

Cómo encontrar los mejores artículos

Jorge Navarro-Pérez

INCLIVA. Universitat de València. CIBERESP. Centro de Salud Salvador Pau. Valencia

El desarrollo de la medicina basada en la evidencia (MBE) ha sido un factor determinante en el acercamiento de los clínicos a la información biomédica. Tres son los problemas relacionados con la bibliografía científico médica: la obsolescencia, la dispersión y el exceso de información. El crecimiento de las publicaciones periódicas no es capaz de absorber el propio crecimiento de la actividad investigadora y su consiguiente oferta de manuscritos. A ello hay que sumar el crecimiento paralelo de la literatura secundaria (índices, resúmenes y bibliografía) y de la denominada «literatura gris» u opaca, de difícil acceso comercial (informes de investigación, prepublicaciones, actas de congresos y tesis doctorales).

Los sistemas de recuperación de la información permiten evitar el acceso a información ya superada o de escasa relevancia. En la actualidad, el acceso se realiza a través de bases de datos, a su vez gestionadas por distribuidores, algunos de ellos gratuitos. Existen también gestores de bases de datos bibliográficas (como Reference Manager) que posibilitan la completa recuperación de la información, así como un posterior tratamiento informático, incluido el análisis bibliométrico.

Los principales repertorios biomédicos son Index Medicus (editado por la National Library of Medicine de Estados Unidos), Excerpta Medica (editado por la Excerpta Medica Foundation, en Ámsterdam) y Current Contents y Science Citation Index (ambos editados por el Institute for Scientific Information de Filadelfia). Sus bases de datos más conocidas

son MEDLINE (Index Medicus), EMBASE (Excerpta Medica) y SciSearch (Science Citation Index y parte de Current Contents). Según el tipo de estudio buscado utilizaremos distintas fuentes: Biblioteca Cochrane para revisiones sistemáticas, Trip Database para guías de práctica clínica o Clinical Evidence para preguntas terapéuticas.

MEDLINE (popularizada por PubMed) y EMBASE son las bases de datos más utilizadas, pues contienen casi el 100 % de las publicaciones biomédicas, complementándose: EMBASE cubre el 72 % del total de revistas biomédicas y MEDLINE el 65 % de este, solapándose en un 35 %. Así como muchos buscadores bibliográficos incorporan filtros predefinidos (estrategia avanzada), el manejo de MEDLINE o EMBASE requiere un entrenamiento adecuado (tabla 1).

ESTRUCTURA DE UNA REFERENCIA BIBLIOGRÁFICA

La búsqueda bibliográfica se apoya en la correcta delimitación de las distintas partes de una referencia bibliográfica. Cada registro (*record*) se divide en categorías de información denominadas campos (*fields*). La mayoría de los campos son de longitud fija. Los campos básicos son: resumen (*abstract*, ab), dirección (*address*, ad), autor (*author*, au), país de publicación (*country of publication*, cp), idioma original (*language of article*, la), tipo de publicación (*publication type*, pt), año

Tabla 1. Fuentes de información y tipos de estudio

	GPC	RS	Tratamiento	Diagnóstico	Pronóstico	Etiología
Biblioteca Cochrane/Biblioteca Cochrane Plus	-	++++	++++	+	+	-
Trip Database	++++	+++	++	++	++	++
Clinical Evidence	-	++	++++	-	+	-
MEDLINE/EMBASE	+	+++	+++	+++	+++	+++
DARE	+	+++	+++	++	+	+

DARE: Database of Abstracts of Reviews of Effects; EMBASE: Excerpta Medica; GPC: guías de práctica clínica; MEDLINE: Index Medicus; RS: revisiones sistemáticas.

de publicación (*publication year*, py), descriptor de materia (*medical headings*, mh), fuente (*source*, so) y título (*title*, ti). Los campos LA (*language* [idioma]), PT (*publication type* [tipo de publicación]) y PY (*publication year* [año de publicación]) se conocen como «campos límite», pues posibilitan una recuperación más precisa. En PubMed se consigna la identificación de cada campo entre corchetes (por ejemplo, [pt]; si son varios, separados por punto y coma: [ti;ab;mesh]). Un ejemplo sería «diabetic neuropathy» [ti;ab;mesh] (que el término «neuropatía diabética» esté en el título o en el resumen o en el tesoro).

ETAPAS DE LA BÚSQUEDA BIBLIOGRÁFICA

Toda búsqueda no estructurada es una frivolidad que puede conllevar una gran pérdida de tiempo, y por tanto debe evitarse. Solo MEDLINE contiene más de 10 millones de referencias, con un crecimiento exponencial (en el año 2013 se registraron 1 157 397, frente a las 199 903 del año 1966). Si en la búsqueda aparecen demasiados registros se habla de «ruido», y si aparecen demasiado pocos, de «silencio». Nuestro objetivo es la precisión. Para ello, es aconsejable seguir los pasos que describimos a continuación.

Delimitación del tema

Un protocolo adecuado de búsqueda bibliográfica comienza con una pregunta clínica. Dicha pregunta clínica puede ser de varios tipos (etiología, prevención, diagnóstico, pronóstico, tratamiento o impacto económico), cada uno de los cuales requiere diferentes fuentes y estrategias. La pregunta de investigación se debe traducir a un lenguaje documental, elaborando un perfil o una ecuación de búsqueda; es decir, identificando los términos específicos y combinándolos adecuadamente en una fórmula que el sistema de recuperación de la información de la base de datos utilizada pueda descifrar. En algunos metabuscadores como Trip Database (www.tripdatabase.com) se emplea la estrategia de búsqueda basada en la MBE de formato PICO.

Para pasar de una pregunta clínica genérica a una formulada de forma específica con el método PICO, se deben tener en cuenta los siguientes componentes:

- Paciente: grupos de edad, estadio de la enfermedad, comorbilidad, etc.
- Intervención: intervención, factor pronóstico, agente etiológico, prueba diagnóstica, etc.
- Comparación: se refiere a la alternativa a la intervención a estudio, tales como tratamiento habitual o placebo, ausencia de un factor de riesgo, ausencia de

agente etiológico, patrón de oro o de referencia de una prueba diagnóstica, etc.

- Resultados (*outcomes*): variables de resultado clínicamente importantes en el caso de estudios sobre eficacia, pronóstico o etiología, y estimadores de rendimiento o validez diagnóstica (sensibilidad, especificidad, coeficientes de probabilidad y valores predictivos) en el caso de los estudios de pruebas diagnósticas.

Selección de fuentes

Debido a su gran popularidad, hay que enfatizar que un acceso exclusivo a MEDLINE es insuficiente (con una estrategia perfectamente diseñada, localizaríamos el 65 % de la información pertinente). Acudiendo a MEDLINE/PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>), EMBASE (<http://www.embase.com>) y Current Contents (www.accesowok.fecyt.es), y realizando posteriormente una búsqueda manual (en la bibliografía de los artículos localizados), posiblemente podamos localizar casi el 100 % de las referencias pertinentes.

Formulación del perfil de búsqueda

Según el lenguaje booleano, se conocen como enlaces lógicos de intersección, unión y exclusión las expresiones AND (y), OR (o) y NOT (no), respectivamente. El primero busca lo común de dos subconjuntos, el segundo busca todo lo comprendido en dos subconjuntos y el tercero incluye solo un subconjunto, eliminando de él otro subconjunto (figura 1). También existen enlaces sintácticos (WITH [con], NEAR [cerca], etc.) y los llamados comodines de búsqueda (los signos ? y *).

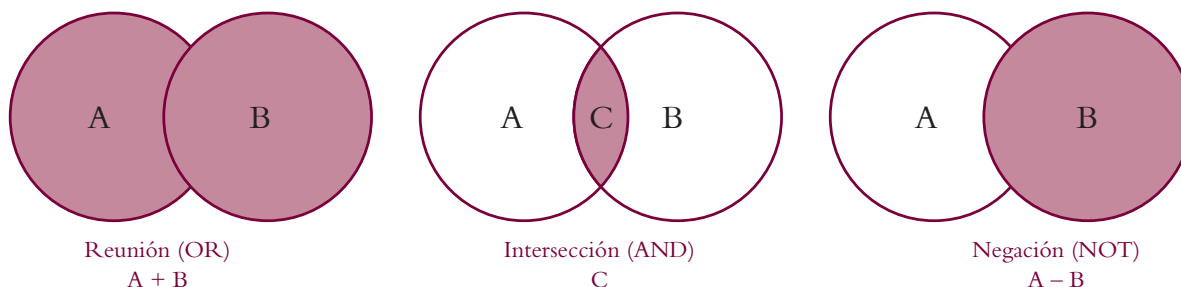
Corrección de la estrategia

La tasa de precisión (porcentaje de artículos pertinentes del total de artículos localizados) nos permite saber si el perfil de búsqueda fue adecuado. Debe ser superior al 80 %. En caso negativo, hay que desechar la estrategia y reformularla.

Publicación de la estrategia

En la edición de todo trabajo de investigación en el que sea relevante la bibliografía (revisiones sistemáticas, sobre todo), debe especificarse con detalle la estrategia de búsqueda empleada, así como la fecha de actualización de cada base de datos (en MEDLINE, el código UD [*update code*] define el mes y el año de actualización de cada referencia).

Figura 1. Operadores booleanos según los diagramas de Venn



ESTRUCTURA DE UNA ESTRATEGIA DE BÚSQUEDA

Si la base de datos contiene un tesoro (como MEDLINE: MeSH, *medical subject headings*), lo adecuado es seleccionar los términos a partir de este y no mediante una búsqueda libre. Por ejemplo, si queremos localizar referencias sobre el tratamiento de la nefropatía diabética, a través del MeSH de MEDLINE obtenemos, como un descriptor reconocido en el tesoro, el término «diabetic nephropathies». Asimismo, nos indica la evolución de su indización y las relaciones semánticas con otros términos. Para acceder al ámbito del tratamiento de nefropatía diabética, especificaríamos en MEDLINE (Ovid) que localice solo lo relativo a tratamiento: «diabetic nephropathies.th» (en PubMed, entre corchetes: «diabetic nephropathies [th]»); «th» es un filtro temático de MEDLINE que corresponde a *therapy* (tratamiento). «Diabetic nephropathies» aparece dentro de «diseases categories», en «kidney diseases», y también en «diabetes complications». El término «diabetic nephropathies» fue indizado en 1965.

Otros *subheadings* incluidos en MEDLINE son: efectos adversos (*adverse effects*, [ae]; complicaciones (*complications*, [co]); contraindicaciones (*contraindications*, [ct]); diagnóstico (*diagnosis* [dt]); tratamiento farmacológico (*drug therapy*, [dt]); educación (*education*, [ed]); enfermería (*nursing*, [nu]); prevención (*prevention and control*, [pc]) o uso terapéutico (*therapeutic use of drug*, [tu]).

Una vez seleccionados los descriptores que más específicamente reflejan lo que buscamos, de forma que cada uno de ellos constituya un subconjunto de referencias, se eligen adecuadamente los enlaces lógicos y sintácticos. Existen enlaces que depuran la búsqueda (*and*, *with*, *near*, *not* y el paréntesis) y enlaces que la amplían (*or*, *?* y ***). No es lo mismo poner «hypertension OR obesity AND diabetes» que «hypertension OR (obesity AND diabetes)». «Hypertension», «obesity» y «diabetes» son tres subconjuntos de referencias. En el primer caso, la base de datos puede, aleatoriamente, realizar primero la unión de «hypertension» y «obesity» o la intersección de «obesity» y «diabetes». En el segundo caso, primero realiza la

intersección de «obesity» y «diabetes» (dado que está entre paréntesis) y, finalmente, el nuevo subconjunto es sumado con «hypertension».

Tres son los campos que poseen mayor especificidad en la búsqueda: el título (TI, *title*), el resumen (AB, *abstract*) y los descriptores (en MEDLINE: MeSH; en Ovid: SH; si especificamos MeSH Major Topic (MJME), la recuperación es más precisa). Por ejemplo: «diabetes in ti,ab,mesh».

Existen los denominados campos límite: idioma (LA o LG [Ovid], *language*), año de publicación (PY, *publication year*) y tipo de publicación (PT, *publication type*). Este último campo es muy importante desde el punto de vista de la MBE, pues permite localizar ensayos clínicos (*clinical trial*), guías de práctica clínica (*practice-guideline*), revisiones (*review*) o metaanálisis (*meta-analysis*), entre otros tipos de documentos. Muchos repertorios permiten elaborar automáticamente estrategias avanzadas, incluyendo el uso de campos límite.

Una forma sencilla de realizar una búsqueda bibliográfica es cruzar primero una enfermedad («type 2 diabetes»), un tratamiento («GLP-1 receptor agonists») y un tipo de documento («meta-analysis») mediante el operador AND.

Pongamos por caso que deseamos saber qué hay publicado sobre nefropatía diabética en español entre los años 2010 y 2014. Con el símbolo # identificamos el número de búsqueda:

- #1 Diabetic nephropathies [ti;ab;sh]
- #2 Spanish [lg]
- #3 2010-2014 [py]
- #4 #1 AND #2 AND #3

Ahora, en dicho subconjunto, deseamos saber si hay ensayos clínicos, revisiones o metaanálisis:

- #5 Clinical trial [pt]
- #6 Review [pt]
- #7 Meta-analysis [pt]
- #8 #5 OR #6 OR #7
- #9 #4 AND #8

FILTROS METODOLÓGICOS

Podemos, pues, construir una estrategia de búsqueda, simple o avanzada (operadores booleanos), utilizando los filtros predefinidos en la base de datos MeSH de PubMed. Tras la localización de un conjunto de referencias sobre un determinado término, MEDLINE dispone de unos filtros temáticos (*subheadings*) que permiten una mayor precisión en la búsqueda. Pero la MBE, dado que plantea un análisis de la literatura biomédica mediante la evaluación de estudios específicos o mediante revisiones sistemáticas, propone el uso de «filtros metodológicos» o «filtros de calidad» como estrategias de búsqueda adaptadas (procedimientos diagnósticos, intervenciones terapéuticas, etiología, pronóstico, revisiones sistemáticas, ensayos clínicos controlados, guías de práctica clínica, etc.). Cada perfil de búsqueda requiere un tipo de estudio diferente:

- Intervención: ensayo clínico aleatorizado.
- Diagnóstico: estudio transversal o de cohortes de pruebas diagnósticas.
- Pronóstico: estudio de cohortes o rama control del ensayo clínico.
- Etiología: estudio de casos y controles o estudio de cohortes.

Desde PubMed Clinical Queries se accede a un motor de búsqueda especializado dirigido a médicos clínicos que permite recuperar los artículos más relevantes en relación con la etiología, el diagnóstico, el tratamiento y el pronóstico de las enfermedades, a partir de filtros metodológicos.

Además, proporciona otras dos categorías o filtros: sensibilidad y especificidad. La primera permite recuperar un buen número de artículos relevantes, con menor grado de pertinencia; la segunda facilita la recuperación de una menor cantidad de artículos, pero con un grado alto de pertinencia (tabla 2).

Describimos a continuación las más relevantes estrategias de búsqueda.

Estudios controlados

El mejor término simple es «clinical-trial in pt» (en Ovid: «clinical-trial.pt»).

Se obtiene un máximo de sensibilidad con:

- «randomized-controlled-trial in pt».
- «drug therapy in mesh».
- «therapeutic use in mesh».
- «random * in ti,ab,mesh».

Se obtiene un máximo de especificidad con:

- «(double and blind*) in ti,ab,mesh».
- «placebo * in ti,ab,mesh».

Pruebas diagnósticas

El mejor término simple es «sensitivity in ti,ab,mesh» (en Ovid: «sensitivity.ti,ab,sh»).

Tabla 2. Filtros metodológicos para la búsqueda de información biomédica

Categoría	Optimizada para...	Sensibilidad/especificidad	PubMed equivalente
Terapia	Sensibilidad	99/74 %	"randomized controlled trial" [PT] OR "drug therapy" [SH] OR "therapeutic use" [SH:NOEXP] OR "random*" [WORD]
	Especificidad	57/97 %	(double [WORD] AND blind* [WORD]) OR placebo [WORD]
Diagnóstico	Sensibilidad	92/73 %	"sensitivity and specificity" [MESH] OR "sensitivity" [WORD] OR "diagnosis" [SH] OR "diagnosis use" [SH] OR "specificity" [WORD]
	Especificidad	55/98 %	"sensitivity and specificity" [MESH] OR ("predictive" [WORD] AND "value*" [WORD])
Etiología	Sensibilidad	82/70 %	"cohort studies" [MESH] OR "risk" [MESH] OR ("odds" [WORD] AND "ratio*" [WORD]) OR ("relative" [WORD] AND "risk" [WORD]) OR "case" control*" [WORD] OR case-control studies [MESH]
	Especificidad	40/98 %	"case-control studies" [MH:NOEXP] OR "cohort studies" [MH:NOEXP]
Pronóstico	Sensibilidad	92/73 %	"incidence" [MESH] OR "mortality" [MESH] OR "follow-up studies" [MESH] OR "mortality" [SH] OR prognos* [WORD] OR predict* [WORD] OR course [WORD]
	Especificidad	49/97 %	prognosis [MH:NOEXP] OR "survival analysis" [MH:NOEXP]

MH: MJME (MeSH Major Topic); PT: publication type; SH: MeSH (medical subject headings); WORD: texto libre.

Se obtiene un máximo de sensibilidad con:

- «sensitivity and specificity».
- «sensitivity in ti,ab,mesh».
- «diagnosis in mesh».
- «radionuclide imaging in mesh».
- «diagnostic use in mesh».
- «specificity city in ti,ab,mesh».

Se obtiene un máximo de especificidad con:

- «exp sensitivity and specificity».
- «(predictive and value*) in ti,ab,mesh».

Se obtiene un máximo de sensibilidad con:

- «incidence in mesh».
- «exp mortality».
- «follow-up studies».
- «mortality in mesh».
- «prognos* in ti,ab,mesh».
- «predict* in ti,ab,mesh».
- «course in ti,ab,mesh».

Se obtiene un máximo de especificidad con:

- «prognosis in mesh».
- «survival-analysis».

Estudios etiológicos

El mejor término simple es «risk in ti,ab,mesh».

Se obtiene un máximo de sensibilidad con:

- «exp cohort-studies».
- «exp risk».
- «(odds and ratio*) in ti,ab,mesh».
- «(relative and risk) in ti,ab,mesh».
- «(case and control*) in ti,ab,mesh».

Se obtiene un máximo de especificidad con:

- «case-control-studies».
- «cohort-studies».

Estudios pronósticos

El mejor término simple es «exp cohort-studies».

Revisiones sistemáticas y metaanálisis

Una estrategia con gran precisión, aunque con alto rango de sensibilidad (93,6 %) es la siguiente:

1. review.ab.
2. review.pt.
3. meta-analysis.ab.
4. meta-analysis.pt.
5. meta-analysis.ti.
6. or/1-5.
7. letter.pt.
8. comment.pt.
9. editorial.pt.
10. or/7-9.
11. (Estrategia de búsqueda principal.)
12. 6 not 10.
13. 11 and 12.

BIBLIOGRAFÍA RECOMENDADA

- Navarro J. Técnicas para la realización de una búsqueda bibliográfica. *Av Diabetol* 2009;25:376-81.
- Navarro J. Búsqueda y análisis de la literatura biomédica. Madrid: 1999.
- Aleixandre-Benavent R, Fuentes de información en ciencias de la salud en Internet. *Panace@* 2011;12(33):112-20.
- Haynes RB, McKibbin KA, Wilczynski NL, Walter SD, Werre SR; the Hedges Team. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005;330:1179.
- Elaboración de guías de práctica clínica en el Sistema Nacional de Salud. Manual metodológico. Madrid: Ministerio de Sanidad y Consumo; 2007. Disponible en: <http://www.guiasalud.es>.
- Haynes RB, Wilczynski N, McKibbin KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc* 1994;1(6):447-58.
- McKibbin KA, Walker-Dilks CJ, Wilczynski N, Haynes RB. Beyond ACP Journal Club: how to harness MEDLINE for review articles. *ACP J Club* 1996;124(3):A12-3.
- White VJ, Glanville JM, Lefebvre C, Sheldon TA. A statistical approach to designing search filters to find systematic reviews: objectivity enhances accuracy. *J Inform Sci* 2001;27(6):357-70.
- Boynton J, Glanville J, McDavid D, Lefebvre C. Identifying systematic reviews in MEDLINE: developing an objective approach strategy design. *J Inform Sci* 1998;24:137-57.