

Construcción de un filtro geográfico para la identificación en PubMed de estudios realizados en España

José M. Valderas^{a,b}, Joan Mendivil^{a,c}, Antoni Parada^d, Marina Losada-Yáñez^e y Jordi Alonso^{a,f}

^aUnidad de Investigación en Servicios Sanitarios. Institut Municipal d'Investigació Mèdica (IMIM). Barcelona. España.

^bDepartment of Health Policy and Management Department. Johns Hopkins School of Public Health. Baltimore. Maryland. Estados Unidos.

^cAgència de Salut Pública de Barcelona (ASPB). Barcelona. España.

^dAgència d'Avaluació de Tecnologia i Recerca Mèdiques (AATRM). Barcelona. España.

^eBiblioteca Universidad Pompeu Fabra (UPF). Barcelona. España.

^fUniversitat Autònoma de Barcelona (UAB). Barcelona. España.

Introducción y objetivos. El volumen de PubMed (Medline) hace necesarias herramientas específicas para la recuperación de información (filtros). La actual configuración de la opción «Limits» en PubMed no permite restringir las búsquedas según criterios geográficos. Por ello, se plantearon como objetivos: *a)* la construcción de un filtro geográfico para la recuperación en PubMed de estudios realizados en la población española, y *b)* la evaluación de su rendimiento.

Métodos. Se ha construido un filtro de segunda generación, basado en la revisión de estrategias previas, para la selección de 3 bloques: *a)* el término «España» y sus variantes en diversos idiomas; *b)* otros topónimos correspondientes principalmente a comunidades autónomas y provincias, y *c)* acrónimos correspondientes a los servicios autonómicos de salud. Se evaluó el rendimiento del filtro (sensibilidad y especificidad) en una muestra de conveniencia (ensayos clínicos indizados en PubMed bajo el término MeSH «Myocardial Infarction»), para lo cual se tomó como patrón de referencia la revisión manual de las referencias. También se comparó el rendimiento del filtro con el de la búsqueda simple «spain[ad]».

Resultados. El filtro ha recuperado 74 (98,7%) de las referencias pertinentes y ha demostrado una sensibilidad (88,1%) muy superior a la de la búsqueda simple

(45,8%), alcanzándose una especificidad excelente en uno y otro caso (100%).

Conclusiones. Se dispone de un filtro eficiente para la identificación de estudios realizados en población española en PubMed.

Palabras clave: Almacenamiento de información/estándares. Recuperación de información/estándares. Bibliometría. Bases de datos. Bibliografía/estándares. Filiación institucional. Publicación/estándares. España. Medicina basada en la evidencia. Sistemas de información.

Development of a Geographic Filter for PubMed to Identify Studies Performed in Spain

Introduction and objectives. The increasing size of the PubMed (Medline) database has made it necessary to develop specific tools for information retrieval (i.e., filters). The current configuration of the Limits option in PubMed does not enable searches to be limited according to geographic criteria. Our objectives, therefore, were *a)* to develop a geographic filter for the retrieval from PubMed of studies performed in the Spanish population, and *b)* to assess its performance.

Methods. A second-generation filter was constructed by reviewing previous strategies. It involved the selection of three types of record that included: *a)* the term Spain in different languages; *b)* Spanish place names (i.e., mainly those of autonomous regions and provinces), and *c)* acronyms for health services in autonomous regions. The performance of the filter (i.e., its sensitivity and specificity) was assessed in a representative sample (i.e., clinical trials indexed in PubMed under the MeSH term Myocardial Infarction) by comparison with the results of a manual search (the gold standard). In addition, the filter's performance was also evaluated by comparing it with a simple search for the term Spain[ad].

Results. Use of the filter retrieved 74 (98.7%) of the relevant references. The sensitivity (88.1%) was higher than that of the simple search strategy (45.8%), and excellent specificity (100%) was achieved in both cases.

VÉASE EDITORIAL EN PÁGS. 1221-4

Una versión previa de este trabajo fue presentada en la XXIII Reunión de la Sociedad Española de Epidemiología.

Este estudio ha sido financiado parcialmente por el Instituto de Salud Carlos III (Red de Investigación Cooperativa en Investigación de Resultados y Servicios Sanitarios [IRYSS] G03/202 y Exp N.º 3000118).

Correspondencia: Dr. J.M. Valderas Martínez.
Unidad de Investigación en Servicios Sanitarios. Institut Municipal d'Investigació Mèdica (IMIM).
Dr. Aiguader, 88. Edifici PRBB. 08003 Barcelona. España.
Correo electrónico: jmvalderas@imim.es

Recibido el 29 de agosto de 2005.

Aceptado para su publicación el 2 de agosto de 2006.

ABREVIATURAS

MeSH: Medical Subject Heading.

Conclusions. A geographic filter has been developed for identifying studies in PubMed that involve the Spanish population.

Key words: *Information storage standards. Information retrieval standards. Bibliometrics. Databases. Bibliographic standards. Organizational affiliation. Publication standards. Spain. Evidence-based medicine. Information systems.*

Full English text available from: www.revespcardiol.org

INTRODUCCIÓN

Con cerca de 15 millones de referencias, Medline, la base de datos biomédica creada y mantenida por la Biblioteca Nacional de Medicina de Estados Unidos, es una de las más completas del mundo¹. En buena medida su éxito es atribuible a PubMed, su popular interfaz de acceso libre y gratuito². Su enorme volumen, que se duplica cada 5 años, hace necesario el recurso a herramientas específicas para la recuperación de información (estrategias de búsqueda). Algunas estrategias o determinadas partes de ella pueden servir a propósitos generales, y entonces se habla de «filtros». Su uso permite ahorrar tiempo en la construcción de las búsquedas e incrementa su reproducibilidad, además de alcanzar una mayor efectividad, habida cuenta de ciertas limitaciones de la indización de PubMed³.

La aparición del paradigma conocido como «medicina basada en la evidencia»^{4,5} ha estimulado el desarrollo de los métodos para la identificación de referencias en las bases de datos biomédicas y ha impulsado la construcción de filtros de búsqueda llamados metodológicos. Su nombre deriva de que están orientados hacia la identificación de estudios con una alta validez interna⁶, sobre todo ensayos clínicos aleatorizados o revisiones sistemáticas, entre otros³. Dada su utilidad, algunos de ellos incluso se han incorporado a PubMed, a través de la opción «Clinical queries»¹.

Hasta el momento no se ha abordado la construcción de filtros geográficos, esto es, orientados a la identificación de información con una ubicación geográfica específica, a pesar de que las opciones actuales de PubMed no permiten identificar fácilmente el origen geográfico de la población estudiada. Las principales alternativas disponibles son la entrada en el cajón de búsqueda del criterio de restricción «Spain», ya sea como término libre, como filiación del primer autor o como lugar de publicación, y la restricción por el idioma español. Todas ellas se han revelado de forma independiente insuficientes para este propósito. Si las

restricciones por idioma o por lugar de publicación ignoran las numerosas publicaciones españolas en revistas internacionales, el uso de búsquedas centradas en la filiación de los autores presenta la limitación de que ese campo de PubMed recoge sólo la filiación del primer firmante. Además, se ha observado que no resulta particularmente eficaz, sobre todo debido a la falta de estandarización, lo que se traduce en una gran variabilidad en la información contenida en este campo, e incluso es frecuente que, al coincidir el país en el que se publica la revista y el de filiación del autor, este último termina por ser eliminado⁷⁻¹⁰.

Por ello, nos propusimos los siguientes objetivos: *a)* construir un filtro geográfico para la identificación en PubMed de estudios en los que haya participado población española, y *b)* evaluar su rendimiento de acuerdo con estándares metodológicos aceptados y respecto a criterios simples de búsqueda geográfica.

MÉTODOS

Construcción del filtro geográfico

Para la construcción del filtro geográfico se ha seguido un procedimiento de segunda generación, según la terminología propuesta por Jenkins³, caracterizado porque su construcción se basa en: *a)* la aplicación de conocimiento experto de la estructura de la base de datos (PubMed) y del tesoro o vocabulario controlado (MeSH), y *b)* la validación de la estrategia obtenida, mediante su comparación con un patrón de referencia³.

Para la identificación de los términos relevantes se revisaron las estrategias de búsqueda utilizadas por otros autores para identificar estudios realizados en España, en su mayoría estudios bibliométricos orientados a la evaluación de la producción científica española. Se contactó con los autores para la obtención de las estrategias de búsqueda originales¹¹⁻¹⁷. A partir de los elementos identificados, la estructura sintáctica del filtro y sus términos se definieron en 3 bloques: *a)* el término «España» en los principales idiomas usados en PubMed (inglés, español, alemán, francés, italiano); *b)* topónimos correspondientes a comunidades autónomas, provincias y sus capitales y otros de interés especial, excluidos los que pudieran corresponder a otros países, y *c)* acrónimos oficiales de los servicios de salud autonómicos.

Se especificó la búsqueda de los términos del primer bloque como texto libre o como filiación ([ad]). El motor de búsqueda de PubMed aplica al texto libre un proceso de mapeo automático de términos que localiza equivalencias de forma consecutiva en tablas de términos MeSH, revistas y autores, por este orden, y en el que no está incluido el campo de filiación¹. Los demás bloques se aplicaron únicamente al campo de filiación. Los términos se truncaron para permitir la recupera-

ción de gentilicios. Los distintos bloques se asociaron con el conector OR (anexo 1).

Evaluación del rendimiento del filtro

La evaluación del rendimiento del filtro geográfico se realizó mediante su comparación con una búsqueda manual (patrón de referencia) en una muestra de validación ad hoc³.

Identificación de la muestra y búsqueda manual

Para la obtención de la muestra se seleccionaron todas las referencias indizadas en PubMed bajo la palabra clave (MeSH) correspondiente a «infarto de miocardio» («myocardial infarction»), que en España supuso 39.400 muertes sólo en el año 2002 y cuya incidencia se encuentra en aumento en nuestro medio^{18,19}. Se seleccionaron sólo estudios clínicos (Limits: «Clinical Trials») y opciones por defecto (activación de «Explode» y desactivación de «Restrict to Major Subject Headings»). El formato seleccionado fue «Medline», que detalla entre 20 y 30 características para cada referencia en función de la información que proporcione la publicación (título, resumen, autores, filiación y términos MeSH, entre otros)¹. Las muestras de validación se exportaron y gestionaron utilizando el programa Reference Manager 10²⁰ para garantizar que el procedimiento fuera auditable.

Dos revisores (JM, JMV) llevaron a cabo la clasificación manual de las referencias de la muestra de validación («pertinentes» frente a «no pertinentes») a partir de la siguiente definición operativa: se consideraron «pertinentes» todas las referencias en las que: *a*) se indicaba específicamente que el estudio se había realizado en España; *b*) se mencionaba la población española como tal, *c*) el estudio había sido realizado por investigadores adscritos a centros de investigación españoles. Las referencias relacionadas con estudios multicéntricos (internacionales o no) en los que participara población española se consideraron también pertinentes. Todas las referencias que no cumplieran ninguno de los criterios anteriores se consideraron «no pertinentes». En caso de duda, cuando la información sobre la filiación de los autores, el lugar de intervención o la población era insuficiente, se identificaba, además, el texto completo del artículo para su evaluación. En algunos casos fue necesario identificar otros estudios realizados por los mismos autores u otras referencias externas a la muestra para poder clasificarlas como «pertinentes».

Esta clasificación se comparó con la obtenida tras la aplicación del filtro a la muestra (mediante el conector «AND»). Se incluyeron todas las referencias que se encontraban en PubMed desde su inicio, en 1966, hasta la fecha de realización de las búsquedas electrónicas, el 1 de agosto de 2005.

Estrategia simple

Para evaluar la utilidad relativa del filtro se comparó su rendimiento con el de una estrategia mucho más sencilla. De forma intuitiva, cabría suponer que ésta correspondería a la utilización del término libre «Spain». Se trata en realidad de un término ineficiente que podría incrementar artificialmente la bondad relativa del filtro geográfico²¹ y se seleccionó la estrategia «Spain[ad]», que recupera un volumen de referencias prácticamente 3 veces mayor¹.

Análisis

La evaluación de las búsquedas en PubMed guarda una estrecha similitud operacional con el estudio del rendimiento de las pruebas diagnósticas, en el que también se aplican los conceptos de sensibilidad (proporción de artículos pertinentes en la muestra de validación que son recuperados por el filtro) y especificidad (proporción de artículos no pertinentes en la muestra de validación que no son recuperados por el filtro) (fig. 1)^{22,23}. Se calcularon estos índices para la aplicación del filtro geográfico y de la estrategia simple en la muestra de validación.

Se comprobó la estabilidad de estas estimaciones mediante un análisis de sensibilidad en una segunda muestra de validación, independiente de la anterior. Se escogió una enfermedad diferenciada, pero también de elevada importancia sanitaria, y se seleccionaron ensayos clínicos indizados en PubMed bajo el término MeSH correspondiente a la enfermedad pulmonar obstructiva crónica («Pulmonary Disease, Chronic Obstructive»).

RESULTADOS

La aplicación del criterio Spain[ad] al total de la base de datos proporcionó 109.344 referencias. La aplicación del filtro proporcionó 59.142 referencias adicionales para un total de 168.486 referencias (un incremento del 54,1%). El bloque «España» y sus traducciones (texto libre y campo [ad]) proporcionaron 141.961 (84,3%) referencias, el correspondiente a otros topónimos proporcionó 26.499 referencias no incluidas entre las anteriores (15,7%) y el correspondiente a los servicios de salud otras 26 (0,1%).

Bajo el término MeSH «Myocardial Infarction» se encontraron indizadas 7.312 referencias correspondientes a estudios clínicos, todas ellas distintas excepto 2 que correspondían al mismo documento, indizado en una ocasión como ensayo clínico y en otra como revisión. La revisión manual de estas referencias clasificó 84 (1,15%) como pertinentes y 7.228 como no pertinentes (98,85%). Poco más de la mitad de todas ellas se publicó en español y en revistas españolas. Siete de las referencias pertinentes correspondieron a estudios

		Búsqueda manual	
		Referencias pertinentes	Referencias no pertinentes
Estrategia de búsqueda	Referencias recuperadas	a	b
	Referencias no recuperadas	c	d
		a+c	b+d

$$\text{Sensibilidad} = \frac{\text{Número de referencias pertinentes recuperadas por la estrategia de búsqueda}}{\text{Número total de referencias pertinentes en la muestra de validación}} \quad 100 = \frac{a}{(a+c)} \quad 100$$

$$\text{Especificidad} = \frac{\text{Número de referencias no pertinentes no recuperadas por la estrategia de búsqueda}}{\text{Número total de referencias no pertinentes en la muestra de validación}} \quad 100 = \frac{d}{(d+b)} \quad 100$$

Fig. 1. Índices para la evaluación del rendimiento de filtros y estrategias de búsqueda.

TABLA 1. Características seleccionadas de los estudios identificados (n = 84)

<i>Idioma</i>	
Español	46 (54,76%)
Inglés	36 (42,85%)
Portugués	2 (2,38%)
<i>Lugar de publicación</i>	
España	45 (53,57%)
Rev Esp Cardiol	32
Med Clin (Barc)	5
Rev Clin Esp	3
Rev Esp Anestesiología Reanim	2
Otras: Aten Prim, Nefrología, Rev Esp Med Nuclear	3
Estados Unidos	20 (23,81%)
J Am Coll Cardiol	5
Circulation	3
Am J Cardiol	2
Otras: Am Heart J, Anesthesiology, Arterioscler Thromb Vasc Biol, Catheter Cardiovasc Interv, Circ Res, Clin Cardiol, Clin Chem, IEEE Trans Biomed Eng, JAMA, Stroke	10
Europa	18 (21,43%)
European Heart J	5
Haemostasis	2
Internacional J Cardiology	2
Rev Port Cardiol	2
Otros: Ann Nutr Metab, Cardiology, Coron Artery Dis, Eur J Echocardiogr, Cardiovasc Surg (Torino), Lancet	7
Otros: New Zealand (Pharmacoeconomics)	1 (1,19%)
<i>Estudios multicéntricos</i>	
Estudios internacionales con participación española	7 (8,33%)
Estudios exclusivamente españoles	9 (10,71%)

internacionales en los que se incluyó a población española (tabla 1).

El filtro geográfico recuperó 75 referencias, 74 (98,67%) de ellas clasificadas como pertinentes en la búsqueda manual. La única referencia recuperada que no correspondía a un estudio realizado en España fue obtenida debido a un error de indización de PubMed. En el resumen, su filiación correspondía al hospital milanés «Niguarda Ca'Granda», incorrectamente indizado como «Niguarda Ca'Granada», pero que está correctamente transcrito en la publicación original²⁴. Por separado, el bloque «Spain» y sus traducciones (texto libre y campo [ad]) permitieron recuperar 53 referencias pertinentes (70,67%), el bloque correspondiente a comunidades, provincias y ciudades, 64 (86,47%), y el bloque de servicios de salud, ninguna. De todas ellas, 31 (41,90%) fueron recuperadas exclusivamente por uno de los 3 bloques, mayoritariamente por el de topónimos (n = 21; 67,74%), seguido del bloque «Spain» (n = 10; 32,26%). Las 43 restantes (58,2%) fueron recuperadas simultáneamente por más de un bloque.

El filtro geográfico no fue capaz de recuperar 10 referencias pertinentes, en 2 casos por causas atribuibles a una insuficiente definición de los términos del filtro (topónimos no incluidos en la estrategia): en una se mencionaba «Catalonia» en el resumen como el lugar en el que se realiza el estudio²⁵, y en otra se menciona en el campo filiación del autor un hospital ubicado en Cartagena²⁶. Las 8 restantes carecían de información suficiente, la mayoría de ellas (n = 6) publicadas entre 1970 y 1990, para las que sólo se disponía de la referencia (título, autores y publicación). Todas ellas se habían clasificado en la revisión manual como pertinentes sólo después de la búsqueda secundaria de los datos del estudio y/o la filiación de los autores a través

TABLA 2. Rendimiento diagnóstico de las estrategias electrónicas en relación con la búsqueda manual (patrón de referencia)

Rendimiento del filtro geográfico			
«Myocardial Infarction» [MeSH] Limits: Clinical Trial		Revisión manual	
		Pertinentes	No pertinentes
Filtro geográfico	Recuperados	74	1
	No recuperados	10	7.227
Sensibilidad del 88,10%		Especificidad del 99,99%	
Rendimiento de la estrategia simple			
«Myocardial Infarction» [MeSH] Limits: Clinical Trial		Revisión manual	
		Pertinentes	No pertinentes
Filtro geográfico	Recuperados	38	0
	No recuperados	45	7.229
Sensibilidad del 45,78%		Especificidad del 100%	

de otras referencias externas a la muestra de validación.

En conjunto, el filtro demostró una sensibilidad del 88,10% y una especificidad del 99,99% (tabla 2). La estrategia simple «Spain[ad]», en cambio, recuperó tan sólo 38 referencias pertinentes, con unos valores de sensibilidad del 45,8% y de especificidad del 100% (tabla 2). La bondad de la estimación del rendimiento del filtro geográfico en la segunda muestra de validación (478 referencias) incluso resultó óptima (sensibilidad y especificidad del 100%).

Al finalizar, se informó de todos los errores de indización identificados a la Biblioteca Nacional de Medicina de Estados Unidos.

DISCUSIÓN

Se ha construido un filtro geográfico de uso sencillo para la identificación de estudios realizados en España en la base de datos Medline (PubMed). La evaluación de su rendimiento ha demostrado que tiene una especificidad excelente y que ha alcanzado una sensibilidad muy elevada, superior a la de un criterio de búsqueda simple.

Fortalezas y debilidades del procedimiento de evaluación

Las inconsistencias en la indización y la presencia de errores de transcripción afectan a la recuperación de las referencias en Medline²⁷⁻³¹. El procedimiento de validación del filtro se ha diseñado precisamente teniendo en cuenta las limitaciones de la indización y superando las limitaciones inherentes a los métodos empleados hasta el momento, basados en la realiza-

ción de una búsqueda manual en una única muestra de validación ad hoc, habitualmente definida a partir de las referencias publicadas en revistas seleccionadas³. Esta práctica es controvertida por la arbitrariedad que introduce esta selección, pero especialmente porque utiliza una referencia externa a la base de datos, de forma que los errores de indización afectan a la estimación del rendimiento del filtro³². Tomando un caso extremo, una referencia que no se haya incluido en la base de datos no podrá ser recuperada por el filtro evaluado, ni por ningún otro, por lo que no parece razonable atribuir esta limitación a la estrategia de búsqueda. El procedimiento de evaluación ha tomado la muestra de validación directamente de PubMed, lo que permite, además, que el procedimiento de búsqueda manual sea fácilmente reproducible, algo que en cualquier caso garantiza la disponibilidad de las muestras de validación en un archivo de Reference Manager. La construcción de una muestra suplementaria no relacionada con la anterior y basada en una entidad clínica diferenciada para evaluar la estabilidad de la evaluación del rendimiento del filtro también incrementa la confianza en la estimación.

Debe señalarse que por el momento no hay estándares para la selección de filtros de comparación o de muestras de evaluación o validación de filtros. La estrategia simple fue seleccionada por tratarse de una estrategia intuitiva, que triplica el rendimiento de una estrategia más sencilla (Spain como texto libre), y por ser la más descrita en la literatura científica, mientras que la interpretabilidad de los resultados ha guiado la selección de las muestras de validación.

Fortalezas y debilidades del filtro propuesto

Los términos se han incluido a partir de un procedimiento sistemático basado en la definición de criterios explícitos: geográficos, con localizaciones geográficas de tamaño decreciente, lingüísticos, con idiomas que cubren un porcentaje superior al 99,99% de la producción científica española³³, y administrativos, con la inclusión de los diversos servicios de salud. Esta misma característica hizo que no se pudiera recuperar electrónicamente 2 referencias pertinentes, cuyo criterio de localización no coincidía con ninguno de los criterios predefinidos. Ambos términos se han incorporado al filtro con posterioridad (anexo 1) y a buen seguro se irán enriqueciendo con las propias aportaciones y sugerencias de los usuarios.

El filtro geográfico propuesto incluye un bloque con el nombre de los servicios de salud autonómicos que, tras ser analizado, no incrementó pero tampoco disminuyó la sensibilidad ni la especificidad en las muestras utilizadas. Atendiendo a su efecto neutro, su pequeño tamaño, el hecho de que cuando se emplea este bloque por separado permite la identificación de 1.519 referencias hasta la fecha de estudio y la experiencia pre-

via de los autores con estos términos, se ha considerado conveniente mantener su inclusión. La construcción modular del filtro permite, en cualquier caso, que dicha parte se pueda eliminar sin dificultad, si así se desea.

Las bases de datos tienden hacia una mayor normalización de los datos y han mejorado su indización con el tiempo. Ello puede suponer que mejore el rendimiento de algunas búsquedas más sencillas que, en cualquier caso, están ya insertadas en el filtro actual.

Las estrategias de búsqueda alternativas basadas en criterios de idioma o de lugar de publicación, si bien son útiles en estudios bibliométricos, tienen escasa aplicabilidad al objetivo de identificar estudios realizados en la población española. Acotar las búsquedas a las entradas en español obviaría la enorme producción española en inglés (66.855 documentos, sólo en el período 1994-2002³³). Inversamente, más de la mitad de todas las referencias contenidas en PubMed en español corresponden a publicaciones de fuera de España, que en su gran mayoría presentan estudios realizados en el continente americano.

Cada vez con mayor fuerza, los investigadores publican sus resultados en revistas internacionales (en el presente estudio, en más del 45% de los casos)²⁷. Paralelamente, algunas de las revistas que se publican en España tienen un marcado carácter internacional, como *Methods and Findings in Experimental and Clinical Pharmacology* o *Test*. Otras, que tuvieron en origen un carácter local, aspiran cada vez más a su reconocimiento como revista internacional, como la propia REVISTA ESPAÑOLA DE CARDIOLOGÍA, en la que se ha publicado más de un tercio de todos los estudios. En cualquier caso, parece claro que el idioma y el lugar de publicación no son fiables para la identificación de estudios realizados en España.

Aplicaciones del filtro geográfico

Sin duda, una de las aplicaciones más útiles de este filtro geográfico es la de permitir identificar los estudios con una mayor validez externa para la aplicación de las observaciones a la población española, entendida como generabilidad o semejanza proximal^{6,34-36}. Es bien sabido que la generabilidad de los hallazgos depende de aspectos muy diversos de los estudios, que incluyen desde las características clínicas, sociodemográficas e incluso genéticas de la propia población estudiada hasta aspectos relacionados con el proceso de atención, como los estilos de práctica clínica y los cuidados y tratamientos complementarios administrados⁶. Si los anteriormente indicados podrían ser de especial relevancia en estudios relacionados con el riesgo cardiovascular (piénsese, p. ej., en la controvertida aplicación de las diferentes tablas y ecuaciones obtenidas en distintas poblaciones), otros aspectos, como los legislativos o los relacionados con la organización de los

servicios sanitarios, podrían ser aún más relevantes en estudios sobre el abandono del hábito tabáquico³⁷ (legislación y modelos sociales de comportamiento) o que evalúen la efectividad de una intervención domiciliaria (accesibilidad y modelo de atención)³⁸. Tanto en uno como otro caso, la ubicación geográfica de un estudio se puede considerar como una variable indirecta (*proxy*) de muchas de estas variables y puede ser de gran utilidad para la evaluación de su validez externa.

Específicamente, la disponibilidad de un filtro geográfico de elevada sensibilidad y especificidad para la identificación en PubMed de estudios realizados en la población española abre la puerta a su aplicación sistemática en la práctica clínica. Su aplicación permitirá identificar de forma rápida y fiable información pertinente sobre los efectos de los procedimientos y tratamientos en condiciones más próximas a las de su uso real. Esta información también es especialmente pertinente para la definición de trayectorias clínicas, que despiertan un creciente y justificado interés y en las que la consideración de las peculiaridades asistenciales locales son determinantes para una implementación efectiva³⁹⁻⁴². Desde el punto de vista cognitivo, la recuperación de información pertinente facilita el aprendizaje, así como la mejora de la práctica clínica tras la aplicación de la nueva información^{43,44}.

Tal vez por ello, sorprende la falta de apartados específicos en las guías de práctica clínica en los que se sintetice la información disponible sobre el uso de los procedimientos diagnósticos y terapéuticos en el propio medio en el que se quiere aplicar^{45,46} y en el que los filtros geográficos como el que aquí se ha presentado pueden tener un papel destacado³⁸. Esta contextualización de la información, antes que contradecir los postulados de la medicina basada en la evidencia, los reafirma, pues la consideración de la validez externa de la información identificada no es si no el paso inmediato a su aplicación a la práctica clínica⁴.

Finalmente, este filtro geográfico brinda una herramienta útil para aplicaciones bibliométricas, hasta el momento centradas predominantemente en la base de datos del Institute of Scientific Information de Philadelphia (ISI) y que de esta manera se podrían complementar a través de PubMed^{15,33}. De hecho, una publicación reciente describe el uso independiente de un filtro geográfico similar al presentado en nuestro artículo para este fin⁴⁷.

CONCLUSIONES

Se dispone de un filtro eficiente para la identificación de estudios realizados en la población española en PubMed, cuyo aplicación permite hallar las referencias con mayor validez externa para la práctica clínica en España y que, empleado de forma sistemática, puede ser de gran utilidad en la elaboración de guías de práctica y en la definición de trayectorias clínicas.

AGRADECIMIENTOS

Los autores quieren agradecer al Centro Cochrane Iberoamericano la orientación recibida sobre la estructura del filtro, a los Dres. Alberto Miguel-Dasit y Luis Martí-Bonmatí habernos facilitado generosamente la estrategia de búsqueda diseñada por ellos y a Olatz Garín sus valiosos comentarios.

BIBLIOGRAFÍA

1. National Library of Medicine (US). Medline [Internet]. Bethesda: National Library of Medicine (US); 2005.
2. Drazen JM, Curfman GD. Public access to biomedical research. *N Engl J Med*. 2004;351:1343.
3. Jenkins M. Evaluation of methodological search filters—a review. *Health Info Libr J*. 2004;21:148-63.
4. Guyatt G, Rennie D, editors. *Users guides: essentials of evidence-based clinical practice*. Chicago: American Medical Association; 2002.
5. Evidence-Based Medicine Working Group. Evidence-based medicine. A new approach to teaching the practice of medicine. *JAMA*. 1992;268:2420-5.
6. Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. *BMJ*. 2001;323:42-6.
7. Loria A, Arroyo P. Language and country preponderance trends in MEDLINE and its causes. *J Med Libr Assoc*. 2005;93:381-5.
8. Rosselli D. Geography of biomedical publications. *Lancet*. 1999;354:517.
9. Takei N. Geography of biomedical publications. *Lancet*. 1999;354:516.
10. Tutarel O. Geographical distribution of publications in the field of medical education. *BMC Med Educ*. 2002;2:3.
11. Baños JE, Bosch F, Bigorra J, Guardiola E. Difusión internacional de los ensayos clínicos realizados en España: un análisis a través de su publicación en revistas científicas. *Med Clin (Barc)*. 1994;102:441-5.
12. Barenys M, Recasens MA, Camps J, Salas-Salvado J. La producción científica española sobre alimentación y nutrición de 1987 a 1990. *Med Clin (Barc)*. 1992;99:9-12.
13. Guardiola E. Algunas consideraciones sobre la búsqueda de estudios españoles en bases de datos internacionales. *Med Clin (Barc)*. 1992;99:795-6.
14. Martí J, Bonfill X, Urrutia G, Lacalle JR, Bravo R. Identificación y descripción de ensayos clínicos publicados en revistas españolas de medicina general y de medicina interna durante el período 1971-1995. *Med Clin (Barc)*. 1999;112 Supl 1:28-34.
15. Pestaña A. El MedLine como fuente de información bibliométrica de la producción española en biomedicina y ciencias médicas. Comparación con el Science Citation Index. *Med Clin (Barc)*. 1997;109:506-11.
16. Ramos Rincon JM, Belinchon RI, Gutiérrez RF. La producción científica española respecto a la infección por el virus de la inmunodeficiencia humana/sida. Un estudio a través de MedLine (1991-1999). *Med Clin (Barc)*. 2001;117:645-53.
17. Miguel-Dasit A, Martí-Bonmatí L, Aleixandre R, Sanfeliu P, Valderrama JC. Producción española sobre diagnóstico por la imagen en cardiología y radiología (1994-1998). *Rev Esp Cardiol*. 2004;57:806-14.
18. Instituto de Información Sanitaria. Ministerio de Sanidad y Consumo. Mortalidad por causa de muerte 1991-2002. España y comunidades autónomas. Documento en línea [accedido 10 Ago 2005]. Disponible en: http://www.msc.es/Diseno/sns/sns_sistemas_informacion.htm
19. Sans S, Puigdefabregas A, Paluzie G, Monterde D, Balaguer-Vintro I. Increasing trends of acute myocardial infarction in Spain: the MONICA-Catalonia Study. *Eur Heart J*. 2005;26:505-15.
20. Reference Manager Professional Edition (versión 10). Stamford: Thompson ResearchSoft; 2002.
21. Castellote Varona FJ. Idioma y Medline. *An Med Interna*. 2002;19:152-3.
22. Allison JJ, Kiefe CI, Weissman NW, Carter J, Centor RM. The art and science of searching MEDLINE to answer clinical questions. Finding the right number of articles. *Int J Technol Assess Health Care*. 1999;15:281-96.
23. Kagolovsky Y, Moehr JR. Current status of the evaluation of information retrieval. *J Med Syst*. 2003;27:409-24.
24. Merlini PA, Bauer KA, Oltrona L, Ardissino D, Spinola A, Cattaneo M, et al. Thrombin generation and activity during thrombolysis and concomitant heparin therapy in patients with acute myocardial infarction. *J Am Coll Cardiol*. 1995;25:203-9.
25. Grupo de Investigación del estudio PREMISE. Prevención secundaria del infarto de miocardio y calidad de vida relacionada con la salud. *Med Clin (Barc)*. 2002;119:9-12.
26. Melgarejo-Moreno A, Galcera-Tomas J, García-Alberola A. Prognostic significance of bundle-branch block in acute myocardial infarction: the importance of location and time of appearance. *Clin Cardiol*. 2001;24:371-6.
27. Jiménez-Contreras E, Delgado López-Cózar E, Ruiz-Pérez R, Fernández VM. Impact-factor rewards affect Spanish research. *Nature*. 2002;417:898.
28. Bador P, Lafouge T. Rédaction des adresses sur les publications. Un manque de rigueur défavorable aux universités françaises dans les classements internationaux. *Presse Med*. 2005;34:633-6.
29. Fernandez E, García AM. Accuracy of referencing of Spanish names in Medline. *Lancet*. 2003;361:351-2.
30. Ruiz-Pérez R, Delgado López-Cózar E, Jiménez-Contreras E. Spanish personal name variations in national and international biomedical databases: implications for information retrieval and bibliometric studies. *J Med Libr Assoc*. 2002;90:411-30.
31. Sezer O, Bohme A. Hazards of translation of non-English medical publications in Medline. *Lancet*. 2000;355:1280.
32. Sampson M, Zhang L, Morrison A, Barrowman NJ, Clifford TJ, Platt RW, et al. An alternative to the hand searching gold standard: Validating methodological search filters using relative recall. *BMC Med Res Method*. 2006;6:33.
33. Cami J, Suñén-Pinol E, Méndez-Vásquez R. Mapa bibliométrico de España 1994-2002: biomedicina y ciencias de la salud. *Med Clin (Barc)*. 2005;124:93-101.
34. Campbell D. Relabeling internal and external validity for applied social scientists. En: Trochim W, editor. *Advances in quasi-experimental design and analysis*. San Francisco: Jossey-Bass; 1986. p. 67-77.
35. Campbell DT. Relabeling internal and external validity for applied social scientists. *New Directions for Program Evaluation*. 1986;31:67-77.
36. Rothman KJ, Greenland S. *Modern epidemiology*. 2nd ed. Philadelphia: Lippincot Williams & Wilkins; 1998.
37. Tappin DM, Lumsden MA, Gilmour WH, Crawford F, McIntyre D, Stone DH, et al. Randomised controlled trial of home based motivational interviewing by midwives to help pregnant smokers quit or cut down. *BMJ*. 2005;331:373-7.
38. Morcillo C, Valderas JM, Aguado O, Delas J, Sort D, Pujadas R, et al. Evaluación de una intervención domiciliaria en pacientes con insuficiencia cardíaca. Resultados de un estudio aleatorizado. *Rev Esp Cardiol*. 2005;58:618-25.
39. Haynes RB, McKibbon KA, Wilczynski NL, Walter SD, Werre SR. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ*. 2005;330:1179.
40. Every NR, Hochman J, Becker R, Kopecky S, Cannon CP. Critical pathways. A review. *Circulation*. 2000;101:461-5.
41. Cardozo L, Aherns S. Assessing the efficacy of a clinical pathway in the management of older patients hospitalized with congestive heart failure. *J Health Qual*. 1999;21:12-7.

42. Cluzeau FA, Burgers JS, Brouwers M, Grol R, Mäkelä M, Littlejohns P, et al. The AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. *Quality Safety Health Care*. 2003;12:18-23.
43. Grad R, Pluye P, Segal B. Assessing the impact of clinical information-retrieval technology in a family practice residency. *J Eval Clin Pract*. 2005;11:576-86.
44. Pluye P, Grad RM, Dunikowski LG, Stephenson R. Impact of clinical information-retrieval technology on physicians: a literature review of quantitative, qualitative and mixed methods studies. *Int J Med Inform*. 2005;74:745-68.
45. Schattner A, Fletcher RH. Research evidence and the individual patient. *QJM*. 2003;96:1-5.
46. Rothwell P. External validity of randomised controlled trials: to whom do the results of this trial apply? *Lancet*. 2005;365:82-93.
47. Escudero C, Milán I, Posada de la Paz M. Análisis de la producción científica española sobre enfermedades raras: 1990-2000. *Med Clin (Barc)*. 2005;125:329-32.

ANEXO 1. Uso del filtro geográfico

El filtro geográfico que aquí se presenta incorpora las enmiendas indicadas en la discusión para mayor facilidad de uso (para obtener el empleado en la evaluación, basta con suprimir los términos << OR catalonia[ad] >> y << OR (cartagen* [ad] NOT indias [ad])>>). Para su utilización práctica basta con insertarlo en el cajón de búsqueda estándar de PubMed (<http://www.pubmed.org>). Recomendamos copiarlo desde un documento electrónico, como el presente, porque la transcripción es muy poco eficiente y probablemente generará numerosos errores. Próximamente será accesible desde la página web de Red IRYSS (<http://www.rediryss.net>). Una vez introducido, bastará con combinarlo con cualquier otro término que se desee para acotar la búsqueda.

Conviene notar que la construcción del filtro geográfico se ha basado en criterios a priori, por lo que su uso en PubMed genera actualmente un mensaje de error que no interfiere en absoluto con el proceso de búsqueda. Se debe a que no reconoce algunos términos como «Iruna» (Pamplona), sencillamente porque aún no se ha indizado ninguna referencia que lo incluya (PubMed no reconoce la letra «ñ», que sustituye de forma automática por «n»).

(spain OR espagne OR espana OR spagna OR (spain[ad] OR espagne[ad] OR espana[ad] OR spanien[ad] OR spagna[ad]) OR (catalunya[ad] OR catalonia[ad] OR catalogne[ad] OR cataluna[ad] OR catala[ad] OR barcelon*[ad] OR tarragona[ad] OR lleida[ad] OR lerida[ad] OR girona[ad] OR gerona[ad] OR sabadell[ad] OR hospitalet[ad] OR l'hospitalet[ad]) OR (valencia*[ad] OR castello*[ad] OR alacant[ad] OR alicant*[ad] OR murcia*[ad] OR cartagen*[ad] NOT indias[ad])) OR (andalu*[ad] OR sevilla*[ad] OR granad*[ad] OR huelva[ad] OR almeria[ad] OR cadiz[ad] OR jaen[ad] OR malaga[ad] OR cordoba[ad] NOT argentin*[ad])) OR (extremadura[ad] OR caceres[ad] OR badajoz[ad] OR madrid[ad]) OR (castilla[ad] OR salamanca[ad] OR zamora[ad] OR valladolid[ad] OR segovia[ad] OR soria[ad] OR palencia[ad] OR avila[ad] OR burgos[ad]) OR (leon[ad] NOT (france[ad] OR clermont[ad] OR rennes[ad] OR lyon[ad] OR USA[ad] OR mexic*[ad])) OR (galicia[ad] OR gallego[ad] OR compostela[ad] OR vigo[ad] OR corun*[ad] OR ferrol[ad] OR orense[ad] OR ourense[ad] OR pontevedra[ad] OR lugo[ad]) OR (oviedo[ad] OR gijon[ad] OR asturia*[ad]) OR (cantabr*[ad] OR santander[ad]) OR (vasco[ad] OR euskadi[ad] OR basque[ad] OR bilbao[ad] OR bilbo[ad] OR donosti*[ad] OR san sebastian[ad] OR vizcaya[ad] OR bizkaia[ad] OR guipuzcoa[ad] OR gipuzkoa[ad] OR alava[ad] OR araba[ad] OR vitoria[ad] OR gasteiz[AD]) OR (navarr*[ad] OR nafarroa[ad] OR pamplona[ad] OR iruna[ad] OR irunea[ad] OR (logron*[ad] OR rioj*[ad]) OR (aragon*[ad] OR zaragoza[ad] OR teruel[ad] OR huesca[ad]) OR (mancha[ad] OR ciudad real[ad] OR albacete[ad] OR cuenca[ad]) OR (toledo[ad] NOT (ohio[ad] OR us[ad] OR usa[ad] OR OH[ad])) OR (guadalajara[ad] NOT mexic*[ad]) OR (balear*[ad] OR mallorca[ad] OR menorca[ad] OR ibiza[ad] OR eivissa[ad]) OR (palmas[ad] OR lanzarote[ad] OR canari*[ad] OR tenerif*[ad]) OR (ceuta[ad] OR melilla[ad])) OR (osasunbide*[ad] OR osakidetza[ad] OR insalud[ad] OR sergas[ad] OR catsalut[ad] OR sespa[ad] OR osasunbidea[ad] OR imsalud[ad] OR sescam[ad] OR ib-salut[ad])