

## Searching the Medical Literature Using PubMed: A Tutorial

JON O. EBBERT, MD; DENISE M. DUPRAS, MD, PhD; AND PATRICIA J. ERWIN, MLS

Staying current with advances in medicine is becoming a major challenge for clinicians. Access to updated repositories of medical information and the expertise to locate relevant information within them are becoming necessary clinical skills. PubMed (<http://www.pubmed.gov>) provides free access to the largest biomedical resource available and is updated daily. Clinicians can use this resource to find answers to focused clinical questions quickly and effi-

ciently. The purpose of this article is to assist clinicians in the development of the basic skills required to use PubMed to make informed clinical decisions.

*Mayo Clin Proc.* 2003;78:87-91

av = arteriovenous; MeSH = Medical Subject Heading; NLM = National Library of Medicine

The amount of medical information doubles every 5 years.<sup>1</sup> For practicing clinicians, staying abreast of advances in medicine is becoming a major challenge. To continue to provide the best possible patient care, clinicians must have access to resources that will assist them in making clinical decisions. Although several commercial products are available that synthesize large amounts of medical information into easily digestible reviews, these services frequently rely on overviews conducted by experts and may not be current. Furthermore, clinicians may have specific clinical questions that are better answered with a single article or systematic review rather than a broad topical review.

PubMed, a resource developed by the National Center for Biotechnology Information at the National Library of Medicine (NLM), provides free access to MEDLINE (MEDLARS Online) and other databases to any person who has Internet access.<sup>2</sup> MEDLINE is the NLM database of all published biomedical literature relevant to clinical practice. It is updated daily and contains more than 11 million references dating back to the 1960s from more than 4000 selected journals from 70 countries. Clinicians seeking information can search PubMed through Entrez, a text-based search and retrieval system at <http://www.pubmed.gov>. Entrez also provides access to other research-oriented databases.

The purpose of this article is to introduce clinicians to this powerful resource. We use clinical cases to highlight

several important features, and we encourage readers to conduct the searches online as they work through the examples. All the actions that readers should complete are in italics. The names of the features in PubMed are capitalized.

### CASE 1

A 45-year-old man presents with a 10-month history of burning epigastric pain that has not improved with a proton-pump inhibitor. Findings on upper endoscopy are normal, but *Helicobacter pylori* serology is positive. You ask yourself, "In patients with non-ulcer dyspepsia, does the treatment of *Helicobacter pylori* improve symptoms?"

### Developing Well-Built Clinical Questions: Key to Successful Searching

The most successful searches are accomplished if users select the most precise terms for the search. Finding numerous, apparently irrelevant articles is likely to discourage the clinician from looking further. These terms are best defined by learning to ask "well-built clinical questions."<sup>3</sup> Such questions are constructed by breaking down questions that arise from a clinical encounter into 4 components: (1) patient/problem, (2) intervention, (3) comparison intervention, and (4) outcome. For the aforementioned clinical question, the patient/problem would be "non-ulcer dyspepsia," the intervention would be "treatment of *Helicobacter pylori* infection," the comparison intervention would be "no treatment," and the outcome would be to "improve symptoms." These phrases can be used as PubMed search terms to retrieve relevant citations that will assist in making the appropriate clinical decision. At least 2 terms should usually be entered to begin the search (ie, "*helicobacter pylori*" and "non-ulcer dyspepsia"), but more can be entered to increase search relevance.

From the Division of Community Internal Medicine (J.O.E., D.M.D.) and Mayo Medical Library (P.J.E.), Mayo Clinic, Rochester, Minn.

This work was supported in part by grant R01 CA96881-01 from the National Cancer Institute.

A question-and-answer section appears at the end of this article.

Address reprint requests and correspondence to Jon O. Ebbert, MD, Division of Community Internal Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905 (e-mail: [ebbert.jon@mayo.edu](mailto:ebbert.jon@mayo.edu)).



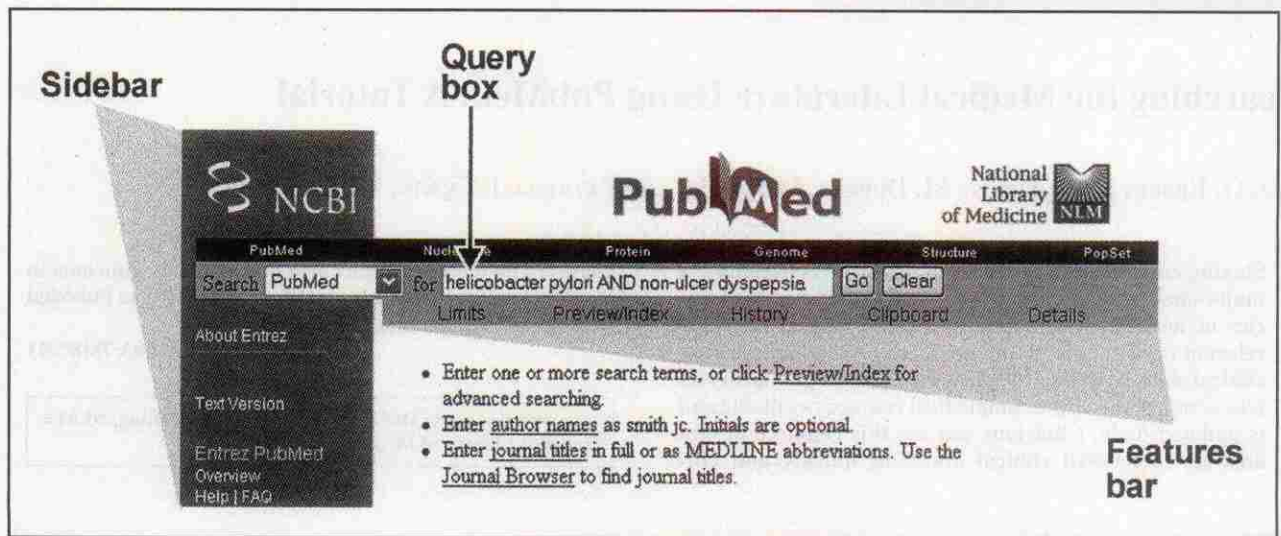


Figure 1. Entrez at <http://www.pubmed.gov>.

### PubMed Homepage

Logon to Pubmed (<http://www.pubmed.gov>).

The PubMed homepage (<http://www.pubmed.gov>) provides users with a Query box, a text box into which search terms are entered (Figure 1). The black menu bar above the Query box provides access to other databases. The homepage provides a Features bar that allows users to set Limits on searches and Preview the number of results that a search will retrieve before displaying the citations. 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 1 hour of inactivity. 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 the blue sidebar on the left-hand side of each screen.

Type "helicobacter pylori" in the Query box and select the Go button.

Citations are displayed from the newest to the oldest in a default Summary format. PubMed will search back to 1966 unless a date limit is specified. The drop-down menu Display allows the user the option of viewing the retrieved citations in a variety of formats. The number of citations are shown on the gray bar, and the first 20 are shown by default. More citations can be displayed by selecting the desired number with the drop-down menu on the left. The number of pages of citations is also displayed on the gray bar, and specific pages can be reviewed by selecting the blue numbers on the right. Users can save the references and display them as a

text file. This text file can be imported easily into commercially available reference manager software.

### 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. 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. For example, the query "heart attack NOT smoke" might eliminate articles in which the subjects are smoking, and it would also eliminate articles dealing with secondhand smoke, nonsmoking, or smoking cessation. PubMed processes all Boolean operators in a left-to-right sequence unless closed off by parenthesis, eg, helicobacter pylori AND (non-ulcer dyspepsia OR gastritis).

Enter "helicobacter pylori AND non-ulcer dyspepsia" in the Query box and select Go.

### Limits

PubMed allows users to refine searches further by imposing Limits, including publication type, age and sex of research subjects, date of publication, language, human or animal, or journal subsets.

Select the following limits from the drop-down menus: randomized-controlled trial; middle aged, 45-64; English; human; and male. Select Go. Look for an article that answers your clinical question.

A check mark will be displayed in front of Limits to signify that limits are imposed for the results of this search.



**Search filter** ➔  **Clinical Queries using Research Methodology Filters**

This specialized search is intended for clinicians and has built-in search "filters" based largely on [Haynes RB et al.](#) Four study categories are provided, and the emphasis may be more sensitive (i.e., most relevant articles but probably some less relevant ones) or more specific (i.e., mostly relevant articles but probably omitting a few). See [filter table](#) for details.

Indicate the category and emphasis below:

**Article type** ➔ Category:  therapy  diagnosis  etiology  prognosis

**Search emphasis** ➔ Emphasis:  sensitivity  specificity

**Systematic Reviews**

This feature retrieves systematic reviews and meta-analysis studies for your search topic(s). For more information, see [Help](#). [Related sources](#) are also provided.

**Query box** ➔ Enter subject search:

Note: If you want to retrieve everything on a subject area, you should not use this screen. The objective of filtering is to reduce the retrieval to articles that report research conducted with specific methodologies.

Figure 2. Clinical Queries search filters.

The specific Limits are displayed beneath the Features bar. Placing Limits on the search decreases the number of citations that need to be reviewed to answer your clinical question. Clicking on a citation title "hyperlinks" (ie, different-colored text that links to other Web pages) to the abstract. 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).

## CASE 2

A 47-year-old woman with cachexia and a history of chronic alcoholism presents with a 3-day episode of epigastric pain, anorexia, and nausea. Findings on a diagnostic evaluation are normal except for elevated pancreatic enzymes and an albumin level of 1.5 g/dL. After 24 hours, she feels hungry, but a general surgeon recommends continued bowel rest. You are concerned about her nutritional status and wish to start feedings. You ask yourself, "In patients with acute pancreatitis, does enteral versus parenteral nutrition decrease the length of hospitalization?"

## Search Filters

PubMed allows users to search for articles using Clinical Queries and Systematic Reviews search filters. These

can be accessed through the Clinical Queries hyperlink on the sidebar on the left-hand side of the screen (Figure 2). Clinical Queries are validated search strategies to locate articles on therapy, diagnosis, etiology, or prognosis.<sup>4</sup> These "filters" or "hedges" are designed to allow clinicians to search by maximizing either the specificity or the sensitivity of the search. The Systematic Reviews search filter will locate articles identified as systematic reviews, meta-analyses, reviews of clinical trials, clinical reviews, guidelines, evidence-based medicine, and consensus development conferences. Either filter can be used by selecting the button in front of it. Search terms need to be entered in the Query box on the bottom of the Clinical Queries page. Searches can be further refined by imposing Limits on the retrieved citations.

*On the Clinical Queries search filter page, select Clinical Queries, "therapy," and "sensitivity." Type "acute pancreatitis AND enteral nutrition AND hospital" in the Query box and select Go. Look for an article that answers your clinical question.*

For this search we entered an additional term for outcome (ie, hospital) instead of 2 terms as we have done before. Entering more search terms may assist in increasing the search relevance.

## CASE 3

An 82-year-old man with a history of gastrointestinal bleeding from arteriovenous (av) malformations presents again with hematochezia, anemia, and orthostatism. He was previously found not to be a surgical candidate, and you wonder about additional therapeutic options when your colleague mentions estrogen therapy. You are un-



aware of the evidence and ask yourself, "In patients with recurrent gastrointestinal bleeding due to *av* malformations, does hormonal therapy result in fewer episodes of bleeding and fewer transfusions than treatment with placebo?"

*Type in "av malformations AND estrogens" and select Go. As you look through the retrieved references, you will note that the articles do not answer your clinical question.*

### Using MeSH Terms to Construct Searches

The language of medicine is rich. Concepts can be described by using a wide range of synonyms. Case 3 illustrates the pitfalls of depending on familiar key words alone. MeSH (*Medical Subject Heading*) is NLM's hierarchical vocabulary that is used for indexing articles in PubMed, with more specific terms organized underneath more general terms. MeSH contains more than 19,000 terms and is updated annually to adapt to changes in medicine and medical terminology. Subject analysts examine each journal article indexed in PubMed and assign the most specific MeSH terms applicable, usually 10 to 12. This provides a reliable way to retrieve citations that use different terminology for the same concept because it indexes articles based on content rather than on the words of the authors.

Two additional features of PubMed capitalize on the MeSH vocabulary and increase the search yield, providing distinct advantages over other commonly used search engines. One of these features of PubMed is "automatic term mapping." For each phrase that users enter into the query line, PubMed attempts to translate the terms by matching them against term libraries in a consistent manner. PubMed first checks them against a library of MeSH terms, then journal names, then commonly used phrases, and finally against author names. If PubMed finds the term in one of these libraries, it conducts the search by using this phrase in the appropriate search field (ie, author name field, MeSH term field, etc) and also searches for it as a text word in all search fields.

The other feature of PubMed that increases search yield is "automatic term explosion." Journal articles are entered into PubMed with both MeSH headings (eg, estrogens) and Subheadings. For example, for an article on the use of estrogen as a therapeutic intervention, the subheading "therapeutic use" would be used to characterize a publication more completely. MeSH Subheadings are further grouped into families of subheadings (to see MeSH Subheading categories, hyperlink to Help through the blue sidebar and see MeSH Subheadings and Families of MeSH Subheading Explosions). When PubMed searches for a term, it automatically "explodes" or includes the more specific terms (eg, "estrone" and "estriol" for the term

"estrogen") and all the MeSH Subheadings in the family of subheadings. This effectively increases the search yield. Any time during a search, users can see how PubMed has executed the search by selecting Details in the features menu.

*Hyperlink through the MeSH Browser on the blue sidebar on the left-hand side of the screen. Type in "av malformation" in the Query box, select Go. Select "arteriovenous malformation" from the list of provided terms and select Browse this term.*

PubMed was not able to translate your term but instead generated a list of other possible terms to use in the search. Additionally, PubMed defined the term "arteriovenous malformation" and displayed where it appears in the MeSH hierarchy.

*Hyperlink to "vascular diseases" to explore the MeSH terms included under this MeSH heading.*

Notice that 2 terms could be used to obtain evidence to answer your clinical question: "arteriovenous malformation" and "angiodyplasia."

*Hyperlink to "angiodyplasia." Select Add.*

PubMed begins to construct the search in the window (Figure 3).

*Hyperlink to "vascular diseases" and hyperlink to "arteriovenous malformations." Add this term but change the Boolean operator to OR. Type "estrogens" in the top Query box and select Go. Select the Boolean operator AND, select Add. Press PubMed Search to execute the search.*

If you are interested in a higher level of evidence (ie, randomized-controlled clinical trial), you can set Limits to clinical trial. At the right of the screen is a link entitled Related Articles. This PubMed feature searches for similar citations by using a complex algorithm that scans titles, abstracts, and MeSH terms.

### ADDITIONAL FEATURES

PubMed provides all users with the option to register for the Cubby, which is free of charge. This feature allow 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.

### ADVANCED FEATURES

PubMed provides many other advanced features that will assist interested clinicians in finding the evidence necessary to make informed clinical decisions. We encourage users who become more comfortable with basic features to challenge themselves to use the Help, Overview, and Tutorials that are available through the blue sidebar on the left-hand side of the screen.

**Search strategy** "Angiodysplasia"[MESH] OR "Arteriovenous Malformations"[MESH] AND "estrogens"[MESH]

**Term definition** Estrogens [Detailed display]  
A class of sex hormones associated with the development and maintenance of secondary female sex characteristics and control of the cyclical changes in the reproductive cycle. They are also required for pregnancy maintenance and have an anabolic effect on protein metabolism and water retention.

**Boolean operators** Add this term to the Search using operator: AND

Term **Estrogens** appears in more than one place in the MeSH tree.

**MeSH hierarchy**

```

All MeSH Categories
  Chemicals and Drugs Category
    Hormones, Hormone Substitutes, and Hormone Antagonists
      Hormones
        Estrogens
          Equilenin
          Equilin
          Estradiol
          Estriol
            Estetrol
          Estrogens, Catechol
  
```

Figure 3. Constructing search strategies with MeSH (Medical Subject Heading) terms.

## REFERENCES

1. Mattox DE. Welcome to ARCHIVES CME. *Arch Otolaryngol Head Neck Surg.* 2000;126:914.
2. 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.
3. Oxman AD, Sackett DL, Guyatt GH. Evidence-Based Medicine Working Group. Users' guides to the medical literature, I: how to get started. *JAMA.* 1993;270:2093-2095.
4. 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;1:447-458.

## Questions About Using PubMed

1. Which one of the following is not a component of a well-built clinical question?
  - a. Patient/problem
  - b. Setting
  - c. Intervention
  - d. Comparison intervention
  - e. Outcome
2. Which one of the following search strategies is the most correct if the user were interested in obtaining all citations that included information on lung cancer caused by cigars and pipes?
  - a. Lung neoplasm OR smoking
  - b. Lung neoplasm AND smoking
  - c. Lung neoplasm AND (cigars AND pipes)
  - d. Lung neoplasm AND (cigars OR pipes)
  - e. Lung neoplasm AND cigars AND pipes
3. Which one of the following Clinical Queries will not provide the user the ability to employ search filters to assist in locating articles that address the following topics?
  - a. Therapy
  - b. Diagnosis
  - c. Harm
  - d. Prognosis
  - e. Systematic reviews
4. Which one of the following hierarchical languages is used by NLM to index articles and increase the power of search strategies?
  - a. Boolean
  - b. NLM subheadings
  - c. Automatic term mapping
  - d. Automatic term explosion
  - e. MeSH
5. Which one of the following buttons on the Features bar should be selected for a quick and efficient way to view the most searches conducted but not saved?
  - a. History
  - b. Clipboard
  - c. Limits
  - d. Preview/Index
  - e. Details

Correct answers:

1. b, 2. d, 3. c, 4. e, 5. a

