

Searching the “Wild, Wild Web”

by Ellen Drake, CMT, AHDI-F, and
Georgia Green, CMT, AHDI-F

Note: This article is based on multiple presentations on Internet research given by Georgia Green and Ellen Drake over the past several years.

In order to meet the demands of our profession, we must educate ourselves on the job and do so throughout our careers. We learn by doing research. It's the life blood of what we do, and in a perfect world we would be free to research as much as we felt was necessary. But the reality of the production environment intrudes on our perfect world. We must find the information we need quickly, with as little effort as possible. We must be efficient researchers.

The Right Tool for the Job

We must use the tools available to us in the most effective way possible. There isn't always one best tool for every circumstance. But you can narrow your choices to the ones most likely to yield results—and use those resources in the most efficient way.

It's probably pretty obvious that the single largest factor that inhibits your productivity as a medical transcriptionist is when you take your hands off the keyboard. Efficiency in research impacts productivity more than how well you use your expander, word processor, or anything else. If you are an efficient researcher, you can find what you need in 60 seconds. That's Georgia Green's rule. For Georgia's guidelines on keeping your research time down, see her article “Feeling the Need for Speed” on the Downloads page at <http://www.hpisum.com>.

So, where do you go when your books let you down?

The Internet: The “Wild, Wild Web”

Yes, as filled as it is with inaccurate and questionable content, the Internet *can* be a useful tool—if you know its limitations, how to evaluate your sources, and how to use it efficiently. Note my translation of *www*—it's only half in jest as you will see in a moment.

In 1998, the first Google index already had 26 million pages; by 2000 it had reached the one billion mark. Last year, Google hit a milestone: 1 trillion (as in 1,000,000,000,000)

unique URLs on the Web at once! Add in duplicates and related links, and the count is over 8 trillion. It's hard to really understand that number, but imagine if you clicked on a new link every second of every day around the clock for a lifetime of 80 years.

Advantages of the Internet

- It's cheaper.
- It can be faster.
- Information can be more up-to-date.
- Man, there's a lot of information there.
- It's electronic, you can copy and paste.
- You can access the Internet from almost anywhere.

Limitations of the Internet

- It can be slower.
- It's contradictory.
- It can be a big distraction.
- It's unedited, unreviewed.
- It takes a lot more discernment.
- Man, there's a lot of information there.
- Information can be outdated and irrelevant.

Perhaps the Biggest Disadvantage of All Is . . .

You lose the chance for serendipity. That, is, you lose the ability to browse and learn other words and related concepts to the term you're looking for. Many MTs can't look up a word in a dictionary without devouring the other entries on the page and then end up moving from page to page, looking up other terms. Sometimes, when you're searching for a difficult spelling, browsing is the only way to find it.

Of course, you can also find interesting rabbits to chase on the Internet; it's just that more often than not, those rabbits adversely affect your productivity and don't improve your knowledge for job-related tasks.

Don't get us wrong. We love, love, love the Internet. We're both pretty good at evaluating the content and using it efficiently. But let's face it, many MTs think the Internet is a wonderful, free resource and they no longer need any reference books. How cheap is it, though, if it sucks up productive time? And for transcriptionists, time is money—right?

And if you fail to evaluate your source and plug in a wrong word, you can have a far bigger problem than just loss of productivity.

Do you remember the gossip game from when you were a kid? You whisper a message into someone's ear, and they whisper it in the next person's ear, and so on down the row, and by the time it has gone through a dozen people, someone says it aloud and it no longer bears any resemblance to the message you started with. Well, the Internet has a similar problem.

If you found a ton of errors in a reference book, would you really trust it? What if you were handed a reference and told that nine-tenths of what it contained was incorrect—wouldn't you just toss it rather than waste your time? But what if you were told that the other one-tenth of the material in that reference actually contained almost everything you would ever need to know? Oh, boy . . . that reference would be invaluable, wouldn't it? But how would you know which information was accurate and which should be ignored? And how could you find the valuable and correct information quickly?

Let's look at an example of an Internet research dilemma that was posted on MT Chat a while back.

Dorland's spellchecker flags *hyperhomocystinemia*.
Stedman's Medical Terms & Phrases and *Stedman's Medical Speller*, 3rd edition, have *homocystinemia*.
Stedman's GI/GU, 2nd edition, has *hyperhomocystinemia*.
Stedman's Dermatology/Immunology has *cysteine*.
Stedman's Neurology, 2nd edition, has *cysteine* and *cystine*.

Both spellings appear in credible Web sites (PubMed, AHA Journals, Wikipedia, Blackwell-Synergy). The poster who answered the question was able to distinguish between the definitions for *cysteine* and *cystine* using a Wikipedia entry. This dilemma probably took an excessive amount of the original poster's time, not to mention the time of all the other posters who tried to help.

Misspellings of drug names abound. If you use the Web as a reference source, you can usually find multiple ways to spell just about anything. The problem is that you can't tell which is right and which is wrong. As authors of *Saunders Pharmaceutical Word Book*, Randy and Ellen Drake frequently get e-mail from readers writing something like the following: "Your book has Lanacane. Google has it spelled Lanacaine," immediately assuming that the spelling in the Drake drug book was an error.

First of all, Google doesn't "spell" anything. Google is a search engine that indexes words found on Web pages. And yes, there are 1840 pages on the Web that have the product misspelled as Lanacaine, which Google indexes right along with the 62,600 pages that have it spelled correctly. There were 269 pages that had Lanacane misspelled and 7860 pages that had it spelled correctly in 2004 when we were first contacted.

You can find "Levothyroid" all over the Web, but nobody makes "Levothyroid"—anywhere in the world, either now or in the past. At least 28,500 Web pages represents a lot of misspelling perpetrated on the Web, doesn't it? The correct spelling is *Levothroid* (without a "y"). What's more concerning is that some of the "hits" are what would be considered reputable sources.

Just one more example, and we'll move on. There are 15,000 pages that contain *zithromycin*, 7350 that have *azithromicin*, and 224 that have *zithromicin*; the correct spelling is *azithromycin* with over 2 million pages having the drug correctly spelled. Google asks, "Did you mean *azithromycin*?" for all of the misspelled search terms, but clearly, not everyone notices that question. And, sometimes, Google's first results can be correctly spelled, and the "Did you mean . . ." message below the search box can have the *incorrect* spelling!

How much time did all this research resulting in inaccurate results take? How do you decide who's right? Randy Drake, coauthor of *Saunders Pharmaceutical Word Book*, candidly points out: "The Internet can be a wonderful source of entertainment, but if you want to find the correct spelling of a drug quickly and with certainty, you should stick to high-quality medical reference books that have a reputation for accuracy. Can you find information on the Internet? Sure you can, but you can't always trust it to be accurate. *Professional* transcriptionists use *professional* medical references."

Even so, there are times when books aren't enough, so it's important to learn when the Internet is your best resource, how to evaluate your search results, and how to be an efficient researcher.

Evaluating the Web

There is a famous Steiner cartoon published in the *New Yorker* (July 5, 1993) with two dogs sitting before a terminal looking at a computer screen; one says to the other, "On the Internet, nobody knows you're a dog." In a similar vein, Andrew Keen has written a book called *The Cult of the Amateur: How Today's Internet is Killing Our Culture*. The Internet has opened up the world to millions of people. It has given them a voice, a way to connect with others who think like them (or don't), and it has made available a wealth of information and misinformation. There are some real "dogs" out there, but there's also great treasure.

What does this teach us? *Discernment is critical*. Books can have errors and be inconsistent, but books are generally proofread and edited. Their errors should be fewer and more minor. Medical journals are peer-reviewed. There is no editing or peer review for most of what you'll find on the Internet. It's important to be able to research more than just spelling.

We're not trying to belabor the point; we merely feel that being forewarned is forearmed. Knowing the limitations of the Internet will make you a more informed user of all that's good about it. Performing a Google search or using any search engine is somewhat like reaching your hand into a huge grab bag full of all kinds of goodies. You may get what you want—

you're just as likely to get what you don't want unless you know what you're doing. To put it another way, using and citing information found over the Internet is a little like swimming without a lifeguard. You shouldn't do it unless you're a really good swimmer.

There is an extremely wide variety of material on the Internet, ranging in its accuracy, reliability, and value. Unlike most traditional information media (books, magazines, organizational documents), no one has to approve the content before it is made public. It's your job as a researcher, then, to evaluate what you locate, in order to determine whether it suits your needs.

In evaluating information on the Internet, it is important to take into account the following factors.

Authority: What are the credentials of the person(s) providing the information?

Affiliation: Who is the sponsor? What is their agenda?

Currency: How current is the information? When was the site last updated? Copyright date is not relevant.

Purpose: Is the information being provided for entertainment or education?

Audience: Is it prepared for patients or clinicians?

Accuracy: Does information on the site contradict itself or other sites? Is the site well developed, free of typos and English usage errors? And, finally.

Verifiability: Is there a bibliography or other resources provided? Where did the provider get its information?

Rather than take up a lot of space going into detail about how to evaluate Web resources, let me point you to a couple of excellent resources that will take you through the process quickly and easily. The first is Georgia Green's article, "Critical Literacy," which goes into detail about how to evaluate the integrity of a Web site

([http://www.hpisum.com/ Downloads.aspx](http://www.hpisum.com/Downloads.aspx)).

Another is "Evaluating Web Pages: Techniques to Apply & Questions to Ask" from the UC Berkeley—Teaching Library Internet Workshops

([http:// www. lib. berkeley.edu/ TeachingLib/ Guides/ Internet/ Evaluate. html#Why](http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/Evaluate.html#Why)).

To summarize, information exists on many levels of quality or reliability. You may have heard that "knowledge is power," or that information, the raw material of knowledge, is power. But the truth is that only some information is power: reliable information. The determination of information quality is something of an art. That is, there is no single perfect indicator.

Whenever a Web site requires that you register in order to access more detailed information, always register as a professional. You don't have to register as a doctor, although Ellen has, on occasion, registered as a physician on some Web sites that provided no other option. One site required that you click on an anatomical structure in the brain on an MRI scan; Ellen simply opened another tab, Googled for an image of what was being asked for, went back to the original site, clicked in the correct spot on the brain scan and got access to

the information she wanted. Usually, however, we register as an "other health professional" or "educator." You may want to set up a separate e-mail account through Google mail (gmail), Yahoo, etc., to use exclusively for registering on Web sites.

Internet Research Techniques

Choosing a search engine is akin to knowing which book to use when. Different search engines are appropriate for different tasks. Google is currently the most popular search engine, and to its credit, it generally lists the most worthwhile pages near the top of its search results—but not always. Google searches blog posts, wiki pages, group discussion threads, and various document formats that are not Web pages per se (e.g., PDFs, Word or Excel documents, PowerPoints). Most of these resources, with the exception perhaps of PDFs which are often journal articles or pure research, are not the sources you want to use for reliable information.

Other useful Google searches include Image search. An Image search brings up photos, drawings, graphics—helpful if you need to see where the navicular bone is, for example. You can avoid certain types of results (like blog posts, wiki pages, and discussion threads) by using Google's Scholar search. In addition, Google "Health" is being beta-tested now. These specialized Google searches will help you not only to limit your searches but also to qualify their reliability and accuracy.

There are other search engines. **Ask.com** is trying hard to compete with Google by giving you a pop-up preview of the page(s) cited before you actually click on the link. You can also Google "most popular search engines" and see what features each touts and which one appeals to you. *Scirus* is a scientific search engine and *Intute* has a health and life sciences specialized search; it's based in the UK, so watch for British spellings.

Research is not one of those skills that just comes naturally. Even if you are very familiar with an English dictionary and a variety of nonmedical electronic references, there are some differences with respect to medical references that still must be learned. This is even more true for medical information on the Internet. It's like searching for that elusive needle in a haystack. You can't begin to evaluate the integrity of information, however, until you are dealing with a manageable amount of information. You can avoid this problem by learning to use a search engine judiciously.

Customize: You can customize your **preferences** and save them, so that, for example, only Web pages in English are searched, and when you click on a link, it always opens in a new window or tab.

To maximize your search, you should learn how to construct an advanced search using boolean logic. Boolean refers to a system that combines key words and certain "operators" (connecting terms) that show relationships between your key words, enabling you to quickly search a large database, which is what the Internet is. When multiple terms are entered for a search using no Boolean operators, a default operation takes

place. Whether results include documents with all the words or any of the words in a search string depends on the search engine's default settings. The search engine's "help" feature or FAQs pages will tell you what the defaults are and teach you how to construct effective searches. Briefly, here are the boolean operators you should be familiar with.

Boolean Operators: These terms are generally typed in all caps to distinguish them from the search string.

AND narrows a search by retrieving only those documents that contain all the keywords.

OR retrieves documents containing any of the keywords.

NOT or the minus sign will retrieve documents that contain the key word(s) but without the word preceded by this operator.

Proximity Searching refers to the ability to specify how close within a record multiple terms should be to each other, such as a phrase search that requires terms to be in the exact order specified within the phrase markings (usually quotation marks). Other proximity operators specify how close terms should be to each other or the order of the search terms. Each search engine can define them differently and use various operator names such as NEAR, ADJ[acent], W[ith], or AFTER.

Nesting: You can nest your search terms within parentheses. With nesting, keywords and operators included in parentheses will be searched for first, then terms and operators outside the parentheses. A search for: **(CVA OR cerebrovascular accident) AND women** will search for documents containing either the acronym *CVA* or the phrase *cerebrovascular accident*, then narrow the search results only to those documents which also contain *women*.

Truncation refers to the ability to search just a portion of a word, typically using an asterisk to represent the rest of the term or a ? to represent a missing letter(s). This may also be referred to as a "wild card search." End truncation is where several letters at the beginning of a word are specified but the ending can vary. **Stemming**, related to truncation, usually refers to the ability of a search engine to find word variants such as plurals, singular forms, past tense, present tense, etc.

Case: Most search engines are NOT case sensitive, so next finds next, Next, NeXT, neXT. If they are case sensitive, you will generally be told so.

Fields: Rather than searching for words anywhere on a Web page, fields define specific structural units of a document. The title, the URL, an image tag, or a hypertext link are common fields on a Web page. Fields searching is often facilitated by check boxes or drop-down menus in an advanced search form.

Limits: Commonly available limits are the date limit and the language limit. In PubMed, which we will discuss later, one of the limits is human which will restrict the search to omit studies done in vitro or on animals.

Stop Words: Frequently occurring words that are not searchable. These are words that are ignored by most search engines unless you have them in quotes or have clicked a box that says "exact phrase."

Stop Words

about, again, all, almost, also, although, always, among, an, and, another, any, are, as at

be, because, been, before, being, between, both, but, by

can, could

did, do, does, done, due, during, . . . etc.

Sorting: The ability to organize the results of a search. For example, by relevance or date.

Advanced: Clicking on Advanced will allow you to use check boxes and drop-down menus to restrict your search, without having to understand Boolean operators.

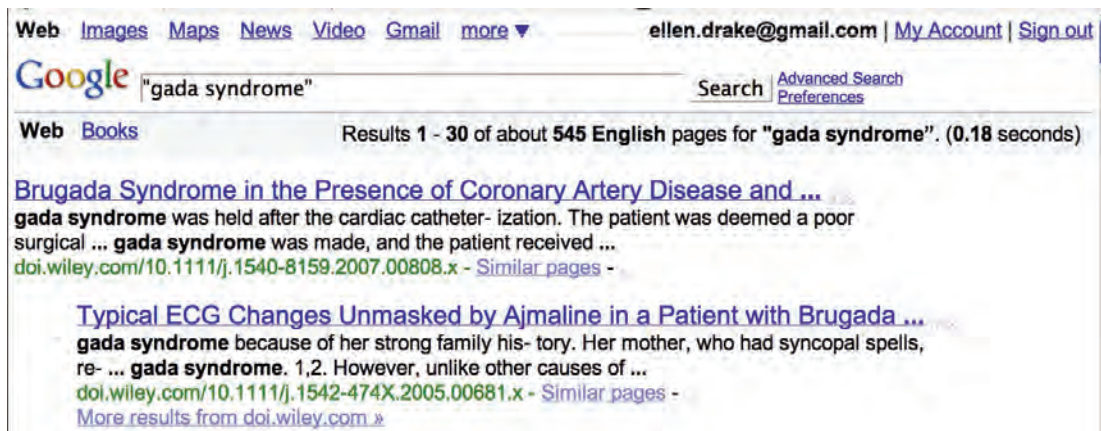
Always, always look for a "help" or FAQs link for the search engine to nail down the best way to construct a search for that particular site.

Specific Searching Tips

Google: Google is a good search engine to use when you don't have enough information to construct a more specific search or when you're not sure how to spell a term. For example, an MT heard "senopalatine block." Putting that string into the Google search box resulted in Google asking "Did you mean: 'sphenopalatine block'?" Truncation is not available on Google but sometimes happens by accident and is worth trying. An MT heard "___gada syndrome." (See Figure 1, next page.) What happened here is that the text on the Web page broke "Bru-gada" at the end of a line. It's a long shot but sometimes works.

Google Scholar provides a simple way to broadly search for scholarly literature. From one place, you can search across many disciplines and sources: peer-reviewed papers, theses, books, abstracts and articles, from academic publishers, professional societies, preprint repositories, universities and other scholarly organizations. It's still a better place to search than regular Google in some ways. You can eliminate a lot of irrelevant hits if, for example, if what you're searching for is an eponym, like Brugada. You won't retrieve genealogy, political, community news, obituaries, and business sites that may also carry the eponym. But Google Scholar doesn't take the place of PubMed because it isn't medicine specific.

Figure 1



You can also use Google to search for a missing word in a phrase by entering several key words surrounding the missing term into the Google search box. For example, suppose you couldn't hear the word "upstroke" in the following sentence: "Neck exam reveals no jugular venous distention, normal carotid upstroke, no carotid bruits." Entering the sentence without the missing word and no quotation marks can bring up a sentence containing the missing word. (See Figure 2.) You have to play around with this type of search. Sometimes, the more key words you use, the better. Other times, the fewer key terms yields better results.

Caveat: There are lots of sample medical reports on the Internet; don't assume that what you find is necessarily correctly spelled.

PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>) is a service of the U.S. National Library of Medicine that includes over 16 million citations from MEDLINE and other life science journals for biomedical articles back to the 1950s. PubMed includes links to full text articles and other related resources. Key terms, called MeSH terms, along with their synonyms are

indexed in PubMed. Every MeSH term is also automatically exploded as well as any subheading that is the top of a "subheading tree." Untagged terms entered in the PubMed search box are automatically mapped to the MeSH vocabulary when a match is found.

If you register (it's free), you can set certain filters that will be "remembered" each time you log on, you can have PubMed keep you signed in, and you can save searches. This is all done under MyNCBI. After registering and logging in, scroll down the left pane to MyNCBI. There are instructions for setting up your preferences that are too detailed to go into here.

PubMed has a very sophisticated search engine and allows all kinds of filters or limits to be placed on a search. We strongly suggest that you read the Help and FAQs files as well as complete a few of the tutorials. Even without registering, you can choose to set limits using the Limits button just under the word Search on the left.

At a minimum, you will probably want to set the language as English and choose Human in the "Humans or Animals" field. (See Figure 3, next page.)

Figure 2

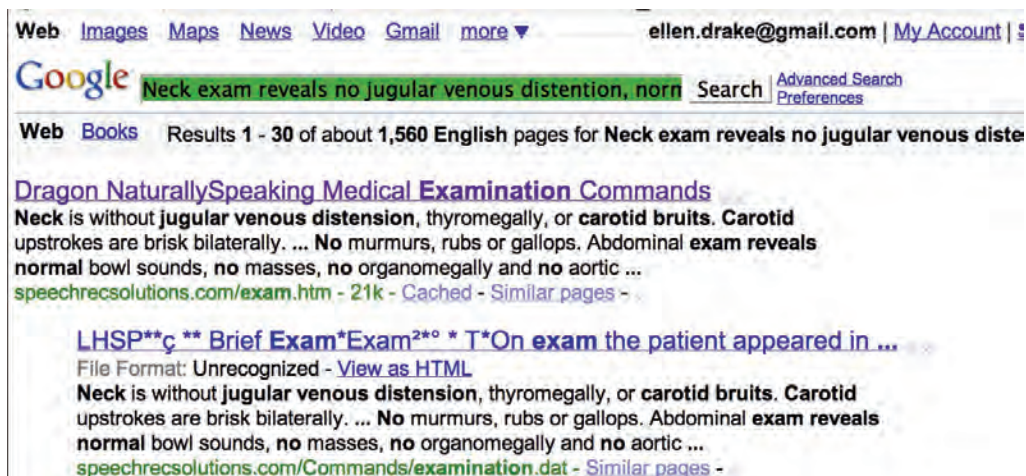


Figure 3

The screenshot shows the 'Limits' section of a PubMed search interface. It includes several filter categories: 'Dates' with 'Published in the Last' set to '5 years' and 'Added to PubMed in the Last' set to 'Any date'; 'Humans or Animals' with 'Humans' selected; 'Gender' with 'Male' and 'Female' options; 'Languages' with 'English' selected; and 'Subsets' with 'Journal Groups' and 'Core clinical journals' options. Each filter has a 'CLEAR' button.

If you're not setting up a MyNCBI account, you will probably also want to limit your searches to a specific date range; anything older than 10 years is probably not relevant to your report. In the examples below, you can see that the results can be displayed in a number of ways (All, English & Humans, Items with Abstracts) . . . Published in the last 10 years); these are MyNCBI preferences. (See Figure 4, with 441 citations for metaphyseal spelling.)

In addition, several filters have been used. The "Field: Text Word" is a tag term and is the last option when you click on the Limits button. Choosing that means that the term appears somewhere in the text and is more likely also to appear in the abstract. You can use this to eliminate citations that contain your search term in the journal name or authors' locations.

One use of a PubMed search is to determine the preferred spelling for a term. Suppose you have two references that disagree. See examples in Figures 4 and 5 for determining whether *metaphyseal* or *metaphysial* is preferred.

You can do a similar "style" search—for example, if you want to know if PT-INR or PT/INR or PT INR is more common in literature. For this type of search, however, you will have to scan the abstracts and simply count the instances you see of each format. In Figure 5 (with only 4 citations for *metaphysial* spelling), the Display drop-down menu has been clicked to choose Abstract. You will probably also want to choose the Show menu next to it and choose a higher number of abstracts to view so that you can scroll through a significant number of abstracts without having to go to another page.

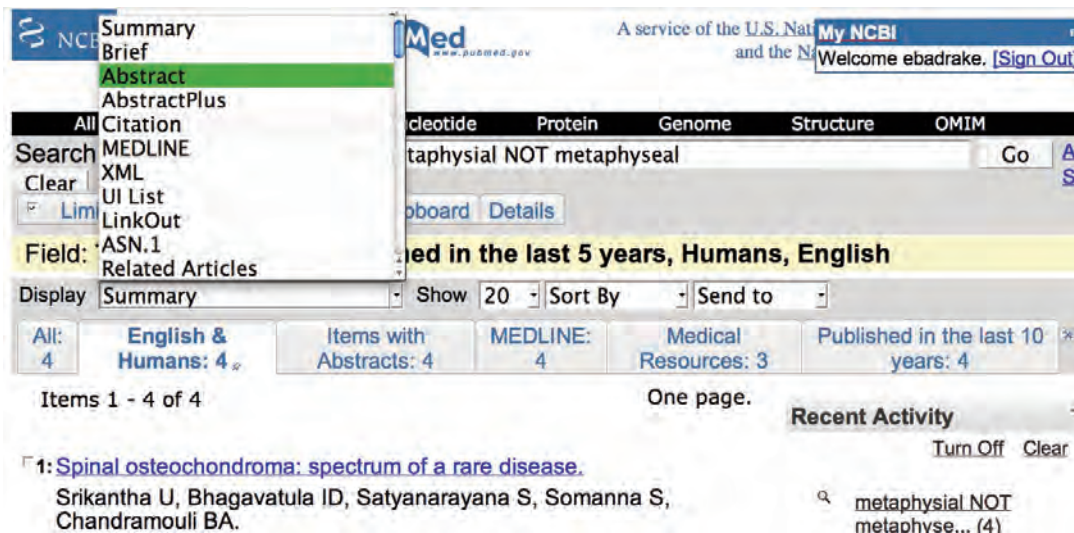
PubMed searches are recommended for eponyms, techniques, methods, procedures, disease entities, signs, phenomena, phrases, and acronyms or abbreviations. Do not assume, however, that acronyms and abbreviations used in journal articles necessarily have the same translation as what you are hearing. Make sure the context matches. You can search for new or investigational terms, but generally after a trademarked device or drug has been on the market for awhile, it will be referenced by its generic name if it is referenced at all. When you do find devices mentioned, you'll almost always be given the manufacturer, which is great because you can then go to the manufacturer's Web site to get more information. PubMed searches are generally useful for slang and coined terms, although not always. You can also find many of those colorful, descriptive terms like "hand-in-the-bucket" sign.

Eponym searches: Eponym searches can be tricky, but PubMed is a far better resource than Google where you are likely to have more unrelated hits than related ones. Putting the eponym-noun phrase in quotation marks will often bring up the phrase, but sometimes, you'll get a "no results" or "phrase not found" message. If that happens, don't give up just yet. Just put in the eponym without its accompanying noun, and you will get any article with an author by that name as well as any article with that eponym associated with any

Figure 4

The screenshot shows the PubMed search results page for the query 'metaphyseal NOT metaphysial'. The search was performed in the 'PubMed' database. The results are displayed in a table with columns for 'All: 441', 'English & Humans: 441', 'Items with Abstracts: 431', 'MEDLINE: 441', 'Medical Resources: 227', and 'Published in the last 10 years: 441'. The first result is titled 'Fixation of extra-articular distal humerus fractures using one locking plate versus two reconstruction plates: a laboratory study' by Tejwani NC, Murthy A, Park J, McLaurin TM, Egol KA, Kummer FJ. The search filters are shown as 'Field: Text Word, Limits: published in the last 5 years, Humans, English'. The page shows 'Items 1 - 20 of 441' and 'Page 1 of 23'. A 'Recent Activity' section is also visible.

Figure 5



noun. If there aren't too many citations, you can scan the abstracts for the phrase you're seeking. You can limit the search to authors only by using the Limits option or by simply typing [au] after the name in the search box. If you want to exclude authors because they seem irrelevant to your search, use Name NOT Name[au], and it will exclude authors (Name being the eponym you're looking for).

As a learning technique, start out by using the Limits check boxes to limit your searches, but pay attention to the search box when the results come up. It will insert the appropriate tags that you can later just type in for those limits you use most often.

OneLook.com. OneLook is an online search engine that indexes, as of this writing, 13,549,061 words in 1009 dictionaries. You can set preferences for the results in OneLook by clicking on Customize. You can choose whether you want the results displayed as verbose or compact; set the category for the type of dictionaries you want to be displayed first to medicine rather than English dictionaries; tell it whether to display results in the same window, another window, or a separate frame; and choose to have results include single words or words and phrases among other preferences. Often the definition for the term you need is displayed on the results page, and you need look no further.

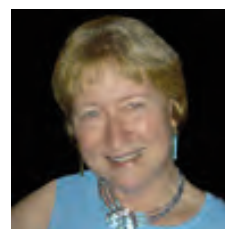
OneLook has a very robust wild card search feature that is clearly explained on the home page. You can also do a reverse search and search for key words in a definition to find a term you're looking for. OneLook searches a number of medical dictionaries and glossaries including *Dorland's* and *Stedman's*. With its wild card feature, it's an invaluable resource. And, like Google, if you spell a term incorrectly, it will suggest an alternative spelling.

Recommended Web sites

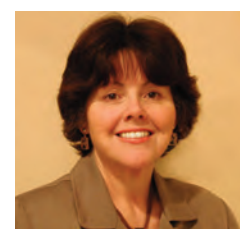
You can find **Ellen Drake's Useful Internet Links** on the <http://www.hpisum.com> Downloads page. Links were working at the time of this writing.

Books: Who Needs 'Em?

We hope your answer to this question is, "We do!" Let's not throw away our books just yet. You may prefer electronic references over paper and ink references, but either way, there's still a place for books. But books are not always the best resource, as noted above. We hope this article has helped you make better choices about which resource to choose when and to be a more efficient and astute researcher when the Internet is your chosen resource.



Ellen Drake



Georgia Green

Ellen Drake, CMT, AHDI-F, and **Georgia Green**, CMT, AHDI-F, are medical transcription educators with decades of experience in teaching, writing, research, and seminar presentations. Ellen Drake is Development Editor for Health Professions Institute. She is also coauthor of the annual *Saunders Pharmaceutical Word Book* and has more years than she likes to admit in the MT industry. E-mail: ellen.drake@gmail.com.