



Similitudes y diferencias entre *PubMed*, *Embase* y *Scopus*

Lic. Rubén Cañedo Andalia ^I

Dr. Mario Nodarse Rodríguez ^{II}

^I Licenciado en Información Científico-Técnica y Bibliotecología. Centro Provincial de Información de Ciencias Médicas. Universidad de Ciencias Médicas de Holguín. Holguín. Cuba.

^{II} Doctor en Estomatología. Diplomado en Gestión de Información en Salud. Centro Nacional de Información de Ciencias Médicas-Infomed. La Habana, Cuba.



Fecha de actualización

13 de junio de 2014

Correspondencia

Lic. *Rubén Cañedo Andalia*. Grupo de Alfabetización Informacional. Centro Provincial de Información de Ciencias Médicas. Universidad de Ciencias Médicas Holguín. Ave. V. I. Lenin No. 4 e/ Aguilera y Agramonte. CP. 80100. Holguín. Cuba. Correo electrónico: ruben@infomed.sld.cu

Derechos de copia



Copyright: © Centro Provincial de Información de Ciencias Médicas. Universidad de Ciencias Médicas Holguín. Contribución de acceso abierto, distribuida bajo los términos de la Licencia Creative Commons Reconocimiento-No Comercial-Compartir Igual 3.0, que permite consultar, reproducir, distribuir, comunicar públicamente y utilizar los resultados del trabajo en la práctica, así como todos sus derivados, sin propósitos comerciales y con licencia idéntica, siempre que se cite adecuadamente el autor o los autores y su fuente original. La imagen del cabezal de la portada se tomó de: <http://www.ncbi.nlm.nih.gov/pubmed/>, www.embase.com y www.scopus.com

Síntesis

La búsqueda y recuperación de información es una etapa fundamental del quehacer científico. La búsqueda de información en bases de datos bibliográficas internacionales prominentes es una práctica común en este contexto. Pero, antes de realizar una búsqueda de información en una de estas, es necesario conocer qué contienen y cuán similares y diferentes son entre ellas, sus facilidades, herramientas, alcance, tipología documental procesada, entre otros aspectos de interés. Ahora bien, obtener esta información es ciertamente difícil, porque existen múltiples fuentes de información con datos dispersos, parciales y faltos de exactitud. Por ello, se realizó una profunda investigación documental y se consultó a los expertos que laboran en el desarrollo de bases de datos como *PubMed-Medline*, *Embase* y *Scopus* con vistas a ofrecer a los profesionales de la salud del país, y en especial a los investigadores, una caracterización precisa de ellas.

Introducción

Un conocimiento profundo de las características de las bases de datos más prominentes en un sector o área de la actividad científica no es solo un requisito indispensable para los profesionales de la información sino también de científicos, académicos y tecnólogos, porque ello es un prerrequisito importante para la obtención de la información especializada que cada uno de ellos requiere para su quehacer profesional, sus proyectos e investigaciones, así como para su actualización sistemática.

Saber dónde buscar qué forma parte del conjunto de competencias en información que debe poseer un profesional. Sin embargo, esta es una de las competencias más difíciles de adquirir, debido a que solo un examen profundo de los principales recursos de información disponibles en un área del conocimiento, su uso extenso e intenso y la comparación de sus resultados obtenidos, permite desarrollar el conocimiento necesario sobre las fortalezas y debilidades de cada uno de ellos. El logro de este conocimiento a partir de la experiencia personal es un proceso complejo y costoso en tiempo y esfuerzo.

A causa de esta necesidad, y como resultado de ciertas condiciones favorables que permitieron a los autores de la presente contribución, acceder a bases de datos como *Embase* y *Scopus*, se decidió desarrollar un visión panorámica y sintética de dichas bases en comparación con *Medline* y *PubMed* con vistas a proveer a los profesionales de la salud con un conocimiento detallado y actual sobre sus principales similitudes y diferencias. Presentar la información obtenida en este difícil análisis es el objetivo del presente artículo.

Métodos

Se realizaron múltiples búsquedas, tanto en bases de datos estudiadas como en las plataformas de productos del National Council for Biotechnology Information de la National Library of Medicine of United States of America; y de Elsevier, ésta última productora de *Embase* y *Scopus*. Se revisaron numerosos catálogos promocionales y documentos no publicados sobre el tema. Se comparó la información suministrada por cada fuente y se sintetizó. Los aspectos contradictorios se aclararon mediante nuevas búsquedas de información y en no pocos casos por medio de la consulta a los expertos que atienden los productos analizados.

Resultados

PubMed y Medline

PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>), accesible por medio de *Entrez* -el sistema integrado de búsqueda y recuperación de información textual, desarrollado por el Centro Nacional para la Información Biotecnológica (NCBI por sus siglas en inglés)- de la Biblioteca Nacional de Medicina de los Estados Unidos, una entidad con rango de instituto nacional de salud-, es el resultado de un proyecto conjunto con un grupo de casas editoras en el campo de la biomedicina para proporcionar a sus usuarios una herramienta de búsqueda capaz ofrecer acceso, tanto a las referencias bibliográficas como a los textos completos de los materiales recuperados, es el recurso bibliográfico gratuito más utilizado en el área de la salud en Internet. El anuncio oficial de la apertura de *PubMed* se realizó el 26 de junio de 1997.

Cubre los campos de la medicina, la enfermería, la estomatología, la veterinaria, la gestión de salud, las ciencias preclínicas y algunas áreas de las ciencias de la vida. Sus archivos contienen más de 23 millones de registros procedentes de casi 25 mil revistas desde el año 1809 hasta la fecha (6 de febrero de 2014); *Medline* (*Medline* 1966-hasta el presente + *OldMedline* 1946-1965)-, su componente más emblemático, atesora más de 20 millones y medio de registros desde 1946 hasta el presente -el 89,52% del total reunido por *PubMed* en su conjunto- procedente de más de 5 600 revistas, seleccionadas mediante un riguroso proceso de evaluación.¹

Contiene referencias en su colección de alrededor de 220 000 revisiones sistemáticas (más de 8 000 realizadas por los grupos de revisión de la Colaboración Cochrane); de más de 18 000 guías para la práctica clínica, más de 700 000 informes de ensayos clínicos y más de 1,8 millones de artículos de revisión. Presenta una tendencia marcada a procesar materiales publicados por revistas de los Estados Unidos, así como en idioma inglés. Entre 2005 y 2009, casi el 45% de los registros que ingresaron a esta base procedieron de revistas estadounidenses, y el 91% se encontraban en inglés.²

Una búsqueda en *PubMed* comprende: referencias bibliográficas "en proceso" de inclusión en *Medline*; referencias a fuentes (revistas esencialmente) que preceden a su fecha de ingreso a la base de datos; referencias a trabajos no cubiertos en *Medline*; referencias a manuscritos de autores pertenecientes a los Institutos

Nacionales de Salud de los Estados Unidos, publicados en revistas no procesadas por *Medline*; así como de algunas revistas en ciencias de la vida; etcétera.

Los artículos procesados se indizan mediante el tesoro conocido como MeSH, una lista de alrededor de 29 000 encabezamientos de materias y unas 213 000 entradas temáticas que facilitan encontrar los términos autorizados en el tesoro, además de 43 calificadores para aumentar la precisión de la búsqueda.^{1,3}

Embase

Embase (<http://www.elsevier.com/online-tools/embase/>), producido por Elsevier - la casa editorial líder en Ciencia, Tecnología y Medicina a escala mundial- es la mayor base de datos de resúmenes con información biomédica. Su colección completa supera los 28 millones de registros bibliográficos -*Embase* (1974 hasta el presente) + *Embase Classic*: 1947-1973)- y procesa más de 8 300 revistas arbitradas en el campo de la biomedicina y las ciencias de la vida, desde 1947 hasta la fecha. Contiene los más de 20 millones de los registros reunidos por *Medline* y unos 6 millones de registros únicos, publicados por unas 2 600 revistas (31%), que no procesa la base de datos estadounidense, donde abundan fuentes europeas, asiáticas, latinoamericanas y otras, así como en idiomas diferentes al inglés. *Medline* y *Embase* comparten aproximadamente el 71% de referencias comunes, aunque existen diferencias importantes entre temáticas y regiones. Su cobertura alcanza más de 90 países y 40 idiomas, y su crecimiento anual es de alrededor de 1,3 millón de registros. Es particularmente fuerte en información sobre medicamentos y equipos médicos.⁴

Para la realización de revisiones sistemáticas de tratamientos con medicamentos es recomendable utilizar junto a bases de datos como *Medline*, *Embase* y CENTRAL, esta última de la Colaboración Cochrane. Es una fuente importante de evidencias clínicas. A partir del año 2009, *Embase* comenzó a incorporar a su colección, materiales de reuniones científicas hasta alcanzar un registro anual de unos 300 000 resúmenes procedentes de unas 1 000 actas de eventos, algo que la diferencia notablemente de otras bases de datos.

La información que se procesa en *Embase* se indiza por medio de su tesoro denominado Emtree, una lista de más de 65 000 términos de materia (más de 30 000 son nombres de medicamentos y compuestos químicos) y 280 000 sinónimos, además de 78 calificadores para aumentar la precisión de la búsqueda. El tesoro de *Embase* proporciona facilidades particulares para la indización profunda de

literatura sobre medicamentos y equipos médicos (existen más de 3 000 términos de indización para equipos y dispositivos médicos). Las referencias importadas de *Medline* que forman parte de las fuentes que procesa *Embase* adquieren automáticamente la indización realizada por los expertos de esta última base de datos; los materiales únicos, es decir, que no proceden de fuentes procesadas por *Embase*, no se indizan manualmente sino que obtienen sus términos índices a partir de un sistema de equivalencias temáticas inter-tesauros MeSH-EMTREE.^{5,6}

En 2008, Elsevier se asoció con QUOSA para potenciar la gestión de los textos completos de las referencias recuperadas en esta base de datos como resultado de una exploración bibliográfica. QUOSA Information Manager empodera la gestión de la información existente en los artículos recuperados a través de la post-búsqueda automatizada de alto rendimiento en los textos completos de estos. Se distribuye mediante suscripción.⁷

Scopus

Scopus (<http://www.elsevier.com/online-tools/scopus>), creada en 2004 por Elsevier B. V., es la mayor base de datos de citas y resúmenes de literatura arbitrada y de fuentes de alta calidad en el Web. Cubre más de 21 000 títulos de publicaciones seriadas activas de más de 5 000 casas editoras; 20 000 son revistas arbitradas; más de 2 600 del total presentan acceso abierto. Contiene más de 50 millones de registros: 29 millones de registros desde el presente hasta 1995 (el 84% de los artículos poseen resumen; hasta 1996, presentan también sus referencias bibliográficas, procedentes de publicaciones académicas -revistas y series monográficas) y comerciales; 21 millones de registros corresponden a contribuciones publicadas antes de 1996. Presenta además, una extensa cobertura de materiales de conferencias (más de 5,5 millones), páginas científicas en Internet (unos 376 millones) y patentes (más de 25,2 millones) y procesa los artículos "en prensa" de más de 3 850 revistas.⁸

A pesar del carácter multidisciplinario de su colección, el 48% de sus títulos clasifican en las categorías ciencias de la vida y de la salud. *Scopus* procesa el ciento por ciento de los artículos procesados por *Medline* —que constituye aproximadamente el 90% de los artículos procesados por *PubMed*— y más del 97% del total de títulos procesados por *Embase*. La retrospectividad del procesamiento de los artículos y sus referencias (necesarias para los análisis de citación) se remonta al año 1996, aunque existe una gran cantidad de artículos fuentes (es decir, sin sus referencias) de fechas anteriores hasta el 1823. Su actualización es

diaria, cada día ingresan unas 5 500 registros nuevos, aproximadamente 2 millones al año. En 2013 inició un esfuerzo particular por extender el número de libros disponibles en la base de datos que se extenderá hasta 2015.⁸

Scopus es la alternativa europea al monopolio que durante más de 40 años ejercieron las bases de datos del antiguo Institute for Scientific Information (ISI - actual Thomson Reuters) en el área de los estudios de citación en el contexto científico internacional. Posee herramientas inteligentes para seguir, analizar y representar el comportamiento de la actividad en la ciencia, en especial con respecto a su consumo, a partir del empleo de los datos de citación de las obras y los autores. Además de la búsqueda bibliográfica, *Scopus* permite la obtención de una serie de indicadores métricos que permiten evaluar el comportamiento de las fuentes de información procesadas por la base de datos. Entre los abundantes indicadores métricos utilizados para la evaluación de las fuentes procesadas por esta base de datos se encuentran el *Source Normalized Impact per Paper* (SNIP) que mide el impacto de citación de una revista en su contexto, y el *SCImago Journal Rank* (SJR), desarrollado por el Grupo SCImago, y que es una medida del prestigio científico de una revista, que se basa en la ponderación del valor de las citas recibidas.⁹

Algo más del 50% de los títulos de revistas y actas de conferencias que procesa *Scopus* son únicos, es decir, no aparecen en el *Web of Science*, el por ciento estimado de títulos únicos en *Web of Science* es del 25%, aunque existen diferencias según temáticas y regiones que se acentúan en particular en los territorios ubicados fuera de Norteamérica y Europa occidental. Las revistas solapadas, es decir, aquellas que se procesan por ambos recursos tienden a ser las más citadas. *Scopus* abarca el 80% de las citas emitidas por las revistas procesadas en su colección. Aunque posibilita la búsqueda por materias según términos comprendidos en el MeSH y el DeCS siempre que los registros procesados los posean, no permite el uso de las facilidades de búsqueda propias de cada uno de los tesauros en sus respectivas bases de datos, como la construcción de la estrategia de búsqueda desde ellos o la llamada explosión de términos subordinados. *Scopus* dispone de las facilidades del sistema para gestión documental y de la información recuperada denominado QUOSA. Se distribuye mediante suscripción.¹⁰

Discusión

PubMed-Medline y *Embase* son bases de datos especializadas en biomedicina y salud. No obstante, *Embase* posee un marcado interés en la literatura sobre farmacología y equipos médicos, que se manifiesta de manera particular en su tesoro de búsqueda. *Embase* comprende *Medline* y este constituye alrededor del 90% del total de *PubMed*. Sin embargo, es importante precisar que este por ciento puede variar considerablemente en áreas del conocimiento como las referidas donde existen diferencias importantes con respecto al volumen de contribuciones procesados por una y otra base de datos. También existen diferencias importantes en la preferencia por ciertos tipos de documentos para su procesamiento; así, por ejemplo, las actas de eventos poseen un interés especial para *Embase* a diferencia de *PubMed-Medline*. Las interfaces de búsqueda de ambas bases de datos presentan facilidades específicas para la exploración bibliográfica en el sector de la medicina y la salud a diferencia de *Scopus* que es una base de datos multidisciplinaria aunque con un amplio cubrimiento de la literatura en estas áreas. La distribución gratuita de *PubMed-Medline*, al contrario de *Embase*, que se adquiere mediante suscripción, aconseja el inicio de la búsqueda de información en la primera de estas. En *PubMed-Medline*, la cobertura de títulos se concentra en los Estados Unidos y otros países anglófonos con un alto nivel de desarrollo; *Embase* procesa una cantidad mucho mayor de títulos procedentes de América Latina, Europa, África y Asia.

Scopus, sin duda, registra la mayor cantidad de artículos en el área de la salud. Comprende la totalidad de las contribuciones procesadas por *Medline* y la inmensa mayoría de las registradas en *Embase*, además de incorporar a sus colecciones un número alto de artículos afines o de interés relacionados con los campos de las ciencias de la salud en general. Debido a su característica de base de datos multidisciplinaria, carece de una interfaz de búsqueda apropiada para las búsquedas principalmente de información de carácter clínico. Como ventaja sobre las bases de datos referidas, presenta herramientas particularmente útiles para la evaluación de artículos, autores, organizaciones, países y regiones desde una perspectiva cuantitativa.

Es muy importante comprender que la cobertura de títulos y la retrospectividad del procesamiento de la literatura en una base de datos no son los únicos aspectos de importancia para su evaluación, porque, elementos como la profundidad de la indexación mediante vocabularios controlados y la existencia de interfaces de

búsqueda diseñadas de acuerdo con las particularidades del contenido temático de la base y las necesidades de sus usuarios como sucede con *PubMed-Medline* y *Embase*, que presentan interfaces elaboradas "a la medida" de las necesidades de la búsqueda de la información biomédica, clínica y de salud, son factores importantes para aumentar la precisión de los resultados de una exploración bibliográfica.

Referencias bibliográficas

1. National Center for Biotechnology Information (NCBI). National Library of Medicine. United States. PubMed. [citado: 14 de noviembre de 2013]. Disponible en: <http://www.ncbi.nlm.nih.gov/pubmed/>
2. National Library of Medicine. United States. Medline. [citado: 14 de noviembre de 2013]. Disponible en: <http://www.nlm.nih.gov/pubs/factsheets/medline.html>
3. National Library of Medicine. United States. MeSH. [citado: 14 de noviembre de 2013]. Disponible en: <http://www.nlm.nih.gov/pubs/factsheets/mesh.html>
4. Elsevier. Embase about. [citado: 14 de noviembre de 2013]. Disponible en: <http://www.elsevier.com/online-tools/embase/about>
5. Elsevier. Embase: Indexing with Life Science Thesaurus Emtree. [citado: 14 de noviembre de 2013]. Disponible en: <http://www.elsevier.com/online-tools/embase/emtree>
6. Rochie AM. Embase: Tips and tricks 2013. [citado: 14 de noviembre de 2013]. Disponible en: <http://www.slideshare.net/rocheam/embase-tips-and-tricks-for-trainers-webinar-26-june-2013>
7. Elsevier. Elsevier's EMBASE.com Partners with QUOSA. [citado: 23 de noviembre de 2013]. Disponible en: <http://www.elsevier.com/about/press-releases/science-and-technology/elseviers-embase.com-partners-with-quosa>
8. Elsevier. Scopus. Facts and figures. [citado: 14 de noviembre de 2013]. Disponible en: http://cdn.elsevier.com/assets/pdf_file/0007/148714/scopus_facts_and_figures.pdf
9. Cañedo Andalia, Cruz Font J. Tendencias, limitaciones y perspectivas en la evaluación de las publicaciones científicas y académicas mediante indicadores cuantitativos. En: Cañedo Andalia R, Rodríguez Labrada R, Fernández Valdés MM, Zayas Mujica R, Nodarse Rodríguez M, Sánchez Tarragó N, et al. Lecturas avanzadas para la alfabetización informacional en salud. Holguín: Centro Provincial de Información de Ciencias Médicas. Universidad de Ciencias Médicas de Holguín; 2011. [citado: 26 de noviembre de 2013]. Disponible en: http://www.hlg.sld.cu/alfin/download/lecturas_avanzadas/5.1.factor_de_impacto_listo.pdf

10. Center for Research Libraries. Global Resources Network. Academic Database Assessment Tool. About Scopus. [citado: 14 de noviembre de 2013]. Disponible en: <http://adat.crl.edu/databases/about/scopus>

¿Cómo citar este documento?

Cita (Vancouver): Cañedo Andalia R, Nodarse Rodríguez M. *Similitudes y diferencias entre PubMed, Embase y Scopus*. En: Cañedo Andalia R, Rodríguez Labrada R, Fernández Valdés MM, Zayas Mujica R, Nodarse Rodríguez M, Sánchez Tarragó N, *et al*. Lecturas avanzadas para la alfabetización informacional en salud. Holguín: Centro Provincial de Información de Ciencias Médicas. Universidad de Ciencias Médicas de Holguín; 2014. [citado día mes año]. Disponible en: http://www.hlg.sld.cu/alfin/download/lecturas_avanzadas/Similitudes%20PubMed%20Embase%20Scopus.pdf