

# NCI Metathesaurus User Guide 2.2



## TABLE OF CONTENTS

<b>ABOUT THIS GUIDE.....</b>	<b>1</b>
<b>ABOUT THE NCI METATHESAURUS .....</b>	<b>1</b>
NCI LOCAL SOURCES .....	2
<b>THE NCI THESAURUS VERSUS THE NCI METATHESAURUS .....</b>	<b>3</b>
<b>USING NCI METAPHRASE .....</b>	<b>3</b>
HINT ABOUT DISPLAY RESOLUTION .....	3
SEARCHING .....	3
TIPS ON SEARCHING FOR WORDS OR PHRASES IN NCI METATHESAURUS .....	4
TIPS TO IMPROVE SEARCH SPEED .....	5
TIPS WHEN SEARCHING FOR CODES .....	5
<b>SELECTING AMONG SEARCH RESULTS.....</b>	<b>5</b>
HINTS ON NAVIGATION .....	8
<b>TYPE OF INFORMATION.....</b>	<b>8</b>
<b>TIPS ON USING THE INFORMATION IN THE NCI METATHESAURUS .....</b>	<b>9</b>
THE NCI THESAURUS AND OTHER NCI LOCAL SOURCES .....	9
CONCEPTS VERSUS TERMS AND THEIR ROLE IN VOCABULARY MAPPING .....	9
THE TWO TYPES OF CONCEPT NUMBER AND CODING OR KEYWORDING .....	10
DEFINITIONS .....	10
SYNONYMS .....	11
SOURCES .....	11
CONCEPTS MORE GENERAL THAN.....	15
CONCEPTS MORE SPECIFIC THAN .....	15
RELATED CONCEPTS .....	15
NAVIGATION IN THE MORE GENERAL, MORE SPECIFIC AND RELATED CONCEPTS LISTS	15
VIEWING THE NCI THESAURUS HIERARCHY .....	17
VIEWING THE NCI THESAURUS HIERARCHY.....	18
<b>GLOSSARY.....</b>	<b>19</b>

Send comments and inquiries to

Frank Hartel  
301 435 3869  
[hartel@nih.gov](mailto:hartel@nih.gov)

## **About This Guide**

This Guide describes the contents of the NCI Metathesaurus and explains how to use Metaphrase, the web-based software system that allows these contents to be examined over the Internet. Users will find this service a useful interactive desk tool.

Software developers will find this Guide useful for the insights it provides about the intended use of the NCI Metathesaurus data. NCI Metathesaurus is a database service that is available through a Java Application-Programming Interface (API). Software developers should refer to the NCI EVS Web site, <http://ncicb.nci.nih.gov/core/EVS> for documentation, technical advice and sample code showing how to use the API.

Screen shots in this document are current as of August 11, 2004.

In this document the terms Metathesaurus Browser, NCI Metathesaurus Browser, Metaphrase and Metaphrase Browser are used interchangeably.

## **About the NCI Metathesaurus**

The NCI Metathesaurus is intended to provide the National Cancer Institute (NCI) with the complete universe of terminology relevant to NCI's operations in a single, integrated resource. Basic, translational, clinical and population-based researchers have distinct terminologies and different ways of organizing their terminologies. Therefore NCI Metathesaurus contains vocabulary from each of these domains, including a great deal of general biological, medical and epidemiological terminology. The terminologies of these four communities partially overlap, but terminological differences impede collaboration. The NCI Metathesaurus "maps" terminology used in these communities to corresponding terms used in the others, thereby facilitating collaboration and sharing of information among disciplines.

The NCI Metathesaurus contains vocabularies that are part of the National Library of Medicine's Unified Medical Language System (UMLS) Metathesaurus (UMLS sources), vocabularies that NCI creates and maintains (NCI local sources) and sources that NCI has licensed.

The UMLS vocabularies are included to provide vocabulary that broadly covers four communities: medicine, nursing, certain other clinical disciplines, and biological science.

The NCI local sources are included to address specific needs of the four communities of the Institute. The local sources as of July 2002 are enumerated in the table on the following page.

## **NCI Local Sources**

The NCI local sources are updated on a monthly basis. To find the current list of local sources, go to the main NCI Metathesaurus Browser page, <http://ncimeta.nci.nih.gov/indexMetaphrase.html> and click the **Sources** link.

## The NCI Thesaurus versus the NCI Metathesaurus

The NCI Thesaurus is intended to become the reference terminology for cancer research. It exists as a stand-alone description logic vocabulary. In its stand-alone form, it is intended to support the Institute's software systems used for database coding and key wording, database search, data mining, text indexing and natural language understanding. The stand-alone NCI Thesaurus is provided to NCI via the Distributed Terminology Server (DTS) Browser, a separate service that is not part of NCI Metaphrase. The July 2004 version of the NCI Thesaurus contains about 36,700 concepts that are represented by about 112,000 terms. The concepts are organized into hierarchies that are up to 15 levels deep and are also linked by about 41,000 semantic relationships.

The NCI Metathesaurus is a great deal larger than the NCI Thesaurus, containing about 975,000 concepts represented by 2,303,000 terms. There are about 5,600,000 relationships in the NCI Metathesaurus. Not all of the sophisticated semantic relationships that exist in NCI Thesaurus are available in NCI Metathesaurus. Much of the content of the NCI Thesaurus is included in the NCI Metathesaurus so that NCI-specific terms will be mapped to corresponding terms used in standard biomedical vocabularies.

## Using NCI Metaphrase

NCI Metaphrase (<http://ncimeta.nci.nih.gov>), a software system running in front of the NCI Metathesaurus database, generates the display pages discussed in this document. The NCI Metaphrase is an interactive tool that lets you search the contents of the NCI Metathesaurus with your Web browser and navigate among vocabularies and other resources.

## Hint about display resolution

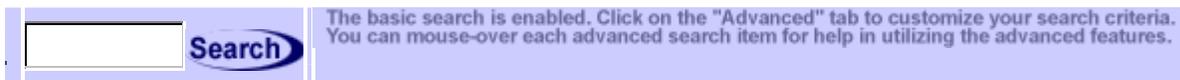
The NCI Metaphrase Browser's display looks best with your monitor resolution set to 1024 X 768 or greater. You can display the page at lower resolution, but there will be too little room horizontally for the layout to appear without some "line wrapping". That makes the layout look messy, but all the information will still be there.

## Searching

The NCI Metaphrase Browser searches using *lexical matching*, meaning that it searches for one or more authoritative terms that share "significant" *lexemes*, i.e., words or word bases, with the word or phrase you entered in the search box. We say that such terms are *lexically related* to the search string (and vice versa). For example, "degenerative joint disease" is lexically related to "Joints, Knee", since they share the lexeme "joint". Of course, a closer lexical match would be "Joint disease, NOS", and an even better one would be "Disease, joint, degenerative", which is *lexically equivalent*--it contains exactly

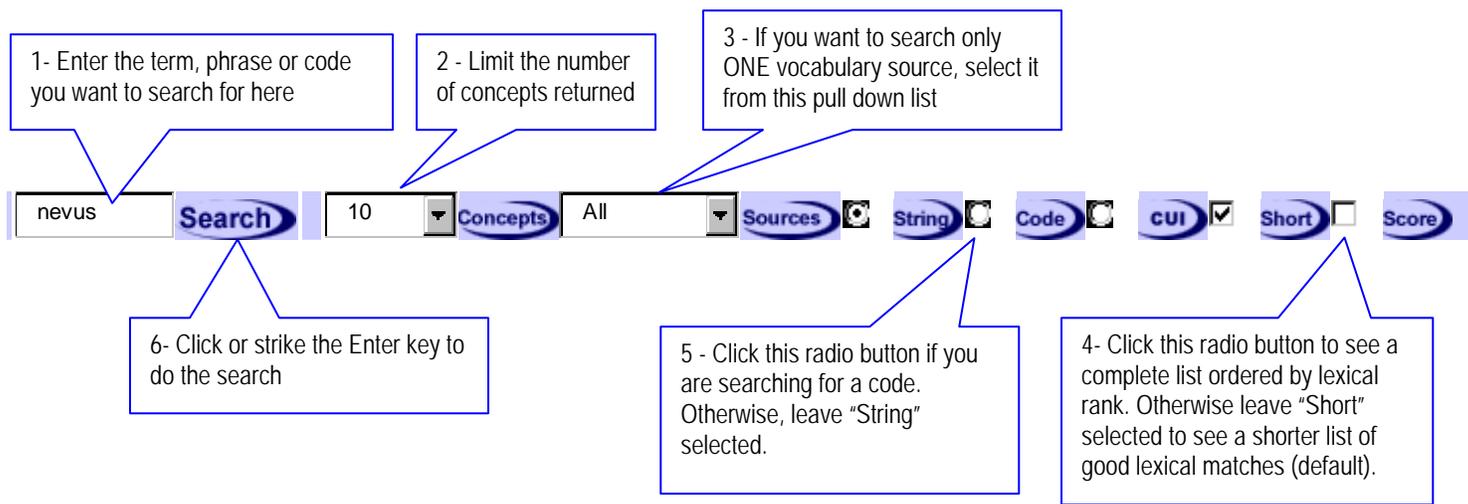
the same set of lexemes. NCI Metaphrase suggests authoritative terms for a given string by calculating a list of the most lexically related terms.

The search controls are at the top of the display page. There are two search choices, Basic and Advanced; each can be selected by clicking the appropriate tab in the upper right hand corner of the main panel. The basic search simply allows you to enter the word or phrase you wish to search for:



Most of the time, just type the word or phrase you want to find and click “Search” or strike the Enter key.

Additional functions are available under the “advanced” alternative. To invoke this alternative, click the “Advanced” tab at the top of the page.



### Tips on Searching For Words or Phrases in NCI Metathesaurus

- Case is ignored; do not worry about capitalizing names or acronyms.
- Punctuation (except for ".", see below) is treated as white space (separating lexemes), and so may be replaced with a <space>, or may be omitted entirely if not in the middle of a "word".
- The words "of", "and", "with", "for", "nos", "to", "in", "by", "on" and "the" are ignored and can be omitted.
- Word order is largely ignored, although adjacent sets of lexemes are looked up before non-adjacent sets. In general, though, don't give any thought to it.
- Rather than typing out long words, abbreviate with a period. A "." following two or more letters is treated as a *wildcard*, matching any word fragment which completes the given string of letters to form a complete lexeme. For example, typing " car. pn." will quickly return a list of matches, the first being " Pneumocystis carini (Pneumocystis carinii)".

- You can usually omit common or redundant words, e.g., "addison" will typically return "Addison's disease" and "non insulin dependent diabetes" will retrieve terms like "Non-insulin dependent diabetes (Non-insulin dependent diabetes mellitus)" if there is no "exact" match.
- Metaphrase has spell-correction built-in, and so can usually correct simple spelling and typing errors; you can usually get the same results without going back to edit the errors in what you've already typed. Of course, in cases where there are many possible corrections, this will tend to bog down the server (since it will have to look up all of them), so you might want to correct it yourself if you don't get a speedy response.

### **Tips To Improve Search Speed**

- When searching for general words like "cancer" it can take a long time for the server to download all the information in the database. By limiting the search to one source vocabulary you can speed things up.
- When making heavy use of spell-correction and/or abbreviation, especially with many-word queries, setting the return limit to one or at most 10 may speed server response, sometimes dramatically.

### **Tips When Searching For Codes**

- Generally you will search by code when you know the code number of a term or phrase in one source vocabulary and you want to find its corresponding number in another source vocabulary.
- Enter the code you know.
- Select the "Code" radio button in Advanced searching.
- Select the source vocabulary from which the code came from from the pull down list of "Sources".
- The spell checking and wild card features do not work with codes. You have to enter the code correctly.

### **Selecting Among Search Results**

The NCI Metaphrase server suggests authoritative terms for a given string by calculating a list of the most lexically related terms.

This lexical matching may result in search results that are exact matches, partial matches and occasionally complete misses. If all sources are searched for the phrase "knock out mouse", the results will include the partial list below. "Knock-out Mouse" is a complete lexically equivalent match, and as such is shown first. "Knock-out" matches two out of three lexemes, and so is shown next, even though its meaning is off the mark. Finally the individual words in the phrase are retrieved. This pattern of showing the most complete lexical matches first and the successively less inclusive matches in order is a characteristic of Metaphrase's search results display.

Matching Concepts 	
<a href="#">Knock-Out Mouse</a>	Mammal; Experimental Model of Disease
<a href="#">Knock-out</a>	Injury or Poisoning
<a href="#">Knock-in</a>	Intellectual Product
<a href="#">Out</a>	Spatial Concept
<a href="#">Mouse</a>	Mammal
<a href="#">mice (Laboratory mice)</a>	Mammal

If you did not find the concept you're looking for, and you're sure your spelling is correct, please click:

*Annotations:*  
 - "List of Search Results" points to the search results table.  
 - "Semantic Type of each search result" points to the second column of the table.  
 - "Brings up a dialog box for you to suggest additions or changes to NCI Metathesaurus" points to the "Suggest New Term" button.

However, perfect lexical matching does not guarantee that the results match the meaning in which you are interested. Take the term “mole” for example.

Matching Concepts 	
<a href="#">Moles (Mole the mammal)</a>	Mammal
<a href="#">Mole (Hairy Nevus)</a>	Disease or Syndrome
<a href="#">Mole (Melanocytic Nevus)</a>	Neoplastic Process
<a href="#">MOLE (MOLE Mouse)</a>	Mammal
<a href="#">Mole (Talpidae)</a>	Mammal
<a href="#">Moles (Mole the mammal)</a>	Mammal

If you did not find the concept you're looking for, and you're sure your spelling is correct, please click:

If you have searched all sources in the NCI Metathesaurus for the term “mole” you would have gotten the results shown above. These results will appear in the display page just below the search controls.

The term mole is semantically ambiguous<sup>1</sup>. That is, “mole” can mean several different things. In the above example, you should look at the semantic type of each search result, and click on the search result with the meaning you are interested in.

When you click on a search result, the information in the NCI Metathesaurus relating to that result is displayed on the right side of the page. The first part of the information about “Mole (Melanocytic Nevus)” is shown below.

The screenshot shows the NCI Metathesaurus interface for the concept 'Melanocytic Nevus'. At the top, there is a navigation bar with links: [Concept](#) | [Definitions](#) | [Synonyms](#) | [Sources](#) | [Broader Concepts](#) | [Narrower Concepts](#) | [Related Concepts](#) | [Medications](#) | [Procedures](#) | [Laboratory](#) | [Diagnosis](#) | [Open NCI Hierarchy](#) | [View Hierarchy Location](#). Below this is a header bar for 'C0027960: Melanocytic Nevus' with a question mark icon. Underneath, the semantic type 'Neoplastic Process' is displayed. A second header bar shows 'Melanocytic Nevus Definitions' with another question mark icon. Below this is a table with two columns: 'Source' and 'Definition'. The table contains five rows of definitions from different sources. Callout boxes point to various parts of the interface: 'Links to additional information' points to the navigation bar; 'Concept number and NCI Preferred Term' points to the header bar; 'Semantic Type of Concept' points to 'Neoplastic Process'; 'Definitions for Concept' points to the table.

Source	Definition
<a href="#">MSH2004_2003_12_12</a>	A circumscribed stable malformation of the skin and occasionally of the oral mucosa, which is not due to external ...
<a href="#">ICDO3</a>	C44._
<a href="#">NCI</a>	Also known as mole, this is a neoplasm composed of melanocytes that usually appears as a dark spot on the skin ...
<a href="#">NCI-GLOSS</a>	(NEE-vus) A benign growth on the skin, such as a mole. A mole is a cluster of melanocytes ...
<a href="#">NCI-GLOSS</a>	A benign growth on the skin (usually tan, brown, or flesh-colored) that contains a cluster of melanocytes ...

(Some definitions in the above sample have been truncated.)

<sup>1</sup> **Semantic** navigation means getting from a set of terms that refer to one meaning, or *concept*, to a set of terms that refer to a similar concept. All of the terms given in the example are *semantically related* to each other; however, the term "mole", while lexically related, is semantically unrelated. Conversely, the term "Osteoarthritis, NOS", while lexically unrelated to "Disease, joint, degenerative", is *semantically equivalent* to "Disease, joint, degenerative". That is, they both refer to the same concept. Semantic navigation in the NCI Metathesaurus is supported by an extensive, though incomplete, knowledge base of binary relationships between concepts. Concepts are usually represented for the user as a single, "NCI preferred" term.

## Hints on Navigation

NCI Metaphrase depends heavily on the browser's forward and back buttons (or <ALT> Left and Right Arrow if your browser's buttons are not visible). These buttons affect the content of the right frame of the display page. They will allow you to view related information, while preserving the ability to return to the prior display. Think of the information as a series of loose-leaf pages stacked up on your desk. The stack gets higher as you view new information. The display changes as you use the back and forward buttons to "flip" through the stack of loose-leaf pages.

You can also use the "Go To" feature of your browser to navigate this way, but most people prefer to use the back and forward buttons.

The following table summarizes the sort of information that the NCI Metathesaurus contains. Each row in the table corresponds to one of the hyperlinks that run across the top of the right side of the display page. Not every concept in the NCI Metathesaurus has each type of information. Tips on using each type of information are provided below.

	<b>Type of Information</b>	<b>Comment</b>
<b>Definitions</b>	Indicates the current concept's meaning	Sources vary in their focus, affecting definitions
<b>Synonyms</b>	Approximate synonyms	
<b>Sources</b>	Sources that contain the concept	
<b>Concepts More General Than _____</b>	Parents of the concept	Not necessarily true is_a
<b>Concepts More Specific Than _____</b>	Children of the concept	Not necessarily true is_a
<b>Concepts Related to _____</b>	Concepts with a semantic relationship to the concept	
<b>Medications</b>	Drugs co-occurring with concept in last three years of Medline	For concepts with a MeSH headings only
<b>Procedures</b>	Clinical procedures co-occurring with concept in last 3 years of Medline	For concepts with a MeSH headings only
<b>Laboratory</b>	Lab tests co-occurring with concept in last 3 years of Medline	For concepts with a MeSH headings only
<b>Diagnosis</b>	Diagnoses co-occurring with concept in last 3 years of Medline	For concepts with a MeSH headings only

Generally there is far more information about a concept in the NCI Metathesaurus than can be displayed at one time on the right side of the display page. You can use the slider bar to the right of the information to scroll down to see the whole body of information, or you can use the links to additional information to jump directly to each sort of information.

## Tips on using the information in the NCI Metathesaurus

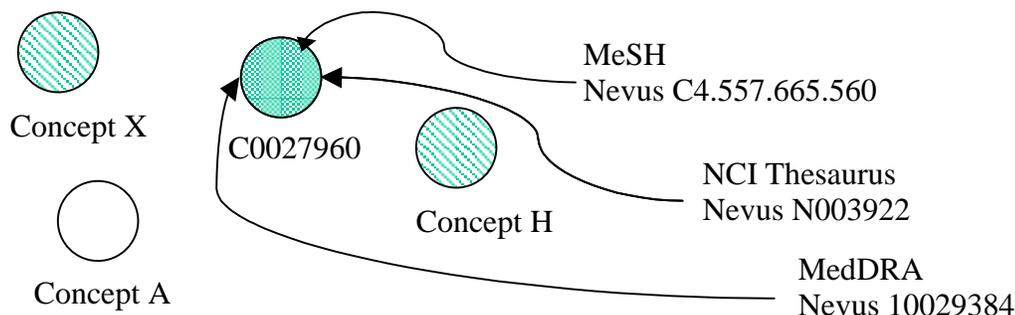
### The NCI Thesaurus and other NCI Local Sources

The version of the NCI Thesaurus in the NCI Metathesaurus has sub-sources. This is unique among source vocabularies in the NCI Metathesaurus. These sub-sources represent information that is specific to only a part of the Institute, for example RAEB (the Research Analysis and Evaluation Branch). The major effect of these sub-sources is to provide a way to identify definitions that are “owned” by RAEB or some other part of the Institute.

### Concepts versus Terms and their Role in Vocabulary Mapping

The mechanism used in the NCI Metathesaurus to map between a term in one source and corresponding terms in other sources is to assign each term to a concept. All the terms and phrases in the NCI Metathesaurus are linked to one or more concepts. A concept in the NCI Metathesaurus is simply a unique reference number. In the NCI Metathesaurus concepts have meaning because of the information associated with them. At a minimum, all concepts have a semantic type and an associated “term” in addition to a unique concept number, called the *Concept Unique Identifier*, or CUI.

In the example below, the concept C0027960 has the semantic type “Neoplastic Process,” denoted by the solid green fill color. “Atoms,” or terms from three sources, are mapped to the concept C0027960. (An “atom” is simply a term in a specific source). This means that the NCI EVS editors or the NLM UMLS editors have determined that the MeSH atom Nevus, the NCI Thesaurus atom Nevus and the MedDRA atom Nevus all have the same meaning. Within a source, each term or “atom” has an “atom ID number.” In the example, the MeSH atom ID is C4.557.665.560.



The Concepts X and H in the above drawing have the same semantic type, which is not the same as either Concept A or Concept C0027960. For example, Concept A might have the semantic type “Mammal” while the other two might be semantic type “Quantitative Process.”

In summary, terms within sources have atom ID numbers that uniquely identify the term within that source. Atoms are associated with concept numbers, which are unique across the entire NCI Metathesaurus. Finally, all concepts are of at least one semantic type. It is these formal properties of the NCI Metathesaurus database, along with the expertise of the NCI EVS editors, which enable you to translate from one source to another.

## **The Two Types of Concept Number and Coding or Keywording**

One major difference between the NLM UMLS Metathesaurus and the NCI Metathesaurus is that in the UMLS Metathesaurus all concepts have concept numbers of the form C#####. In the NCI Metathesaurus there are sources that do not appear in the UMLS, such as the NCI local sources. Therefore in the NCI Metathesaurus you will see concept numbers of the form CL##### as well as C#####. Both concept number formats function identically, but the CL indicates that the concept occurs only in a NCI local source.

If you need to store codes in an information system to provide links to vocabulary concepts, we recommend using the concept codes in the NCI Thesaurus for this purpose. The codes in the Thesaurus are guaranteed to be stable over time, and the Thesaurus also provides a concept history facility to track changes made to a concept over time. Metathesaurus codes, derived as they are from the NLM's UMLS, provide neither concept history or guaranteed long-term stability.

## **Definitions**

Many concepts in the NCI Metathesaurus have definitions from multiple sources. In general they will differ in wording or emphasis but not in meaning. If you encounter a concept with definitions that differ in meaning it is an error on our part. Please call it to our attention.

Regarding the NCI local sources, most NCI Thesaurus terms have a definition that is labeled "NCI". In addition, the term may have other definitions labeled, for example, "RAEB" or "NCI-GLOSS".

All NCI atoms will eventually have a consensus definition, which will be labeled "NCI". These "NCI" definitions are intended to be scientifically correct, precise definitions. Some NCI atoms will have a second definition labeled "NCI-GLOSS". NCI-GLOSS is short for NCI Glossary. These definitions are generally correct definitions aimed at a lay audience, rather than those in the health science professions. A few NCI atoms will also have "RAEB" definitions. RAEB definitions are intended to be useful in internal NCI operations. These may not be true definitions but rather guidance intended to assist coders in using the RAEB coding terms correctly.

## Synonyms

Synonyms are displayed in the main panel of the display page below the definitions. In many cases there are fine distinctions among the synonyms shown, and in some cases there are “synonyms” that are arguably incorrect. This lack of precision is due to the limitations of our semantic tools for mapping among atoms. Nevertheless, the synonyms have value because they enable one to search using any one of many terms and eventually find the right underlying concept. (See Mapping, below.)

## Sources

Below the list of synonyms is a list of the sources that contain one or more atoms mapped to the concept for which you searched. These sources are denoted by their acronyms. (You can check the meaning of a source acronym by clicking on the “Sources” link in the navigation bar to the extreme left of the browser window. A separate window will open containing all of the sources.)

The source acronyms are all links. If you click on one, the right frame of the display page will show the preferred term, synonyms and acronyms (and the acronyms’ atom ID numbers), and in some sources, part of the hierarchy from the source you clicked on<sup>2</sup>. (Use the browser’s back button or ALT\_Left\_Arrow (on a PC) to restore the display of sources.)

Below the list of sources you will see “View Neighborhood in” followed by a pull down list of sources and an “OK” Button. If you select a source in the pull down list and click OK, you will see the same list of synonyms as you would see had you clicked on the source’s hyperlink, but below the synonym list you would see a (usually long) list of “semantically related” atoms from the source. However since in NCI Metaphrase the meaning of the links differs depending on the source, the neighborhood displayed may be perfectly rational and easy to understand in one source and puzzling in another. The notion of neighborhood is easy to define; it is the list of all atoms within the source you pick from the pull down that are a distance of one link from the atom you are displaying when you click “OK”.

The screenshot shows the following interface elements:

- Melanocytic Nevus Sources** (Section Header)
- List of source acronyms: [NCI](#), [MDBCAC](#), [NCIPDQ](#), [ICDO3](#), [BI98](#), [CCPSS99](#), [COS89](#), [CSP2000](#), [CST](#), [MDR33](#), [MSH2001](#), [MTH](#), [RCD99](#), [SNMI98](#), [WHO97](#)
- View neighborhood in:
- Table View Of [Sources](#)

Callout boxes provide the following instructions:

- Click on a source to see the preferred name, synonyms and hierarchy of the concept in that source
- Choose a source and click OK to see the synonyms and network neighborhood of the concept in that source

<sup>2</sup> See **Viewing the NCI Thesaurus Hierarchy** at the end of this document.

The NCI Thesaurus neighborhoods typically consist of the NCI atoms associated with concepts that are immediately more general and more specific than the concept you were displaying when you clicked “OK”.

Concepts More General Than <u>Melanocytic Nevus</u> 	
(inverse_isa)	<a href="#">Benign Melanocytic Nevus</a>
	<a href="#">Benign Neoplasm</a>
	<a href="#">Benign Skin Neoplasm</a>
(inverse_isa)	<a href="#">Common Neoplasm</a>
	<a href="#">Disorder of skin appendage</a>
(inverse_isa)	<a href="#">Hairy Nevus</a>
(inverse_isa)	<a href="#">Halo Nevus</a>
	<a href="#">Melanocytic Neoplasm</a>
(inverse_isa)	<a href="#">Melanocytic Skin Neoplasm</a>
	<a href="#">NEVI AND MELANOMAS</a>
	<a href="#">Skin and subcutaneous tissue disorders congenital NEC</a>
	<a href="#">Skin Neoplasm</a>
	<a href="#">Skin neoplasms malignant and unspecified (excl melanoma)</a>

Click on a link to see that concept's information

Concepts More Specific Than <u>Melanocytic Nevus</u> 	
(isa)	<a href="#">Benign Melanocytic Nevus</a>
	<a href="#">Benign melanocytic nevus of skin</a>
(isa)	<a href="#">Benign nevus</a>
	<a href="#">Blue Nevus</a>
	<a href="#">Compound Nevus</a>
(isa)	<a href="#">Dysplastic Nevus</a>
	<a href="#">Dysplastic Nevus Syndrome</a>
(isa)	<a href="#">Eyelid Nevus</a>
	<a href="#">Hairy Nevus</a>
	<a href="#">Intradermal Nevus</a>
	<a href="#">Nevus, Halo</a>
(isa)	<a href="#">Nonpigmented Nevus</a>
(isa)	<a href="#">Pigmented Nevus</a>

The neighborhoods from sources such as ICDO3, which has atom ID numbers that specify hierarchy location, often provide the more interesting neighborhoods.

In the example below, ICDO3 was chosen as the source, and Melanocytic Nevus as the atom. By default the atoms in the neighborhood are displayed in alphabetical order. However you can display them ordered either by the numerical order of their atom ID numbers, or by the alphabetical order of their names. In the example below the Atoms are displayed by numeric order. By clicking on the blue underlined “Name” link one could display them in alphabetic order. If you did sort by name, the bold “Code” would turn into a link, and by clicking it, the display would be sorted code.

## Neighborhood of `Melanocytic Nevus' in ICDO3

### Synonyms

ICDO3/	SY	<a href="#">8720/0</a>	<a href="#">Melanocytic nevus</a>
ICDO3/	SY	<a href="#">8720/0</a>	<a href="#">Nevus, NOS</a>

Click Code and Name to change neighborhood display back and forth from numeric to alphabetic order

### Neighborhood

		<a href="#">Code</a>	<a href="#">Name</a>
ICDO3/	SY	<a href="#">8730/0</a>	<a href="#">Achromic nevus</a>
ICDO3/	PT	<a href="#">8780/0</a>	<a href="#">Blue nevus, NOS</a>
ICDO3/	PT	<a href="#">8760/0</a>	<a href="#">Compound nevus</a>
ICDO3/	SY	<a href="#">8760/0</a>	<a href="#">Dermal and epidermal nevus</a>
ICDO3/	SY	<a href="#">8750/0</a>	<a href="#">Dermal nevus</a>
ICDO3/	PT	<a href="#">8727/0</a>	<a href="#">Dysplastic nevus</a>
ICDO3/	ET	<a href="#">8720/0</a>	<a href="#">Hairy nevus</a>
ICDO3/	PT	<a href="#">8723/0</a>	<a href="#">Halo nevus</a>
ICDO3/	PT	<a href="#">8750/0</a>	<a href="#">Intradermal nevus</a>
ICDO3/	SY	<a href="#">8780/0</a>	<a href="#">Jadassohn blue nevus</a>
ICDO3/	PT	<a href="#">8000/0</a>	<a href="#">Neoplasm, benign</a>
ICDO3/	HT	<a href="#">872-879</a>	<a href="#">NEVI AND MELANOMAS</a>
ICDO3/	PT	<a href="#">8730/0</a>	<a href="#">Nonpigmented nevus</a>
ICDO3/	PT	<a href="#">8720/0</a>	<a href="#">Pigmented nevus, NOS</a>
ICDO3/	SY	<a href="#">8000/0</a>	<a href="#">Tumor, benign</a>
ICDO3/	SY	<a href="#">8000/0</a>	<a href="#">Unclassified tumor, benign</a>

List of all atoms in the chosen source that are directly linked to the target atom

### **Concepts More General Than**

The list of more general concepts shows the next more general concept from every source that contains the current concept. Put another way, if a source contains an atom that is linked to the current concept, and if the source is a hierarchical source, there will most likely be a atom in the source that the source authors consider more general than the atom linked to the current concept. The broader concepts list shows each of these more general atoms.

### **Concepts More Specific Than**

The list of more specific concepts shows the next more specific concept from every source that contains the current concept. Again, if a source contains an atom that is linked to the current concept, and if the source is a hierarchical source, there will most likely be an atom in the source that the source authors consider more specific than the atom linked to the current concept. The more specific concepts list shows each of these atoms.

### **Related Concepts**

The list of related concepts shows the semantically related concepts from every source that contains the current concept. Again, if a source contains an atom that is linked to the current concept, and if the source has semantic relationships, there may be an atom in the source that the source authors consider related in some way to the atom linked to the current concept. The narrower concepts list shows each of these semantically related atoms.

If broader and narrower relationships organize concepts vertically, semantic relations can be thought of as relating them horizontally. Semantic relationships in some sources are very sophisticated and specific, representing ontologic information such as the fact that a specific bacterium is the etiologic agent that causes a specific disease. In other sources there is only some primitive indication that some sort of undefined relationship exists among two or more concepts.

The NCI Metathesaurus cannot differentiate among these relationships, but lumps the sophisticated precise relationships from sources such as the NCI Thesaurus with the vague ones. The way to think of the related concepts in the NCI Metathesaurus is that in one or more source(s), some sort of horizontal relationship is stated between the current concept and the concepts whose atoms appear in the list of related concepts.

### **Navigation in the More General, More Specific and Related Concepts Lists**

All the atoms listed in the more general, more specific and related lists are hyperlinks. If you click on one, the contents of the right panel of the display page changes to show information about the concept corresponding to the atom on which you clicked. Use the page back commands of your Web browser to restore the information that was displayed before you clicked.

Some of the atoms in the broader, narrower, and related lists have a notation appended to them that describes the specific type of broader relationship (for example, inverse is\_a, is\_associated\_with), narrower relationship (is\_a, part\_of) or related relationship (is\_mapped\_from). We display this information when it is available in the source. It is intended to help you determine if you want to follow the link to view the information on the concept pointed to by the link.

Some of the atoms in the three lists are **highlighted in blue**. These are atoms from the NCI local sources.

Metaphrase software supports *lexical* and *semantic* navigation of the NCI Metathesaurus terminology database.

All concepts in the NCI Metathesaurus that have atoms in MeSH that are MeSH headings will have lists of links in each of the medications, procedures, laboratory and diagnosis sections at the bottom of the right side of the display page.

We index the most recent three years of Medline to create a list of citations. In each citation, the current concept co-occurred with one or more concepts contained in MeSH that denote a medication, a diagnostic procedure, a clinical laboratory test, or a clinical diagnosis.

There are blue hyperlinks and green hyperlinks displayed. The green links are links to Medline. If you follow the green link, a new browser window opens, the National Center for Biotechnology Information Entrez site is contacted, and a query is executed that results in a list of all citations in Medline that mention both the current concept and the blue linked concept to the right. If you follow the blue link, the right frame of the display page will show the concept information for the atom you click on.

This feature of Metaphrase is intended to assist those who want to obtain a quick bibliography of clinically relevant articles, or to find out more about the meaning of concepts and how they fit into the clinical enterprise.

Medications		?
<a href="#">MedLine</a>	<a href="#">Sunscreen</a>	Link to information about the concept represented by these atoms
<a href="#">MedLine</a>	<a href="#">Collagen</a>	
<a href="#">MedLine</a>	<a href="#">Anticarcinogenic Agents</a>	
<a href="#">MedLine</a>	<a href="#">Antineoplastic Agents, Phytogetic</a>	
<a href="#">MedLine</a>	<a href="#">Cytokine (Therapeutic)</a>	
<a href="#">MedLine</a>	<a href="#">Fibroblast Growth Factor-2 (Therapeutic)</a>	
<a href="#">MedLine</a>	<a href="#">Lymphokine (Therapeutic)</a>	
<a href="#">MedLine</a>	<a href="#">Nerve Growth Factor (Therapeutic)</a>	
<a href="#">MedLine</a>	<a href="#">Transforming Growth Factor-Beta (Therapeutic)</a>	
Procedures		?
<a href="#">MedLine</a>	<a href="#">Laser Surgery</a>	Link to Entrez. Displays bibliography of articles mentioning both the current concept and the concept listed in blue
<a href="#">MedLine</a>	<a href="#">Skin Transplantation</a>	
<a href="#">MedLine</a>	<a href="#">Body Surface Area</a>	
<a href="#">MedLine</a>	<a href="#">Curettage Procedure</a>	
<a href="#">MedLine</a>	<a href="#">Diagnostic Imaging</a>	
<a href="#">MedLine</a>	<a href="#">Image Interpretation, Computer-Assisted</a>	
<a href="#">MedLine</a>	<a href="#">Magnetic Resonance Imaging</a>	
<a href="#">MedLine</a>	<a href="#">Radiation Brachytherapy</a>	
Laboratory		?
<a href="#">MedLine</a>	<a href="#">Colorimetry</a>	
<a href="#">MedLine</a>	<a href="#">Confocal Microscopy</a>	
<a href="#">MedLine</a>	<a href="#">Image Cytometry</a>	
<a href="#">MedLine</a>	<a href="#">Microtomy</a>	

## Viewing the NCI Thesaurus Hierarchy

Some sources have atom numbers that specify the location of each atom in the source's hierarchy. Examples include PDF 2000, MeSH and Snomed Version 3. When you display the current concept in one of these sources<sup>3</sup>, you will see the fragment of the source hierarchy from the current concept back to the top, or "root" of the source's hierarchy tree. This is useful in seeing where a concept fits in the hierarchy of the specific source.

There are at least two issues that you should be aware of, one related to a limitation of NCI Metaphrase and the other related to a change in the philosophy and practice of vocabulary construction.

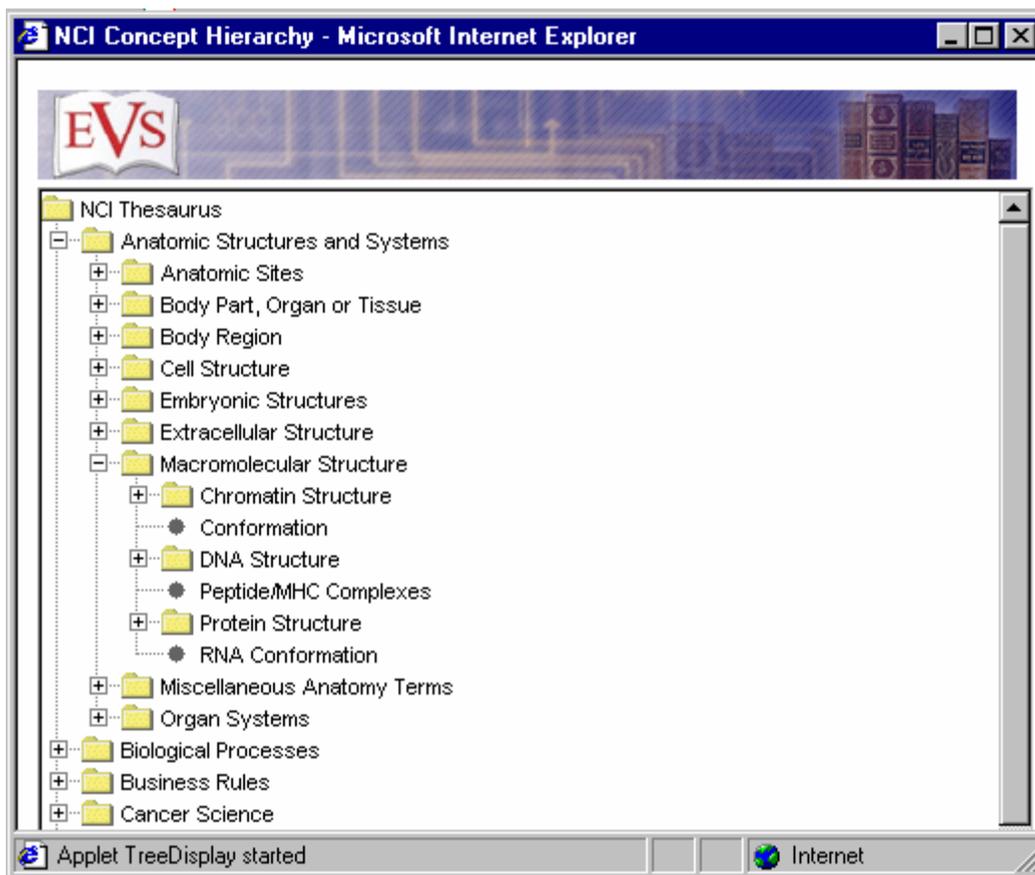
When you display the current concept in a source, you will see a hierarchy fragment only for sources that use atom numbers that specify hierarchical location. Furthermore, the NCI Metaphrase has no means of letting you browse the entire hierarchy of any source. At best you will see the fragment from the current concept to the root.

The sources that use hierarchical numbering have found that this practice causes them serious problems. If a lot of new information is discovered that results in a large number of new atoms being added to an already heavily populated part of their hierarchy, they run out of numbers to assign to the new atoms. The labor expense involved in renumbering is too great to be tolerable. Therefore most major vocabulary publishers are choosing to use numbers that do not specify tree position.

The NCI Thesaurus and our other local sources use numbers that are unique, but do not specify hierarchy location. Nevertheless it is possible for Metaphrase users to browse the NCI Thesaurus from within the NCI Metaphrase Web page. You can invoke the function by clicking on the link "Browse" in the navigator bar to the left of the browser display screen. A new window opens for the NCI Thesaurus tree display. Double-click the "NCI Thesaurus" folder that appears at the top left of the window. The NCI tree display allows you to expand the hierarchy by clicking on the plus icons. If you wish to see the Metathesaurus entry for a concept you find in the tree, click on the concept, do an EDIT/COPY (CTRL-C), then EDIT/PASTE (CTRL-V) into the Search window in the Metathesaurus Browser main page and click Search.

---

<sup>3</sup> See **Sources**, above



## Glossary

### Atom

The specific term for a concept from a specific source.

### Atom ID number

The unique identifier for a specific atom.

### Binary relationship

A relationship between two concepts, e.g., a [specific] protein is a marker for a [specific] disease.

### Concept

A general idea derived or inferred from specific instances or occurrences.

### Description logic

A subset of first-order predicate calculus. *Description logic* serves as the theoretical basis for the NCI Thesaurus.

### CUI (Concept Unique Identifier)

The unique code assigned to a concept in the UMLS and the NCI Metathesaurus.

**Code number**

A unique code associated with a specific concept.

**Hyperlink**

A string of characters that can be used by software to locate and access a page on the worldwide web. A *hyperlink* is used by a web browser to locate and display the contents of a web page.

**inverse is\_a**

The complementary relationship to *is\_a, q.v.* That is, if A *is\_a* B, then B *inverse-is\_a* A.

**is\_a**

A binary relationship asserting that one concept is a subcategory, or instance, of another, more general concept.

**Lexeme**

The fundamental unit of the lexicon of a language. *Find, finds, found, and finding* are forms of the English lexeme *find*.

**Lexically equivalent**

Two atoms are lexically equivalent if they are lexemes of the same term. Two phrases are lexically equivalent if they are composed of lexically equivalent atoms. (“Stop” words, such as “a”, “of”, “the”, etc. are ignored in determining whether phrases are lexically equivalent.) Note that lexical equivalence does not imply that the atoms or phrases are semantically equivalent, i.e., mean the same thing. The terms mole (a little furry animal), mole (a measure of the amount of a pure substance) and mole (a skin nevus) are all lexically equivalent even though they mean different things.

**Local source**

A local source is an NCI organization which has had its terminology inserted as a separate source in UMLS. It can be edited, and it can have its own hierarchy.) “Local source” is also used to describe the set of terms contributed by the organization.

**Neighborhood**

the list of all atoms within the given source that are a distance of one link from the atom you are displaying.

**NOS (Not Otherwise Specified)**

The "catchall" subcategory of a parent concept to subsume all concepts not subsumed by a sibling subcategory.

**RAEB (Research Analysis and Evaluation Branch)**

Research Analysis and Evaluation Branch, Division of Extramural Activities, National Cancer Institute. The organization charged with assigning content codes to NCI grants and grant applications.

**Semantic type**

A categorical grouping of concepts. A list of the current semantic types contained in UMLS can be found at [http://www.nlm.nih.gov/research/umls/META3\\_current\\_semantic\\_types.html](http://www.nlm.nih.gov/research/umls/META3_current_semantic_types.html) (beware of the underscores in this URI).

**Semantically equivalent**

Two terms are said to be semantically equivalent if and only if they have the same meaning.

**Sibling**

Concept B is a sibling of concept A only if both concepts have the same *is\_a* parent.

**Source**

An organization that has contributed a set of terminology to the NCI Metathesaurus. The term *source* is also used to denote the contributed terminology itself.

**Subsource**

A subsource is a subset of the terminology in NCI source. It is attributed to the organization that contributed the terminology. It can be edited (because the NCI source can be edited), but it cannot have its own hierarchy distinct from the NCI hierarchy.

**Subsume**

*vt.* to include as a subset. To classify, include or incorporate in a more comprehensive category or under a general principle.

**Subsumption**

*n.* the act of subsuming

**Term**

*n.* a word or phrase denoting a concept.

---

Parts of the foregoing are included in the Metaphrase distribution, ©1997-2001 Lexical Technology Inc. and Apelon Inc. All rights reserved.

For more information, visit <http://www.apelon.com> or <http://www.lexical.com/>, or email [info@lexical.com](mailto:info@lexical.com).

Lexical, Metaphrase, Metaphrase Enabled, and the Metaphrase and Metaphrase Enabled logos are trademarks of Lexical Technology, Inc.