EURA MEDICOPHYS 2007;43:563-9

Beyond PubMed

Other free-access biomedical databases

E. GIGLIA

Central Library of Medicine
University of Turin, Turin, Italy

This article presents several popular free-access biomedical databases (general, genetic, pharmacological, evidence-based oriented, or dealing with health technology assessment) from which researchers can select the best tool for the purpose and context of a specific research. Some gateways for searching across multiple databases that provide a "one-stop shopping" point of access are also presented.

KEY WORDS: PubMed - Databases - Libraries, digital.

Medline/PubMed is by far the most commonly used and best known point of access to the biomedical literature. For some it may be their only research tool. Undoubtedly a first-rate source, but technically speaking, it is a "bibliographical database", i.e. it contains only references from journal articles with an ever-increasing number of free full texts. In future articles in this series we'll explore other bibliographical databases, together with biomedical lists of resources, dedicated search engines, and institutional Web sites. But there are also many other databases, technically known as "factual", that contain primary information, raw data, images, but not literature. In this article we present a selection of these databases, from which researchers can select the best tool for the purpose and context of the research. Owing to the richness and the variety of these databases, metasearch tools have been developed to provide a "one-stop shopping" search page: they, too, will be explored.

Primary, quick information: Medline Plus, a patient-oriented database, also useful for physicians

Medline Plus (http://medlineplus.gov/) is the National Library of Medicine's web site tailored for consumer health information. Although patient-oriented, it provides quick references, definitions, practical overviews and useful links also for physicians. The "NLM" brand is a guarantee that all information and links have been evaluated and selected by the NLM's information specialists: in other words, if you need a sort of fact sheet, you can get trusted information here.

Medline Plus is divided into many sections of interest:

— **Health Topics**: covering 740 topics on conditions, diseases and wellness.

— **Drugs and Supplements**: searchable by generic or brand name. Formatted like a patient leaflet, these cards give information on side effects, interactions, precautions and so on.

— **Medical Encyclopaedia**: includes over 4000 articles about diseases, tests, symptoms, injuries, and surgeries. It also contains an extensive library of medical photographs and illustrations.*

— **Dictionary**: with spelling and definitions of medical words

— **News**: press releases on biomedical topics

— **Interactive tutorials**: 165 tutorials on diseases and conditions (e.g. cerebral palsy), tests and diagnostic procedures (e.g. knee arthroscopy), surgery and treatment procedures (*e.g.*

Address reprint requests to: E. Giglia, Biblioteca Centralizzata di Medicina, Università degli Studi di Torino, Torino. E-mail: elena.giglia@unito.it

*There are Web sites specialized in medical images, like HonMedia (http://www.hon.ch/HONmedia/) or TRIP medical images (http://www.tripdatabase.com/ index.html, first run the search for your term, then open the "Medical images" label (no options for direct search): these will be explored in a future article. When downloading an image, remember the copyright holders' rights.

Figure 1.—Medline Plus: results for "Hip replacement" (pathway: browsing from Health Topics, Body location, then Bones, joints and muscles).

Figure 2.—NLM Gateway results for "carpal tunnel syndrome".

hip replacement – physical therapy), and prevention and wellness (e.g. back exercises)

— **Other resources**: organizations (also by Health topic, like "Disability"), government health sites, other databases (including the National Rehabilitation Information Center, http://www.naric.com/public/ default.cfm, or ABLEDATA, for assistive technology, http://www.abledata.com/abledata.cfm?pageid=19327&sectionid=19327).

In Medline Plus you can run a search by typing your keyword in the search box (the system retrieves the references of this term in each section) or you can browse section by section, exploring the resources and information to get a basic orientation.

Let's explore the **Health Topics** section, a treasure chest of useful data. You can browse by Body Location/System, Disorders and conditions, Diagnosis and therapy, Demographic groups, Health and wellness. Once you have found your topic, the system opens a real "hub", like the one you find in Figure 1, with an overview, the latest news, specific conditions and related issues, financial issues, information on clinical trials, selected journal articles, resources for specific age range or gender (children, seniors, women, etc.). As you can see, the links are accredited associations, like the American Academy of Orthopaedic Surgeons. For diseases, in the right-hand column there are links to the US National Institute of competence, for instance, the National Institute of Arthritis and Musculoskeletal and Skin Diseases (http://www.niams.nih.gov/) or the National Institute of Neurological Disorders and Stroke (http://www.ninds.nih.gov/), whose Web sites are other precious mines of information.

Medline Plus, jointly with PubMed and other NLM databases, is the core of the NLM Gateway (http://gateway.nlm.nih.gov/gw/Cmd), a Web-based system that lets users search simultaneously. The NLM Gateway integrates the search in PubMed with several toxicological databases and genetic databases illustrated below, and gives unique access to the Meeting Abstract database, the National Library of Medicine's online collection of abstracts from medical meetings. As with other metasearch applications, the user enters a query that is automatically sent to multiple retrieval systems with different characteristics but potentially useful results. The results from the target systems are presented in broad categories (bibliographic resources, consumer health resources, other information resources) rather than by database, as shown in Figure 2, with a search for "carpal tunnel syndrome": please note the 13 references from Meeting Abstracts.

Searching a chemical substance: properties, pharmacokinetics, interactions, toxicity

If you are looking for information about a chemical substance, e.g. Naproxen, you can search two databases: HSDB (Hazardous Substances Data Base) or PubChem.**

HSDB (http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB) is the National Library of Medicine's database on the toxicology of potentially hazardous chemicals. It is organized into over 5000 individual chemical records and is enhanced with information on human exposure, industrial hygiene, emergency handling procedures, environmental fate, regulatory requirements, and related areas. The data are fully referenced and peer-reviewed by a scientific review panel composed of expert scientists.

You can search by chemical or other name, chemical name fragment, Chemical Abstracts Service Registry Number (RN), and/or subject terms. Search results present data about Human Health Effects, Emergency Medical Treatment, Metabolism and Pharmacokinetics, Pharmacology (therapeutic use, interactions, etc.), Environmental Fate/Exposure, Chemical/Physical Properties (solubility, molecular formula, etc.), and Occupational Exposure Standards.

HSDB can be accessed directly from this URL and is searched by the NLM Gateway. It is also part of the TOXNET – TOXicology Data NETwork (http://toxnet. nlm.nih.gov/), which is a cluster of databases covering toxicology, hazardous chemicals, environmental health and related areas. It is managed by the Toxicology and Environmental Health Information Program (TEHIP) of the National Library of Medicine. TOXNET's web interface is designed as an easy and integrated way to launch a metasearch of databases of various formats and content, such as:

— IRIS – Integrated Risk Information System

(http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?IRIS): carcinogenic and non-carcinogenic health risk information on over 500 chemicals, scientifically reviewed by the EPA (US Environmental Protection Agency).

— CCRIS – Chemical Carcinogenesis Research Information System

(http://toxnet.nlm.nih.gov/cgi-bin/sis/ htmlgen?CCRIS): over 9000 chemical records with carcinogenicity, mutagenicity, tumor promotion, and tumor inhibition test results.

— GENE-TOX

(http://toxnet.nlm.nih.gov/cgi-bin/sis/ htmlgen?GENETOX): genetic toxicology (mutagenicity) test data from expert peer review of the open scientific literature on over 3000 chemicals.

— LACTAMED

http://toxnet.nlm.nih.gov/cgi-bin/sis/ htmlgen?LACT): database of drugs and other chemicals to which breastfeeding mothers may be exposed. It includes information on the levels of such substances in breast milk and infant blood and their possible adverse effects on the nursing infant.

— HOUSEHOLD PRODUCT DATABASE

(http://hpd.nlm.nih.gov/): cards on the potential health effects of chemicals contained in more than 6000 common household products used inside and around the home.

**Here we present the factual databases on properties and chemical aspects. If you need to search other practical or technical issues, you can visit the Ministero della salute [Italian Ministry of Health] site for the Prontuario Farmaceutico Nazionale [National Formulary] (http://www.ministerosalute.it/prontuario/decreto.jsp) or for the Medicinali equivalenti [Equivalent Drugs] list (http://www.aifa.gov.it/ PREZ_RIMB_MER/pagina00012.html). For Europe, you can access the European Pharmacopoeia at http://www.edqm.eu/ site/page_581.php. For an international perspective, the US Food and Drug Administration site (http://www.fda.gov/) and the section "Medicines" in the World Health Organization site at http://www.who.int/medicines/en/ are both valuable sources. The WHO gives access to the International Pharmacopoeia at http://www.who.int/medicines/publications/pharmacopoeia/overview/en/index.html. Again, we'll explore these sites in the article dealing with biomedical and institutional Web sites.

Figure 3.—TOXNET, integrated results for "Ibuprofen".

Figure 4.—Entrez search results for "Tramadol".

— TOXLINE
(http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?TOXLINE): bibliographical database with comprehensive coverage of the biochemical, pharmacological, physiological, and toxicological effects of drugs and other chemicals from 1965 to the present. TOXLINE contains over 3 million citations, almost all with abstracts and/or index terms and CAS Registry Numbers.
— DART/ETIC - Development and Reproductive Toxicology/Environmental Teratology Information Center
(http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?DARTETIC): bibliographical database with references to reproductive and developmental toxicology literature published since 1965.

Submitting a keyword in the TOXNET search box, the "Search All Databases" feature provides a list of TOXNET databases in which the search term is found, each paired with the retrieval record count and links to the Search Results page of those databases, as shown in Figure 3, with a search for "Ibuprofen":

PubChem is the US National Institutes of Health's database that provides information on the biological activities of small molecules. It includes substance information, compound structures, and bioactivity data in three primary databases:
— PubChem Substance
(http://www.ncbi.nlm.nih.gov/sites/entrez?db=pcsubstance): searches deposited chemical substance records using names, synonyms or keywords for more than 19 million records. Links to biological property information and depositor web sites are provided.
— PubChem Compounds
(http://www.ncbi.nlm.nih.gov/sites/entrez?db=pccompound): searches more than 10 million unique chemical structures using names, synonyms or keywords. Links to available biological property information are provided for each compound.
— PubChem BioAssay
http://www.ncbi.nlm.nih.gov/sites/entrez?db=pcassay): presents the results of biological screening (more than 600 as of October 2007).

You can search each database from its own search box, or from the PubMed home page, by opening the Search pull down menu on the top left corner of the PubMed homepage. But if you want to conduct a cross search, you can access the Entrez system (http://www.ncbi.nlm.nih.gov/sites/gquery).

Entrez is the integrated, text-based search and retrieval system used at the National Center for Biotechnology Information (NCBI) for major databases, including PubMed, Nucleotide and Protein Sequences, Protein Structures, Complete Genomes, Taxonomy, and the three PubChem databases mentioned above.

By using the Entrez "Global query", a search across all Entrez databases is performed by entering a simple search term or phrase in the "Search across databases" query box. Select the Go button to launch the search. The results found in each database are displayed on the Global Query page, as shown in Figure 4, with a query for "Tramadol". Click on the result number or its adjacent database name to visualize the specific results: for example, the PubChem BioAssay result displays Tramadol bioactivity, and data on the MTRD (Maximum therapeutic recommended dose); the first record on the Gene database shows the genetic implications of the cytochrome P450 superfamily of enzymes in the metabolism of many drugs, including Tramadol.

Genetic implications of diseases, conditions, drugs

If your research deals with genetic implications, there are several possibilities to perform a query, going from quick information to ultra-specialist genetic databases.

The referring database is OMIM – Online Mendelian Inheritance in Man, (http://www.ncbi.nlm.nih.gov/sites/entrez?db=OMIM). It is a catalogue of human genes and genetic disorders, with links to literature references, sequence records, maps, and related databases. It is edited by Dr. Victor A. McKusick and his colleagues at Johns Hopkins University and elsewhere, and was developed for the Web by the National Center for Biotechnology Information (NCBI). The OMIM Gene Map presents a cytogenetic map location of disease genes and other expressed genes described in the OMIM organized by chromosome, while the OMIM Morbid Map is organized by disease. Being integrated in the Entrez gateway, the OMIM provides for all the possible Entrez related records if associated with the cited OMIM record, such as Genome links, Nucleotide links, Protein links, going deeper and deeper into specialist data. Basically, entering a query, e.g. for "Parkinson's disease", the OMIM provides a description of the disorder, information on its clinical features, inheritance, clinical management, pathogenesis, and, of course, a map location for the genes involved and, if present in the literature, pertinent reviews of the genetic causes of the disease. The OMIM runs with an interface similar to PubMed, so you can use the Limits, History and Detail labels as you learned to do in PubMed to refine your search.

Gene (http://www.ncbi.nlm.nih.gov/sites/entrez? db=gene) is the Entrez database for gene-centered information. It focuses on the genomes that have been completely sequenced or are scheduled for intense sequence analysis. The content of Entrez Gene represents the result of curation and automated integration of data from NCBI's Reference Sequence project (RefSeq), from collaborating model organism databases, and from many other databases available from NCBI. You can run a search by keyword, chromosome, gene or other identifiers. If, for instance, you run a search for "Human muscular dystrophy", the system retrieves all the genes involved, and for each the nomenclature, map location, gene products and their attributes, markers, phenotypes, and links to citations, sequences, variation details, maps, expression, homologs, protein domains and external databases.

Entrez also searches other more specialist databases like Genome, UniGene, and Homologene: they are all integrated in the Global query feature.

Two practical genetic databases not included in Entrez are Genetics Home Reference (http://ghr.nlm. nih.gov/), a guide to understanding genetic conditions, diseases, and syndromes with basilar description, incidence, and related genes. Since it is more patient-oriented, it is useful for quick consultation; GeneTest (http://www.genetests.org/) for a requested disease or syndrome indicates "Research" with a link to the testing laboratories, "Reviews" with an overview, causes, evaluation strategy, genetic counselling, and "Resources", a link to associations or sites with useful information. A handy tool for finding more European-oriented testing laboratories is Orphanet (http:// www.orpha.

Figure 5.—TRiP results for "disk herniation". On the right, the clustering results.

Figure 6.—INATHA database search results for "knee replacement".

net/consor/cgi-bin/home. php?Lng=IT), which provides the testing laboratories for a rare disease or syndrome, besides other information and news.

Evidence-based Medicine databases

There are many EBM databases, most are bibliographical and, unfortunately, require payment for access to them.

Netting the evidence is a British project committed to facilitate evidence-based healthcare by providing support and access to helpful organisations and useful learning resources. It edits a Web site (http://www.shef.ac.uk/scharr/ir/netting/) which is a sort of EBM virtual library that collect links about Searching, Appraising, Implementing, and provides lists of Software (statistical or appraisal), Journals, Databases. Interesting databases are PEDro, Physiotherapy Evidence database (http://www.pedro.fhs.usyd.edu.au/index.html) and RebTrials.org (http://www.rehabtrials.org/), specific on rehabilitation: we will explore them, as said.

A very interesting user-friendly tool is TRiP, Turning Research into Practice, (http://www.tripdatabase. com/index.html), aimed at developing a sophisticated tool for locating the highest possible evidence with which to inform clinical decisions, using the principles of Evidence-Based Medicine.

The database contains systematic reviews, guidelines, evidence-based synopses, e-textbooks, clinical calculators, medical images. Results are clustered in each category, as shown in Figure 5, a search for "disk herniation": you can recognize your topic of interest at a glance.

Assessing outcomes, assessing costs: HTA databases

For those who deal with Health Technology Assessment, there are two referring sites.

INATHA – International Network of Agencies for Health Technology Assessment (http://www.inahta.org/Home/) is a global network for effective healthcare, whose mission is to provide a forum for the identification and pursuit of interests common to health technology assessment agencies –45 agencies from 23 countries to date. In the INHATA database (http://www.crd.york.ac.uk/crdweb/), edited in collaboration with the UK Centre for Reviews and Dissemination, merge many different types of research: systematic reviews, ongoing and completed research based on trials, questionnaires, and economic evaluations. The interface is a metasearch engine that also includes DARE (Database of Abstracts and Reviews of Effects), which contains over 4000 abstracts of quality assessed and critically appraised systematic reviews, and focuses on the effects of interventions used in health and social care. The UK National Health Service NHS Economic Evaluation Database contains over 6000 abstracts of quality assessed economic evaluations and assists decision-makers by systematically identifying and describing economic evaluations, appraising their quality and highlighting their relative strengths and weaknesses. Figure 6 displays the results for "knee replacement": you can see the indication of the record type and source; in the top left corner is a search window for refining the search by year or Boolean operators AND-OR-NOT.

The US National Information Center on Health Services Research and Health Care Technology (http://www.nlm.nih.gov/nichsr/) is committed to improving the collection, storage, analysis, retrieval, and dissemination of health services research, and lists:

— Databases, including the HSTAT – Health Services Technology Assessment Text (http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat), is a free, Web-based resource of full-text documents that provide health information and support health care decision making. Also very useful is the HSRR – Health Services Research Resources (http://www.nlm.nih.gov/ nichsr/hsrr_search/) to search datasets, instruments, software; and the Health Services Research PubMed Searches Project (http://www.nlm.nih.gov/nichsr/hedges/search.html), which provides specialized PubMed searches on healthcare quality and costs with pre-set filters, and a successive choice between categories like Appropriateness, Process Assessment, Outcomes Assessment, Costs, Economics, Qualitative Research.

— HSR Information Central (http://www.nlm.nih.gov/hsrinfo/) gives a complete overview of guidelines, statistics, meeting and conferences, discussion lists and others.

— Outreach and training activities

— Presentations and Publications

Catalogs of databases

BioMed Central, the Open Access publisher, maintains a catalog of more than 1000 databases. Some databases contain experimental data, others provide synopses of published information. Most are freely accessible. You can browse them (http://databases.biomedcentral.com/browsecatalog) or can run a search by keyword (http://databases.biomedcentral.com/search).