# **INTERNET BOOKSHELF**

EURA MEDICOPHYS 2007;43:427-37

Starting with this contribution, a new column will be run on these pages. The "Internet bookshelf" will present, like a traditional library's bookshelf, the last acquisitions for biomedical researchers: databases, search engines, web sites, free resources, to help to orient oneself in the scenario of a growing information overload, and to profit by the boundless and seamless resources of the Net, the real "library without walls".

# **PubMed at its best** Useful skills to perform an effective search, save it and automatically receive updates

E. GIGLIA

## PubMed: the free MEDLINE on the web

P ubMed (http://www.pubmed.gov) is the worldwide best known and accessed service for biomedical researchers. Since 1997 MEDLINE, the bibliographical database edited by the U.S. National Library of Medicine, with its over 17 million citations from about 5 000 journals in 37 languages from all over the world, became freely accessible via web. MEDLINE is the core of PubMed, but in the latest years PubMed has been integrating in-process items, free full texts, value-added services and other useful tools like the instant retrieval of "Related articles" for any citation. The Entrez interface, developed by the National Center for Biotechnology Information (NCBI) provides now an integrated search and retrieval system for all the NCBI databases: beyond MEDLINE, running a query by accessing the All Database option on the black naviUniversity of Turin, Turin, Italy

gation toolbar from PubMed homepage opens the gate to Nucleotide, Protein, Genome, Taxonomy, Chemical databases. We'll explore them. Let's focalize on PubMed.

Like every easy, user-friendly application, PubMed is often used in a partial way, not exploiting all its potentiality: you type a word in the search box, and click the "Go" button. Often thousands of items – not all relevant - are displayed, and the reader immediately gets discouraged. There are few devices to avoid it. Aim of this contribution is to offer a short, basic, easyto-use guide to PubMed and its main features. It does not pretend to be exhaustive, just would like to provide a working tool.

### How to read between the lines of the PubMed homepage

Let's have a look at the left navigation toolbar (in blue) and to the five grey labels just under the search box.

Acknowledgements.—I'm deeply indebted with Valentina Comba (University of Bologna, Centro E-Learning di Ateneo), Annalisa Jourdan and Anna Marchino (University of Turin, Biblioteca Centralizzata di Medicina - polo Clinico) and Professor Gian Piero Pescarmona (University of Turin, Dept. of Genetics, Biology and biochemistry), not only for their precious suggestions. The know why.

Address reprint requests to: E. Giglia, Head Librarian, Biblioteca Centralizzata di Medicina, Polo Clinico, University of Turin, Corso Dogliotti 14, 10126 Turin, Italy. E-mail: elena.giglia@unito.it

#### A blue navigation toolbar to explore an unknown PubMed

On the left are listed:

ENTREZ PUBMED, which includes among the others:

— Overview, interesting about the coverage of the material included in MEDLINE (*e.g.*, do you know that MEDLINE contains only references from journal articles since 1950, never from books?) and the steps an item is submitted to before becoming a PubMed record: *e.g.* when you find an item [In-process] it means that only basic information are given, the article being not yet indexed, *i.e.* it hasn't received its MeSH terms yet: it's important when you perform a search within the MeSH database.

— *Help*, which shows you how to use PubMed in a "quick-start" way. It also lists all the Search Fields Descriptions and Tags, in square brackets: you can add them to your search word in order to build an effective research strategy: *e.g.* "Affiliation" [AD] allows you to retrieve all the articles coming from an institution (University of Turin [AD]), or "Author" [AU] all the works of an author (Smith G [AU]), or "Journal Title" [TA] all the articles in a specific journal (Europa Medicophysica [TA]).

— *Tutorials*, which explain how to run a search for subject, how to use the MeSH database, how to save a search, the use of the fiters, how to apply subheadings. They take only two minutes each, and they're very very useful, really a "recommended reading" (direct link at http://www.nml.nih.gov/bsd/disted/pubmed.html).

PUBMED SERVICES, among which we shall look at:

— *Journal database*: allows you to search by topic, journal title or abbreviation. Note that by clicking this headline the search box changes, and on the upper left in the first box you read now "Search Journals" instead of "Search PubMed" (by default).

— *MeSH database*: runs your search only in the Medical Subject Headings (MeSH) database, as indicated now near the search box: "Search MeSH". MeSH is the U.S. National Library of Medicine's controlled vocabulary used for indexing articles for MED-LINE/PubMed, with over 23 000 terms, constantly updated. It has a tree structure: each term like a branch has narrower branches below and broader above. It is important to notice that PubMed search engine automatically explodes the term, *i.e.* it catches the term –

and the articles this term was assigned to - you enter and all its narrower terms, like in the example below:

All MeSH Categories

Analytical, Diagnostic and Therapeutic Techniques and Equipment Category

Therapeutics Rehabilitation Activities of Daily Living Art Therapy Bibliotherapy Dance Therapy Early Ambulation **Exercise Therapy** Motion Therapy, Continuous Passive Muscle Stretching Exercises Music Therapy Occupational Therapy Rehabilitation of Hearing Impaired Communication Methods, Total Lipreading Manual Communication + Rehabilitation of Speech and Language Disorders Language Therapy Speech, Alaryngeal + Speech Therapy Voice Training Rehabilitation, Vocational

The system helps you giving the whole tree, so, you can navigate among the terms and find out the one that is more fitting to your search (maybe you can find a narrower term that describes your search more precisely). When you type your initial search, the system maps it to the existent MeSH term (if this is the case), and then offers you a list of MeSH terms. They can contain your keyword or they can stand for the concept. Each term has a brief definition - to make you sure this is really what you were searching for (*e.g.* remember that there is also a "Rehabilitation, Vocational" that belongs to neurosciences) - and the date of its addition to the MeSH vocabulary: it's obvious that if you decide to use a MeSH term introduced in 2003, you won't retrieve all the previous record, you need to use the related terms. Once you have selected your term, the system shows you also its specific Subheadings (Figure 1). Using the Subheadings by flagging them is a first, useful way to limit your search: e.g. if you are interested in "Rehabilitation" but only in the side of "Adverse effects", "Methods", "Standards",



Figure 1.-Navigation toolbars and the MeSH subHeadings.

"Supply and distribution" and "Utilization", you flag them and you sensibly reduce your results (100 120 without using Subheadings, 12 252 with Subheadings).

In the same page you find other two check boxes:

— *Restrict Search to Major Topic headings only:* it means that the system retrieves only the items in which your term is the first for relevance (in the cited example, from 12 252 to 9 000).

— Do not Explode this term means that the system will not include terms found below in the MeSH tree. It could be dangerous in some cases, because some MeSH term could include synonyms that otherwise you'll never find. Anyway: not exploding our term (catching "rehabilitation" records only, not art therapy, not exercise therapy and so on) you'll find only 818 items.

To run your search after having set all your preferences in this page, you need to go at the very end of the page, open the "Send to" window, and select "Search box with OR". Then click on the "Search PubMed" button just below the new window that the system shows to you, containing the translation of your query in the engine language.

Just a complementary note about the Boolean operators AND, OR, NOT: be careful that AND narrows down your search (it looks for a term AND another all together, in the same record), whereas OR expands it by finding all the records which contain the first term, all the records which contain the second one, and all the records that contain the two together. In our example, if we use the operator AND with our flagged Subheadings we'll find records that deal together with Adverse effects, Methods, Standards, and so on (low useful strategy, too narrow, better to use OR). Remember also that the system reads your query from left to right. For a complex strategy, you need to use brackets to «nest» groups of terms and that must be



Figure 2.-MeSH terms with their definitions.

processed first: if you are searching the implications of both ligament injuries and reconstruction with osteoarthritis, the right strategy is "osteoarthritis AND (ligament injuries OR reconstruction)".

Two last information about MeSH terms: first of all, librarians working at the National Library of Medicine always choose, in indexing an article, the most specific term (an article dealing with "Muscle stretching exercise" will be indexed with this MeSH term, not with the broader "Exercise therapy"). Take it into account in deciding to explode or not your term. To see the MeSH terms assigned to the article you've found, when you're on the article page just open the window "Display" and select "Citation": at the end you'll be able to read and navigate the assigned MeSH terms, and it is very useful because it can suggest terms to improve or deepen your search.

Last, but not least, notice that the system always

tries to match your entry word to the appropriate MeSH term: *e.g.* typing "physiotherapy" – which is not a MeSH term – the system offers you the choice between five possibilities, and the relative definitions to help you to choose the term that better describes your search (Figure 2);

— *Single citation matcher*: this is a very useful feature which allows you to find an article when you have just a few bibliographic information – *e.g.* journal title, year, volume -; it gives you also the Table of Contents of a single issue of a journal, by typing the title of the journal, year, volume.

*— Clinical queries:* provides specialized PubMed searches for clinicians, with built-in search filters, to find out:

— Search by Clinical Study Category: enter your term in the search box, you are then requested to select a category between Therapy, Diagnosis, Etiology, Prognosis, Clinical Prediction Guides and a scope between Narrow or Broad: the system automatically will run a search specifically designed for clinicians (*e.g.* typing "scoliosis" and selecting "Diagnosis" it displays 57 very relevant records, whereas if you simply search "scoliosis AND diagnosis" in PubMed you retrieve 7 632 items, or if you search "Scoliosis" as MeSH term with "/diagnosis" as Subheading you find 3 553 items).

— *Find Systematic reviews:* entering a term in this search box you will be provided by a specific list of results covering a broad set of articles – systematic reviews, meta-analyses, reviews of clinical trials, evidence-based medicine, consensus development conferences, guidelines, citations from journals specialized in clinical review studies – and using a very complex customized search strategy. Typing "scoliosis" you'll find 79 reviews, all relevant, whereas a simple search in PubMed with "scoliosis AND systematic review" returns 865 records, most of which are too far from the aim, or a search for "scoliosis" with the Limits for Publication Type as Reviews gives 862 items.

— *Medical genetic searches:* enter a term in the search box, then select a category between Diagnosis, Differential Diagnosis, Clinical description, Manage-ment, Genetic Counselling, Molecular genetics, Genetic testing: *e.g.* with "scoliosis" you'll find 744 Genetic testing, or 73 Differential Diagnosis.

— *Special queries:* another useful way to limit your search is to use the filters and subject limits listed in this section: *e.g.* Health services research (for assessment), Cancer Topic Searches, Bioethics, Complementary Medicine...

— *My NCBI*: this feature is so important that we shall see it below.

RELATED RESOURCES: links to useful tools or web sites for researchers, such as:

— *NLM gateway*: a metasearch engine that we'll explore in the next contribution to this column; it runs searches within PubMed and about other 20 medical databases.

— *TOXNET*, a metasearch engine for several databases on toxicology, hazardous chemicals, environmental health, and toxic release.

*— Consumer health*, a link to the «Medline Plus» database dealing with tutorials, medical encyclopaedia, dictionary and so on. It has been designed for patients, but it is very useful for quick reference searches. It is scheduled to be a topic on a next contribution.

— *Clinical alerts* are provided to expedite the release of findings from the U.S. National Institutes of Health-funded clinical trials where such release could significantly affect morbidity and mortality.

— *Clinical trials.gov*, link to the official site that provides regularly updated information about federally and privately supported clinical research in human volunteers. ClinicalTrials.gov gives you information about a trial's purpose, who may participate, locations, and more details.

#### Five labels to tame millions of records

On the PubMed homepage, you can see five grey helpful labels – or better Tabs:

– LIMITS: these are parameters to refine your search by Author, Journal title, Full text or Free Full Text (only citations that include a link to full text, useful for whom has no access to an institution subscribing library), Languages, Type of article (e.g. Clinical trials, Guidelines, Reviews...), Dates (you can select a date or a range both of publication or of addition to PubMed), Subsets of journal groups or Topics, and, concerning the design of the study, Gender, Humans or Animals, and Ages. If we run a simple search with "adolescent idiopathic scoliosis" we get 2229 results, but if we limit the same search to Ages - Child 6-12 years we find 1 335. If we limit also to articles published in the last three years, we get 239 items. When Limits are in use a check will appear in the Limits tab and your selected limits will be displayed in a yellow bar. To turn off limits click off the Limits tab check box and run a new search. Beware that for "[In-process]" items citations may be excluded from some limit selection because they have not yet completed the indexing process.

— PREVIEW/INDEX: these are two features that provide an alphabetical display of all search terms in each PubMed search field. They are useful to combine searches. You can browse by all fields or within a specific field such as MeSH Terms. It shows you also all the MeSH term Subheadings, and the rate of occurrences. If you type "scoliosis" and run for Preview, the system replies that there are 12 544 items in PubMed: you can display the result list clicking on the number. If you run for Index, it lists all the terms in which it finds the word "scoliosis", such as the MeSH term "scoliosis/etiology and pathogenesis" as Subheading,



Figure 3.-How to send items to the Clipboard.

with 40 articles, or "scoliosis curve progression" with 7 items.

— HISTORY: it tracks the history of your previous searches (to a maximum of 100, expires after 8 hours or if you close your work session). To view the results of a search, click on the number of results. To operate with the options – Delete, Go to re-run a search, Save to save the search in MyNCBI, we'll see how to - simply click on the search statement number. This feature is very useful to combine searches, without loosing information. If we type "back pain" we get 26 046 results, search #1. If we type "posture" we get 54 799 items, search #2. If we click on #1 and select AND, we tell to the system to combine our first and second search with the Boolean operator AND, that means we're asking to get only records that contain both the concepts of "back pain" AND "posture": we get 1 397 records. If for instance we use (if this is the case) the LIMIT tab to set the age "Child 6-12" and "Adolescent 13-18", we reduce to 192 records.

- CLIPBOARD gives you a place to store the selected citations retrieved in one or more searches, to a maximum of 500, and expires after 8 hours of inactivity or if you close your work session. Collecting items in the Clipboard is also a fixed route to save them indefinitely to My NCBI Collections, as we'll see. To add citations to the Clipboard, select them from the results display page by flagging the check box, or select no one if you want them all, then open the "Send to" window that you find at the top of your result list, and select "Clipboard" (Figure 3). The figure shows that we entered "cerebral palsy" AND "chronic pain", we get 69 items and we want to save them all. An asterisk is displayed on the tab if items have been added. Notice that the system uses cookies to add selection to the Clipboard, so your browser must be set to accept cookies.

— DETAILS is the most useful tab. It is very impor-



Figure 4.—The "Details" box and the translation of a query.

tant because it reveals you the translation of your query, that is, how the system understood your request. In an information retrieval system you ask for something, but if you don't know how the engine translates your words, often the results can be misleading: it is only a lack, a noise of communication. Imagine to type "physiotherapy" in the search box. You think the system will retrieve... what? All the items containing the word "physiotherapy"? Where? In the title? In the abstract? As a MeSH term assigned to articles dealing with the subject? If you click on the DETAILS tab the system shows you the translation of the query, and in the box you can operate by adding or deleting anything.

To start a query, the system first of all tries to automatically map the word in the search box with the MeSH translation table. If a perfect match is found in this translation table, the term will be searched as MeSH term. The system also adds a query with the same word as a "Text Word": it means that it is searched in all the words and numbers in the Title, Abstract, Other abstracts, MeSH terms, MeSH Subheadings, Publication Types, Substance Names, Personal Name as Subject (more details can be found in the PubMed Help). If the entered word doesn't match with a MeSH term, the system will also automatically include a Title/Abstract search for all the records not yet indexed in Medline, because they haven't yet received their MeSH term, and this would exclude a lot of articles.

In Figure 4 we read the DETAILS of our query. We entered "physiotherapy" and got 86 432 results. It is clear, and it will be clearer from the next example, that "physiotherapy" is not a MeSH term. The system mapped it to "physical therapy modalities", a narrow MeSH term. But if you remember, in Figure 3 searching for "physiotherapy" in the MeSH database we found 5 terms, between which "Exercise movement

techniques": reading the DETAILS Tab you learn that all the records associated with this MeSH term have not been included in this search. If you want to add them, simply type in the DETAILS box «OR "exercise movement techniques" [MH]».

Let's go on reading the DETAILS box: the system searched for "physiotherapy" as "physical therapy modalities" in the Title and Abstract [TIAB] for all the records not yet indexed in Medline [NOT Medline subset], searched for "physical therapy modalities" as MeSH term retrieving all the records associated with it, and searched also "physiotherapy" as Text word [TW]. All the Fields and their abbreviation are shown in the PubMed Help,

(http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=help-pubmed.section.pubmedhelp.Search\_Field\_Descrip).

Other examples: if we type "sciatica", we find 4 413 items. If we open the DETAILS tab, we read:

The system translated our search mapping with the corresponding MeSH term – in this case, you see, there is no search in the Title/Abstract, because it doesn't help – and searching also the term as Text Word. If you want to exclude it, simply delete it in the box: you'll find 3 439 items.

# Searching by MeSH terms or by a free keyword?

There are differences between a search run only by MeSH terms and a search run by a free keyword. The system assists you: if you simply type a word, as we've just seen, it runs a search both for MeSH terms and for Text Word. There are advantages in both choices: it depends on the topic or the aim of your search. But to make up one's mind, one has to know the two different criteria.

The Medical Subject Headings (MeSH) is a controlled vocabulary of biomedical terms that is used to describe the subject of each journal article in MED-LINE. As we saw, it has a tree structure with narrower and broader branches. Skilled subject analysts examine journal articles and assign to each the most specific MeSH terms applicable - typically ten to twelve. The main feature of a MeSH term search lays exactly upon this: running a search with a MeSH term ensures a more specific list of results, as the system recognizes the heading and retrieves the articles that mostly deal with the requested subject remind that you can also activate the "Major topic" button to further restrict the results.

The search by keyword gives you back a broader set of items, not so specific, because the system looks for the word even if it appears few times. As the system when you just type a word in the search box automatically runs the two searches together, if you want to search only by keyword you have to type [TIAB] after your keyword, so you will exclude the MeSH term search.

Applying the MeSH vocabulary ensures that articles are uniformly indexed by subject, whatever the author's words, or whatever synonyms exist. A clear example is "Bone Cancer": if you type it, the system automatically maps it to the correct MeSH term "Bone Neoplasms". Running a search with "Bone neoplasms" you will retrieve all the records, including those in which the concept appears in the title or in the abstract as "bone cancer" or "bone tumor" or "bone tumour". This is the great advantage of the MeSH term search. The problem rises about the records "[In-progress]", or the other ones not yet fully indexed. That's why the system, by default, translates your query both to the MeSH term search and the Text Word search: to ensure you the full coverage. Remember to add in the DETAILS box all the synonyms, because the system uses only what you typed in the search box (in our example, just "bone cancer", not tumor, tumour and so on), as you can see below:

> "bone neoplasms" [MeSH Terms] OR bone cancer [Text Word]

On the other side, running a search with a free keyword is useful to catch also the most recent contribution not yet indexed, as we have just seen, or if the topic you're looking for is very specific, or if there is a particular meaning, or if the subject is so unexplored that a specific MeSH term doesn't exist yet. In this case, the suggestion is to type the word and add [TW] to search within the fields cited above, or [ALL]

<sup>&</sup>quot;sciatica" [MeSH Terms] OR sciatica [Text Word]



Figure 5.—Saving a Search strategy in the My NCBI pop-up window.

to search within all fields (but it could generate a lot of noise, and too many results).

### My NCBI: saving searches and results has never been so easy

MyNCBI is a feature tailored to customize PubMed services for each user. As you see in all the figures, you find MyNCBI box on the upper right side of PubMed pages. The first time you have to register, choosing a username and a password. It is completely free and takes only few minutes. Then, when you connect again to PubMed, you only need to "Sign in" to MyNCBI: you will see all your saved searches and collections, and you'll be able to change some preferences to personalize PubMed labels on your screen. Once in MyNCBI, click on the blue toolbar on the left "Filters", choose PubMed database in the grid that appears, then the label "Quick Pick" to select which type of material you want the system to point out in a specific tab in displaying the results: for instance, you can choose Reviews, Clinical trials, and Free Full text. The results of your next search will be broken up into your choice, and they will be immediately visible with their labels, as shown in Figure 5 (below, about collections): 119 items, 11 of which are reviews, 13 clinical trials and 4 free full text. You can also choose to highlight the term of your search within the context, to have an immediate sight.

But the most performing feature is that MyNCBI provides a free place a user can store both his search strategies (and receive update, or re-run the search) and his selected results.

#### Search strategies

Imagine to have set a research strategy for a therapy with exercises in the adolescent idiopathic scol-



Figure 6.—Saving items in a Collection in the My NCBI pop-up window.

iosis, and to have limited it by Age to Child and Adolescents. The answer is 119 items. In the DETAILS tab you will find:

("adolescent" [MeSH Terms] OR adolescent [Text Word]) AND idiopathic [All Fields] AND ("scoliosis" [MeSH Terms] OR scoliosis [Text Word]) AND ("Physical Therapy Modalities" [MH] OR "Exercise Movement Techniques" [MH]) AND ("child" [MeSH Terms:noexp] OR "adolescent" [MeSH Terms])

If you don't want to type this all again the next time you will connect to PubMed, the easiest way is to register in MyNCBI and save your search strategy. Once you have set your strategy, and you are logged to MyNCBI, simply click on the right of the search box on "Save search". As shown in Figure 5 a pop-up appears (your browser needs to allow pop-ups at least for this time), in which you can give your search a name, you can select if you want e-mail updates and how often (every day, week, month): the system automatically re-runs the search and sends the new results to your e-mail address. Now the search strategy is saved in MyNCBI. If you didn't choose the automatic e-mail update, signing in to MyNCBI you can select the desired search and click the button "What's New for Selected". The search will be rerun.

Another innovative way to make the system run a determinate search strategy is to create a RSS feed. RSS (Really Simple Syndication) is an XML-based format used to send new items or information to recipients who use RSS feed readers.\* Once run a search, click on "Send to" window and select "RSS feed". Following the instructions, you have only to cut and paste the URL that the system gives to you in your RSS reader. You will receive update results in your RSS reader.

#### Collections

For each search strategy, or for each search run in PubMed, you can save results to a Collection. It is a two-step process: results are first sent to the Clipboard (temporary storage) and then to MyNCBI Collections (indefinitely).

Once run a search, from the result page select the items you're interested in – if you don't select any, all up to 500 will be sent to the Clipboard -, click on "Send to" window and select "Clipboard". Then open the Clipboard Tab and you'll be shown your selected items. You can operate again (maybe you can store in the Clipboard 10 items, then you read the abstracts with more attention and find that you need only 6 items), select the items you want to save

to" window and select "MyNCBI Collections". On the "Save Collection" pop-up window that appears, you may create a new collection or append to an existing collection, enter a collection name, and delete the saved items from the clipboard (suggested, so you'll find the Clipboard empty, ready for the next search). In Figure 6 you can see our 13 selected items (2 of which are clinical trials, 1 free full text, 6 reviews, as we set these preferences in MyNCBI to display our screen) and the pop-up window to create a new collection. Whenever you sign in to MyNCBI you can open your collection, read, print, or delete any item.