Data-mining

Some useful resources
Data Mining? “Yeah, I could use a nap...”
What We’ll Discuss

- Why search
- Who should do it
- Sources
  - National Library of Medicine
    - Includes the NLM Gateway, PubMed, ClinicalTrials.gov, MEDLINE Plus, and more
  - Google
  - FDA website
  - National Guidelines Clearinghouse website
- Obtaining articles
Why Search?

- Find product data and competitive information
  - Clients often provide incomplete information
- Find new data
  - Clients may not even be aware it exists
- Learn about the latest research and where the market is headed
  - Can you ever have too much information about your client’s or competitor's products?
Who should do it

- You
  - Consider as a value added service in any and all work you do?
  - Or you could request a budget for it upfront -- it can be very time consuming

- The client
  - But will you get the information you need?
National Library of Medicine

- The NLM Gateway
  - Simultaneous multiple retrieval system for “one-stop searching”
- PubMed
  - Enormous database of citations and abstracts
- MEDLINE Plus
  - Physician and patient oriented information
- ClinicalTrials.gov
  - A great resource to check on existing clinical trials
- Health Services Research Projects in Progress - HSRProj
  - Provides research information before results are published
The NLM Gateway

- Simultaneous “one-stop searching”
  - Journal citations from 1951 to present
  - Books, Serials, AVs
  - Consumer health
  - Meeting abstracts
- Other collections
  - HSRProj
  - OMIM
  - HSDB
Gateway – Synercid Search Results
Gateway – Thrombin-JMI Search Results

![Gateway screen shot](image)

**Results Summary**

<table>
<thead>
<tr>
<th>Category</th>
<th>Items Found</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journal Citations</td>
<td>3</td>
<td>Display Results</td>
</tr>
<tr>
<td>Books / Serials / AVs</td>
<td>0</td>
<td>Display Results</td>
</tr>
<tr>
<td>Consumer Health</td>
<td>0</td>
<td>Display Results</td>
</tr>
<tr>
<td>Meeting Abstracts</td>
<td>0</td>
<td>Display Results</td>
</tr>
<tr>
<td>Other Collections</td>
<td>71</td>
<td>Display Results</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>74</strong></td>
<td>Display Results</td>
</tr>
</tbody>
</table>

Note: There may be a misspelling in your query.
PubMed

- Where we search for research from medical journals
- 13 million citations/abstracts
- 4,500 journals
  - From 1966

- Covers
  - Medicine
  - Nursing
  - Dentistry
  - Veterinary medicine
  - Healthcare systems
  - Preclinical sciences
More About PubMed

- Links to other databases
  - OLDMEDLINE (1950-1965)
  - International biomedical journals
    - 1,760,000 citations
- Not all journals are strictly scientific or medical
  - We recently found a Newsweek article in PubMed
- Links (LinkOut) to the full-text of articles at participating publishers' web sites
What Can You Do in PubMed?

- Search for articles (usually abstracts)
  - By keyword
  - By author
  - By Journal, etc
- Combine searches
- Link to related articles
- Link to outside sources
  - To purchase the full article
  - Look at related books (including pages in the books)
- Clinical queries
The PubMed Homepage

PubMed – Search Overview

- Search by keyword
  - E.g., thrombin, but you get 31,365 results
- Use a more specific keyword
  - Like “Thrombin-JMI”, but now you only get 2 hits
- Additional tools
  - Limits
  - Search Field Tags
  - meSH
PubMed – Additional Tools

- Limits
  - Age, groups, gender, type of study, etc

- Search Field tags
  - Qualify terms, used in brackets [ta]

- Boolean Statements
  - And, OR, NOT

- MeSH
  - The “vocabulary” of PubMed
  - Stands for Medical Subjects Headings
  - Used for indexing biomedical literature
PubMed -- Limits

- Use All Fields pull-down menu to specify a field.
- Boolean operators AND, OR, NOT must be in upper case.
- If search fields tags are used enclose in square brackets, e.g., rubella [ti]
- Search limits may exclude in process and publisher supplied citations.

Limited to:
- All Fields
- Publication Types
- Languages
- Humans or Animals
- Gender

Publication Date: From [ ] To [ ]
Use the format YYYY/MM/DD, month and day are optional.
PubMed -- Search Field Tags

- Search Field Tags also qualify terms
- Letters in brackets
  - Nature [ta]
  - Watson JD [au]
  - 1953 [dp]
  - “Fever of unknown origin” or [ti]

Boolean Statements

- Example: citations on DNA authored by Crick in 1993
- Effect of heat or humidity on surgical hemostasis
  - (heat OR humidity) AND surgical hemostasis
meSH organizes terms in a hierarchical manner
PubMed -- meSH Tutorial


www.nlm.nih.gov/bsd/viewlet/mesh/subheadings/mesh3_launcher.html
Real world meSH example

Thrombin-JMI use in bony ingrowth (cementless implants) in hip arthroscopy

- On a noncemented hip, can Thrombin JMI be injected into the medullary canal of the femur and will it affect the bony ingrowth?
- Can Thrombin JMI be put in the acetabulum and will it affect the bony ingrowth?
MEDLINE Plus

- Information on over 650 diseases and conditions
- Medical encyclopedia and dictionary
- Information on prescription and nonprescription drugs
- Links to ClinicalTrials.gov
- Links to news
- Sponsored by the NIH – no ads

www.medlineplus.gov
The MEDLINE Plus Homepage
ClinicalTrials.gov

- Information about federally funded and private human clinical trials
- Includes the trial’s
  - Purpose
  - Locations
  - Participant requirements
  - Phone number
ClinicalTrials.gov Homepage

ClinicalTrials.gov provides regularly updated information about federally and privately supported clinical research in human volunteers. ClinicalTrials.gov gives you information about a trial's purpose, who may participate, locations, and phone numbers for more details. Before searching, you may want to learn more about clinical trials.

Search Clinical Trials

Example: heart attack, Los Angeles

Search by Specific Information

Focused Search - search by disease, location, treatment, sponsor...

Browse

Browse by Condition - studies listed by disease or condition
Browse by Sponsor - studies listed by funding organization
Browse by Status - studies listed by recruitment status

Resource Information

Understanding Clinical Trials - information explaining and describing clinical trials
HSRProj

- Provides research-in-progress before results are available
- From ongoing grants and contracts from government or academic studies
- Retrieve:
  - Name of sponsoring agency
  - Principal investigator
  - Dates of projects, etc.

http://www.nlm.nih.gov/hsrproj/
HSRProj Homepage
Google

- Images
- Alerts
- Groups
Images From Google?

- Yes, you really can find images in Google.

http://www.google.com/imghp?hl=en&tab=wi&q=
Google Alerts

- Currently in Beta tests
- Enter keywords
- Get email updates on the keyword(s)
Google Groups

- A “forum” dedicated to a specific topic and usually open to the general public.
- Sometimes moderated and open only to members chosen by the moderator.
FDA Webpage – Tons of Info.

- Contains MedWatch
  - product safety alerts
    - Alerts be sent to your email address as they happen

- Find
  - Warning Letters
    - [http://www.fda.gov/foi/warning.htm](http://www.fda.gov/foi/warning.htm)
  - Approved drugs and patent information
    - Using the Electronic Orange Book database
      - [http://www.fda.gov/cder/ob/default.htm](http://www.fda.gov/cder/ob/default.htm)
  - Devices
    - [http://www.fda.gov/cdrh/index.html](http://www.fda.gov/cdrh/index.html)
  - Guidance documents
Guidance Documents

Guidance documents represent the Agency's current thinking on a particular subject. They do not create or confer any rights for or on any person and do not operate to bind FDA or the public. An alternative approach may be used if such approach satisfies the requirements of the applicable statute, regulations, or both. For information on a specific guidance document, please contact the originating office.

Most of these documents are in Adobe Acrobat format, also known as PDF. The free upgrade to Adobe Acrobat 5.0 or higher is recommended, especially if you have difficulty opening any of the documents below. Another method of obtaining guidance documents is through the Division of Drug Information.

http://www.fda.gov/cder/guidance/guidance.htm
National Guidelines Clearinghouse

- A comprehensive database of evidence-based clinical practice guidelines and related documents
- Sponsored by the Agency for Healthcare Research and Quality (AHRQ)
- Updated weekly
National Guidelines Clearinghouse

http://www.guideline.gov/search/detailedsearch.aspx
Obtaining Articles for Clients

Free
- Check online – sometimes PubMed has links to free articles
- Freemedicaljournals.com

Pay
- Through Emory
  - $12 normal; $24 rush
  - Fax versions
  - Can be generated into PDF
- External sources (e.g., publisher web sites)
  - Prices vary ($20-30)
  - PDF versions (usually)