de valoración sustitutos para efectividad/beneficios.

Revisión de la calidad metodológica

Los estudios incluidos en esta sección fueron valorados frente a la guía del *British Medical Journal* para revisores de evaluaciones económicas (Cuadro 3)²⁰. Estas guías están diseñadas

Cuadro 3. Criterios para valorar la calidad de las evaluaciones económicas y de los análisis de costes

Diseño de evaluación económica

1. Se expresa la pregunta planteada en la investigación
2. Se expresa la importancia de la pregunta
3. Se expresa y se define el punto de vista del análisis
4. Se expresa el fundamento de la elección de programas o intervenciones alternativos en la comparación
5. Se definen claramente las alternativas que se están comparando
6. Se expresa la forma de evaluación económica usada
7. Se justifica la forma de la evaluación económica en relación con la pregunta planteada

Recogida de datos

8. Se indica(n) la(s) fuente(s) de las estimaciones de efectividad
9. Se facilitan detalles del diseño y resultados del estudio de efectividad
10. Se facilitan detalles de los métodos de síntesis de las estimaciones de los efectos sobre la salud
11. Se indica(n) claramente la(s) medida(s) primaria(s) del resultado para la evaluación económica
12. Se declaran los métodos usados para valorar estados de salud y otros beneficios
13. Se facilitan detalles de los sujetos de los que se obtuvieron valoraciones
14. Se comunican por separado los cambios de productividad (si se incluyen)
15. Se aborda la relevancia de los cambios de productividad para la cuestión planteada en el estudio
16. Las cantidades de recursos se comunican separadamente de sus costes unitarios
17. Se describen métodos para la estimación de cantidades y costes unitarios
18. Se documentan datos sobre moneda y precio
19. Se facilitan detalles sobre los ajustes de moneda y precio para tener en cuenta la inflación o la conversión de moneda
20. Se facilitan detalles de todo modelo usado
21. Se justifica la elección del modelo usado y los parámetros clave en los que se basa la elección

Análisis

22. Se indica el horizonte temporal de costes y beneficios
23. Se indica el tipo (o tipos) de descuento
24. Se justifica la elección de tipo(s)
25. Se explica el motivo si no se descuentan costes y beneficios
26. Se facilitan detalles de las pruebas estadísticas y del intervalo de confianza de los datos estocásticos
27. Se indica el enfoque del análisis de sensibilidad
28. Se justifica la variable elegida para el análisis de sensibilidad
29. Se indican los límites de las variables que varían
30. Se comparan alternativas pertinentes
31. Se explicita el análisis incremental
32. Los principales resultados se presentan de forma agregada y desagregada
33. Se da respuesta a la cuestión planteada en el estudio
34. Se comunican las conclusiones derivadas de los datos
35. Las conclusiones se acompañan de las advertencias apropiadas

para mejorar la calidad de las evaluaciones económicas, y sus recomendaciones abarcan tres campos amplios: diseño del estudio, recogida de datos y análisis e interpretación de los resultados. Si los estudios sólo analizaban los costes, no se valoraron frente a los criterios relacionados con los beneficios. Los criterios no se usaron como sistema de puntuación sino como un medio para resumir aquellos aspectos de una evaluación económica que generalmente se consideran importantes.

Revisión de resultados

Se extrajeron datos sobre el uso de recursos y el coste del desarrollo, la difusión y la implementación de directrices, y se resumieron en función del tipo de estrategia de difusión e implementación adoptada.

RESULTADOS

Estudios que realizan análisis de costes o evaluaciones económicas

Sesenta y tres de 235 (29,4%) estudios (que incluían 78 comparaciones) realizaron análisis de costes o evaluaciones económicas que intentaban valorar los costes o la relación coste-efectividad de distintas estrategias de implantación de directrices (A3, A8, A10, A11, A12, A18 – 3 comparaciones, A23, A29, A30, A32 – 2 comparaciones, A41, A51, A52, A53, A55 – 3 comparaciones, A57, A61 – 3 comparaciones, A63, A64, A72, A73, A82, A87, A93, A96, A100, A101 – 2 comparaciones, A104, A106, A107, A109, A110, A112, A114, A120, A123, A129 – 3 comparaciones, A133, A142, A143, A144, A145, A147 – 2 comparaciones, A154 – 2 comparaciones, A158, A159, A164, A169, A179 – 3 comparaciones, A189, A196, A200, A203, A207, A208, A217, A221 – 2 comparaciones, A223, A224, A226, A228, A229, A234).

Características generales

Las características de los estudios incluidos se describen en la Tabla 19. La mayoría de los estudios se llevaron a cabo en Estados Unidos, y su objetivo era mejorar el tratamiento. Treinta y seis estudios intentaron cambiar un comportamiento, y el resto estudió varios comportamientos (hasta un máximo de seis). La población de estudio siempre era la médica, aunque en 14 estudios también se estudiaba a otros profesionales de la salud (mayormente enfermeras).

Intervenciones evaluadas

La Tabla 20 describe las intervenciones evaluadas. La mayoría de las comparaciones evaluaron intervenciones múltiples —58 de las 78 comparaciones (74%)— y el número máximo de intervenciones de cambio conductual empleadas fue de 7. Diez estudios evaluaron los recordatorios como intervención única.

Tabla 19. Características generales de los estudios incluidos

Característica de estudio	Tipo	Nº de estudios
Diseño de estudio	C-ECA	24
	P-ECA	10
	C-ECC	2
	P-ECC	5
	CPP	4
	STI	18
País de origen	Estados Unidos	45
	Reino Unido	11
	Canadá	3
	Otros	4
Fundamento de estudio	Mejora del tratamiento	36
	Contención de costes	19
	Mejora del tratamiento y contención de costes	8
Comportamiento estudiado	Enfoque global	24
	Educación y asesoramiento del paciente	14
	Expedición de recetas	28
	Prevención	13
	Derivación de pacientes	8
	Petición de pruebas	24
	Otros	18
	Población de estudio	Médicos
Médicos, enfermeras		3
Médicos, enfermeras, farmacéuticos		1
Médicos, enfermeras, farmacéuticos, otros		1
Médicos, enfermeras, otros		2
Médicos, no claro		1
Médicos, otros		6

C-ECA: ensayo controlado aleatorizado con conglomerados; P-ECA: ensayo controlado aleatorizado con pacientes; C-ECC: ensayo controlado con asignación de conglomerados; P-ECC: ensayo controlado con asignación de pacientes; CPP: ensayo controlado pre-post; STI: ensayo con series temporales interrumpidas.

Tabla 20. Intervenciones evaluadas en evaluaciones económicas y análisis de costes

Tipo de intervención	Nº de veces evaluada
Material educativo	29
Encuentros educativos	25
Proceso de consenso	5
Visitas educativas	10
Mediada por el paciente	3
Auditoría y retroalimentación	24
Recordatorios	20
Otra	17
Financiera	4
Organizativa	10
Estructural	11

Calidad metodológica de las evaluaciones económicas

Criterios de diseño

El punto de vista adoptado en la evaluación económica se especificó en 10 estudios (A3, A8, A23, A32, A57, A123, A145, A159, A208, A217) y se justificó en 5 (A23, A57, A159, A208, A217). Cuatro estudios tomaron la perspectiva del servicio de salud (A23, A57, A159, A217), 5 estudios tomaron la perspectiva del hospital o de los profesionales (A3, A8, A32, A217, A123) y 1 estudio tomó una perspectiva social (A208).

Como resultado de los criterios de inclusión de la revisión sistemática, se definieron claramente las alternativas que se estaban comparando.

Sólo 12 de los 63 estudios (19%) especificaron la forma de evaluación económica usada; 9 estudios comunicaron que habían analizado la relación coste-efectividad (A11, A23, A32, A104, A109, A114, A144, A145, A154) y 3 comparaciones comunicaron que habían analizado la relación coste-beneficio (A12, A208). Al examinarlos más detenidamente, 2 de los estudios que declararon haber analizado la relación coste-beneficio parecían haber analizado la relación coste-consecuencia (presentando las diferencias de costes frente a las diferencias en varias medidas de la efectividad) (A221, A208), y 1 estudio parecía haber analizado los costes (A12). Un estudio comunicó que había analizado los costes; sin embargo, al examinarlo más detenidamente, parecía ser un análisis de la relación coste-consecuencia (A57).

Los estudios restantes no definieron la forma de evaluación económica emprendida. Treinta y cinco parecían ser análisis de la relación coste-consecuencia, 2 parecían ser análisis de la relación coste-efectividad y 13 eran análisis de costes, ya que aportaron algún aspecto del coste (por ejemplo, costes de personal o material) pero no intentaron relacionar costes con beneficios.

Treinta y ocho estudios comunicaron costes de cambios de tratamiento como consecuencia de la difusión y la implementación de directrices, y 11 comunicaron costes de difusión e implementación de directrices (A8, A10, A23, A32, A55, A120, A142, A158, A179, A228, A229). Dos estudios comunicaron costes de desarrollo, difusión e implementación de directrices y del tratamiento (A87, A208). Doce estudios comunicaron costes de difusión e implementación de directrices y del tratamiento (A11, A29, A61, A64, A72, A114, A133, A196, A207, A221, A224, A226).

Trece estudios que aportaron más que sólo los costes del tratamiento pudieron definirse como análisis de la relación coste-consecuencia; 5 como análisis de la relación coste-efectividad y 7 como análisis de costes.

Criterios de recogida de datos

Fuentes de las estimaciones de efectividad

Como resultado de los criterios de inclusión en la revisión sistemática, todos los estudios precisaban las fuentes de las

estimaciones de efectividad, así como detalles del diseño del estudio de efectividad. Sin embargo, las deficiencias metodológicas de los estudios primarios muchas veces desvirtuaban los resultados de efectividad. Por ejemplo, la significación estadística de los beneficios fue incierta en 20 C-ECA, C-ECC y estudios controlados pre-post que tenían errores potenciales de la unidad de análisis (A10, A23, A34, A51, A57, A61, A72, A77, A93, A96, A101, A106, A147, A154, A158, A179, A196, A207, A208, A228); y 11 estudios de series temporales interrumpidas se analizaron de forma inapropiada en los informes publicados (A11, A12, A41, A52, A73, A107, A143, A159, A164, A203, A234).

Criterio principal de valoración en la evaluación económica

En muchos estudios no se expuso claramente el criterio principal de valoración en la evaluación económica. Sólo 5 estudios intentaron medir el resultado para el paciente o aportar otras medidas económicas del beneficio (A104, A145, A208, A224, A226). El resto usaron medidas del proceso como criterio principal de valoración. Como ya se ha indicado, es justificable enfocarse en estas variables de valoración intermedias si se ha demostrado que la directriz que se está evaluando es eficiente en la implementación seleccionada. Sin embargo, la base de pruebas de efectividad para la gran mayoría de las directrices evaluadas en la revisión sistemática fue incierta (véase el Capítulo 3).

Métodos para la estimación de costes

Los métodos utilizados para la estimación de costes fueron bastante completos en 29 estudios. Se indicó el año del precio en sólo 13 estudios (A10, A30, A32, A41, A52, A57, A73, A87, A101, A110, A158, A221, A234), y se efectuaron ajustes por la inflación en sólo 5 (A29, A32, A57, A101, A226) —aunque en 4 estudios no fue necesario debido a la escasa duración de los análisis (A57, A73, A110, A221)—. Sólo 3 aportaron datos sobre métodos, año del precio y ajustes por la inflación (A32, A57, A101). Uno de estos 3 estudios examinó sólo el coste de implementación (A32) y los otros 2 examinaron sólo algunos de los costes del tratamiento.

Sólo 9 estudios facilitaron datos sobre el uso de recursos para los costes incluidos (A11, A30, A73, A87, A145, A147, A179, A208, A229). Cinco de estos estudios sólo se centraron en los costes del tratamiento (A11, A30, A73, A145, A147), 2 examinaron sólo los costes de implementación (A179, A229) y 2 examinaron los costes de las 3 fases (A87, A208). Estos análisis no fueron completos en términos de los costes considerados, pero 1 estudio sí intentó incorporar en el proceso de la decisión el uso de los recursos cuyo coste no había sido determinado y los beneficios no cuantificados mediante la aplicación de un enfoque de balance (A208).

Globalmente, ningún trabajo estudió las tres fases y aportó información razonablemente completa sobre la estimación del coste.

Detalles del modelo usado

Todos los estudios se basaron en datos primarios y no se realizaron procedimientos de modelado.

Criterios de análisis

No se utilizó o no se mencionó ningún método de rebajar costes para ninguna de las evaluaciones económicas en las que la recogida de datos sobre costes y beneficios duró más de 12 meses. Sólo 16 estudios comunicaron algún tipo de análisis de sensibilidad del coste y de la relación coste-efectividad (A3, A23, A29, A41, A82, A87, A101, A107, A114, A169, A184, A189, A208, A221, A224, A234). En la mayoría de los casos, el efecto sobre los resultados se valoró a partir de los cambios en una sola variable (por ejemplo, del coste de un procedimiento) en un análisis de sensibilidad unidireccional (A23, A29, A41, A82, A87, A101, A107, A114, A169, A189, A221). Un estudio examinó el efecto de los cambios en el personal y del nivel salarial sobre la relación coste-efectividad (A179); un estudio investigó el efecto de cambiar cada aspecto del coste de forma independiente (análisis múltiple de sensibilidad unidireccional) (A208) y 3 estudios proporcionaron un intervalo de confianza alrededor de una estimación del ahorro de costes (A3, A224, A234).

No se realizó un análisis incremental de costes y efectos o de la relación coste-efectividad en 9 de los 46 análisis [los 3 estudios restantes no precisaron un análisis incremental, ya que en 1 era desconocido (A73) y 2, en los que sí procedía, no detectaron diferencias ni en los costes ni en los efectos (A100, A112)] (excluyendo los 14 análisis de costes de este análisis) (A23, A63, A72, A104, A169, A179, A207).

En un estudio planteado inicialmente como un análisis de la relación coste-efectividad, la estrategia de implementación fue el ahorro de costes (A145). Salvo en 11 de los 38 análisis de la relación coste-consecuencia, se concluyó que la estrategia de implementación era eficiente (A30, A53, A63, A100, A112, A129, A158, A208, A221, A226, A229). En 5 de estos estudios, la estrategia de implementación fue más costosa pero no fue más efectiva (A100, A112, A221, A226, A229). Estas conclusiones, al igual que las relativas al coste y la eficiencia, deben considerarse con cautela, dadas las limitaciones en la metodología y en la información sobre los datos indicados arriba.

Resumen de las estimaciones de recursos y costes para el desarrollo, la difusión y la implementación de directrices

Debido a la calidad generalmente deficiente de la información facilitada sobre las evaluaciones económicas, en la mayoría de ellas no se dispone de datos sobre el uso de recursos y el coste del desarrollo, la difusión y la implementación de directrices. Cuatro análisis aportaron información útil sobre el uso de recursos y sobre los métodos de estimación de costes (A87, A179, A208, A229). Dos de estos análisis informaron del uso

y el coste de recursos en la fase de implementación; los otros 2 análisis abarcaron las tres fases (A87, A208). Los detalles de estos análisis se resumen en las Tablas 21a (detalles de los estudios), 21b (uso de recursos en la fase de desarrollo de directrices) y 21c (uso de recursos en la fase de difusión e implementación de directrices). Dos estudios utilizaron la preparación y la difusión de material educativo junto con encuentros educativos (A87, A208). El último estudio informó sobre el uso de recursos para la auditoría y la retroalimentación.

Los datos sobre recursos de los que se informa para la auditoría y la retroalimentación pueden no ser transferibles a otro marco, ya que se refiere a tiempo de ordenadores y de personal, y el estudio se llevó a cabo a finales de los años 70 (A229). Desde entonces, la tecnología informática ha evolucionado mucho.

DISCUSIÓN

La finalidad de la evaluación económica es aportar información para facilitar la toma de decisiones respecto al uso de recursos escasos. Al identificar y valorar sistemáticamente las pruebas disponibles sobre las estrategias de implementación de directrices, este capítulo ha procurado aportar información para facilitar la toma de decisiones sobre la mejor manera de llevar las directrices a la práctica. Aunque el número de estudios incluidos en la revisión es elevado, el carácter múltiple de muchas de las estrategias de implementación adoptadas y la multitud de aspectos estudiados hicieron imposible la presentación de los resultados sobre la relación coste-efectividad de estrategias de implementación alternativas de una forma inteligible. Este problema ya se preveía al inicio de este estudio y, por este motivo, se buscaron datos sobre el uso de recursos y los costes de desarrollo, difusión e implementación de directrices. Se creyó que estos datos serían más útiles para aquellas personas que estaban planificando la implementación de una pauta en su propio entorno.

La cantidad de datos utilizables sobre el uso y el coste de recursos fue comparativamente pequeña dado el número de estudios incluidos en la revisión. Sin embargo, estos datos siguen constituyendo una fuente valiosa para los investigadores, y han demostrado proceder de análisis con una calidad razonable.

En general, la calidad metodológica de los estudios incluidos era deficiente. Este hallazgo concuerda con otras revisiones de evaluaciones económicas²¹, y en parte era de esperar dada la amplia interpretación de lo que define una evaluación económica. En otra revisión de evaluaciones económicas finalizada recientemente, el criterio de inclusión estipuló que todo estudio, para ser incluido, debía presentar la metodología de determinación de costes en la sección de métodos, y los resultados de la evaluación económica en la sección de resultados²². Si se hubiera aplicado este criterio a esta revisión, sin duda habría reducido el número de estudios incluidos, pero es poco probable que hubiera cambiado la conclusión general sobre el predominio de una metodología deficiente.

En una revisión de estudios de implementación de directrices, quizás sorprende constatar que relativamente pocos

Tabla 21 a. Detalle de los estudios que comunicaron las consecuencias en los recursos de las estrategias de desarrollo, difusión e implementación de directrices

Estudio	Intervención	Comportamientos estudiados	Marco
Gurwitz [A87]	Distribución de material educativo, encuentros educativos, otros cambios conductuales profesionales (lista de pacientes que reciben el tratamiento seleccionado), intervenciones financieras (cambio a productos reembolsables)	Prescripción de antagonistas de los receptores H ² (tratamiento de alteraciones gastrointestinales) en pacientes ingresados en residencias de ancianos	Boston, Estados Unidos Residencia para enfermos crónicos Período de estudio: 1988-1990 Costes en dólares de Estados Unidos de 1989 No se especifica coste de obtener cambio financiero
Thomas [A208]	Distribución de material educativo, encuentros educativos, cambios organizativos (consultorio de acceso abierto)	Directrices para dos problemas urológicos: prostatismo por hiperplasia prostática benigna y hematuria microscópica en la población general	Grampian, Escocia Consultorio de medicina general Período de estudio: 1995-1996 Costes en libras esterlinas Año no indicado, posiblemente 1995/96 No se especifica coste de obtener la intervención de cambio organizativo
Rosser [A179]	Recordatorios sobre la necesidad de realizar un procedimiento/prueba	Vacuna contra la gripe Cribado cervicouterino Cribado para hipertensión Dosis de refuerzo de la vacuna antitetánica	Ontario, Canadá Consultorio de medicina general Período de estudio: 1984/85 No está claro si los costes están en dólares de Estados Unidos o de Canadá No se indica el año al que están referidos los costes
Wínickoff [A229]	Recordatorios, auditoría y retroalimentación, proceso de consenso local	Tratamiento de la hipertensión	Boston, Estados Unidos Atención primaria Período de estudio no claro, pero el estudio duró 18 meses Costes en dólares de Estados Unidos de 1979 Sólo se comunicaron los costes de la intervención de auditoría y retroalimentación

Los datos en **negrita** identifican la intervención sobre la que se facilitan datos sobre el uso de recursos

Tabla 21 b. Recursos utilizados en la fase de desarrollo de directrices

Estudio	Área de uso de recursos	Cantidad de recursos usados	Coste económico de los costes de desarrollo
Gurwitz [A87]	Revisión de la bibliografía por médico / farmacéutico	20 horas (también incluye preparación de la documentación que debería considerarse parte de los costes de difusión/implementación)	650 dólares (a 32,50 dólares/hora)
Thomas [A208]	Tiempo de médico/enfermera/investigador clínico dedicado a reuniones (incluyendo tiempo de desplazamiento)	217 horas (163 horas de tiempo libre; 54 horas de trabajo)	1944 libras esterlinas (sólo horas en tiempo de trabajo); 9029 libras con valoración de tiempo libre
	Preparación de reuniones de desarrollo por personal de investigación	182 horas	2676 libras esterlinas
	Costes de desplazamiento	No se especifica precio por kilómetro o coste unitario, pero se usó el coste por milla especificado por la Automobile Association británica en 1998	462 libras esterlinas
	Material fungible	No detallado	3329 libras esterlinas

Tabla 21 c. Recursos utilizados en la fase de difusión e implementación

Estudio	Evento generador de costes	Cantidad de recursos usados	Coste
Gurwitz [A87]	<i>Material educativo</i> Documentación preparada por médico / farmacéutico Revisión de historias clínicas para los documentos por médico / farmacéutico Costes de imprenta	Véase la Tabla 21b 25 horas Uso de recursos no documentado pero se necesita material para 16 profesionales	Véase la Tabla 21b 812,50 dólares 200 dólares
	<i>Encuentros educativos</i> Tiempo de preparación de la presentación por médico / farmacéutico Grupos de debate a los que asisten 16 miembros entre personal y profesionales médicos o farmacéuticos	5 horas Tiempo no indicado	32,50 dólares Coste no especificado
Thomas [A208]	<i>Material educativo</i> Material fungible (imprimir carpetas) Tiempo dedicado a organizar y enviar por correo las directrices Franqueo de directrices y cartas / recordatorios	No se especifican los recursos utilizados, pero se difundieron las directrices a unos 300 médicos de cabecera y a 74 gerencias de centros sanitarios 24 horas No se especifican los recursos utilizados, pero se difundieron las directrices a unos 300 médicos de cabecera y a 74 gerencias de centros sanitarios	2484 libras esterlinas 265 libras esterlinas 430,85 libras esterlinas

Continúa

Tabla 21 c. Recursos usados en la fase de difusión e implementación (Cont.).

Estudio	Evento generador de costes	Cantidad de recursos usados	Coste
	<i>Encuentros educativos</i> Tiempo de médico/enfermera/investigador clínico dedicado a las reuniones (incluyendo tiempo de desplazamiento y beca de educación de posgrado para médicos)	111 horas, todas ellas fuera del horario laboral. Este coste se repartió entre 74 consultorios y unos 300 médicos de cabecera	0 libras ya que no se saca de tiempo de trabajo, pero 7024 libras esterlinas si se valora el tiempo de ocio
	Preparación de la reunión del personal de investigación Costes de desplazamiento	40 horas No se especifica precio por kilómetro o coste unitario, pero se usó el coste por milla especificado por la Automobile Association británica en 1998	517 libras esterlinas 304 libras esterlinas
Rosser [A179] ^a	<i>Carta recordatorio enviada a pacientes</i> Tiempo de administrativa para preparar la carta Sellos Papel <i>Repetición de costes anteriores para 2ª carta</i> Recordatorio telefonico a pacientes Tiempo de administrativa para preparar la documentación del paciente Tiempo de enfermera para ponerse en contacto con el paciente Coste de las llamadas Repetición de llamadas a pacientes <i>Tiempo de médico para explicar la necesidad de la prueba/procedimiento</i>	1,7-5,77 minutos (según el distrito clínico) 10 segundos No especificado No especificado 84% de pacientes precisaron 2ª carta 0,33-2,8 minutos 2,8 minutos No especificado 1,7-2,4 llamadas por paciente 0,25-1,70 minutos	0,28-0,96 dólares 0,16 dólares 0,32 dólares 0,06 dólares 0,06-0,48 dólares 0,70 dólares No especificado No especificado 0,25-1,70 dólares
Winickoff [A229]	Tiempo de ordenador dedicado a crear informes a partir de datos sistemáticos de 2216 pacientes para 16 equipos de médico/enfermera Tiempo del personal para generar informes	90 horas 5-10 horas	2300 dólares 27,50-55,00 dólares

^aCantidades de uso de recursos y costes expresadas por paciente.

estudios consideraron costes que no fueran los del tratamiento y sus consecuencias. En varios casos, dada la magnitud del ahorro de costes con la práctica recomendada, sería de esperar que esta limitación no cambiara las conclusiones (A52, A101, A143, A189, A203). En estos casos, por tanto, dejando al margen otras deficiencias metodológicas, la evaluación habría cumplido la finalidad para la que se diseñó. Sin embargo, los resultados de estas evaluaciones son específicos para sus contextos y se caracterizan porque son poco transferibles. Por otro lado, en muchos casos en los que se concluyó que la implementación suponía un ahorro de costes, la inclusión de los costes de desarrollo, difusión e implementación de las directrices podría haber cambiado el signo de los resultados.

La estrategia de búsqueda que se utilizó se elaboró con la intención de identificar los estudios basados en diseños robustos. Los diseños de estudio se eligieron porque se consideró que, sometidos a un análisis adecuado, podían proporcionar los datos más robustos sobre la efectividad. Puesto que estos datos muchas veces ayudan a determinar los costes totales y forman parte de las estimaciones de la eficiencia, se entiende que estos estudios deberían aportar los mejores datos en los que basar al menos algunas partes de un análisis económico. Es posible que la estrategia de búsqueda no identificara algunos estudios que aportaban datos económicos por separado

o que utilizaban datos procedentes de diseños de estudio robustos para modelar la eficiencia.

CONCLUSIONES

Ha quedado demostrado en esta revisión la escasez de datos sobre el coste y la eficiencia de las estrategias de implementación de directrices. En general, los estudios tenían una deficiente calidad metodológica y no incluyeron todas las fases de la introducción de directrices que podrían ser relevantes. Aunque es tentador recomendar que se realicen más estudios a gran escala capaces de aportar estimaciones no sesgadas de los costes y los efectos de todas las fases de la implementación de directrices, en muchos casos es poco probable que estos estudios sean prácticos, debido a los problemas estadísticos relacionados con los tamaños de muestra²³. Más realista es el enfoque presentado por Mason y sus colaboradores²⁴; según este enfoque, los estudios primarios deberían concentrarse en evaluar el cambio conductual y los costes de desarrollo, difusión e implementación de las directrices, mientras que los ejercicios de modelado deberían utilizar estos datos para determinar si la directriz es eficiente al respecto del comportamiento obtenido o deseado.

Estimación de la factibilidad y de los recursos que son necesarios para las estrategias de difusión e implementación de directrices en el marco asistencial británico

MÉTODOS

Se llevó a cabo una encuesta telefónica semiestructurada de informantes clave en centros de asistencia primaria y de asistencia especializada para valorar la factibilidad y los recursos que serían necesarios para las estrategias de difusión e implementación de directrices en el Reino Unido. La primera intención había sido la de formar grupos focales con informantes clave; sin embargo, resultó imposible organizar los encuentros dentro del plazo del estudio.

Se seleccionaron informantes para representar una gama de puntos de vista aplicables a la implementación de directrices en centros de asistencia primaria, en hospitales generales de distrito y en hospitales universitarios. Se seleccionaron a partir de una muestra de conveniencia de médicos con experiencia en la implementación de directrices y que trabajaban mayormente en el norte de Inglaterra. Inicialmente se entró en contacto con ocho personas. Un médico (de cabecera) no contestó la encuesta. Otro no pudo efectuar la entrevista debido a otros compromisos, pero fue sustituido por un colega que cumplía los mismos criterios. Los siete informantes finalmente elegidos fueron dos médicos de cabecera a dedicación completa, uno de los cuales también era presidente de un grupo de asistencia primaria (MC1, MC2), un profesor de medicina general y médico de cabecera a dedicación parcial

(MC3), un médico especialista de un hospital general de distrito (EH1), un profesor de medicina y especialista de un hospital general de distrito (EH2), un subdirector médico de un hospital general de distrito (SDH1) y un subdirector médico/jefe de servicio clínico de un hospital universitario (JSH1).

Las entrevistas telefónicas semiestructuradas las realizó PW, y duraron entre 1 y 1,5 horas cada una. Las entrevistas, de las que también se sacaron apuntes, fueron grabadas y transcritas. Antes de que realizaran la entrevista se envió a los encuestados un resumen (véase el Anexo 7) que perfilaba las preguntas e incluía descripciones de las estrategias de difusión e implementación de directrices que se iban a considerar (basadas en las definiciones del grupo Cochrane EPOC; véase el Cuadro 1). El resumen también daba ejemplos del uso de recursos referido en la revisión sistemática de estrategias de difusión e implementación de directrices. Se preguntó a los encuestados su opinión sobre los factores que podían influir en el uso de recursos de cada estrategia de intervención, en qué medida pensaba el encuestado que las intervenciones eran factibles con los recursos disponibles en su propio entorno y su opinión sobre la factibilidad a escala nacional en el Reino Unido. También se les pidió que clasificaran las intervenciones por orden de intensidad en el uso de recursos. A medida que se fueron haciendo las entrevistas, se puso de manifiesto que habría que adoptar una definición amplia de

la expresión 'uso de recursos', y se exploró todo el espectro de problemas relacionado con la factibilidad de las intervenciones.

Las entrevistas se analizaron desde un punto de vista temático (a partir de las notas que se habían tomado). Los temas se respaldan con un número reducido de citas relacionadas que se tomaron de las entrevistas. Se envió a los encuestados el borrador del capítulo para que hicieran los cambios que consideraran pertinentes en sus aportaciones y como validación de los resultados expuestos. La presentación de resultados se estructura en torno a las seis extensas intervenciones de difusión/implementación consideradas.

RESULTADOS

Distribución de material educativo

Se pidió a los encuestados que consideraran la distribución de tres tipos distintos de material educativo:

- envió a los médicos pertinentes de una directriz escrita en una sola hoja plastificada con una carta de presentación
- envió de las directrices por correo a los miembros de las asociaciones de profesionales especialistas a los que afectan y a asociaciones de consumidores; publicación en revistas médicas nacionales y en boletines especializados
- envió de un vídeo educativo a centros y a especialistas pertinentes.

Los médicos de cabecera encuestados no preveían problemas de recursos para la distribución de hojas de tamaño DIN A4 plastificadas o para las copias del texto completo de las directrices (aunque todos los encuestados comentaron que no veían adecuada la distribución del texto completo). Todos los encuestados veían factible la distribución de vídeos si se podía contar con el patrocinio de un laboratorio farmacéutico. Excepto un médico de cabecera (MC2, que tenía experiencia en la filmación de un vídeo que había sido financiado por una beca de investigación), los demás no veían posible (no tenían los conocimientos necesarios) ni asequible hacer ellos mismos un vídeo.

Al menos dos de los cuatro encuestados de asistencia hospitalaria consideraron que todas las estrategias necesitarían financiación privada (un laboratorio farmacéutico o una organización benéfica especializada en temas de salud).

"Los hospitales generales de distrito no tienen presupuesto para esa clase de cosas" (EH1).

Un encuestado de asistencia especializada (EH2) tenía experiencia en la producción de un vídeo de ámbito regional, financiado por fondos de investigación y con la colaboración de una universidad cercana, pero reconoció que la mayoría de los especialistas no tendrían esas facilidades.

El principal recurso identificado por los encuestados fue el de los costes de franqueo si se enviaba por correo externo. Los encuestados identificaron una serie de factores que también podrían incidir en la necesidad de recursos (Cuadro 4).

Cuadro 4. Factores que influyen en la necesidad de recursos para la distribución de material educativo

Material impreso

- Número de páginas (p.ej., texto completo de las directrices frente a una sola hoja DIN A4 escrita por las dos caras)
- Número de copias
- Frecuencia de revisión de las directrices; por lo tanto, frecuencia con que tienen que distribuirse
- Costes de personal para fotocopiar, montar y ensobrar
- Costes de oportunidad del tiempo de lectura
- Disponibilidad y coste de una máquina plastificadora (unas 50 libras esterlinas) para enviar hojas plastificadas escritas por las dos caras
- Disponibilidad y coste de programas informáticos para producir material profesional y atractivo, de una sola hoja; y de las habilidades necesarias para usar los programas

Compra de vídeos educativos

- Coste (una media de unas 12 libras esterlinas)
- Cantidad necesaria (distribución individual o a grupos)
- Disponibilidad de instalaciones audiovisuales para ver el vídeo
- Costes de oportunidad del visionado en grupo

Desarrollo local de vídeos educativos

- Conocimientos clínicos
- Competencia en la redacción de guiones
- Competencia para hacer la filmación
- Competencia para el montaje o edición técnica

Encuentros educativos

Se pidió a los encuestados que consideraran cuatro tipos distintos de encuentro educativo:

- una serie de charlas para el departamento médico de 1 hora de duración
- una conferencia de un 'experto' local de media jornada para médicos
- varias sesiones educativas intensivas (más de 2 horas) para grupos reducidos de médicos
- sesión de dos horas para médicos (que pueden elegir entre las celebradas en distintos lugares) con un experto local y con colegas que trabajan en el desarrollo de las directrices de interés

Un tema clave para los encuestados de asistencia primaria era la hora y la duración de la reunión. Sugirieron que las reuniones durante la hora de almuerzo o por la tarde serían mucho más económicas, ya que no sería necesario pagar a sustitutos de médicos asistentes; las reuniones programadas a otras horas o de mayor duración (media jornada o más) sí que supondrían ese gasto. Cuando es necesario pagar a un sustituto, el número de médicos de cabecera que asisten al evento es el determinante clave de los recursos que se necesitan. MC2 lo ilustró con una estimación de que un solo evento de media jornada para 38 médicos de cabecera costaría 4.000 libras a su grupo de atención primaria con cargo a un presupuesto anual para la for-

mación de médicos de cabecera de 20.000 libras. Un encuestado comentó que el método tradicional de charlas didácticas de 1 hora impartidas por un experto local muchas veces se consideraba relativamente barato, pero “esto sólo es cierto en cuanto al conferenciante, no en cuanto al público. Los ‘costes de oportunidad’ de la asistencia de una gran cantidad de médicos de cabecera a una conferencia (en horario laboral) son enormes” (MC1).

La hora del día, la duración de la reunión y el tiempo disponible también son temas importantes para los especialistas. Los encuestados opinaban que si las reuniones eran durante la hora de comer, era más fácil que pudieran asistir más médicos; en cambio, si la reunión era de media jornada era más difícil que se pudieran liberar, ya que podría suponer la cancelación de visitas ambulatorias o de intervenciones quirúrgicas. Según el servicio al que afecte, también puede haber problemas para cubrir los servicios de urgencias. Cuando las directrices afectan potencialmente a un número elevado de personas, con distintas profesiones y especialidades (por ejemplo, prevención y tratamiento de trombosis venosa profunda), estos problemas se ven agravados por la necesidad de organizar varias sesiones para acoger distintos turnos. Un encuestado observaba: “no disponemos de tiempo reservado para nosotros... Y no existe algo como una media jornada para asistir a un curso. La única forma de organizar sesiones de cierta duración es cancelar visitas” (EH1).

En todos los marcos asistenciales podría ser problemático encontrar a ponentes expertos, con las habilidades docentes y de presentación adecuadas, o formadores preparados. MC2 observó que en su grupo de asistencia primaria sólo había tres médicos con habilidades para realizar esa tarea. En el caso de programas educativos locales, la disponibilidad de formadores locales y el coste de prepararlos pueden suponer un problema.

Una diferencia clave para los especialistas era el número relativamente pequeño de personas que participarían en un evento educativo local centrado en su especialidad, por lo que opinaban que podría ser más adecuado que fuera de ámbito regional o nacional. Sin embargo, los encuestados observaron que los médicos especialistas solían trabajar en equipos multidisciplinarios, por lo que podrían esperar celebrar encuentros educativos regulares dentro del propio centro.

En resumen, se consideraron factibles charlas de 1 hora (en una sola reunión didáctica o en una serie de ellas) tanto en el marco de la atención primaria como en el de la especializada. En el caso de la asistencia primaria, otro tipo de actividad educativa incurriría en gastos adicionales, sobre todo debido al coste de los sustitutos.

MC3 opinaba que todas estas actividades eran factibles si se utilizaban los presupuestos para formación. Uno de los encuestados, que trabajaba tanto en la atención primaria como en la especializada (SDH1), opinaba que era más fácil disponer de recursos para estas actividades en la atención primaria: “bajo epígrafes como gobierno clínico, formación permanente, etc. En cambio, resulta mucho más difícil en la asistencia especializada, sobre todo convencer a los médicos de que prescindan de media jornada de trabajo.”

Los encuestados identificaron una serie de factores, comunes a la asistencia primaria y a la especializada, que influían en los recursos que se necesitaban (Cuadro 5).

Cuadro 5. Factores que influyen en los recursos que se necesitan para los encuentros educativos

- Situación: interna (en el propio centro) frente a externa
- Distancia del lugar de trabajo de los participantes; coste del local y del *catering* (en el caso de instalaciones externas)
- Frecuencia de los encuentros
- Duración de los encuentros
- Honorarios del formador del grupo
- Disponibilidad local de formadores de grupo experimentados
- Costes de administración

Visitas educativas

Se pidió a los encuestados que consideraran tres tipos distintos de visitas educativas:

- visitas individuales hechas por farmacéuticos (de la comunidad o del hospital, según proceda) formados como educadores (> 2 días de formación). No más de dos visitas breves (de menos de media hora) a los médicos pertinentes.
- visitas a equipos clínicos o de consultorios hechas por un/a enfermero/a formado como educador (>2 días de formación). Se hacen varias reuniones de por lo menos una hora.
- un médico formado como educador (>2 días de formación). El médico expone las directrices y los mensajes educativos específicos en presentaciones de 1 hora a grupos de médicos o a servicios. Cuando sean en hospitales, el médico formador puede acompañar al grupo en la visita a la planta a modo de complemento de las reuniones.

Los encuestados consideraron que las visitas educativas probablemente serían una estrategia cara; pensaban que un aspecto clave era si el tema se consideraba prioritario, en cuyo caso se podría disponer de otros recursos.

Los médicos de atención primaria encuestados reconocieron que las visitas educativas eran una estrategia menos cara para el personal de asistencia primaria, ya que se podían realizar en el propio consultorio. Algunos sugirieron que los nuevos farmacéuticos de enlace de los grupos de atención primaria podrían realizar visitas educativas como parte de su función, aunque un encuestado (SDH1) ponía en duda la factibilidad de esta sugerencia si sólo podían visitar muy de vez en cuando los consultorios de su zona. Un encuestado manifestó que su grupo de atención primaria local pensaba que sería necesario crear más plazas de farmacéutico para llevar a cabo este tipo de trabajo: se había estimado que harían falta hasta tres plazas, lo que supondría un coste aproximado de 40.000 libras por plaza y año. Les parecía mucho menos factible que el formador fuera un médico, ya que podría suponer un coste de hasta 60.000 libras al año (a dedicación completa). El profesor universitario encuestado (MC3), que había investigado la relación coste-efectividad de las visitas educativas realizadas por médicos, refirió un coste de 583 libras por consultorio en 1995, para implementar un conjunto completo de pautas mediante formadores

médicos que cobraban el sueldo de un profesor de medicina y teniendo en cuenta los costes de una beca de posgrado.

Había una discrepancia significativa en la opinión de los encuestados respecto a si debían ocuparse de estas funciones médicos o enfermeras. Dos encuestados de atención hospitalaria (SDH1, JSH1) consideraban que las enfermeras de sala, en particular, eran demasiado valiosas para que se dedicaran a estas actividades, sobre todo teniendo en cuenta los problemas para encontrar personal para muchas especialidades de enfermería hospitalaria aguda. SDH1 también cuestionó la factibilidad de la estrategia señalando que el personal de enfermería, para esta función, podía tener problemas de credibilidad para algunos médicos. No obstante, otros encuestados (EH1, EH2) señalaron que esta función generalmente la desempeñan enfermeras especialistas (por ejemplo, en diabetes, atención renal, enlace respiratorio, accidente vascular cerebral...), que generalmente son respetadas por los médicos. SDH1 consideraba una mera fantasía la posibilidad de asignar médicos a esta función, mientras que otros encuestados de atención hospitalaria consideraban más factible asignar a médicos residentes que a enfermeras.

Entre los médicos de asistencia primaria encuestados parecía predominar la opinión de que, en general, no se dispone de recursos para hacer visitas educativas, pero que se podría acceder a ellos a través de fondos para proyectos específicos o a través de fondos de las autoridades sanitarias (siempre que la idea resultara elegida en un concurso público). En esta área posiblemente había más recursos en atención especializada, al menos en cuanto a los farmacéuticos. Esto podría constituir un ejemplo de la reasignación de los recursos existentes ya que, EH1, por ejemplo, comentaba que no se les había ocurrido que este papel lo pudieran llevar a cabo farmacéuticos clínicos.

Uno de los encuestados de atención especializada (EH2) opinaba que, para esta estrategia y la siguiente (uso de líderes de opinión), los recursos no sólo estaban disponibles sino que se usaban. EH2 refirió que en el hospital en el que él tra-

bajaba, en cada especialidad había médicos adjuntos especialmente preparados que participaban en la formación de sus colegas. Sin embargo, EH2 reconoció que puede que no fuera posible hacer extensivo este modelo a otros hospitales: "... para que funcione este modelo de trabajo, hay que ser capaz de aceptar los consejos de otros colegas."

SDH1 consideraba poco probable que las visitas educativas fueran una estrategia económicamente factible: "El único uso factible de las visitas educativas sería para los 'rezagados' que no han asistido a los otros eventos educativos que se han organizado."

El Cuadro 6 enumera los factores que influyen en los recursos que se necesitan para las visitas educativas.

Líderes de opinión locales

Se pidió a los encuestados que consideraran dos posibilidades, y para ambas debían nombrar colegas que consideraran líderes de opinión:

- a través de un número reducido de líderes de opinión que realizan un breve cursillo de formación (< 1 día). Entre las actividades que realizan se incluyen escribir cartas de presentación de unas determinadas directrices para otros médicos, presidir un encuentro educativo o potenciar sus actuales contactos divulgativos con otros colegas
- a través de un gran número de líderes de opinión que realizan un curso de formación de mayor duración (mínimo de 2 días). Entre las actividades que realizan se pueden incluir formar y dirigir equipos de trabajo, con actividades divulgativas y programas de visitas educativas.

Los encuestados pensaban que podía haber un 'efecto umbral', y éste dependería del número de líderes de opinión que se necesitaran y de la cantidad de tiempo que tuvieran que dedicar a la estrategia: "por definición, los líderes de opinión siempre son pocos" (MC1). Los médicos de atención primaria encuestados observaron que una sola persona puede ser considerada un líder de opinión en varias áreas clínicas; lo que tendría un impacto significativo en el tiempo que esta persona tendría que dedicar (MC2). Esto también sucede en el marco de la asistencia especializada, donde puede haber un solo 'líder especialista' para un tema (por ejemplo, diabetes), sobre todo en los hospitales generales. Sólo es posible que se identifiquen varios líderes de opinión en un ámbito regional y no en uno local. El tiempo que habría que dedicar a esa actividad podría determinar si habría que remunerar a los líderes de opinión. Un encuestado opinaba que ser un líder de opinión para otros médicos forma parte de ser un 'líder especialista' (EH2); y uno de los médicos de cabecera encuestados (MC3) lo concebía como una actividad informal que no habría que remunerar. Sin embargo, la necesidad de dedicar una cantidad importante de tiempo y la extensión de su actividad más allá de los colegas y de la zona locales, en opinión de ambos encuestados, aumentaría las probabilidades de que fuera necesario remunerar el tiempo dedicado.

Aunque la primera posibilidad se vio como más factible que la segunda al implicar a pocas personas, los encuestados

Cuadro 6. Factores que influyen en los recursos que se necesitan para las visitas educativas

- Lugar: en la mayoría de los casos, se esperaría que las visitas se hicieran en el propio centro de los médicos.
- Frecuencia de las visitas
- Duración de las visitas (con implicación de tiempo tanto para los que las imparten como para los que las reciben)
- Visitas individuales o dirigidas a grupos
- Número de personas que realizan las visitas
- Profesión de las personas que realizan las visitas (disponibilidad de personal y costes salariales)
- Coste de formación de las personas que realizan las visitas
- Coste de regalos similares a los que dan los representantes de las compañías farmacéuticas (almuerzos, bolígrafos, blocs de notas, etc.)
- Tiempo del personal administrativo para organizar las citas

de atención primaria parecían considerar el uso de líderes de opinión como una estrategia menos factible en general. Un problema específico era la escasez de líderes de opinión fáciles de identificar en asistencia primaria.

EH1 expuso el problema que tendría la aplicación de estas estrategias en su marco: cuando un médico tiene un interés especial, hay que traer un líder de opinión de otro lugar. En cambio, cuando ese mismo médico (por ejemplo, un diabetólogo) actúa de líder de opinión para médicos generalistas, su disponibilidad de tiempo para preparar e impartir actividades divulgativas se convierte en un problema, y muchas veces se realizará fuera de horas de trabajo. EH2 opinó de otro modo: para él, todos los 'líderes especialistas' son líderes de opinión de sus colegas, y todos los médicos son líderes de opinión los unos de los otros. Pero esto sólo es posible cuando existe "un ambiente abierto para pedir y recibir consejos".

Hubo opiniones dispares sobre la identificación y la formación de líderes de opinión. JSH1 comentó que "sería bueno tener una forma sistemática para identificar y usar líderes de opinión, pero es un tema que no tiene la prioridad suficiente". SDH1 sugirió que en la atención primaria se podrían identificar a muchos líderes de opinión si se tomaran los responsables del gobierno clínico en los consultorios, pero luego se preguntó si realmente podrían llegar a ser líderes de opinión. MC3 indicó que, aunque pudiera ser posible identificar a un mayor número de líderes de opinión en un ámbito regional, sería difícil justificar el tiempo necesario para formarlos: "Por definición, se trata de personas muy ocupadas, con poco tiempo disponible". EH2, en contraste, sugirió que no era necesario impartir formación para actividades locales como líderes de opinión: "Los líderes de opinión, ¿no son por naturaleza buenos comunicadores? Si no lo fueran, ¿cómo iban a ser considerados como influyentes por sus colegas?".

Los factores que influyen en los recursos que se necesitan para los líderes de opinión locales se resumen en el Cuadro 7.

Auditoría y retroalimentación

Si pidió a los encuestados que consideraran cuatro posibilidades:

- mensajes de correo electrónico regulares y frecuentes (como mínimo, una vez por semana) enviados a médicos, con informes generados por ordenador de su grado de cumplimiento de las directrices durante las últimas semanas

Cuadro 7. Factores que influyen en los recursos que se necesitan para los líderes de opinión locales

- Número de líderes de opinión que se necesitan
- Costes de selección y formación de los líderes de opinión
- Remuneración a los líderes de opinión del tiempo que hayan dedicado
- Costes de oportunidad del tiempo de los líderes de opinión
- Dedicación del líder de opinión (es decir, tiempo de preparación fuera de horas de trabajo)
- Coste de estrategias de divulgación usadas por los líderes de opinión (encuentros educativos, etc.)

- informes mensuales impresos enviados a los médicos, con datos sobre su grado de cumplimiento de las directrices y una comparación con el de otros colegas anónimos. Los datos se generarían a partir de historias clínicas electrónicas
- seminarios mensuales en los que se entrega a cada médico un informe impreso con su grado de cumplimiento personal de las directrices, una comparación de su cumplimiento con el de colegas anónimos y un comentario personalizado de un experto local. Los datos se obtendrían a partir de auditorías (manuales) de historias clínicas
- reuniones trimestrales con el equipo en las que se entrega un informe sobre su cumplimiento de las directrices, con los comentarios de un experto externo. Los datos se obtendrían a partir de auditorías (manuales) de historias clínicas.

Los principales factores que identificaron que influirían en los recursos que se necesitarían estaban relacionados con los métodos de recogida de información (concretamente, de si los datos debían obtenerse de historias clínicas hechas a mano o informáticas). Todos estuvieron de acuerdo en que la extracción de datos a partir de historias clínicas hechas a mano necesitaba recursos muy intensos en términos de dedicación de personal y de tiempo. El tratamiento posterior de los datos y la elaboración de los comentarios también son muy laboriosos, y es posible que el personal que haya de realizar esas tareas necesite una formación específica. Los encuestados opinaron que podía haber personal que hiciera todo o una parte de estas actividades sin que hubiera que pagarles más (por ejemplo, en atención especializada, médicos residentes como parte de sus proyectos de auditoría obligatorios), pero pudiera ser que estas personas no fueran las más indicadas para llevar a cabo estas actividades, sobre todo sin una formación complementaria (EH2).

Cuando los datos se tienen que obtener de un sistema informatizado, el factor clave es si ya existe un sistema capaz de cumplir todos o algunos de los requisitos. Si el sistema existe, proporcionar los datos de auditoría supone muy pocos costes adicionales. Si no existe, los costes de comprar e instalar un sistema nuevo son considerables. Los encuestados pensaban que en ese momento existían muy pocos sistemas informáticos capaces de aportar una retroalimentación fiable y con información útil. Los encuestados de atención especializada opinaron que la obtención manual de datos era en ese momento el único método posible de generar datos de auditoría.

Los médicos de atención primaria encuestados, aunque reconocían una mayor disponibilidad de sistemas informáticos, destacaron las limitaciones de esos sistemas para hacer funciones de auditoría y retroalimentación. SDH1 observó que "aunque los médicos de cabecera generalmente tienen acceso a un ordenador, no suelen tener sistemas para generar informes automatizados". MC1 y MC2 también expresaron su frustración ante la falta de sistemas con capacidad de generar de forma sistemática informes de auditoría clínica. MC3 identificó algunas excepciones en asistencia primaria basadas en

MIQUEST (que permite la consulta y la extracción de datos de distintos tipos de sistemas de consultorios de medicina general, a través de un lenguaje de consulta común)²⁵, pero matizó que las auditorías de zona con software MIQUEST eran poco comunes y frecuentemente financiadas con fondos privados.

Otros factores relacionados con los recursos que se podrían necesitar fueron la frecuencia de los informes (sobre todo en el caso de las auditorías basadas en historias clínicas hechas a mano y cuando la retroalimentación se realiza a través de una reunión en grupo), y si la retroalimentación se realiza a través de informes que se envían individualmente a los médicos o si precisa una reunión en grupo. Existen los 'costes de oportunidad' relacionados con el tiempo que hay que dedicar para organizar la reunión. Cuando se espera que un experto comente la retroalimentación, puede ser difícil mantener su dedicación de tiempo y entusiasmo si debe participar en varias reuniones (EH2).

Los factores que influyen en los recursos que se necesitarán para la auditoría y la retroalimentación se resumen en el Cuadro 8.

Cuadro 8. Factores que influyen en los recursos que se necesitarán para la auditoría y la retroalimentación

- Costes de la obtención de datos
- Costes de preparación y tratamiento de datos
- Costes de elaborar informes de retroalimentación
- Costes de hacer llegar los informes
- Frecuencia de los informes
- Otras actividades asociadas, por ejemplo, encuentros educativos

Recordatorios

Se pidió a los encuestados que consideraran tres posibilidades:

- Pegatinas (con espacios para escribir acciones apropiadas de gestión clínica) que el personal administrativo pega en las historias clínicas. Marcadores de colores vivos que el personal administrativo puede pegar en las historias y que indicarían que los pacientes cumplen determinados criterios que hacen que sea indicada una intervención específica.
- informes generados por ordenador enviados anualmente a los médicos a través de un sistema administrativo centralizado; en ellos se detallarían las intervenciones o los procedimientos realizados durante el año anterior y los que aún están pendientes. Se dejaría espacio para que los médicos completaran la información sobre los procedimientos llevados a cabo durante el año; una vez completado, el informe debería ser devuelto.
- en el contexto de un sistema de seguimiento informatizado y/o de historia clínica informatizada, 'alertas' y 'mensajes' generados por ordenador enviados a los médicos y

derivados de las directrices. Las 'alertas' se enviarían al médico cuando uno de sus pacientes sufriera un determinado acontecimiento; el 'mensaje' le llegaría al médico en el momento de abrir la historia clínica del paciente, y representaría que hay una indicación para llevar a cabo una determinada acción terapéutica.

Los encuestados plantearon algunos problemas relacionados con los recursos que serían necesarios para poner pegatinas sobre las notas; concretamente, la identificación de los pacientes correspondientes y el tiempo del personal administrativo. También plantearon problemas relacionados con la factibilidad de esta estrategia, sobre todo si se utilizaban pegatinas diferentes: "Te olvidas de para qué son o empiezas a ignorarlas" (MC1). MC3 manifestó que cada consultorio de medicina general puede absorber un máximo de dos directrices al año, en términos de la formación necesaria y del número de recordatorios generados, lo que tiene implicaciones claras para la implantación generalizada de estas estrategias de recordatorios.

Los encuestados de atención especializada dijeron que en su entorno no se habían explorado a fondo las posibilidades que ofrecían las pegatinas sobre las historias clínicas hechas a mano. EH1 comentó que "no sería posible dedicar personal de archivos a esta actividad en nuestro hospital. Podría ser posible llevarla a cabo en algunos servicios clínicos para un número reducido de personas en áreas clínicas seleccionadas". EH2 también destacó que los protocolos asistenciales usan recordatorios con frecuencia en forma de historias clínicas estructuradas y hechas a mano que contienen mensajes recordatorios y casillas de verificación.

Los encuestados plantearon problemas respecto al uso y la funcionalidad de los sistemas de información respecto a los recordatorios similares a los planteados acerca de la auditoría y la retroalimentación. Los encuestados de atención primaria opinaron que generar informes con recordatorios por ordenador, tanto anuales como de menor frecuencia, no precisaría necesariamente la asignación de recursos adicionales, ya que ya existen sistemas para situaciones determinadas (registros de diabéticos, cribado de prevención del cáncer de cuello uterino, vacunaciones...). Estos informes podrían generarse a partir de historias clínicas manuales guardadas en archivos centrales. Sin embargo, MC3 sugirió que no se debía subestimar el tiempo que necesitarían los facultativos para considerar y cumplimentar las listas generadas por estos sistemas, aunque el personal administrativo puede valorar positivamente la ayuda que aportan las listas generadas externamente. Un encuestado de asistencia especializada (EH2) destacó la complejidad adicional que supone un mensaje recordatorio en un informe de tipo resumen, en comparación con una pregunta de 'sí/no' sobre si se ha realizado una intervención determinada.

Todos los encuestados estuvieron de acuerdo en que los sistemas más sofisticados, descritos en la tercera posibilidad, supondrían una gran inversión en ordenadores y en programas informáticos. Además, se necesitarían recursos para formar al personal en la entrada de datos y el uso del sistema. Varios encuestados sugirieron que haría falta un cambio cul-

Cuadro 9. Factores que influyen en los recursos que se necesitarán para los recordatorios*Factores para las pegatinas*

- Tiempo de personal para identificar a los pacientes que cumplen los criterios establecidos y para pegar las pegatinas en las historias clínicas
- Disponibilidad de personal
- Formación que necesita el personal administrativo para identificar a los pacientes [cuanto más complejo el problema clínico, más formación se necesita (MC2)]

tural importante para que fuera viable el uso habitual de un sistema de este tipo (MC3, SDH1, JSH1). Sin embargo, una vez implantados estos sistemas, los recursos que necesitaría esta estrategia de recordatorios serían bastante pocos. A los especialistas encuestados les preocupaba el efecto de estos sistemas sobre el tiempo del médico (concretamente, la diferencia entre el tiempo que necesita el médico para introducir los datos en el ordenador y el que necesita para tomar notas a mano o dictarlas) (EH1, EH2).

Los factores que influyen en los recursos que se necesitan para los recordatorios (pegatinas) se resumen en el Cuadro 9.

USO RELATIVO DE RECURSOS PARA LAS DISTINTAS ESTRATEGIAS

Se pidió a los encuestados que clasificaran las distintas estrategias por orden de la importancia relativa de los recursos que necesitarían (Tabla 22). Se consideró que la que necesitaba más recursos era la de las visitas educativas; la que menos, la distribución de material educativo. En cuanto a las demás intervenciones, hubo una mayor variación en las clasificacio-

nes; esto en parte se explica porque la necesidad de más o menos recursos puede variar con la aplicación individual de la estrategia, así como por las distintas maneras de evaluar los costes de recursos informáticos.

DISCUSIÓN

Los encuestados consideraron que la distribución de material educativo era la estrategia que necesitaba menor intensidad de recursos, y las visitas educativas la que necesitaba mayor intensidad. Consideraron que la disponibilidad de sistemas informáticos adecuados era un aspecto básico para que pudieran aplicarse ampliamente las estrategias de 'auditoría y retroalimentación' y la de 'recordatorios'. Este aspecto era especialmente problemático en la asistencia especializada en hospitales: ninguno de los encuestados trabajaba en un centro con sistemas informáticos capaces de generar recordatorios o de proporcionar una retroalimentación de forma sistemática. Incluso con los sistemas más desarrollados de atención primaria, ninguno parecía tener capacidad para ejecutar las estrategias de recordatorios más sofisticadas.

A los encuestados (especialmente a los médicos especialistas), les resultó difícil identificar recursos ordinarios disponibles para la financiación de estrategias de implementación de directrices, pero reconocieron que era más probable que se dispusiera de recursos para prioridades de ámbito nacional. También parecía haber diferencias en la gestión de historias clínicas de algunas enfermedades crónicas, como la diabetes, ya que en muchas ocasiones estaban informatizadas y, por tanto, podían permitir la implementación mediante las estrategias basadas en recursos informáticos.

En algunos casos, los problemas prácticos de factibilidad pueden tener más importancia que los financieros; por ejemplo, un encuentro educativo entre médicos procedentes de una zona geográfica extensa supone la contratación de sustitutos, lo que puede significar un gasto importante, pero tam-

Tabla 22. Clasificación de las intervenciones de implementación de directrices en función de la intensidad de recursos que necesitan según los encuestados (1 = más intensivo en recursos; 6 = menos intensivo en recursos)

Encuestado	Distribución de material educativo	Encuentros educativos	Visitas educativas	Líderes de opinión locales	Auditoría y retroalimentación	Recordatorios
MC1	4	3	1	2	5	6
MC2	6	5	2	4	1	3 ^a
MC3	6	3	1	2	5	4
EH1	6	4	1	2	3	5
EH2	6	1	3	4	2	5
SDH1	6	5	1	4 ^b	3 ^c	1 ^d
JSH1	6	4	3	2	1	5 ^d

^a Se clasificaría '1' si se utilizaran historias clínicas hechas a mano para la extracción de datos.

^b Habría clasificado '1' para la segunda posibilidad.

^c Las posibilidades 1 y 2 de 'auditoría y retroalimentación' y las posibilidades 2 y 3 de 'recordatorios' tendrían el coste más alto debido a la necesidad de recursos informáticos.

^d Excepto la tercera posibilidad, que se clasificaría como 1.

bién puede plantear un problema de continuidad asistencial de los pacientes. En la asistencia especializada, un problema clave de factibilidad es el coste de oportunidad de cancelar visitas ambulatorias o intervenciones quirúrgicas para participar en actividades en grupo que no se pueden realizar en la hora del almuerzo.

En términos generales, parece haber significativamente más recursos para todas las estrategias en la atención primaria. En la asistencia especializada, incluso la distribución de material educativo puede precisar un grado importante de patrocinio. Sin embargo, parecía haber algunas diferencias culturales entre distintos marcos de atención especializada; estas diferencias indicaban que las estrategias pueden ser más o menos factibles en función del entusiasmo de los médicos líderes y del grado de receptividad de sus colegas.

Estos resultados se ven claramente limitados por el tamaño de la muestra, el análisis realizado hasta la fecha y la naturaleza de la selección de la muestra de conveniencia. Como relativos entusiastas de las estrategias que se planteaban, es posible que subestimaran las dificultades inherentes a su implementación. Por motivos pragmáticos, no estuvieron representados en la encuesta otros servicios ambulatorios o de salud mental, por lo que no se puede hacer ningún comentario respecto al grado de extensibilidad a estos marcos de cualquiera de los problemas planteados.

Al concluir sus entrevistas, no parecía que los médicos de atención primaria encuestados hubieran identificado ningún tema nuevo. En la asistencia especializada, las diferencias que se encontraron en entorno y cultura (entre especialidades, entre diferentes categorías de hospitales, etc.) sugiere que sería necesaria una muestra de mayor tamaño para obtener una imagen completa a partir de estos distintos marcos. Sin embargo, los temas planteados por los especialistas encuestados siguen teniendo implicaciones importantes para el gobierno clínico en muchos marcos hospitalarios.

Los resultados de esta encuesta deben situarse en el contexto del desarrollo de la estrategia de información del

Servicio Nacional de la Salud británico, sobre todo a raíz de la publicación de *Information for Health* en 1998²⁶. Esta publicación fijó el objetivo de informatizar las historias clínicas, tanto en la atención primaria como en la hospitalaria, antes de 2005. Los médicos encuestados en este trabajo tenían opiniones dispares sobre si estos objetivos eran alcanzables y sobre si, suponiendo que se alcanzaran, los sistemas resultantes podrían ser soportes de las estrategias informáticas de recordatorios planeadas aquí. Los médicos de cabecera con experiencia o conocimientos del software MIQUEST para la extracción de datos de los sistemas de asistencia primaria²⁵ o del uso del sistema de recordatorios por ordenador PRODIGY²⁷ fueron los más optimistas respecto a su potencial. Dos de los encuestados de asistencia especializada también pensaban que, con los planes de sus propios centros sanitarios para informatizar las historias clínicas de los pacientes, se podrían llevar a cabo estrategias informáticas de recordatorios. Las opiniones más pesimistas en atención primaria preveían un fracaso mientras la estrategia nacional siguiera fomentando sistemas mixtos de información en lugar de un sistema único para todo el país. En atención especializada había dudas sobre la factibilidad de la informatización de las historias clínicas sin que hubiera un cambio cultural amplio, se diera una formación y, sobre todo, hubiera más personal médico.

Una de las características notables de las entrevistas fue la creatividad con la que los encuestados habían buscado fuentes no habituales de financiación para las iniciativas de implementación. Algunos ejemplos son: tratar de conseguir fondos de la Health Authority, de presupuestos de educación, de los de investigación y desarrollo (I+D)... Pero es posible que esto no sea lo habitual, y puede que la situación en cuanto a financiación sea incluso peor que la descrita aquí. A pesar de ello, era evidente el carácter extraordinario de la mayor parte de los fondos asignados para circunstancias específicas, y varios encuestados mencionaron la necesidad de aumentar la capacidad global del sistema para permitir una expansión progresiva de la formación continuada en la práctica médica habitual.

PRINCIPALES RESULTADOS DEL ESTUDIO

Revisión sistemática de estrategias de difusión e implementación de directrices

En conjunto, la mayoría de las comparaciones que aportaban datos dicotómicos del proceso (86,6%) observaron mejoras en la asistencia, lo que sugiere que la difusión y la implementación de directrices puede fomentar el cumplimiento de las prácticas recomendadas. Sin embargo, aunque la mayoría de las intervenciones obtuvieron mejoras de la asistencia entre modestas y moderadas, hubo una variación considerable en los efectos observados tanto entre diferentes intervenciones como para una misma intervención. Las personas encargadas de hacer la revisión examinaron la reproducibilidad de la orientación y el tamaño del efecto de los resultados entre distintos diseños de estudio, así como entre los estudios en los que pudo determinarse la significación estadística y los estudios en los que no pudo determinarse. Por lo tanto, las conclusiones son provisionales, y deben explorarse en futuras evaluaciones robustas y bien diseñadas.

La intervención que se ha evaluado con más frecuencia es la de los recordatorios. Los resultados sugieren que los recordatorios constituyen una intervención potencialmente eficaz y tienden a producir mejoras moderadas del proceso asistencial. La mayoría de las comparaciones evaluaron los recordatorios en un amplio espectro de marcos asistenciales y de comportamientos.

La intervención a través de visitas educativas fue la segunda más frecuente. En muchas ocasiones formaba parte de una intervención múltiple, y podrían considerarse intrínsecamente polifacéticas. En la mayoría de los estudios se obtuvieron efectos modestos. La combinación de material educativo y visitas educativas parecía ser relativamente ineficaz. Los resultados sugieren que las visitas educativas pueden producir mejoras modestas del proceso asistencial; pero a ello habría que contraponer tanto los recursos que resultan necesarios para conseguir este cambio, como algunas consideraciones de índole práctica.

Las pruebas acerca de la efectividad de material educativo, auditoría y retroalimentación, y de intervenciones dirigidas al paciente fueron menos robustas, ya que hubo menos evalua-

ciones de estos tipos de intervención. El material educativo y la auditoría y la retroalimentación parecían tener efectos modestos, mientras que las intervenciones dirigidas al paciente parecían producir efectos moderados. No obstante, si pudiera lograrse en la práctica habitual el efecto mediano observado para el material educativo como intervención única, representaría un efecto importante (especialmente teniendo en cuenta la factibilidad y el coste relativamente bajo de la difusión de este tipo de material; véase abajo). Sin embargo, la adición de material educativo a otras intervenciones no parecía aumentar la efectividad de éstas.

Aunque la mayoría de las comparaciones evaluaron intervenciones múltiples, hubo algunas replicaciones de intervenciones múltiples específicas frente a un control sin intervención o frente a un grupo de control específico. Este hecho suscitó problemas importantes a la hora de interpretar y analizar las comparaciones. El proyecto original estipulaba la realización de una metarregresión, pero al final no fue posible. Aunque habría sido posible emprender un análisis de este tipo bajo el supuesto de que los efectos de las distintas intervenciones eran aditivos, se consideró poco realista. En su lugar, se describieron todos los resultados de las comparaciones y se resumieron los resultados de las intervenciones múltiples que incluían visitas educativas y de otras intervenciones múltiples con al menos cuatro comparaciones. Resultó difícil extraer conclusiones generalizables a partir de comparaciones de intervenciones múltiples debido al elevado número de distintas combinaciones evaluadas. Sin embargo, considerando todas las combinaciones, las intervenciones múltiples no parecían ser más efectivas que las intervenciones únicas, y los efectos de las intervenciones múltiples no parecían aumentar con el número de intervenciones que las componían.

La mayoría de los estudios se llevaron a cabo en Estados Unidos, y no queda claro que los resultados se puedan aplicar a otros marcos asistenciales. En general, no se sabe mucho acerca de los factores que pueden influir en la aplicabilidad de los resultados de la investigación de implementaciones.

La revisión pone de relieve muchas de las deficiencias metodológicas y de presentación de datos de los estudios existentes, y recalca la importancia de usar diseños robustos y bien aplicados a la hora de evaluar estrategias de difusión e implementación de directrices.

Escaseaban los detalles sobre las intervenciones de estudio y sobre los factores contextuales y, muchas veces, era difícil valorar el fundamento de la elección de una intervención. También se describieron poco los factores que facilitaban u obstaculizaban la práctica. Muchos estudios adolecieron de deficiencias metodológicas (por ejemplo, errores de la unidad de análisis) que dieron lugar a dudas considerables en cuanto a la síntesis y la interpretación de los resultados de la revisión.

Revisión sistemática de las evaluaciones económicas y los análisis de costes en las evaluaciones de estrategias de difusión e implementación de directrices

Para poder tomar decisiones informadas respecto a si vale la pena introducir directrices, los gestores sanitarios deben disponer de información acerca de los beneficios y los costes probables de distintas estrategias de difusión e implementación de directrices. A pesar de esto, sólo el 29% (63) de los estudios aportaron datos económicos. Once analizaron la relación coste-efectividad, 38 analizaron la relación coste-consecuencia (y las diferencias de coste se opusieron a diferencias en varias medidas de la efectividad) y 14 analizaron los costes (aportaron algún aspecto del coste pero no lo relacionaron con los beneficios). La mayoría de los estudios sólo aportaron datos sobre los costes de tratamiento; sólo 25 estudios aportaron datos sobre los costes del desarrollo de las directrices o de su difusión e implementación. La mayoría de los estudios utilizaron medidas del proceso como variable principal de valoración, a pesar del hecho de que sólo 3 directrices estaban basadas explícitamente en evidencias (y pueden no haber sido eficientes).

En conjunto, la calidad metodológica de las evaluaciones económicas y de los análisis de costes era deficiente. Se declaró el punto de vista adoptado en las evaluaciones económicas en sólo 10 estudios. Los métodos que se utilizaron para estimar los costes fueron completos en aproximadamente la mitad de los estudios, y pocos estudios facilitaron detalles sobre el uso de recursos. Debido a la mala calidad de la presentación de los estudios, en la mayoría no se disponía de datos sobre el uso de recursos ni sobre el coste del desarrollo, la difusión y la implementación de directrices. Sólo 4 estudios aportaron datos lo suficientemente robustos como para poder incluirlos en la valoración. Estos estudios demostraron que los costes del desarrollo local de directrices no son insignificantes si se determina explícitamente su coste, aunque consideraron que muchas veces no se explicita el tiempo necesario para muchas actividades y que algunas actividades a menudo se realizan fuera de las horas de trabajo. Antes de que pueda determinarse que los costes que se han aportado pueden hacerse extensivos a otros casos, se necesitan nuevas estimaciones sobre los recursos y los costes de distintos métodos de difusión e implementación. Se trata de un campo de trabajo relativamente nuevo y hay pocas publicaciones que incluyan un análisis sobre los métodos de evaluación económica de los ensayos de investigación sobre implementación.

Sin embargo, para maximizar sus beneficios para los gestores sanitarios, los estudios que se realicen en un futuro necesitan llevar a cabo estos análisis.

Encuesta de informantes clave

La encuesta a informantes clave, aunque de alcance limitado, planteó muchos temas relacionados con la viabilidad y la factibilidad del uso de estrategias de difusión e implementación en el marco de la sanidad pública. En pocas ocasiones los encuestados identificaron la existencia de presupuestos que financiaran las estrategias de difusión e implementación de directrices, y se hicieron muchos comentarios sobre el uso de fondos o recursos privados para iniciativas específicas destinadas a apoyar estas actividades. Algunos encuestados observaron que había recursos no económicos dentro del sistema que se podían aprovechar. En general, los encuestados opinaban que la difusión de material educativo y los encuentros informativos breves (durante la hora del almuerzo) eran factibles con los recursos existentes (aunque los médicos de atención primaria eran conscientes de los costes de oportunidad adicionales que suponía invitar a personal de atención primaria a esos encuentros). Los encuestados plantearon dudas sobre la factibilidad de la auditoría y la retroalimentación y la de los recordatorios; en general, se consideraba que requerían una cantidad importante de recursos para llevarlos a cabo mediante sistemas manuales y que, para hacerlo por medio de la informática, había poca disponibilidad de sistemas informáticos y éstos tenían una funcionalidad limitada (especialmente en marcos de atención especializada). A los encuestados de atención primaria les atrajo la idea de las visitas educativas, ya que eso les podía permitir permanecer en el consultorio, pero tanto los especialistas como los médicos de cabecera expresaron dudas respecto a las posibilidades de obtener los recursos necesarios en el marco de la sanidad pública.

PUNTOS FUERTES Y PUNTOS DÉBILES

La revisión sistemática es la más extensa, sofisticada metodológicamente y actualizada de las estrategias de implementación de directrices. Los revisores que la llevaron a cabo desarrollaron estrategias de búsqueda altamente sensibles, y consideran que los estudios incluidos representan un conjunto relativamente completo del período comprendido entre 1966 y mediados de 1998. Las búsquedas detectaron más estudios que otras revisiones anteriores de períodos similares. Por ejemplo, el boletín *Effective Health Care* sobre la implementación de directrices clínicas⁴ identificó 91 estudios entre 1976 y 1994, mientras que las búsquedas de esta revisión identificaron en el mismo período 165 estudios. Se han examinado según el registro Cochrane EPOC algunos estudios potencialmente relevantes que no cumplieron los criterios de inclusión para esta revisión. Desde entonces, se ha utilizado la estrategia de búsqueda para desarrollar la del EPOC, de forma que en un futuro se podrá reducir el tiempo de búsqueda de estudios pertinentes en el registro EPOC (epoc@uottawa.ca).

Se hizo una extracción de datos detallada sobre la calidad de los estudios, así como sobre sus características y sobre las intervenciones. Como resultado, en esta revisión se han hecho más explícitas las deficiencias metodológicas de los estudios primarios. Cuando fue posible, se intentaron volver a analizar los resultados de las comparaciones con errores metodológicos comunes; así, por ejemplo, se volvieron a analizar 27 comparaciones de series temporales interrumpidas. Por otro lado, también se caracterizaron las intervenciones con mayor detalle que en otras revisiones anteriores, lo que puso de manifiesto el carácter múltiple de la mayoría de las intervenciones evaluadas. Las revisiones anteriores normalmente han descrito las intervenciones basándose en la descripción principal de la intervención que hacían los autores, y éstos a menudo pasan por alto las intervenciones coexistentes.

Las revisiones anteriores han tendido a usar métodos de recuento de votos de estudios de comparaciones estadísticamente significativas, muchas veces pasando por alto errores de la unidad de análisis. Estos métodos pueden ser engañosos debido a la significación estadística incierta de las comparaciones con errores de la unidad de análisis y a la falta de poder estadístico de muchas comparaciones. Asimismo, aportan poca información acerca de los tamaños de efecto potenciales de distintas intervenciones. En la presente revisión, se utilizó un marco más explícito para considerar la calidad metodológica de los estudios (basado en el diseño y la presencia de errores de la unidad de análisis al interpretar los resultados de comparaciones). También permitió aportar información acerca del tamaño del efecto potencial de las intervenciones. Además, se han descrito detalladamente los métodos, los supuestos y los resultados de la revisión para permitir a los lectores explorarlos en mayor profundidad.

La revisión presenta una serie de puntos débiles. Es inevitable que se hayan pasado por alto algunos estudios; la estrategia de búsqueda tenía una sensibilidad estimada del 93%, lo que sugiere que pueden haberse pasado por alto unos 17 estudios. Los autores creen que las búsquedas son completas hasta mediados de 1998. Actualmente se estima que se están publicando unos 35–40 estudios al año, lo que sugeriría que la revisión está ya desfasada. También hay pruebas de que los estudios más recientes tenían menos deficiencias metodológicas, como errores potenciales de la unidad de análisis. Se incluyeron los estudios que, en opinión de los revisores, cumplían la definición de directrices clínicas del Institute of Medicine¹¹, independientemente de si se mencionaban explícitamente términos como guía clínica o directrices. Esto se debe mayormente a razones pragmáticas: el uso de estas palabras ha ido cambiando a lo largo del tiempo, y muchos estudios realizados en los años 70 y 80, que se centraron en determinadas recomendaciones, probablemente hablarían de guía clínica o directrices si se hicieran hoy. Estos juicios inevitablemente supusieron cierto grado de subjetividad. Para reducir al mínimo esta subjetividad, dos revisores valoraban los estudios de forma independiente frente a los criterios de inclusión; la opinión de un tercer revisor resolvía las eventuales discrepancias. Se utilizó la taxonomía de intervenciones del grupo Cochrane EPOC; esta clasificación se basa en descripciones claras y útiles de las intervenciones, es

relativamente fácil de aplicar y se ha utilizado en una gran variedad de revisiones. Sin embargo, es posible que otras formas de clasificar las intervenciones hubieran podido ser más precisas (aunque los autores no saben de ninguna otra taxonomía desarrollada hasta un grado comparable). La amplitud de la revisión también dificultó un examen detallado de los modificadores del efecto de las intervenciones individuales y, aunque estaba previsto, tampoco fue posible realizar una metaregresión. Originalmente se proyectaba entrar en contacto con los autores para recabar información adicional, pero la envergadura inesperada de la tarea hizo que no fuera posible hacerlo con los recursos disponibles. Por último, quizás la limitación más importante de la revisión es la calidad de los estudios primarios, muchos de ellos con deficiencias metodológicas comunes que merman la certeza de las conclusiones.

Esta investigación también incluye la primera revisión sistemática de las evaluaciones económicas y los análisis de costes en evaluaciones de estrategias de difusión e implementación de directrices; para valorar las evaluaciones económicas, se utilizó una lista de verificación estándar. La revisión sólo tuvo en cuenta los estudios identificados que estudiaban la efectividad. Es posible que los revisores hayan pasado por alto evaluaciones económicas aparecidas en publicaciones secundarias (aunque las búsquedas identificaron muchas publicaciones secundarias) o evaluaciones económicas basadas en el modelado. De esta revisión destaca la escasez de evaluaciones económicas de alta calidad, lo que limita la capacidad de sacar conclusiones firmes acerca de los recursos que necesitan y la eficiencia de las distintas estrategias de difusión e implementación.

Para complementar los resultados de ambas revisiones, se hizo una pequeña encuesta a informantes clave para explorar la factibilidad y los recursos que se necesitarían para implementar intervenciones dentro del sistema sanitario británico. Aunque era una encuesta pequeña, puso de manifiesto muchas de las limitaciones prácticas a las que hay que enfrentarse a la hora de considerar distintas estrategias de difusión e implementación. En investigaciones futuras se podría ampliar esta encuesta para explorar hasta qué punto son generalizables sus resultados.

INTERÉS DEL ESTUDIO: POSIBLES MECANISMOS Y REPERCUSIONES PARA MÉDICOS Y PARA GESTORES SANITARIOS

Comparación con revisiones sistemáticas anteriores

Un estudio sobre revisiones sistemáticas de estrategias de cambio conductual profesional publicado en 1999 identificó 44 revisiones que abarcaban un amplio espectro de actividades e intervenciones². De estas revisiones se infiere que la “difusión pasiva (por ejemplo, enviar material educativo por correo a médicos seleccionados) generalmente es ineficaz y tiene pocas probabilidades de producir un cambio conductual cuando se utiliza sola; sin embargo, puede ser útil para

aumentar la conciencia sobre el cambio conductual deseado. Las intervenciones activas tienen más probabilidades de ser eficaces, pero también es probable que sean más caras. También encontramos intervenciones de efectividad variable, como la que se lleva a cabo a través de líderes de opinión locales o de la auditoría y la retroalimentación. Entre las estrategias generalmente eficaces se encuentran las visitas educativas (especialmente en cuanto al comportamiento en la expedición de recetas) y los recordatorios. Las intervenciones múltiples basadas en la evaluación de las barreras potenciales para que se produzca el cambio tienen más probabilidades de ser eficaces que las intervenciones únicas.”

Las conclusiones de la presente revisión pueden parecer en desacuerdo con las del estudio mencionado arriba y con las de revisiones sistemáticas anteriores. Por ejemplo, Freemantle y sus colaboradores²⁹ revisaron 11 estudios que evaluaban los efectos de la difusión de material educativo. Utilizaron métodos de recuento de votos y excluyeron los estudios con errores de la unidad de análisis. Ninguna de las comparaciones que utilizaban técnicas estadísticas apropiadas encontró mejoras estadísticamente significativas en la práctica. En la presente revisión, se utilizó un marco analítico más explícito para explorar la mediana y el intervalo de los tamaños de efecto en estudios con errores de la unidad de análisis y sin ellos. Se identificaron 4 comparaciones C-ECA que aportaron medidas dicotómicas del proceso. Todas ellas observaron mejoras de la asistencia; el efecto mediano fue una mejora absoluta del cumplimiento del +8,1% (límites: +3,6% a +17%). Dos comparaciones tenían errores potenciales de la unidad de análisis, y no se pudo determinar la significación de 1 comparación. La comparación restante, que no tenía error potencial de la unidad de análisis, observó un efecto del +6% (no significativo). Si se hubiera utilizado aquí el mismo enfoque analítico, se habría llegado a conclusiones similares a las de Freemantle y sus colaboradores.

Sin embargo, un enfoque de este tipo no habría detectado que el material educativo impreso producía mejoras en la asistencia en la mayoría de los estudios (aunque la significación estadística de la mayoría de las comparaciones es incierta). Basándose en esta revisión, y dadas las deficiencias metodológicas de los estudios primarios, los autores hubieran sacado conclusiones distintas acerca de la eficacia del material educativo impreso. En cambio, se concluyó que éste puede mejorar la asistencia y que, por tanto, los diseñadores de políticas no deberían desestimarlos, ya que tiene un posible efecto, un coste relativamente bajo y una factibilidad alta en el marco del sistema sanitario público.

Thomson O'Brien y sus colaboradores³⁰ llevaron a cabo una revisión sistemática de las visitas didácticas mediante métodos de recuento de votos; concluyeron que había ciertos indicios que respaldaban las “visitas educativas combinadas con otras intervenciones para reducir la prescripción médica inapropiada”, pero que la eficacia económica de este enfoque no estaba clara. En la presente revisión, la mayoría de los estudios que evaluaron intervenciones múltiples que incluían visitas educativas detectaron direcciones positivas del efecto. De nuevo, si se hubieran usado aquí los mismos métodos de revisión,

se habría llegado a conclusiones similares a las de Thomson O'Brien y sus colaboradores. Sin embargo, el enfoque analítico adoptado en esta revisión destacó los modestos tamaños del efecto observados en los estudios primarios (tamaño del efecto mediano del +6%, límites: -4% a 17,4% en 11 C-ECA). Teniendo en cuenta estos resultados, así como el elevado coste potencial de la intervención y las dudas respecto a la factibilidad de esta intervención de forma habitual en el marco de la sanidad pública, los autores son más cautelosos en sus conclusiones relativas a la eficiencia de esta intervención. Desde que se llevó a cabo esta revisión, se han publicado 2 estudios bien planteados sobre las visitas educativas, ambos realizados en el Reino Unido. Hall y sus colaboradores³¹ observaron un efecto de las visitas educativas sobre la expedición de recetas. Freemantle y sus colaboradores observaron una mejora del +3% en cuatro áreas de la expedición de recetas. Curiosamente, a pesar del reducido tamaño del efecto, un análisis económico²⁴ sugirió que las visitas educativas ofrecían una buena relación coste-efectividad en consultorios pequeños de medicina general para determinados fármacos.

Revisiones anteriores también han sugerido que las intervenciones múltiples son más eficaces que las intervenciones únicas, ya que actúan sobre múltiples barreras para la implementación. La revisión realizada por Davis y sus colaboradores sobre las estrategias de formación médica continuada¹³ llegó a la conclusión de que las intervenciones múltiples tenían más probabilidades de ser eficaces. Wensing y sus colaboradores⁸, que llevaron a cabo una revisión sobre la efectividad de la introducción de directrices en marcos de atención primaria, concluyeron que cuando se combinaba más de una intervención, había una tendencia a obtener una mayor efectividad, aunque estas intervenciones múltiples podían ser más caras. Las dos revisiones mencionadas no especificaban detalles sobre cómo se codificaron las intervenciones ni el método analítico utilizado. En esta revisión, en cambio, se codificaron todos los componentes de la intervención y se utilizaron métodos explícitos para determinar un tamaño del efecto único para cada estudio. El análisis indicó que la efectividad de las intervenciones múltiples no aumentaba proporcionalmente con el número de componentes. Pocos estudios aportaron una base o justificación teórica para la intervención elegida, de forma que no quedaba claro si los investigadores tenían un fundamento previo para la elección de los componentes de las intervenciones múltiples basado en posibles mecanismos causales, o si era una elección con un enfoque más aleatorio. Es plausible que las intervenciones múltiples fundamentadas en una cuidadosa valoración de los obstáculos y con una base teórica coherente sean más efectivas que las intervenciones únicas.

Es importante no malinterpretar estos resultados, ya que hubo pocas comparaciones directas de las distintas intervenciones. En particular, sería inapropiado concluir que una intervención puede ser más eficaz que otra basándose en comparaciones indirectas. Probablemente existe un factor de confusión natural en el sentido de que los investigadores eligieron intervenciones que, en su opinión, podrían ser eficaces en su

contexto de estudio; por lo tanto, puede haber diferencias importantes en el contexto, las barreras o el comportamiento estudiado entre los trabajos que valoraron, por ejemplo, el material educativo impreso y las visitas educativas.

Repercusiones para los gestores sanitarios

Las directrices son cada vez más comunes en sanidad. En el Reino Unido, los responsables locales de política sanitaria que trabajan en estructuras de gobierno clínico deben plantearse cómo responder al número creciente de directrices desarrolladas por agencias nacionales como NICE (National Institute for Clinical Effectiveness) y SIGN (Scottish Intercollegiate Guidelines Network). Idealmente, la decisión de si es conveniente desarrollar, difundir e implementar unas directrices clínicas y la decisión de cómo hacerlo debe basarse en estimaciones hechas cuidadosamente sobre los costes y los beneficios de la estrategia de difusión e implementación, así como sobre los costes y los beneficios de los cambios resultantes en la atención del paciente. Mason y sus colaboradores²⁴ acuñaron el término 'efectividad económica de una política' para describir esta combinación de la efectividad económica del tratamiento y del coste y la magnitud del cambio logrado por un método de implementación. Demostraron con elegancia que las visitas educativas ofrecían una relación coste-efectividad interesante para fomentar el uso de inhibidores de la enzima convertidora de angiotensina en la insuficiencia cardíaca; en cambio, no la ofrecían a la hora de fomentar la sustitución de antidepresivos tricíclicos por inhibidores selectivos de la recaptación de serotonina en la atención primaria británica.

La revisión mencionada ha demostrado que no existe una base de pruebas robusta y generalizable que pueda aportar estimaciones de los costes y los beneficios probables de distintas estrategias. Además, no hay pruebas robustas sobre la eficacia económica de las intervenciones en muchas áreas de la atención clínica, y pocas directrices han incorporado consideraciones económicas al elaborar sus recomendaciones³³. Por ello, es poco probable que los responsables de estas decisiones puedan replicar el sofisticado método de modelado demostrado por Mason y sus colaboradores²⁴. No obstante, se puede usar una estructura similar para guiar sus decisiones sobre cómo usar mejor los limitados recursos de los que disponen para el gobierno clínico y las actividades relacionadas para maximizar los beneficios para la población.

Se elaboraron las conclusiones siguientes para los diseñadores de políticas.

- Dada la limitación de recursos, parece probable que se tenga que decidir qué directrices difundir e implementar activamente; para ello, hay que tener en cuenta la carga local que supone la enfermedad, la disponibilidad de intervenciones sanitarias eficaces y eficientes, y la existencia de pruebas de que la actuación local hasta el momento es subóptima. Todo ello también estará influido por las iniciativas y las prioridades nacionales (varios encuestados observaron que muchas veces es más fácil obtener más fondos para temas relacionados con prioridades nacionales).
- También se debe considerar si se dispone de unas directrices de alta calidad y con una buena previsión de los costes. Si no es así, los recursos que se necesitan para desarrollar una directriz sólida *de novo* son sustanciales.
- Se deben identificar los recursos de los que se disponen para sufragar estrategias de difusión e implementación. A pesar de que los presupuestos disponibles para este objetivo son siempre limitados, es posible acceder a otros recursos a través de la asociación con otros equipos y la coordinación con más interesados (por ejemplo, coordinar la orientación educativa con los profesionales de la formación médica continuada) o dentro del entorno de la organización (por ejemplo, sistemas de información capaces de generar recordatorios).
- La elección de la estrategia de difusión e implementación de directrices debe basarse en los posibles obstáculos o barreras y potenciadores a la hora de intentar conseguir el cambio, así como en los costes y los beneficios probables de distintas estrategias. Varios autores han ofrecido justificaciones teóricas para elegir determinadas intervenciones en función de los distintos tipos de potenciadores y barreras que se den (como Grol³⁴, Mittman y sus colaboradores³⁵, Grimshaw y Eccles³⁶, y Moulding y sus colaboradores³⁷). Sin embargo, no existe una base empírica sólida para estas justificaciones. Por otra parte, la ausencia de una justificación explícita de las intervenciones y la descripción deficiente de las barreras y los potenciadores posibles en los estudios actuales limitan nuestro conocimiento sobre la aplicabilidad de los resultados de los estudios a otros marcos y a otros problemas.
- También se tiene que tener un gran criterio a la hora de decidir qué intervenciones tienen más probabilidades de ser eficaces en una circunstancia determinada, y elegir las intervenciones en función de la factibilidad, los costes y los beneficios asociados a intervenciones potencialmente eficaces.
 - Siempre que sea posible, las intervenciones deben incluir recordatorios escritos en papel o generados por ordenador.
 - Puede ser más eficiente utilizar una intervención más barata y factible aunque menos eficaz (por ejemplo, difusión pasiva de directrices) que una intervención más cara pero potencialmente más eficaz.
 - También se tienen que considerar las implicaciones que pueden tener los cambios en la práctica clínica sobre los recursos o sobre distintos presupuestos.
- Se deben valorar los efectos de cualquier intervención, preferentemente a través de diseños evaluativos robustos.

PREGUNTAS SIN CONTESTAR Y FUTURAS INVESTIGACIONES

Esta revisión destaca el hecho de que, a pesar de los 30 años de investigación en este campo, todavía carecemos de una base de pruebas robusta y generalizable para apoyar la

toma de decisiones sobre las estrategias destinadas a fomentar la aplicación de directrices u otros mensajes comprobados.

El consejo de investigaciones médicas británico (Medical Research Council - MRC) propuso recientemente un marco para evaluar intervenciones complejas (como en el caso de las estrategias de difusión e implementación de directrices)³⁸. Sugiere que la evaluación de intervenciones complejas debe seguir un enfoque secuencial, y debe incluir:

1. el desarrollo de la base teórica de una intervención
2. la definición de los componentes de la intervención (utilizando modelado o técnicas simuladas y métodos cualitativos)
3. estudios exploratorios para definir con mayor profundidad la intervención y planificar un estudio evaluativo definitivo (utilizando diversos métodos)
4. un estudio evaluativo definitivo (con métodos cuantitativos de evaluación, especialmente con diseños aleatorizados).

Este enfoque reconoce los beneficios de establecer la base teórica de las intervenciones y llevar a cabo estudios exploratorios para elegir y perfeccionar intervenciones y así reducir al mínimo el número de costosos ECA 'definitivos'. La mayoría de los estudios incluidos en esta revisión sistemática podrían considerarse evaluaciones definitivas. Se ha investigado relativamente poco para desarrollar la base teórica del comportamiento profesional y organizativo, y del cambio conductual en el ámbito sanitario. En pocos casos se han realizado estudios preparatorios para definir la intervención o para demostrar la factibilidad del estudio definitivo proyectado.

La falta de una base teórica coherente para comprender el cambio conductual profesional y organizativo limita nuestra capacidad para formular hipótesis sobre las intervenciones que tienen más probabilidades de ser efectivas bajo distintas circunstancias, y dificulta que lleguemos a determinar si los resultados de los ensayos definitivos son generalizables. Un asunto primordial para la investigación futura debería ser el desarrollo de un mejor conocimiento teórico del cambio conductual profesional y organizativo. Ferlie y Shortell³⁹ sugirieron cuatro campos en los que podrían aplicarse las intervenciones para mejorar la calidad de la atención sanitaria: el profesional sanitario individual, los equipos o grupos sanitarios, las organizaciones que prestan la atención sanitaria, y el sistema o entorno sanitario más amplio en el que están integradas las organizaciones individuales. Una justificación científica completa de las intervenciones de difusión e implementación de directrices requiere una exploración de las teorías aplicables a cada uno de estos cuatro campos. Muchas teorías educativas, conductuales, sociales y organizativas han demostrado ser útiles para comprender el comportamiento en un amplio espectro de contextos. Sin embargo, en pocas ocasiones se ha estudiado la aplicabilidad de estas teorías al comportamiento profesional y organizativo en la atención sanitaria. Por lo tanto, un primer paso en el desarrollo de las bases teóricas del comportamiento profesional y organizativo podría consistir en estudiar la aplicabilidad de estas teorías en marcos de atención sanitaria. Distintos tipos de teoría serán aplicables en intervenciones a distintos niveles; por ejemplo, las teorías psicológicas serán más aplicables en las interven-

ciones dirigidas a individuos o a equipos; y las teorías del cambio organizativo serán más aplicables a las intervenciones dirigidas a hospitales u organizaciones. Por otra parte, antes de llevar a cabo ensayos definitivos hay que hacer más investigación en los métodos para optimizar las intervenciones (que es equivalente a las fases de modelado y pruebas exploratorias del Medical Research Council).

Al igual que en otras áreas de investigación médica y social, las evaluaciones rigurosas (principalmente ECA) aportarán las pruebas más robustas acerca de la efectividad de estrategias de difusión e implementación de directrices, porque los efectos probablemente serán modestos, hay un riesgo importante de sesgo y nuestro conocimiento sobre los modificadores de efecto posibles y los factores de confusión es incompleta^{40,41}. Sin embargo, a pesar del elevado número de evaluaciones que se han realizado en este campo, todavía falta una base de pruebas coherente en la que fundamentar las decisiones acerca de qué estrategias de difusión e implementación usar. Esto se debe en parte a la mala calidad metodológica de los estudios existentes; por ejemplo, no se pudo determinar la significación estadística de muchas comparaciones a partir de los estudios publicados debido a deficiencias metodológicas comunes. Así pues, se necesitan nuevas evaluaciones rigurosas de distintas estrategias de difusión e implementación. Estas evaluaciones deben ser metodológicamente robustas (evitando los errores metodológicos comunes identificados en la revisión sistemática), deben incorporar evaluaciones económicas y, cuando sea posible, comprobar explícitamente las teorías conductuales para mejorar nuestro conocimiento teórico sobre los factores que influyen en la efectividad de las estrategias de difusión e implementación de directrices.

El marco del Medical Research Council para el desarrollo y la evaluación de intervenciones es secuencial e idealizado. Sin embargo, su modelo se basa en las fases del desarrollo de productos farmacéuticos, que generalmente tardan 15–20 años en pasar por todas las fases y precisan una inversión de recursos importante. Es poco probable que los investigadores de implementaciones tengan el tiempo y los recursos necesarios para pasar secuencialmente por cada fase al evaluar estrategias de difusión e implementación de directrices. Por lo tanto, las futuras investigaciones deben adoptar una variedad de enfoques a la hora de desarrollar y comprobar teorías y de evaluar intervenciones.

Recomendaciones para futuras investigaciones

Se precisa más investigación en los ámbitos teórico, concreto y metodológico.

Teórico

- Se precisa más investigación para desarrollar y confirmar un marco teórico coherente sobre el comportamiento profesional y organizativo y sobre el cambio conductual en el ámbito sanitario; de esta forma se podrá aportar una mejor información para la elección de intervenciones en

entornos de investigación y servicios. Inicialmente, la investigación debe evaluar la aplicabilidad de las teorías predictivas existentes a una serie de niveles (por ejemplo: individuo, equipo, organización, sistema) en vez de intentar desarrollar una teoría *de novo*.

Concreto

- Se necesitan más evaluaciones rigurosas de distintas estrategias de difusión e implementación que:
 - sean metodológicamente robustas (evitando los errores metodológicos comunes encontrados en la revisión sistemática)
 - incorporen evaluaciones económicas
 - siempre que sea posible, comprueben explícitamente teorías conductuales para ampliar nuestro conocimiento teórico sobre los factores que influyen en la efectividad de las estrategias de difusión e implementación de directrices. Si esto no resulta posible, los ensayos deben integrar evaluaciones de proceso para describir e ilustrar mejor sus resultados
 - evalúen un espectro de intervenciones (provistas de un fundamento teórico) en comparaciones directas.
- Se necesitan más evaluaciones rigurosas de intervenciones relativamente baratas y que tienen el potencial de ser eficientes, como la del material educativo impreso.

Metodológico

- Se necesita más investigación para desarrollar métodos para optimizar las intervenciones antes de llevar a cabo los ensayos definitivos.
- Se necesita más investigación para mejorar los métodos de realización de C-ECA sobre intervenciones complejas como son las estrategias de difusión e implementación de directrices. El Medical Research Council ha publicado una monografía⁴² sobre la realización de C-ECA con conglomerados; sin embargo, tiene que haber un conocimiento más generalizado sobre estos aspectos y su aplicación dentro de la comunidad investigadora.

- Aunque no es estrictamente una recomendación para futuras investigaciones, hay una clara necesidad de mejorar la presentación de los datos en los estudios que evalúen intervenciones complejas, como lo son las estrategias de difusión e implementación de directrices. Con ello se podrían abarcar aspectos como, por ejemplo, la mejor forma de aportar detalles sobre la elección y la construcción de las intervenciones y la evaluación de los modificadores del efecto.
- Los resultados de esta revisión sistemática difieren de las revisiones anteriores como consecuencia de los métodos que se han utilizado. Se necesita más investigación para establecer métodos óptimos de revisión de intervenciones complejas como las de las estrategias de difusión e implementación de directrices; deberían abarcar aspectos como:
 - cómo agregar mejor los resultados de estudios que tienen errores de la unidad de análisis
 - cómo enfocar mejor la categorización de intervenciones.

Otros temas

Hay una considerable necesidad de investigación, y ésta requerirá una perspectiva interdisciplinaria en la que se apliquen varios enfoques metodológicos. Se trata de un campo relativamente nuevo de investigación de los servicios sanitarios, con una perspectiva diferenciada y retos metodológicos específicos. Es importante desarrollar un fondo de experiencia y competencia en este campo, básicamente a través de actividades programáticas interdisciplinarias. Para seguir avanzando en esta área, se necesitará el apoyo económico de distintos organismos. Por ejemplo, el desarrollo teórico necesario (que equivaldría a establecer 'la ciencia básica de la implementación') puede costearse mejor a través de organismos que financian la investigación sanitaria fundamental (y no la investigación aplicada). Por otra parte, la necesidad de investigación trasciende posibles diferencias nacionales y culturales en la práctica y en la financiación de la atención sanitaria; su envergadura hace pensar que ningún programa de investigación de servicios sanitarios de un solo país ni ningún grupo solo de investigación, trabajando aisladamente, puede llevarlo a cabo de una forma sistemática y completa. Por tanto, hay una importante oportunidad para la colaboración internacional para mejorar nuestro conocimiento en este campo.

Bibliografía

1. Woolf, S.H., Grol, R., Hutchison, A., Eccles, M., Grimshaw, J. *Clinical practice guidelines. The potential benefits, limitations, and harms of clinical guidelines.* BMJ 1999; 318: 527-530.
2. NHS Centre for Reviews and Dissemination. *Getting evidence into practice.* Effective Health Care 1999; 5: 1-16.
3. Lomas, J. *Words without action? The production, dissemination, and impact of consensus recommendations.* Annu Rev Public Health 1991; 12: 41-65.
4. NHS Centre for Reviews and Dissemination and Nuffield Institute for Health. *Implementing clinical guidelines. Can guidelines be used to improve clinical practice?* Effective Health Care 1994; 1: 1-12.
5. Oxman, A., Thomson, M.A., Davis, D.A., Haynes, R.B. *No magic bullets: a systematic review of 102 trials of interventions to improve professional practice.* CMAJ 1995; 153: 1423-1431.
6. Davis, D.A., Taylor-Vaisey, A. *Translating guidelines into practice. A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines.* CMAJ 1997; 157: 408-416.
7. Worrall, G., Chaulk, P., Freake, D. *The effects of clinical practice guidelines on patient outcomes in primary care: a systematic review.* CMAJ 1997; 156: 1705-1712.
8. Wensing, M., van der Weijden, T., Grol, R. *Implementing guidelines and innovations in general practice: which interventions are effective?* J Gen Pract 1998; 48: 991-997.
9. Thomas, L., Cullum, N., McColl, E., Rousseau, N., Soutter, J., Steen, N. *Clinical guidelines in nursing, midwifery and other professions allied to medicine (Cochrane Review).* The Cochrane Library (Issue 1). Update Software, Oxford 1999.
10. Bero, L., Grilli, R., Grimshaw, J.M., Mowatt, G., Oxman, A., Zwarenstein, M. *Cochrane Effective Practice and Organisation of Care Review Group (Cochrane Group Module).* En: Bero, L., Grilli, R., Grimshaw, J.M., Mowatt, G., Oxman, A., Zwarenstein, M. (Eds.). The Cochrane Library (Issue 3). Update Software, Oxford 2001.
11. Institute of Medicine Committee on Clinical Practice Guidelines. *Guidelines for clinical practice: from development to use.* National Academy Press, Washington, DC 1992.
12. Cook, T.D., Campbell, D.T. *Quasi-experimentation: design and analysis issues for field settings.* Rand McNally, Chicago, IL 1979.
13. Davis, D.A., Thomson, M.A., Oxman, A.D., Haynes, R.B. *Changing physician performance. A systematic review of the effect of continuing medical education strategies.* JAMA 1995; 274: 700-705.
14. Whiting-O'Keefe, Q.E., Henke, C., Simborg, D.W. *Choosing the correct unit of analysis in medical care experiments.* Med Care 1984; 22: 1101-1114.
15. Ukoumunne, O.C., Gulliford, M.C., Chinn, S., Sterne, J.A.C., Burney, P.G.J. *Methods for evaluating area-wide and organisation based interventions in health and health care: a systematic review.* Health Technol Assess 1999; 3(5).
16. Bushman, B.J. *Vote counting methods in metaanalysis.* En: Cooper, H., Hedges, L. (Eds.). The handbook of research synthesis. Russell Sage Foundation, Nueva York 1994; 193-213.
17. Drummond, M., O'Brien, B., Stoddart, G., Torrance, G.W. *Methods of economic evaluation of health care programmes.* Oxford University Press, Oxford 1997.
18. McIntosh, E. *Economic evaluation of guideline implementation studies.* En: Makela, M., Thorsen, T. (Eds.). Changing professional practice. Theory and practice of clinical guidelines implementation. Danish Institute for Health Services Research and Development; Copenhagen 1999; 77-98.
19. Mason, J., Wood, J., Freemantle, N. *Designing evaluations of interventions to change professional practice.* J Health Serv Res Policy 1999; 4: 106-111.
20. Drummond, M.F., Jefferson, T.O., para el BMJ Working Party on guidelines for authors and peer reviewers of economic submissions to the British Medical Journal. *Guidelines for authors and peer reviewers of economic submissions to the British Medical Journal.* BMJ 1996; 313: 275-283.
21. Jefferson, T.O., Demicheli, V., Vale, L. *Methodological reviews of economic evaluations in health care.* JAMA 2002; 287: 2809-2812.
22. Fergusson, D., Van Walraven, C., Coyle, D., Laupacis, A. *Economic evaluations of technologies to minimize perioperative transfusion: a systematic review of published studies.* International Study of Perioperative Transfusion (ISPOT) investigators. Transfus Med Rev 1999; 13: 106-117.

23. Campbell, M.K., Mollison, J., Grimshaw, J.M. *Cluster trials in implementation research: estimation of intracluster correlation coefficients and sample size*. *Stat Med* 2001; 20: 391-399.
24. Mason, J., Freemantle, N., Nazereth, I., Eccles, M., Haines, A., Drummond, M. *When is it cost effective to change the behaviour of health professionals?* *JAMA* 2001; 286: 2988-2992.
25. Neal, R.D., Heywood, P.L., Morley, S. *Real world data – retrieval and validation of consultation data from four general practices*. *Fam Pract* 1996; 13: 455-461.
26. Department of Health. *An information strategy for the modern NHS 1998-2005. A national strategy for local implementation*. Department of Health, Londres 1998.
27. Purves, I.N. *PRODIGY: implementing clinical guidelines using computers*. *Br J Gen Pract* 1998; 48: 1552-1553.
28. Sculpher, M. *Evaluating the cost-effectiveness of interventions designed to increase the utilisation of evidence-based guidelines*. *Fam Pract* 2000; 17: S26-S31.
29. Freemantle, N., Harvey, E.L., Wolf, F., Grimshaw, J.M., Grilli, R., Bero, L. *Printed educational materials: effects on professional practice and health care outcomes (Cochrane Review)*. *Cochrane Database of Systematic Reviews* 2003; 1.
30. Thomson O'Brien, M.A., Oxman, A.D., Davis, D.A., Haynes, R.B., Freemantle, N., Harvey, E.L. *Educational outreach visits: effects on professional practice and health care outcomes (Cochrane Review)*. *Cochrane Database of Systematic Reviews* 2003; 1.
31. Hall, L., Eccles, M.P., Barton, R., Steen, N., Campbell, M. *Is untargeted outreach visiting in primary care effective? A pragmatic randomized controlled trial*. *J Public Health Med* 2001; 23: 109-113.
32. Freemantle, N., Nazereth, I., Eccles, M., Wood, J., Haines, A. *Evidence-based OutReach trialists. A randomised controlled trial of the effect of educational outreach by community pharmacists on prescribing in UK general practice*. *Br J Gen Pract* 2002; 52: 290-295.
33. Eccles, M., Mason, J. *How to develop cost-conscious guidelines*. *Health Technol Assess* 2001; 5(16).
34. Grol, R. *Beliefs and evidence in changing clinical practice*. *BMJ* 1997; 315: 418-421.
35. Mittman, B.S., Tonesk, X., Jacobsen, P.D. *Implementing clinical practice guidelines: social influence strategies and practitioner behavior change*. *Quality Review Bulletin* 1992; 18: 413-422.
36. Grimshaw, J.M., Eccles, M.P. *Clinical practice guidelines*. En: Haines, A., Silagy, C. (Eds.). *Evidence-based practice in primary care*. BMJ Publishing; Londres 2001; 120-134.
37. Moulding, N.T., Silagy, C.A., Weller, D.P. *A framework for effective management of change in clinical practice: dissemination and implementation of clinical practice guidelines*. *Quality in Health Care* 1999; 8: 177-183.
38. Medical Research Council. *A framework for development and evaluation of RCTs for complex interventions to improve health*. MRC Health Services and Public Health Board discussion document. MRC, Londres 2000.
39. Ferlie, E.B., Shortell, S.M. *Improving the quality of health care in the United Kingdom and the United States: a framework for change*. *Milbank Q* 2001; 79: 281-315.
40. Cochrane, A.L. *Effectiveness and efficiency: random reflections on health services*. Nuffield Provincial Hospitals Trust, Londres 1979.
41. Campbell, D.T. *Methodology and epistemology for social science: selected papers*. En: Overman, E.S. (Eds.). *The experimenting society*. University of Chicago Press, Chicago, IL 1988; 290-314.
42. Medical Research Council. *Cluster randomised trials: methodological and ethical considerations*. MRC Clinical Trials Series. MRC, Londres 2002.
43. Grimshaw, J., Campbell, M., Eccles, M., Steen, I. *Experimental and quasi-experimental designs for evaluating guideline implementation strategies*. *Fam Pract* 2000; 17: S11-18.
44. Donner, A., Birkett, N., Buck, C. *Randomization by cluster. Sample size requirements and analysis*. *Am J Epidemiol* 1981; 114: 906-914.
45. Mollison, J., Simpson, J., Campbell, M.K., Grimshaw, J.M. *Comparison of analytical methods for cluster randomised trials: an example from a primary care setting*. *J Epidemiol Biostat* 2000; 5: 339-346.
46. Draper, N., Smith, H. *Applied regression analysis*. New York: Wiley; 1981.
47. Dickersin, K., Scherer, R., Lefebvre, C. *Identifying relevant studies for systematic reviews*. *BMJ* 1994; 309: 1286-1291.

ANEXOS

- 1- Statistical considerations
- 2- Development of search strategies
- 3- HTA guidelines data abstraction checklist
- 4- Bibliographic details of included papers
- 5- Details of included studies
- 6- Results table
- 7- Interview brief: interventions to implement clinical guidelines

Esta versión no incluye los anexos. Las aproximadamente 269 páginas de los anexos están disponibles (en inglés) en la página web:

www.ncchta.org/fullmono/mon806a.pdf

ESTEVE