

Importancia de las bases de datos en la búsqueda bibliográfica. Primer paso de una revisión sistemática

M. SUBIRANA^a, I. SOLÀ^b, J.M. GARCÍA^c, A. GUILLAMET^d, E. PAZ^e, I. GICH^f, G. URRÚTIA^g,
y GRUPO DE TRABAJO IMACE

^aDiplomada en Enfermería. ^bLicenciado en Psicología. Servei d'Epidemiologia i Salut Pública, Hospital de Sant Pau.

^cDiplomado en Enfermería. Centro Cochrane Iberoamericano, Hospital de Sant Pau. ^dDiplomada en Enfermería. Escola Universitària d'Infermeria i Fisioteràpia Blanquerna, Universitat Ramon Llull. ^eDiplomada en Enfermería. Unitat de Cirurgia Cardíaca, Hospital de Sant Pau. ^fDoctor en Medicina y Cirugía. Servei d'Epidemiologia i Salut Pública, Hospital de Sant Pau.

^gLicenciado en Medicina y Cirugía. Centro Cochrane Iberoamericano, Hospital de Sant Pau. Barcelona. España

Resumen

El crecimiento del conocimiento científico y de las bases de datos a disposición del investigador hace que cada vez sean necesarias búsquedas más complejas, para poder recopilar la cantidad de información más completa y de mejor calidad sobre el tema que se pretende investigar.

Este estudio analiza el número de referencias identificadas en Cumulative Index to Nursing and Allied Health Literature (CINAHL) y en Medline y cuantifica su relevancia para llevar a cabo una revisión sistemática. Las variables a analizar son: número de referencias recuperadas con y sin *abstract*, grado de solapamiento entre las bases de datos, accesibilidad y grado de relevancia de las referencias.

Se han recuperado 116 artículos, 49 (42,2%) en CINAHL y 67 (57,8%) en Medline, sin que se haya producido solapamiento. De las referencias recuperadas el 60,3% (70) incluye *abstract*, 13 (26,5%) en CINAHL y 57 (85,1%) en Medline ($p < 0,005$). El 62,9% (73) de las referencias

identificadas puede obtenerse en bibliotecas españolas, 42,5% (31) en CINAHL y 57,5% (42) en Medline ($p = ns$). Se han considerado relevantes 23 (18,9% del total), de las cuales 7 (30,4%) se han identificado en CINAHL y 16 (69,6%) en Medline ($p = ns$), la concordancia entre la clasificación de relevancia en función del título y del *abstract* muestra una coincidencia en el 72,9% de las referencias ($p < 0,005$).

Cuando se plantee una búsqueda bibliográfica relacionada con temas de enfermería es necesario que se consulte tanto CINAHL como Medline.

Palabras Clave: Bases de datos bibliográficas. Medline. CINAHL. Enfermería. Búsqueda electrónica. Revisión sistemática.

Importance of the database in the literature search: the first step in a systematic review

The growth in scientific knowledge and databases available to researchers means that increasingly complex searches are needed to collate information which is both comprehensive and which provides the best quality information on the specific research topic.

This study analyzes the number of references identified in CINAHL and MEDLINE and assesses their relevance for undertaking a systematic review. The variables analyzed were the number of references retrieved with and without abstract, the degree of overlap between the databases, and accessibility and degree of relevance of the references.

Correspondencia: Sra. M. Subirana.

Servei d'Epidemiologia Clínica i Salut Pública.
Hospital de la Santa Creu i Sant Pau (Casa de Convalescència).
St. Antoni Maria Claret 171. 08041 Barcelona. España
Correo electrónico: msubirana@hsp.santpau.es

Grupo IMACE: X. Bonfill, A. Capitán, A. Chamarro, J.M. García, A. Guillamet, E. Paz, J.L.R. Martín, I. Solà, D. Soler, M. Subirana, G. Urrútia.

Proyecto financiado por la Agencia de Evaluación de Tecnologías Sanitarias (AETS). Número de expediente 01/10089.

Aceptado para su publicación el 15-7-2002.

A total of 116 references were retrieved: 49 (42.2%) in CINAHL and 67 (57.8%) in MEDLINE; no overlap was found. Of the retrieved references, 60.3% (70) included abstract, 13 (26.5%) in CINAHL and 57 (85.1%) in MEDLINE ($p < 0.005$). A total of 62.9% (73) of the identified references were available in Spanish libraries: 42.5% (31) in CINAHL and 57.5% (42) in MEDLINE ($p = ns$). Twenty-three references were considered relevant (18.9% of the total), of which seven (30.4%) were identified in CINAHL and 16 (69.6%) in MEDLINE ($p < 0.005$). The agreement between the classification of relevance on the basis of the title and abstract was 72.9% ($p < 0.005$).

To obtain the best results, bibliographical searches on nursing-related subjects should combine CINAHL and MEDLINE.

Key words: Bibliographical databases. MEDLINE. CINAHL. Nursing. Electronic Search. Systematic Review.

Introducción

Con el crecimiento exponencial de la literatura científica y del número de bases de datos al alcance del investigador y del usuario, se ha puesto de manifiesto la necesidad de llevar a cabo búsquedas bibliográficas cada vez más complejas y completas. Estas búsquedas deben servir para recopilar la mayor, y también la mejor, cantidad posible de conocimiento disponible sobre el tema que se pretende investigar. En consecuencia, aparece la pregunta de cuál es la mejor fuente a la que el investigador puede acceder para hacer una búsqueda bibliográfica de mayor calidad. Si bien la bibliografía de enfermería cuenta con el referente de Cumulative Index to Nursing and Allied Health Literature (CINAHL), en muchas ocasiones los temas de interés en el campo de enfermería se extienden a diversos ámbitos y a líneas de investigación generalmente multidisciplinarias, por lo que se hace necesario consultar también otras bases de datos genéricas (mayoritariamente Medline y EMBASE) para mejorar el resultado de las búsquedas.

La base de datos CINAHL, producida por Information Systems en EE.UU., tiene su origen electrónico en 1982. Incluye temática relacionada con enfermería y salud, aparte de otros ámbitos específicos como la bibliotecnia en ciencias de salud, educación, ciencias de la conducta o administración. Desde 1984 también incluye las publicaciones de la National League for Nursing y la American Nurses' Association. A partir del inicio de la década de los noventa comienza a indexar tesis de enfermería, capítulos de libros y material audiovisual relacionadas con este ámbito. CINAHL indexa en la actualidad aproximadamente 1.200 revistas biomédicas, de ciencias sociales y de salud de pa-

cientes y usuarios, de las cuales 700 son títulos de enfermería¹; de éstos, 144 (20%) están también indexados en Medline². Anualmente la base de datos se amplía con unas 300.000 nuevas referencias, e incluye principalmente, aunque no de forma exclusiva, títulos en lengua inglesa. Para facilitar las búsquedas, CINAHL incorpora *Thesaurus* (lenguaje controlado) con descriptores (término que forma parte de un lenguaje controlado), organizados jerárquicamente, de los cuales un 70% coincide con los términos de Medline³.

Medline, iniciada en 1966, está producida por la National Library of Medicine de EE.UU. y constituye la base de datos de referencias bibliográficas de uso más extendido. Incluye las publicaciones del Index Medicus, Index to Dental Literature y del International Nursing Index. Abarca todas las áreas de la biomedicina y contiene referencias exclusivamente de revistas biomédicas, 3.700, de las que aproximadamente el 8% son revistas de enfermería³. La base de datos se actualiza con unas 500.000 nuevas referencias anuales⁵. El 69% de las referencias incluye un *abstract*², y aproximadamente doce descriptores MeSH para cada referencia.

Algunos autores han analizado cuál es el nivel de solapamiento^{2,4,6}, el criterio de asignación de los descriptores² o la cobertura de revistas⁷ entre CINAHL y Medline, y Burnham (1993) incluye además la base de datos EMBASE en la comparación³. Brenner compara la asignación de descriptores para Medline y CINAHL², y llega a la conclusión de que Medline incluye el doble de descriptores que CINAHL, lo que permitiría hacer búsquedas más específicas. Burnham y Shearer comparan las tres bases de datos, y recomiendan el uso de Medline y CINAHL para una búsqueda más minuciosa de la bibliografía³. Okuma abarca la comparación de Medline y CINAHL desde la perspectiva de los estudiantes de enfermería⁷, y recomienda CINAHL como base de datos para estudiantes. Watson y Perrin centran su estudio en el solapamiento de títulos de revistas recuperados en la búsqueda, y concluyen que es necesaria la consulta tanto en CINAHL como en Medline, ya que cada base de datos recupera títulos de revistas únicos⁴. Brazier y Begley, además de comparar las dos bases de datos, analizan la accesibilidad de las referencias recuperadas en Irlanda. Los resultados obtenidos llevan a concluir que para una búsqueda exhaustiva de la bibliografía se debe consultar tanto CINAHL como Medline, y que por otra parte, tres cuartas partes de la bibliografía recuperada puede ser obtenida en un máximo de dos días⁶. Finalmente Coma et al llevan a cabo una comparación de las bases de datos CINAHL y Medline y un análisis de la accesibilidad de los documentos en nuestro medio, y concluyen que CINAHL ofrece un mayor número de registros apropiados para el campo de la enfermería aunque presenta una mayor dificultad de accesibilidad con respecto a Medline¹. Cabe destacar que, pese a la variedad de resultados recogidos en los estudios previos sobre el tema, todos los

autores aconsejan la consulta de ambas bases de datos para una búsqueda bibliográfica más completa y detallada. Otro aspecto importante a tener en cuenta es la necesidad de ajustar la estrategia de búsqueda a los *Thesaurus* específicos de cada base de datos².

La búsqueda bibliográfica es una de las características diferenciales entre una revisión sistemática (basada en una metodología científica que busca minimizar los sesgos) y una revisión narrativa (refleja el punto de vista subjetivo de expertos en una área determinada). En la revisión sistemática la búsqueda de la bibliografía es global, por lo que es muy importante definir una estrategia de búsqueda para identificar los estudios que sean relevantes y seleccionar las fuentes de información necesarias para identificar los estudios.

Siguiendo esta misma línea, y aprovechando la realización de una revisión sistemática de los instrumentos que miden la adecuación de la actividad de enfermería y su repercusión sobre los resultados en salud que están realizando los autores, nuestro objetivo es analizar y comparar el número de referencias identificadas en las bases de datos CINAHL y Medline, así como cuantificar su relevancia.

Método

Se diseñó una estrategia de búsqueda basada en el tema de la revisión que se aplicó a las bases de datos CINAHL y Medline y fue adaptada a los *Thesaurus* específicos de cada una de ellas. Para consultar CINAHL se accedió mediante el distribuidor SilverPlatter disponible en la red de bibliotecas de la Universitat de Barcelona, y el acceso a Medline se realizó a través de PubMed. Las bases de datos han sido consultadas desde su inicio (1982 para CINAHL y 1966 para Medline) hasta mayo de 2002 y en ningún caso se estableció ningún tipo de límite según el idioma o el año de publicación del artículo para la recuperación de las referencias:

Medline

#1 "Nursing Staff, Hospital"[MESH]
#2 "Nursing" [MESH]
#3 [#1 OR #2]
#4 "Workload" [MESH]
#5 [#3 AND #4]
#6 "Measurement" [MESH]
#7 [#5 AND #6]
#8 "Validation"
#9 [#5 AND #8]

CINAHL

#1 "Workload-Measurement"/all subheadings
#2 "Nurses"/all subheadings
#3 "Nursing-Staff-Hospital"/all subheadings
#4 [#2 OR #3]
#5 [#1 AND #4]
#6 "Instrument-Validation"/all subheadings
#7[#4AND#6]

Para cada base de datos se han analizado diversos aspectos. Se han cuantificado el número de referencias recuperadas, el número de referencias coincidentes en ambas bases de datos, y el número de referencias que se recuperan con *abstract*.

Con la finalidad de conocer la accesibilidad a los documentos recuperados a partir de la búsqueda, se ha establecido un criterio de disponibilidad de las re-

vistas en las que se encuentran los artículos identificados en la búsqueda: a) publicaciones disponibles en nuestro centro; b) publicaciones disponibles en Cataluña; c) publicaciones disponibles en España, y d) publicaciones disponibles en el extranjero.

A partir de la clasificación de las revistas en cada una de estas categorías, se ha llevado a cabo una estimación del coste aproximado de los artículos relevantes para la revisión.

Para conocer la relevancia de las referencias identificadas en cada base de datos, se han escogido cinco revisores externos conocedores del tema de la revisión sistemática. Cada revisor recibió un listado con el título de cada una de las referencias recuperadas en la búsqueda, sin identificar la base de datos de procedencia o información sobre la disponibilidad del *abstract*. El revisor debía pronunciarse sobre la relevancia de cada referencia en relación con el tema de la revisión sistemática, estableciéndose la siguiente clasificación: 1: relevante (claramente adecuado al tema); 2: poco relevante (ofrece dudas sobre la adecuación), y 3: irrelevante (claramente inadecuado).

Las diferentes valoraciones de cada revisor se contrastaron para ver cuál era el grado de concordancia entre los revisores dos a dos. Fueron excluidas las valoraciones de los revisores más discordantes. A partir de los datos de los dos revisores que obtuvieron un grado de concordancia más alto, se computaron sus puntuaciones para obtener una valoración final de relevancia para cada referencia basándose solamente en el título.

Dado que ninguna de las dos bases de datos ofrece el *abstract* en todas las referencias, se pretende objetivar si existen diferencias en la valoración de la relevancia a partir del título o del *abstract*. Con esta finalidad, el análisis de la relevancia se repitió con los *abstracts* de las referencias recuperadas. Así, se enviaron los *abstracts* sin ninguna información referente al título del artículo o de su procedencia a los dos revisores, para que juzgaran la relevancia del resumen utilizando el mismo criterio aplicado a los títulos. Se realizaron los mismos análisis de concordancia que con los títulos. También se valoró si alguna de las dos bases de datos obtiene resultados de más relevancia en función del título o del *abstract*.

Análisis estadísticos

Para comparar las dos bases de datos se utilizó la prueba de la χ^2 . La concordancia entre los revisores se analizó mediante el índice kappa (κ ; este índice toma valores desde el 0, total desacuerdo, hasta el 1, total acuerdo). Para saber si una base de datos recupera más referencias que la otra, si tiene una mejor accesibilidad o si las referencias son más relevantes, se llevó a cabo la prueba de Mann-Whitney. Para conocer la posible diferencia entre la clasificación de relevancia

para el título y para el *abstract* se aplicó la prueba de Wilcoxon. Para saber si la clasificación de relevancia entre título y *abstract* difiere en función de la base de datos se realizó la prueba de Kruskal-Wallis.

Resultados

La búsqueda bibliográfica realizada permitió la recuperación de 116 referencias, 49 (42,2%) en CINAHL y 67 (57,8%) en Medline, sin que hubiera ningún solapamiento entre las bases de datos. De las referencias recuperadas en ambas bases de datos el 60,3% (70) incluía el *abstract* en CINAHL, el 26,5% (13) y en Medline el 85,1% (57) ofrecía el *abstract*. Medline recuperó más referencias con *abstract* que CINAHL ($\chi^2 = 40,5$; gl = 1; $p < 0,005$) (tabla 1).

El 62,9% (73) de las referencias recuperadas puede obtenerse en bibliotecas españolas, el 42,5% (31) se recuperó en CINAHL y el 57,7% (42), en Medline ($p = ns$) (tabla 2).

La tabla 3 muestra la clasificación de la relevancia de las referencias según los cinco revisores en cada base de datos. El análisis de concordancia entre los re-

visores se refleja en la tabla 4. Se ha calculado el valor del índice (para cada revisor y se han seleccionado para el análisis definitivo de la clasificación de la relevancia los dos revisores que han obtenido un valor de κ más alto (revisor 3 y revisor 5; $\kappa = 0,50$; $p < 0,05$).

En la clasificación de la relevancia en función de los títulos ($n = 116$) se han obtenido 44 referencias relevantes, 19 que podrían ofrecer duda y 53 irrelevantes. La clasificación detallada para cada base de datos se muestra en la tabla 5 ($p = 0,2$).

La clasificación de la relevancia para las referencias que ofrecen *abstract* ($n = 70$) se distribuye en 32 referencias relevantes, 15 que podrían ofrecer duda y 23 irrelevantes. La clasificación detallada para cada base de datos se muestra en la tabla 6 ($p = 0,2$).

El análisis para conocer la concordancia entre la clasificación de la relevancia en función del título y del *abstract* muestra una coincidencia en el 72,9% de las referencias ($\chi^2 = 47,7$; gl = 4; $p < 0,005$) (tabla 7). El análisis de variabilidad entre las clasificaciones de la relevancia a partir del título y a partir del *abstract* muestra que a partir de los *abstracts* se hacen más valoraciones de relevancia que para los títulos ($Z = -2,7$; $p < 0,05$).

TABLA 1. Referencias recuperadas para cada base de datos

	CINAHL	Medline	Total
Referencias recuperadas	49 (42,2%)	67 (57,8%)	116
Solapamiento	—	—	0
Referencias con abstract	13 (26,5%)*	57 (85,1%)*	70 (60,3%)

*Los porcentajes se refieren a la proporción respecto a la base de datos.

TABLA 2. Accesibilidad a las referencias recuperadas para cada base de datos

	CINAHL	Medline	Total
Disponible en el centro	4 (36,4%)	7 (63,6%)	11 (9,5%)
Disponible en Cataluña	16 (40%)	24 (60%)	40 (34,5%)
Disponible en España	11 (50%)	11 (50%)	22 (18,9%)
Disponible en el extranjero	18 (41,9%)	25 (58,1%)	43 (37,1%)

*Los porcentajes se refieren a la proporción respecto al total de referencias recuperadas.

Discusión

Los resultados demuestran que, para la búsqueda bibliográfica del tema de la revisión sistemática, es imprescindible consultar tanto CINAHL como Medline, ya que la proporción de referencias recuperadas para este tema se acerca a la mitad en ambas bases de datos sin que exista solapamiento entre ellas. En este sentido nuestros resultados son similares a los de otros autores^{2,3,6,7} que también recomiendan el uso de ambas bases de datos para una búsqueda completa. Sin embargo, vale la pena destacar que la disponibilidad de ambas bases de datos no es comparable. Medline es de acceso gratuito y sólo requiere de un ordenador con conexión a Internet, mientras que CINAHL sólo se puede consultar mediante suscripción, lo que puede representar una limitación importante para su consulta. Otro aspecto a tener en cuenta es que el porcentaje de referencias con *abstract* que recupera Medline es mayor (85,1% en Medline frente a sólo 26,5% en CINAHL). Nuestros resultados están en la línea de los de Brazier y

TABLA 3. Clasificación de la relevancia de las referencias para cada revisor y para cada base de datos

	Revisor 1			Revisor 2			Revisor 3			Revisor 4			Revisor 5		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
CINAHL	34	27	6	23	23	21	26	22	19	28	10	29	28	14	25
Medline	21	25	3	15	22	12	12	26	11	10	17	22	13	10	26
Total	55	52	9	38	45	33	38	48	30	38	27	51	42	24	51

1: relevante; 2: poco relevante; 3: irrelevante.

TABLA 4. Grado de concordancia entre los revisores en el análisis de relevancia del título

	Revisor 1	Revisor 2	Revisor 3	Revisor 4
Revisor 2	0,15 ^a			
Revisor 3	0,27 ^a	0,29 ^a		
Revisor 4	0,08	0,12	0,12	
Revisor 5	0,23 ^a	0,21 ^a	0,50 ^a	0,16 ^a

^aSignificativa, $p < 0,05$. En negrita el valor de κ más alto.

TABLA 5. Clasificación de la relevancia en función del título para cada base de datos

	Procedencia del artículo		
	Medline	CINAHL	Total
Relevante	30	14	44
Poco relevante	10	9	19
Irrelevante	27	26	53
Total	67	49	116

TABLA 6. Clasificación de la relevancia en función del abstract para cada base de datos

	Procedencia del artículo		
	Medline	CINAHL	Total
Relevante	28	4	32
Poco relevante	12	3	15
Irrelevante	17	6	23
Total	57	13	70

TABLA 7. Concordancia entre la clasificación de relevancia a partir del título y del abstract

Clasificación de relevancia de los títulos	Clasificación de relevancia de los abstracts		
	Relevante	Poco relevante	Irrelevante
Relevante	24 (34,3%)	1 (1,4%)	1 (1,4%)
Poco relevante	3 (4,3%)	6 (8,6%)	1 (1,4%)
Irrelevante	5 (7,1%)	8 (11,4%)	21 (30%)

Begley, que en su estudio recuperan el 69% de referencias en Medline y el 43% en CINAHL con *abstract*⁶, o los de Coma et al que recuperan referencias con *abstract* en el 74,7% para Medline y en el 47,8% para CINAHL¹. En nuestro estudio no se ha producido ningún solapamiento entre las dos bases de datos, lo que contrasta con los resultados de otros estudios en los que se produce un solapamiento del 16,4%¹, del 18%³ e incluso del 20%⁶, lo que podría explicarse por la especificidad del tema de nuestra revisión sistemática, mientras que en los otros estudios la búsqueda bibliográfica se realiza al mismo tiempo para más de un tema. En cualquier caso, nuestros datos refuerzan la necesidad de consultar las dos bases de datos.

La valoración de la relevancia de las referencias que se recuperan a partir de la búsqueda resulta muy im-

portante para garantizar la calidad de una revisión sistemática, por lo que la información recogida en las bases de datos es esencial. A pesar de que la clasificación hecha a partir de los títulos o de los *abstracts* coincide en un 73%, destaca el hecho de que se producen más valoraciones como relevantes a partir de los *abstracts* (45,7%) que solamente a partir de los títulos (37,1%), probablemente porque los *abstracts* ofrecen información mucho más detallada. Dada la gran utilidad del *abstract* para decidir la inclusión de una referencia en la revisión sistemática, en nuestro estudio, Medline ofrece un mejor rendimiento, sobre todo si se tiene en cuenta que, en nuestra búsqueda, CINAHL sólo permitió recuperar el 26,5% de las referencias con *abstract*.

Para nuestra búsqueda se ha podido obtener una copia de casi dos tercios de las referencias identificadas en bibliotecas españolas, un resultado bastante aceptable si lo comparamos con el de Coma et al, que tan sólo obtenía un 17,8% de las referencias de su búsqueda en bibliotecas españolas¹. A pesar de todo hay que tener en cuenta que el 37% de los artículos se debe obtener en el extranjero, lo que encarece mucho el coste de la investigación, dado que un artículo solicitado en el extranjero cuesta entre unos 20 y 70 euros, mientras que un artículo obtenido en España cuesta 6 euros de media.

En conclusión, aconsejamos que cuando se plantee una búsqueda bibliográfica relacionada con temas de enfermería se consulte tanto CINAHL como Medline. A pesar del atractivo de CINAHL para la investigación específica en enfermería, los problemas que presenta todavía esta base de datos limitan su potencial utilidad, por lo que no debería ser la única base de datos utilizada para una revisión sistemática. Por el contrario, la consulta de Medline proporcionará el complemento que la búsqueda necesita para optimizar su calidad (exhaustividad y relevancia de las referencias recuperadas), sobre todo teniendo en cuenta el carácter cada vez más interdisciplinario de la investigación en enfermería.

Bibliografía

1. Coma I, De Dios R, Montcusí C. Fuentes de información para enfermería: comparación entre las bases de datos CINAHL y Medline. *Metas* 1999;14:21-7.
2. Brenner SH. CINAHL and Medline: a comparison of indexing practices. *Bull Med Libr Assoc* 1989;77:366-71.
3. Burnham J, Shearer B. Comparison of CINAHL, EMBASE and Medline databases for the nurse researcher. *Med Ref Serv Q* 1993;12:45-57.
4. Watson MM, Perrin R. A comparison of CINAHL and Medline CD-ROM in four allied health areas. *Bull Med Libr Assoc* 1994;82:214-6.
5. SilverPlatter International. SilverPlatter Information [revista electrónica]. New Virginia, EE.UU., 2002 [consultado 5/6/2002]. Disponible en: <<http://www.silverplatter.com/catalog>>
6. Brazier H, Begley CM. Selecting a database for literature searches in nursing: Medline or CINAHL? *J Adv Nurs* 1996;24:868-75.
7. Okuma E. Selecting CD-ROM databases for nursing students: a comparison of Medline and CINAHL. *Bull Med Libr Assoc* 1994;82:25-9.