

## EFFICIENT SEARCHING STRATEGIES IN PUBMED

SAEEDA ABDULLAH

### ABSTRACT

*Literature search is a fundamentally important problem in research but become harder as the literature grows at a faster speed and broader scope. PubMed is a free Web literature search service developed and maintained by the National Center for Biotechnology Information (NCBI). It is as part of NCBI's Entrez retrieval system that provides access to a diverse set of 38 databases. PubMed currently includes citations and abstracts from over 5000 life science journals for biomedical articles dated back to 1948, therefore it serves as the primary tool for electronically searching and retrieving biomedical literature.*

*Although PubMed provides a broad, up-to-date and efficient search interface, it has become more and more challenging for its users to quickly identify information relevant to their individual needs, owing mainly to the ever-growing biomedical literature. In present article several tools for efficient search are discussed, like Boolean operators ("AND", "OR", "NOT"), filters, search field tags, use of MeSH terms, wildcard ("\*") and search with key words. Aims: The purpose of this article is to assist general readers in the development of the basic skills required to use PubMed.*

**Key Words:** *information storage and retrieval; PubMed searching, Boolean operators, search field tags, wildcard, use of keywords.*

### INTRODUCTION

Now a days, as the literature grows at a faster speed and broader scope, literature search is harder. Therefore continuous developments and new emerging systems in this field are expected. As of 2010, there are over 20-million citations indexed through PubMed, a free web literature search service developed and maintained by the National Center for Biotechnology Information (NCBI). PubMed is a part of NCBI's Entrez retrieval system that provides access to a diverse set of 38 databases.<sup>1</sup> PubMed currently includes citations and abstracts from over 5000 life science journals for biomedical articles back to 1948. Since its inception, PubMed has served as the primary tool for electronically searching and retrieving biomedical literature.

Although PubMed provides a broad, up-to-date and efficient search interface, it has become more and more challenging for its users to quickly identify information relevant to their individual needs, owing mainly to the ever-growing biomedical literature.

#### **PubMed: the primary tool for searching biomedical literature**

From a search perspective, PubMed takes free-text key words as input natural language and returns a

list of citations that match input keywords (PubMed ignores stop words). Its search strategy has two major characteristics: The First major uniqueness of PubMed is its choice for ranking and displaying search results in reverse chronological order. More specifically, PubMed returns matched citations in the time sequence of when they were first entered in PubMed by default. This date is formally termed as the Entrez Date (EDAT) in PubMed.<sup>1,2</sup>

Second major characteristic is, by default PubMed adds Boolean operators into user queries and uses automatic term mapping (ATM). Specifically, the Boolean operator 'And' is inserted between multi-term user queries to require retrieved documents to contain all the user keywords. For example, if a user issued the query 'pubmed search', the Boolean operator 'AND' would be automatically inserted between the two words as "pubmed AND search". In addition, PubMed automatically compares and maps keywords from a user query to lists of pre-indexed terms (e.g. Medical Subject Headings MeSH) through its ATM process. That is, if a user query can be mapped to one or more MeSH concepts, PubMed will automatically add its MeSH term(s) to the original query. As a result, in addition to retrieving documents containing the query terms, PubMed also retrieves documents indexed with those MeSH terms. Take the earlier example "pubmed search" for illustration, because the word "pubmed" can be mapped to MeSH so the final executed search is ["pubmed" (MeSH terms) or "pubmed" (all fields)] and

<sup>1</sup> Dr Saeeda Abdullah, BDS, MCPS, FCPS, Assistant Professor, Dental Mentor Silver City Plaza, G-II Markaz, Islamabad  
Email: dr\_saeeda@hotmail.com

**Received for Publication:** February 25, 2015

**Revised:** May 26, 2015

**Approved:** May 30, 2015

“search” (all fields)’ where the PubMed search tags (all fields) and (MeSH terms) indicate the preceding word will be searched in all indexed fields or only the MeSH indexing field, respectively.<sup>3-7</sup>

**How to start search on PubMed Homepage**

Log in to Pubmed (www.pubmed.gov), it will open Query box (Fig 1 with red circle), a text box into which search terms are entered.

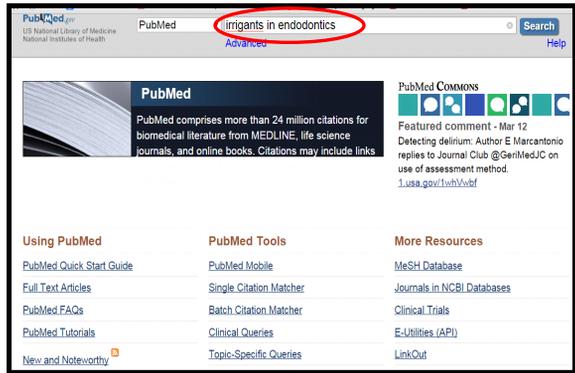


Fig 1: Home page PubMed

Type “irrigants in endodontic” in the Query box and select the Go button.

Citations are displayed from the newest to the oldest in a default summary format (Fig 2 with red circle). PubMed will search back to 1966 unless a date limit is specified. The numbers of citations are shown on the resulted page and the first 20 are shown by



Fig 2: Shows the results

default. More citations can be displayed by selecting the desired number with the drop-down menu (Fig 2 with red arrow).

**PubMed literature Search (Basic Search Strategies)**

1. Boolean operators (“AND”, “OR”, “NOT”)
2. Filters
3. Search field tags (e.g. [au] means “author”, “Rutherford[au]”)

4. Use of MeSH terms
5. Wildcard ( “ \* ”)
6. Search with key words

**1. Boolean Operators**

PubMed searches can be refined by using multiple search terms connected by Boolean operators: AND, OR, and NOT. Boolean operators must be entered in upper case. PubMed uses “AND” as default operator. If any Boolean operator is not used in search, then PubMed will automatically use “AND” between terms. The operator AND selects the references that contain both search terms, OR selects the references that contain either search term, and NOT selects the references that contain the first term but not the second term. NOT should be used carefully, if at all. NOT is an absolute, and it can have unintended effects.<sup>8-10</sup>

For example, the query “dental caries NOT xerostomia” might eliminate articles in which the subjects have xerostomia, and it would also eliminate articles dealing with xerostomia due to other causes like radiation or medicines.

**i)- “OR” (Boolean operator)**

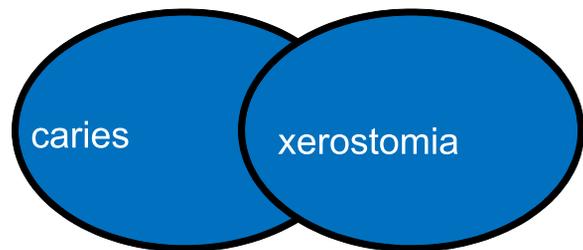


Fig 3: Caries OR xerostomia. It will have all the articles of caries and all articles of xerostomia. It does not matter whether its caries or xerostomia only (Fig 3).

**ii)- “NOT” (Boolean operator)**

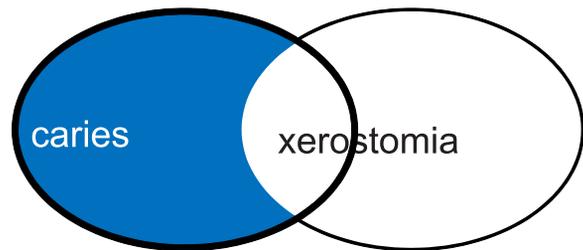


Fig 4: Caries NOT xerostomia

Boolean “NOT” is used when you want to exclude all articles with caries and xerostomia or containing word xerostomia. It will include all articles of caries except with word xerostomia. (Fig 4)

## iii)- “AND” (Boolean operator)

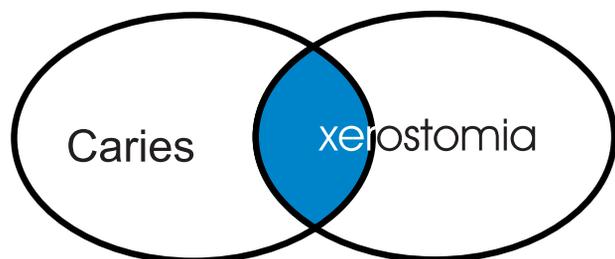


Fig 5: Caries AND xerostomia. Include all articles of caries with word xerostomia. When you want article with both caries and xerostomia, but you don't want any article with only caries or xerostomia alone (Fig 5). The result will be less articles than other Boolean terms “OR” and “NOT”.

## 2- Limits/ Filter

PubMed allows users to refine searches further by imposing Limits including article type, text availability, publication dates, species, languages, sex, subjects, journal categories, ages and search fields. (Fig 2 with purple arrow and purple outlined area). To activate a sidebar filter, click the filter selection. A checkmark will appear next to the activated filters. Subsequent searches will be filtered until the selected filters are cleared.

All these filters may exclude "in process" and "supplied by publisher" citations because they have not yet completed the MEDLINE indexing process.

To add additional filter categories to the sidebar, click the “Show additional filters” link, select the additional categories, and then click Show. To activate the additional filters, click the filter selection. When filters are selected a Filters activated (Fig 6 with orange arrow) message will display on the results page. To turn off filters, click either the “Clear all” link (Fig 3 with orange circle) to remove all the filters, or the “clear” link (Fig 6 with yellow arrows) next to a filter category to clear the individual filter. Only valid filter options for a result set will display as a sidebar selection. Users can also activate additional filters with NCBI filters.

**Article types:** Results can be narrowed by selecting article types, but based on the type of material the article represents, such as: Clinical Trial and Review.

**Text availability:** To filter results to only citations that include a link to full text, a link to free full text, or an abstract, click the appropriate selections.

**Ages:** Age filters restrict results to a specific age group for a human study. To add ages to the sidebar, click the “Show additional filters” link, select Ages, and then click Show (Fig 6 with red arrow).

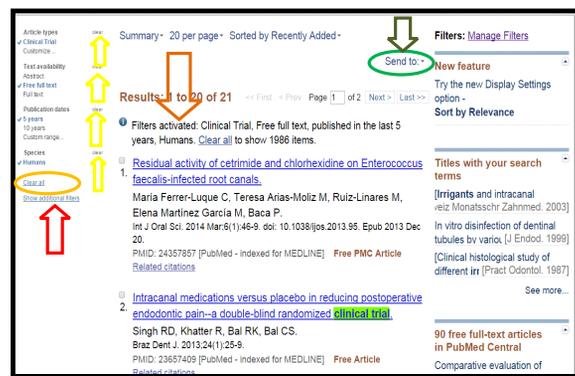


Fig 6: Different activated filters and saving options for the articles

A check mark will be displayed in front of Limits to signify that limits are imposed for the results of this search. The specific Limits are displayed beneath the Features bar. Placing Limits on the search decreases the number of citations that need to be reviewed.

The user can connect to the full-text article online by selecting the journal icon displayed at the top of the abstract display. In some instances, the article is available or will be after a certain period. In most cases, however, access to the full text is limited to those individuals and institutions subscribing to the electronic version of the publication. For individuals without access to a medical or hospital library, articles can be ordered through the NLM through Loansome Doc (explored in more detail through the Order Documents or the NLM Gateway hyperlink on the blue sidebar on the left-hand side of the screen).

## 3- Helpful search field Tags: (also appear in advanced search)

Writing of search tags in query box can also be used for searching, 11 for example [au] for author, [ti] for title, [tiab] for title /abstract, [tw] for text word. Example: cross infection [ti] OR Hepatitis[tw].

## 4- MeSH Terms [MH]

The NLM Medical Subject Headings uses controlled vocabulary of biomedical terms that is used to describe the subject of each journal article in MEDLINE. MeSH contains approximately 26 thousand terms and is updated annually to reflect changes in medicine and medical terminology. MeSH terms can be selected for searching in the MeSH database and from the advanced search builder index.<sup>12</sup>

MeSH terms are arranged hierarchically by subject categories with more specific terms arranged beneath broader terms. PubMed allows to view this hierarchy and select terms for searching in the MeSH database. To turn off this automatic feature, use the search syntax [mh:noexp], e.g., neoplasm [mh:noexp]. To search the

term only as a MeSH term, it must be tagged using the search field, e.g., [mh] for MeSH Terms or [majr] for MeSH Major Topic. A tagged term is checked against the MeSH translation table, and then mapped to the appropriate MeSH term(s). To turn off mapping to multiple MeSH terms, enter the tagged MeSH term in double quotes.

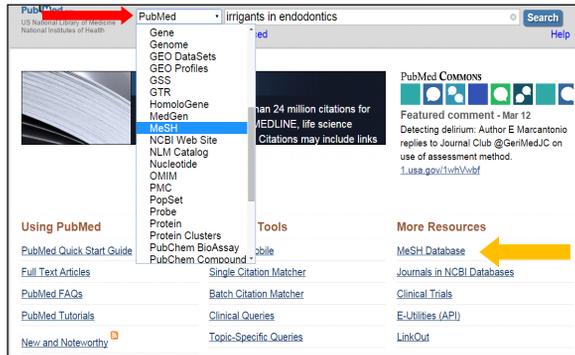


Fig 7: Two ways to go to MeSH database. Either select from drop down menu (Red arrow) or directly (Yellow arrow)

**MeSH Major Topic [MAJR]**

A MeSH term is one of the main topics discussed in the article denoted by an asterisk on the MeSH term or MeSH/Subheading combination, e.g., Cytokines/physiology\*.

**MeSH Subheadings [SH]**

MeSH subheadings are used with MeSH terms to help describe more completely a particular aspect of a subject. PubMed increases search by “automatic term explosion.” Journal articles are entered into PubMed with both MeSH headings (e.g., MTA) and Subheadings. For example, the drug therapy of asthma is displayed as asthma/drug therapy.

The MeSH Subheading field allows users to "free float" Subheadings, e.g., hypertension [mh] AND toxicity [sh]. MeSH Subheadings automatically include the more specific subheading terms under the term in a search. To turn off this automatic feature, use the search syntax [sh:noexp], e.g., therapy [sh:noexp]. In addition, two-letter is entered as MeSH subheading abbreviations rather than spelling out the Subheading, e.g., dh [sh] = diet therapy [sh].

**5- Comprehensive search using wildcard “ \* ”**

If we type word “Prevent \*”, it will finds articles with the words; prevention, preventative or preventable etc. This asterisk will basically substitute for all the later combinations that happened after the word prevent without specifically typing prevention, preventable, preventive separately. Asterisk is also helpful for finding American or British version of words like diarrhea or

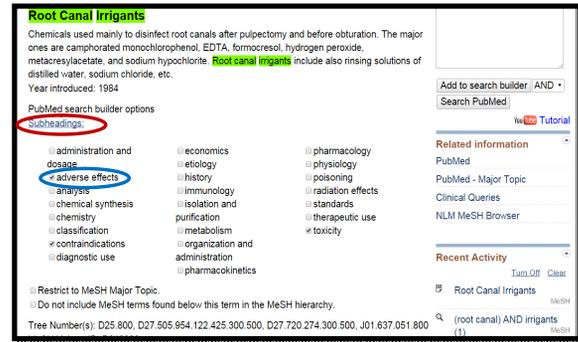


Fig 8: Sub heading of MeSH term

diarrhea, if we write “diarrh \* ” in search box. It will search in all possible words with initial word of “diarrh” e.g. diarrhea, diarrhoea, diarrheal, diarrhoeal etc.<sup>10,13</sup>

**6- Search with key words**

Place searching words in quotation marks (“”) which will make grouping of search word. If search is done without phrasing for example to search for endodontic restoration, then result will show either article of endodontic or restoration but not as one word. Therefore the result will show large numbers of articles. Placing quotation marks around the “endodontic restoration” will be taken as phrase by web and search will just be for that phrase.

**Some more Search tricks**

- 1- When searching word has multiple meanings then add negative or plus sign after that word: for example word Mullet (means hairstyle or a type of fish). Therefore if you don’t want hairstyle articles then write Mullet-hairstyle, which will give results by excluding hair style articles.
2. Regarding the user interface and usability, the “My NCBI” tool was introduced to PubMed, which let users select and create filter options, save search results, apply personal preferences like highlighting search terms in results, and share collections of citations. It also allows users to set automatic emails for receiving updates of saved searches. Additional search help such as a spell checker and query auto-complete have also been deployed in PubMed.<sup>14</sup>
3. If you want to see the abstracts of all resulted results, then click “display heading”, and click on abstract from drop box menu. Abstract format is the only format in PubMed which displays the options of “get it VCU button”, which is important because it gives direct access to all electronic resources. If an article is not available electronically then it can be retrieved from shelf (may be in print or may give option of interlibrary loan request which is free for the staff and students in many libraries).<sup>15</sup>

4. Break the search in concept, so if you are looking for beverages and its influence on caries, then place caries in the search box. Results will show articles on caries only after clicking search. If want to see the relation of beverages with caries then go back to history and bind caries and beverages together. This can be done in advance search, where in search history click on add the search (on any search history, which you want to add in builder, click one by one). One search will be in one query and add Boolean operator to show the relationship between 2 queries. Click the search to get results.<sup>8</sup>

The History feature is available after the first search and reviews the results of up to 100 prior searches. PubMed also provides users with an electronic Clipboard that can be used to compile up to 500 citations from multiple searches. The Clipboard erases the contents after 8 hour of inactivity.<sup>11</sup>

Users can visualize the search techniques employed by PubMed with the Details feature. The Features bar and Query box are available from every screen. PubMed also provides resources such as an Overview, online Help, and frequently asked questions (FAQ) on each screen.

### ADDITIONAL FEATURES

PubMed provides all users with the option to register for the Cubby, which is free of charge. This feature allows users to store and update searches. First-time users must register and select a unique password for access. This resource might be used to conduct systematic reviews because it allows users to update searches quickly, which can be kept in the Cubby indefinitely.<sup>11,12</sup>

### ADVANCED FEATURES

PubMed provides many other advanced features that will assist interested clinicians in finding the evidence necessary to make informed clinical decisions.<sup>11,12</sup>

### REFERENCES

- 1 A. Thomas Agoritsas et al. Sensitivity and Predictive Value of 15 PubMed Search Strategies to Answer Clinical Questions Rated Against Full Systematic Reviews. *J Med Internet Res.* 2012 May-Jun; 14(3): e85.

- 2 Zhiyong Lu. PubMed and beyond: a survey of web tools for searching biomedical literature. *Database*, Vol. 2011, Article ID baq036, doi:10.1093/database/baq036
- 3 Jon o. Ebbert, denise m, Dupras and patricia j. Erwin. Searching the Medical Literature Using PubMed: A Tutorial. *Mayo Clin Proc.* 2003; 78: 87-91.
- 4 McGowan J, Sampson M: Systematic reviews need systematic searchers. *Journal of the Medical Library Association* 2005, 93: 74-80.
- 5 National Library of Medicine. PubMed overview. Available at: <http://www.ncbi.nlm.nih.gov:80/entrez/query/static/overview.html>. Accessibility verified October 31, 2002.
- 6 Oxman AD, Sackett DL, Guyatt GH, Evidence-Based Medicine Working Group. Users' guides to the medical literature, I: how to get started. *JAMA.* 1993; 270: 2093-095.
- 7 Haynes RB, Wilczynski N, McKibbon KA, Walker CJ, Sinclair JC. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc.* 1994; 1: 447-58.
- 8 Zhang L, Ajiferuke I, Sampson M. Optimizing search strategies to identify randomized controlled trials in MEDLINE. *BMC Medical Research Methodology* 2006, 6(23): 10-58.
- 9 A. Thomas Agoritsas et al. Sensitivity and Predictive Value of 15 PubMed Search Strategies to Answer Clinical Questions Rated Against Full Systematic Reviews. *J Med Internet Res.* 2012 May-Jun; 14(3): e85.
- 10 Karimi S1, Pohl S , Scholer F, Cavedon L, Zobel J.: Boolean versus ranked querying for biomedical systematic reviews. *BMC Medical Informatics and Decision Making* 2010: 58.
- 11 F Fatehi, Gray LC, Wootton R. How to improve your PubMed/MEDLINE searches: 1. background and basic searching. *J Telemed Telecare.* 2013 Dec; 19(8): 479-86.
- 12 F Fatehi, Gray LC, Wootton R. How to improve your PubMed/MEDLINE searches: 3. advanced searching, MeSH and My NCBI. *J Telemed Telecare.* 2014 March; 20(2): 102-12.
- 13 BMA: British Medical Association (BMA) Library - MEDLINE Plus, Basic Course - Notes for OvidSP [[http://www.bma.org.uk/images/medlinesp2010\\_tcm41-199937.pdf](http://www.bma.org.uk/images/medlinesp2010_tcm41-199937.pdf)], (accessed 27 September 2010) 2010.
- 14 Stewart MG, Kuppersmith RB, Moore AS. Searching the medical literature on the Internet. *Otolaryngol Clin North Am* 2002; 35: 1163-74.
- 15 National Library of Medicine. PubMed overview. Available at: <http://www.ncbi.nlm.nih.gov:80/entrez/query/static/overview.html>. Accessibility verified October 31, 2002.