

# COMPREHENSIVE LITERATURE SEARCH

Efraín Flores Rivera, MLS, EdD  
Conrado F. Asenjo Library  
UPR- Medical Sciences Campus  
September 20<sup>th</sup>, 2016



# Objectives



1. Identify basic elements for conducting a search in a database
2. Design a search strategy applying principles of boolean logic, filters, controlled vocabulary, etc.
3. Conduct searches in several databases: PubMed, Science Direct, Scopus, OvidSP, ProQuest, EBSCOhost.

# Searching a Database



Different search interfaces do the same things in slightly different ways.

Good search interfaces should provide ability to:

- ▣ Search for a specific item
- ▣ Search for related items to a known item
- ▣ Search in a specific field or fields
- ▣ Combine search terms using Boolean Logic
- ▣ Retrieve search results in a useful way

# Boolean Operators



Standard Boolean Logic for database searching uses  
3 relationships among search terms.

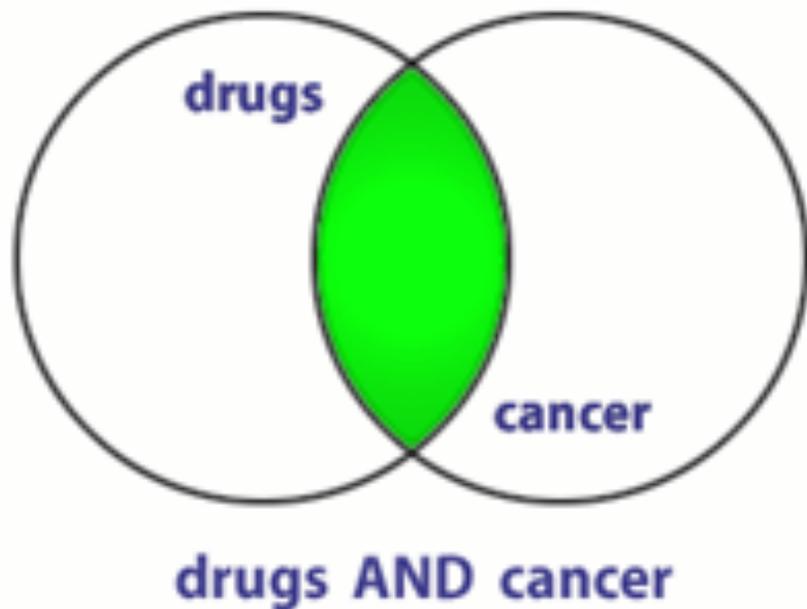
**AND**

**OR**

**NOT**

It is both simple and powerful.

# AND



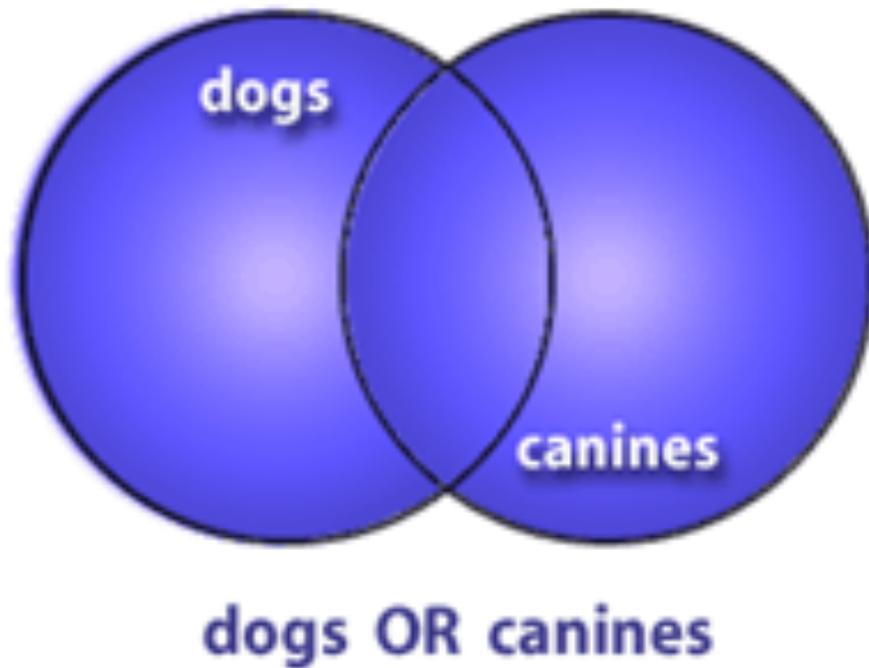
**BOTH** terms included in any results.

If a record has only one of the two terms, it will not be retrieved.

If the record has neither term, it will not be retrieved.

What does this do to the amount of records retrieved?

# OR

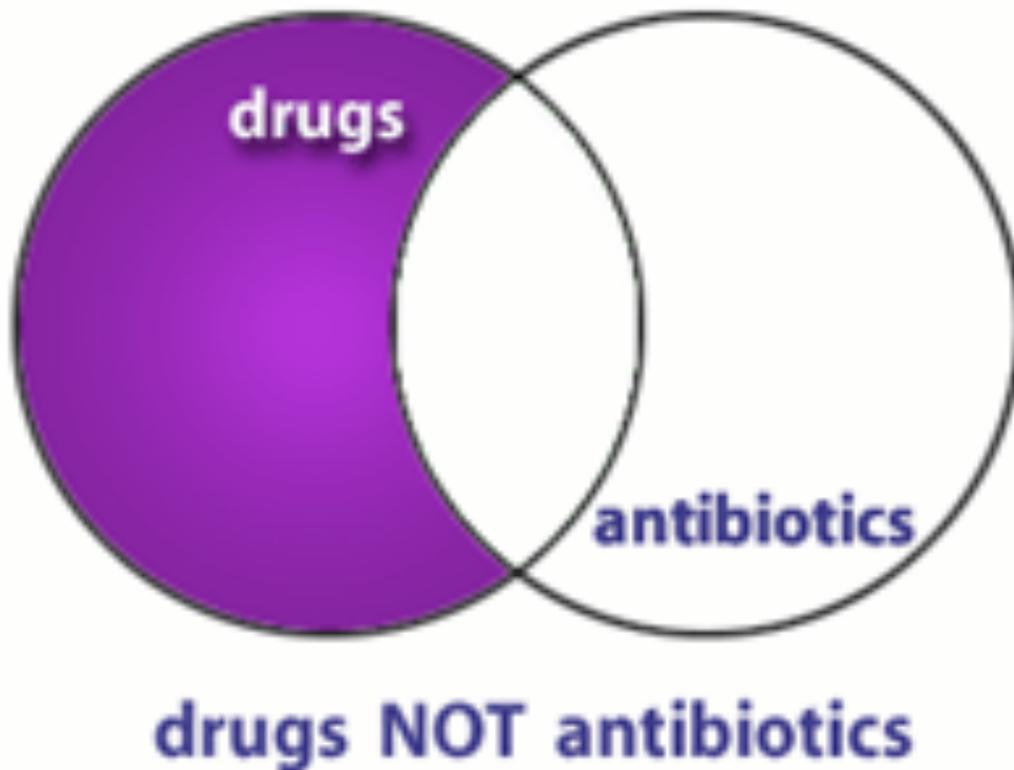


**Either term (or both)** will be in the returned documents.

OR broadens the search.

What does OR do to the amount of records retrieved?

# NOT



Excludes any results containing the term after the operator.

Records containing both will not be retrieved.

What does NOT do to the amount of records retrieved?

# Using OR



“OR” groupings contain terms for the same idea/concept and are usually put in parenthesis

*(term OR term OR term)*

where all terms are difference ways of representing the same concept

**(faculty OR teachers OR professors)**

**(students OR learners OR pupils)**

# Using AND

---

“AND” groupings contain terms for different ideas/concepts and can combine OR groupings

*Term AND (Term OR Term)* where each represents a different concept

**heart attack AND smoking**

**Diabetes AND exercise**

**Cancer AND (treatment OR therapy)**

# Using NOT



“NOT” statements are usually put last and can contain an “OR” grouping; they are often used to get rid of a common subgroup

Students NOT dental

Diabetes NOT juvenile

# Putting Them Together



1. Identify the concepts (break down the question)
2. List specific terms for each concept
3. Put the terms for each concept in an OR statements within parentheses
4. Combine OR statements with AND
5. Add any NOT statements to the end

# Creating a Boolean Search

**QUESTION: Is Vitamin C helpful in treating the flu?**

**1. Identify concepts and list terms**

| <b>Concept 1<br/>Influenza</b> | <b>Concept 2<br/>Vitamin C</b> | <b>Concept 3<br/>Treatment</b> | <b>Concept 4<br/>helpfulness</b> |
|--------------------------------|--------------------------------|--------------------------------|----------------------------------|
| Influenza                      | Vitamin C                      | Treatment                      | Outcome                          |
| Flu                            | Ascorbic acid                  | Therapy                        | Recovery                         |
|                                | Orange Juice                   | Management                     | Success                          |

# Step 2



## 2. Make your OR statements, one per concept

- ▣ (influenza OR flu OR orthomyxovirus)
- ▣ (vitamin C OR ascorbic acid OR ascorbate)
- ▣ (treatment OR therapy OR management)
- ▣ (outcome OR recovery OR success)

# Steps 3 and 4

---

3. Put “AND” between each of the OR statements  
**(influenza OR flu) AND (vitamin C OR ascorbic acid OR orange juice) AND (treatment OR therapy OR management) AND (outcome OR recovery OR success)**
4. Consider any NOT statements you might want to add.

*Note: NOT isn't used very often*

# Parsing a Boolean Search

(emergency **OR** acute **OR** critical) **AND** (treatment **OR** therapy **OR** management **OR** care) **AND** (motor vehicle accident **OR** car crash) **NOT** (pedestrian **OR** walking)

What are the four concepts?

What terms are used for each concept?

Which three concepts must be included in all records found?

Which concept must not be included in any record found?

# Beyond Basic Boolean



- Field Searching
- Controlled Vocabulary
  - ▣ Subject vs. Keyword Searching
- Specialty Features
  - ▣ Truncation
  - ▣ Phrase searching

# Field Searching



Almost all databases will provide you with some ability to search a specific field or fields.

- Allows faster searching
- Allows more accurate searching

Not all databases may make all fields searchable.

Each search system will require a specific format.

# All Field vs. Specific Field Searches



I would like to find articles by John Smith.

- Search all fields: John Smith
- Search Author Field only: John Smith

I would like to find an article published in 1997.

- Search all fields: 1997
- Search Publication Date Field: 1997

Why waste time searching for a date in the author field or an author in the volume field?

# Final Notes on Fields



- Each database provides its own specific fields
- Each database requires a specific format to designate field searching
- When searching a new database, take a moment to read the help documentation; most will provide a list of fields and how to search them.

# Controlled Vocabulary

A controlled vocabulary is a set of established terms where

- every term represents a single concept
- only one term is used for that concept

*Example: Many words can be used to represent the concept "people who teach."*

teachers  
faculty  
instructors  
professors  
tutors  
educators  
lecturers, etc

In a controlled vocabulary  
one of these will be chosen  
to represent the concept

MeSH = Faculty  
Faculty, Dental  
Faculty, Medical  
Faculty, Nursing, etc.

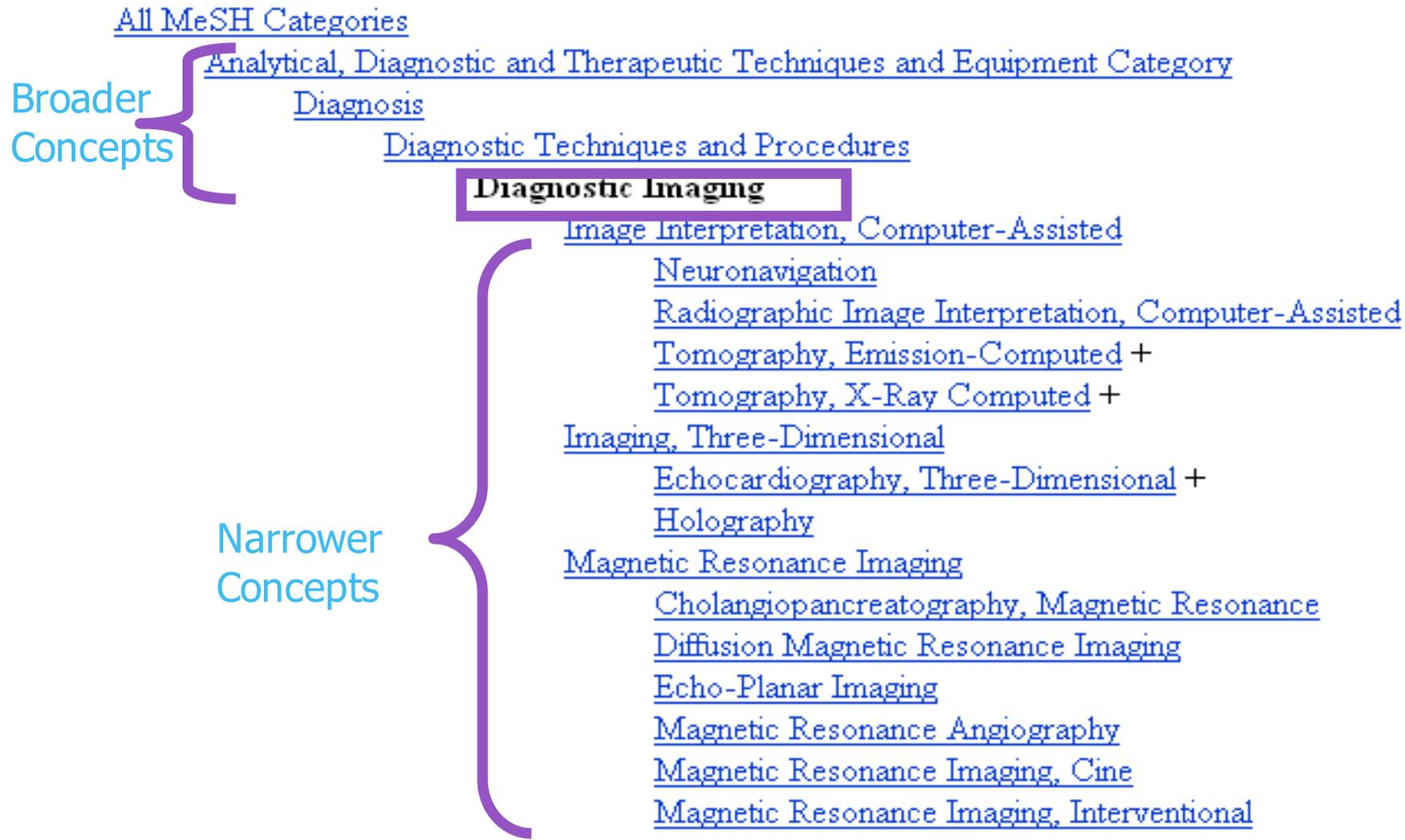
# Another example

- How many words could you think of for the idea of “cancer”?
  - ▣ Cancer, tumor, malignancy, neoplasm, sarcoma...
  
- Articles in a database
  - ▣ Article one: “Breast tumors in young women”
  - ▣ Article two: “Surgery for prostate cancer.”
  - ▣ Article three: “Diagnosing Melanoma.”

*All three articles are about types of cancer but different terms are used in titles.*

- 
- In a controlled vocabulary ONE word (e.g., neoplasm) is chosen and placed in a special field, usually called a subject field.
  
  - For all three articles
    - ▣ Article one: “Breast tumors in young women”
    - ▣ Article two: “Surgery for prostate cancer.”
    - ▣ Article three: “Diagnosing Melanoma.”
  
  - The subject term (concept term) “neoplasm” is placed in the subject field by database indexers.

# Structure of Controlled Vocabulary



# More on Controlled Vocabulary

- “Expanding” = Search includes all narrower terms beneath the searched term
  - ▣ Some databases do it automatically, others don’t
- “Focusing” or “Majoring” = For a given item, some subject terms are considered the major focus; you can select to return only those articles.
  - ▣ Hip fracture[majr] = only give me articles where hip fracture is an important concept

# Subject vs. Keyword Searching

## Controlled vocabulary searching

- Matches terms against a specific field in the record.
- You need to consult a thesaurus (paper or online) to find out what the controlled vocabulary term is for each concept.

## Free-text (keyword) searching

- Some concepts have many synonyms. A free-text search statement would mean "OR"ing all those terms together
- Matches terms against words anywhere in record (abstract, title, etc.).

# Advantages to Controlled Vocabularies



- Using the controlled vocabulary can make your search more precise and easier.
- Increases the relevancy of results (fewer false drops)
- The indexers have already done much of the work for you.
- Searchable tree structures of terms can help you find new terms to use.

# Problems with Controlled Vocabularies

- ❑ NOT all databases use a controlled vocabulary
- ❑ New concepts take time to be added
- ❑ There is often a lag phase during which the newest articles aren't indexed
- ❑ Controlled vocabularies can contain some very strange things and some concepts may not be handled well
- ❑ The controlled vocabulary must be easily searchable

Trying to understand what is and isn't in a particular controlled vocabulary can give you a big headache!

# Combining Subject and Keyword Searching



To be comprehensive, it is often helpful to combine subject and keyword searching

(diabetes mellitus[mh] OR diabetes)

(sickle cell anemia[mh] OR sickle cell anaemia)

# Some Specialty Features



- Truncation
- Phrase searching
- Neighboring and other rarer Boolean operators

# Truncation



What about including the singular and plural versions of words as well as other word variations?

For example: **therapy, therapies, therapeutics,**

You could combine them all in an OR relationship:

**(therapy OR therapies OR therapeutics OR therapeutic)**

But an easier way is by the use of truncation.

**therap\***

Each database handles truncation in a unique way.

The '\*' and '\$' are the most common wildcard symbols.

# More on Truncation

**Some examples:**

**Bacter\$**

**Proc\***

**Vir?**

**Staph?**

Be cautious when truncating!

If the word stem is too short, there may be too many possible variations and you might pick up unrelated terms.

For example, using `proc*` for finding procaine-like drugs will also include words like proceedings and process.

# Phrase Searching

- Sometimes you want to force the database to search for a set of words in exact order  
“fever of unknown origin”

Most databases will accept a phrase in quotes.

BUT...some do not handle phrases well and will automatically break them up – usually ‘AND’-ing the terms

Check how the database handles phrase searching before doing it!

# Limits Options

---

- Many databases provide “limits” pages that make it easier for you to select common options such as language, article type, publication dates, human or animal, gender, age groups, etc.
- Each database’s limits options are unique
- Most limits can be done ‘by hand’ using field tags, but sometimes limit pages save time

# PubMed Limits Page

The screenshot shows the PubMed.gov search results page. At the top, the search query is "Hypoplastic Left Heart Syndrome"[Mesh] AND "Diagnosis"[MeSH]. The page is set to display 20 results per page, sorted by "Recently Added". The results list includes four entries, each with a checkbox, a title link, authors, journal information, and PMID. The left sidebar contains various filter options such as "Text availability", "Publication dates", "Species", "Article types", and "Languages".

PubMed.gov  
US National Library of Medicine  
National Institutes of Health

PubMed ("Hypoplastic Left Heart Syndrome"[Mesh]) AND "Diagnosis"[MeSH]

RSS Save search Advanced

Show additional filters

Display Settings: Summary, 20 per page, Sorted by Recently Added Send to:

Text availability  
Abstract available  
Free full text available  
Full text available

Publication dates  
5 years  
10 years  
Custom range...

Species  
Humans  
Other Animals

Article types  
Clinical Trial  
Practice Guideline  
Randomized Controlled Trial  
Review  
Systematic Reviews  
more ...

Languages  
English  
more ...

Clear all

Show additional filters

Results: 1 to 20 of 793

1. [Successful Fontan completion after cardiac resynchronization therapy.](#)  
Enomoto Y, Aoki M, Nakamura Y, Hagino I, Fujiwara T, Nakajima H.  
Circulation. 2012 May 15;125(19):e655-8. No abstract available.  
PMID: 22586294 [PubMed - indexed for MEDLINE]  
[Related citations](#)

2. [Does initial shunt type for the Norwood procedure affect echocardiographic measures of cardiac size and function during infancy?: the Single Ventricle Reconstruction trial.](#)  
Frommelt PC, Guey LT, Minich LL, Bhat M, Bradley TJ, Colan SD, Ensing G, Gorentz J, Heydarian H, John JB, Lai WW, Levine JC, Mahle WT, Miller SG, Ohye RG, Pearson GD, Shirali GS, Wong PC, Cohen MS; Pediatric Heart Network Investigators.  
Circulation. 2012 May 29;125(21):2630-8. Epub 2012 Apr 21.  
PMID: 22523314 [PubMed - indexed for MEDLINE]  
[Related citations](#)

3. [Pseudoaneurysm following percutaneous balloon angioplasty for aortic arch recoarctation after the Norwood procedure.](#)  
Yamanaka K, Yoshimura N, Higuma T, Ichida F.  
Gen Thorac Cardiovasc Surg. 2012 May;60(5):305-7. Epub 2012 Mar 28.  
PMID: 22453542 [PubMed - indexed for MEDLINE]  
[Related citations](#)

4. [Tricuspid valve repair in single ventricle: timing and techniques.](#)  
Tsang VT, Raja SG.  
Semin Thorac Cardiovasc Surg Pediatr Card Surg Annu. 2012;15(1):61-8. Review.  
PMID: 22424509 [PubMed - indexed for MEDLINE]

# Step-By-Step Search Construction

1. State the question
2. Identify the concepts in the question
3. For each concept, determine keywords and subject terms
4. Combine terms for the same concept with “OR” in parenthesis
5. Combine “OR” statements with AND
6. Put any NOT terms at the end

Keep track of your searches, how many articles were found total, and how many you selected as relevant

# Example

---

- Question: What is the appropriate ED medical management of adult patients with intracranial hemorrhage (either trauma or spontaneous)?