

NLM resources for research and practice: an overview and an R&D application

Dina Demner-Fushman

IGERT: From Data to Solutions

October 3, 2014



U.S. National Library of Medicine



Disclaimer

The views and opinions expressed do not necessarily state or reflect those of the U.S. Government, and they may not be used for advertising or product endorsement purposes.

America's two greatest gifts to the world are jazz and Medline.

Smith R, Chalmers I. Britain's gift: a "Medline" of synthesised evidence. *BMJ*. 2001 Dec 22-29;323(7327):1437-8.

NLM DATABASES

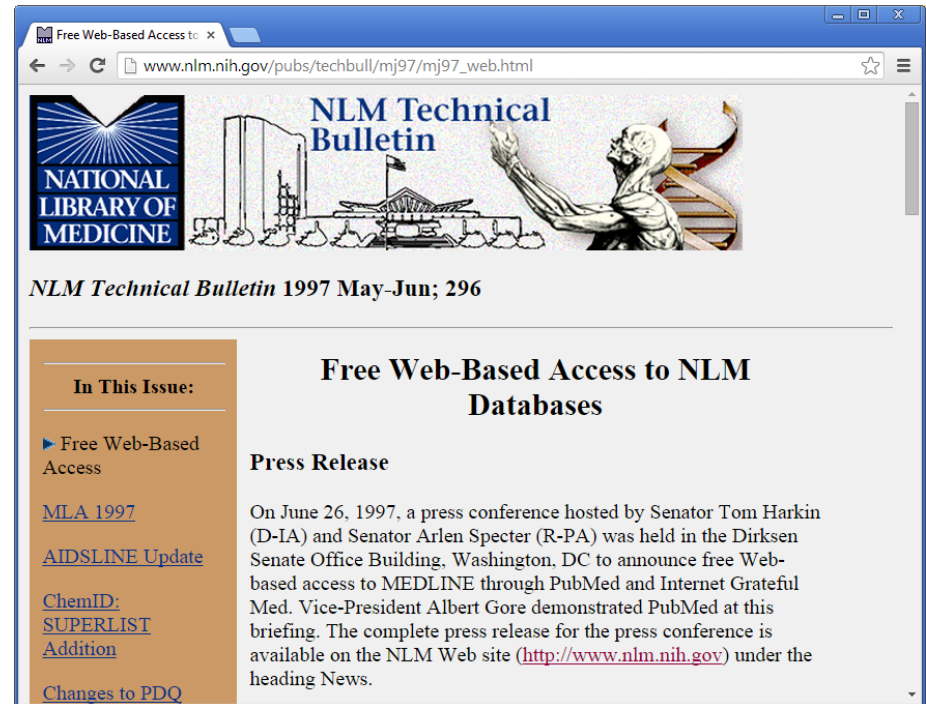




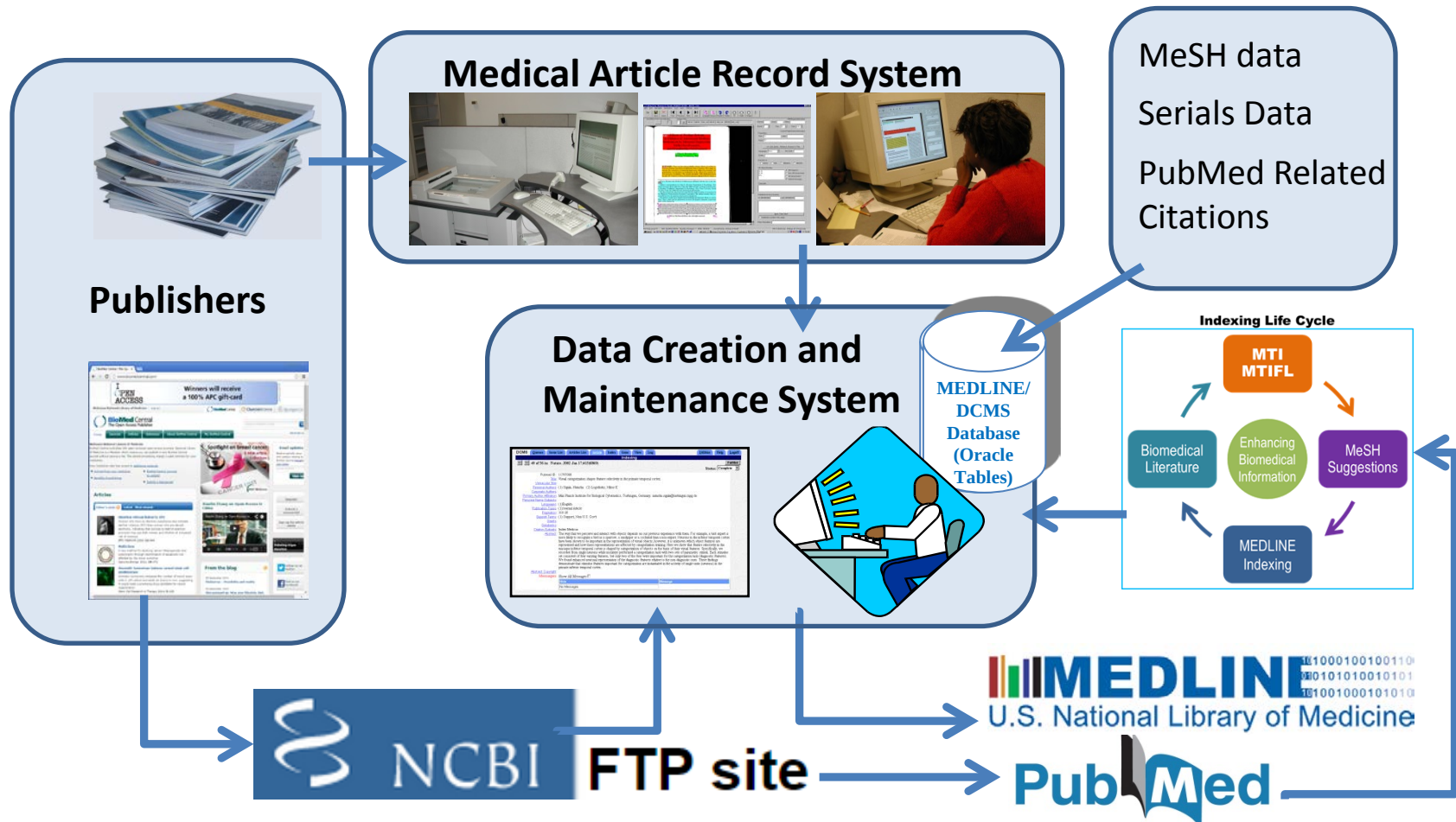
- MEDLINE® is the National Library of Medicine® (NLM®) journal citation database. Started in the 1960s and online since 1971, it now provides over 21 million references to biomedical and life sciences journal articles back to 1946.
- MEDLINE includes citations from over 5,600 worldwide journals in about 40 languages; about 60 languages for older journals.
 - The journals are selected based on the recommendation of the Literature Selection Technical Review Committee (LSTRC), an NIH-chartered advisory committee of external experts analogous to the committees that review NIH grant applications.
 - The LSTRC considers the quality of the scientific content of a journal, including originality and the importance of the content for the MEDLINE global audience.
- Since 2005, between 10,000-20,000 completed references are added each week.



- Available since 1996
- Over 23 million references include
 - the MEDLINE database
 - In-process citations
 - Citations to articles that are out-of-scope (e.g., covering astrophysics) from MEDLINE journals
 - "Ahead of Print" citations that precede the article's final publication in a MEDLINE indexed journal.
 - Citations that precede the date that a journal was selected for MEDLINE indexing (when supplied electronically by the publisher).
 - Citations to author manuscripts of articles published by NIH-funded researchers.
 - Citations for the books available on the NCBI Bookshelf
 - (a citation for the book and in some cases each chapter of the book).



How do Citations get into MEDLINE/PubMed?



Human readable citation formats

The screenshot shows the PubMed web interface for the article 'Duct tape for broken chromosomes' by Gorbisky GJ. The page includes the NCBI logo, search bar, and various navigation links. The abstract text is visible, along with the full text link and related citations. The page is titled 'Duct tape for broken chromosomes' and the author is 'Gorbisky GJ'. The abstract describes a cell undergoing mitosis with a potentially catastrophic situation when a DNA double-strand break creates a chromosome fragment that lacks connection to a centromere. The article is published in Cell, 2010 Jan 22;140(2):178-80. doi: 10.1016/j.cell.2010.01.005.

The screenshot shows the PubMed citation format page for the article 'Duct tape for broken chromosomes' by Gorbisky GJ. The page displays the citation in various formats, including MEDLINE, EMBASE, and others. The citation text is as follows:

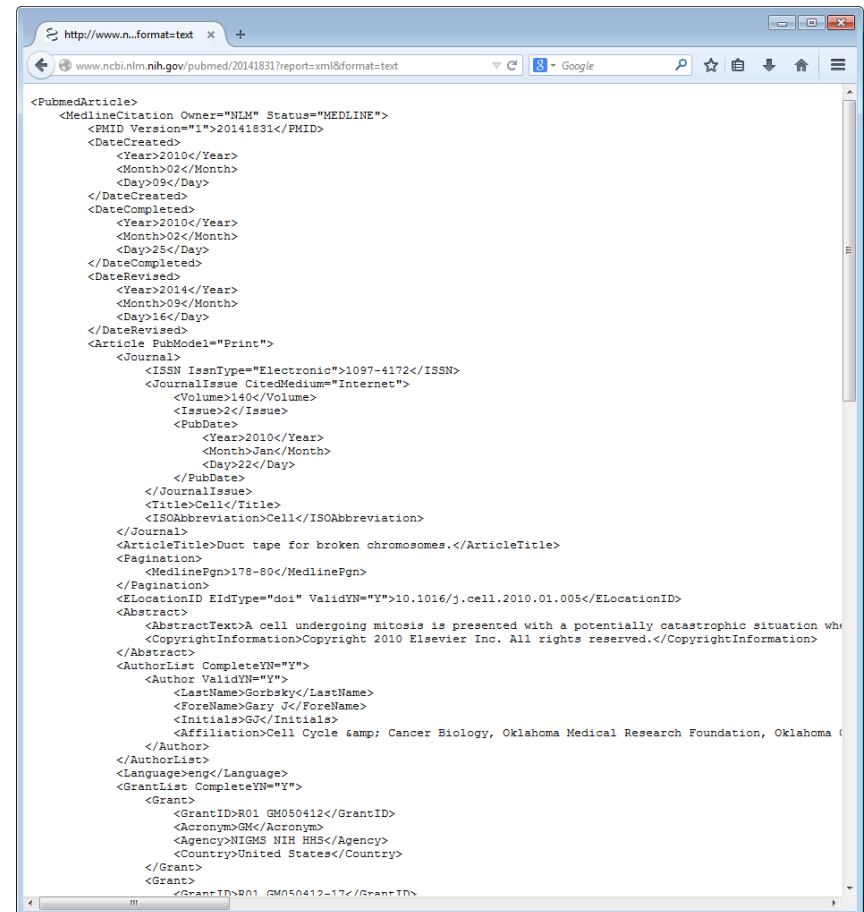
```
PMID- 20141831
OWN - NLM
STAT- MEDLINE
DA - 20100209
DCOM- 20100225
LR - 20140916
IS - 1097-4172 (Electronic)
IS - 0092-8674 (Linking)
VI - 140
IP - 2
DP - 2010 Jan 22
TI - Duct tape for broken chromosomes.
PG - 178-80
LID - 10.1016/j.cell.2010.01.005 [doi]
AB - A cell undergoing mitosis is presented with a potentially catastrophic situation when a DNA double-strand break creates a chromosome fragment that lacks connection to a centromere. Royou et al. (2010) now reveal that this cellular crisis is averted in fruit fly neuroblasts by thin chromatin tethers that hold on to the ends of the broken chromosomes.
CI - Copyright 2010 Elsevier Inc. All rights reserved.
FAU - Gorbisky, Gary J
AU - Gorbisky GJ
AD - Cell Cycle & Cancer Biology, Oklahoma Medical Research Foundation, Oklahoma City, OK 73104, USA. gjg@omrf.org
LA - eng
GR - R01 GM050412/GM/NIGMS NIH HHS/United States
GR - R01 GM050412-17/GM/NIGMS NIH HHS/United States
GR - R01 GM050412-18/GM/NIGMS NIH HHS/United States
PT - Comment
PT - Journal Article
PL - United States
TA - Cell
JT - Cell
JID - 0413066
SB - IM
CON - Cell. 2010 Jan 22;140(2):235-45. PMID: 20141837
MH - Animals
MH - Centromere/*metabolism
MH - Chromosomes/*metabolism
MH - DNA Breaks, Double-Stranded
MH - Drosophila/*cytology/embryology/*metabolism
MH - Mitosis
EDAT- 2010/02/10 06:00
MHDA- 2010/02/26 06:00
CRDT- 2010/02/10 06:00
AID - S0092-8674(10)00006-1 [pii]
AID - 10.1016/j.cell.2010.01.005 [doi]
PST - publish
SO - Cell. 2010 Jan 22;140(2):178-80. doi: 10.1016/j.cell.2010.01.005.
```


XML Publisher DTD tags and XML for the MEDLINE DTD example

The following is a glossary of the tags defined in the [PubMed DTD](#). Click on each of the tag names below for more information. You can also view an [Example of a Standard XML File](#).

Data Tags (**R** = Required, **O** = Optional **O/R** = Optional or Required). Tag names are **case sensitive**. Required tags must be included; optional tags must be included only if the data requested appears in the print or electronic article. Optional or Required tags are dependent on the use of other tags.

File Header (R)	Replaces (O)	Affiliation (O)
ArticleSet (R)	ArticleTitle (O)	Identifier (O)
Article (R)	VernacularTitle (O)	GroupList (O/R)
Journal (R)	FirstPage (O/R)	Group (R)
PublisherName (R)	LastPage (O)	GroupName (R)
JournalTitle (R)	ELocationID (O/R)	IndividualName (O)
Issn (R)	Language (O)	PublicationType (O)
Volume (O/R)	AuthorList (O/R)	ArticleIdList (O/R)
Issue (O/R)	Author (R)	ArticleId (R)
PubDate (R)	FirstName (O/R)	History (O)
Year (R)	MiddleName (O)	Abstract (O)
Month (O/R)	LastName (O/R)	OtherAbstract (O)
Season (O)	Suffix (O)	CopyrightInformation (O)
Day (O)	CollectiveName (O)	ObjectList (O)
		Object (O)
		Param (O)



```
<PubmedArticle>
  <MedlineCitation Owner="NLM" Status="MEDLINE">
    <PMID Version="1">20141831</PMID>
    <DateCreated>
      <Year>2010</Year>
      <Month>02</Month>
      <Day>09</Day>
    </DateCreated>
    <DateCompleted>
      <Year>2010</Year>
      <Month>02</Month>
      <Day>25</Day>
    </DateCompleted>
    <DateRevised>
      <Year>2014</Year>
      <Month>09</Month>
      <Day>16</Day>
    </DateRevised>
    <Article PubModel="Print">
      <Journal>
        <ISSN IssnType="Electronic">1097-4172</ISSN>
        <JournalIssue CitedMedium="Internet">
          <Volume>140</Volume>
          <Issue>2</Issue>
          <PubDate>
            <Year>2010</Year>
            <Month>Jan</Month>
            <Day>22</Day>
          </PubDate>
        </JournalIssue>
        <Title>Cell</Title>
        <ISOAbbreviation>Cell</ISOAbbreviation>
      </Journal>
      <ArticleTitle>Duct tape for broken chromosomes.</ArticleTitle>
      <Pagination>
        <MedlinePgn>178-80</MedlinePgn>
      </Pagination>
      <ELocationID EIdType="doi" ValidYN="Y">10.1016/j.cell.2010.01.005</ELocationID>
      <Abstract>
        <AbstractText>A cell undergoing mitosis is presented with a potentially catastrophic situation wh...
        <CopyrightInformation>Copyright 2010 Elsevier Inc. All rights reserved.</CopyrightInformation>
      </Abstract>
      <AuthorList CompleteYN="Y">
        <Author ValidYN="Y">
          <LastName>Gorbsky</LastName>
          <ForeName>Gary J</ForeName>
          <Initials>GJ</Initials>
          <Affiliation>Cell Cycle & Cancer Biology, Oklahoma Medical Research Foundation, Oklahoma C...
        </Author>
      </AuthorList>
      <Language>eng</Language>
      <GrantList CompleteYN="Y">
        <Grant>
          <GrantID>R01 GM050412</GrantID>
          <Acronym>GM</Acronym>
          <Agency>NIGMS NIH HHS</Agency>
          <Country>United States</Country>
        </Grant>
      </GrantList>
    </Article>
  </MedlineCitation>
</PubmedArticle>
```

Publisher DTD



PubMed Central

- A free archive for full-text biomedical and life sciences journal articles launched in 2000.
- A repository for journal literature deposited by
 - participating publishers
 - author manuscripts submitted in compliance with the NIH Public Access Policy
- Reciprocal links between the full text in PMC and citations in PubMed.
 - PubMed citations are created for content not already in MEDLINE
 - Some PMC content, e.g., book reviews, is not cited in PubMed
- The PMC Open Access Subset under a Creative Commons or similar license
 - FTP service to download a complete set of files — XML, images, PDF, and supplementary data files.
 - Articles: 852,041; Images: 2,618,629; Articles containing no images: 246,066
 - An API to find articles by ID, or those that have been updated recently.

Entrez

Entrez is NCBI's primary text search and retrieval system that integrates the PubMed database of biomedical literature with 39 other literature and molecular databases including DNA and protein sequence, structure, gene, genome, genetic variation and gene expression.



Entrez Programming Utilities (E-utilities)

- A set of nine programs that provide an application programming interface (API) to Entrez
 - **EInfo (database statistics)** Provides the number of records indexed in each field of a given database, the date of the last update of the database, and the available links from the database to other Entrez databases.
 - **ESearch (text searches)** Responds to a text query with the list of matching UIDs in a given database (for later use in ESummary, EFetch or ELink), along with the term translations of the query.
 - **EPost (UID uploads)** Accepts a list of UIDs from a given database, stores the set on the History Server, and responds with a query key and web environment for the uploaded dataset.
 - **ESummary (document summary downloads)** Responds to a list of UIDs from a given database with the corresponding document summaries.
 - **EFetch (data record downloads)** Responds to a list of UIDs in a given database with the corresponding data records in a specified format.
 - **ELink (Entrez links)** Responds to a list of UIDs in a given database with list of external and internal links.
 - **EGQuery (global query)** Responds to a text query with the number of records matching the query in each Entrez database.
 - **ESpell (spelling suggestions)** Retrieves spelling suggestions for a text query in a given database.
 - **ECitMatch (batch citation searching in PubMed)** Retrieves PubMed IDs (PMIDs) corresponding to a set of input citation strings.

Platforms for “crowdsourcing”

- **dbGAP**: the **d**atabase of **G**enotypes and **P**henotypes archives and distributes the results of studies that have investigated the interaction of genotype and phenotype.
 - Open access to some datasets:
<http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/about.html#open>
- **Genetic Testing Registry (GTR)**: provides a central location for voluntary submission of genetic test information by providers. The scope includes the test's purpose, methodology, validity, evidence of the test's usefulness, and laboratory contacts and credentials.
<http://www.ncbi.nlm.nih.gov/gtr/>
- **clinVar**: a freely accessible, public archive of reports of the relationships among human variations and phenotypes, with supporting evidence. <http://www.ncbi.nlm.nih.gov/clinvar/>

DRUG AND CHEMICALS INFORMATION



DailyMed

dailymed.nlm.nih.gov/dailymed/index.cfm

Google

SEARCH

Enter drug, NDC code, drug class, or Set ID

MORE WAYS TO SEARCH:

ADVANCED SEARCH

BROWSE DRUG CLASSES

LABEL ARCHIVES

TABLET/CAPSULE ID TOOL

This website contains 66527 drug listings as submitted to the Food and Drug Administration (FDA).
At the present time, this Web site does not contain a complete listing of labels for approved prescription drugs.

SHARE

NEWS

DailyMed Announcements

Posted: September 26, 2014

On September 18, 2014, NLM launched a newly redesigned DailyMed Web site. The update includes a responsive design and many changes that improve usability and navigation.

MORE INFO

Get RSS News & Updates

 The DailyMed RSS feed provides updates and information about new drug labels approved by the FDA and published on NLM's DailyMed Web site.

ABOUT DAILYMED

DailyMed provides trustworthy information about marketed drugs in the United States. DailyMed is the official provider of FDA label information (package inserts). This Web site provides a standard, comprehensive, up-to-date, look-up and download resource of medication content and labeling found in medication package inserts.

The National Library of Medicine (NLM) provides this as a public service and does not accept advertisements. The drug labeling information on this Web site is the most recent submitted to the Food and Drug Administration (FDA) and currently in use; it may include, for example, strengthened warnings undergoing FDA review or minor editorial changes. These labels have been reformatted to make them easier to read.

MORE INFO

FDA GUIDANCES & INFORMATION

Drug Guidance, Compliance & Regulatory Information

 [View FDA Structured Product Labeling Resources](#)
[View FDA Drug Labeling Guidances](#)
[View All FDA Drug Guidances](#)

NLM SPL RESOURCES

The following Structured Product Labeling (SPL) resources have been created to assist industry professionals.

Download Data

 [All Drug Labels](#)
[All Indexing Files](#)
[All Mapping Files](#)

SPL Image Guidelines

 Guidelines for SPL image files of oral solid dosage forms that are submitted to the FDA with SPL documents.

Presentations & Articles

 Stay informed through recent SPL related articles, online workshops, presentations and more.

APPLICATION DEVELOPMENT SUPPORT

Resources to Get the Most Out of DailyMed

 [Web Services](#)
[Mapping Files](#)

HELP

Pillbox - prototype pill identific...
pillbox.nlm.nih.gov/pillimage/search.php
Google

Pillbox rapid identification, reliable information
United States
National Library of Medicine
National Institutes of Health

Home
About
FAQ
Developer
Contact Us

Identify or Search for a Pill

Imprint:

☐ Pill does not have an imprint.

Drug Name or Ingredient(s):

Shape:

Inactive Ingredient(s):

☐ Find pills WITHOUT this ingredient.

Color:

Label Author:

Size:

DEA Schedule:

Score:
☒ Unknown
☐ 1
☐ 2
☐ 3
☐ 4

Product Code:
 -

Discover

There's more to a pill than how it looks. What's inside the pill other than the drug? Is it a controlled substance?

Connect

Learn more than the pill's name. Pillbox links you to the drug label, clinical trials, breastfeeding safety, and more.

Explore

Pillbox's image explorer is a photo album for pills. Sort by color, shape, size, scoring, and the text printed on the pill.

(requires Adobe Flash and will not run on some mobile devices)

Nat'l Library of Medicine
Nat'l Institutes of Health
U.S. Dept. of Health & Human Services
Copyright
Accessibility

Data updated: November 4, 2013

United States Department of Health & Human Services

sis.nlm.nih.gov/chemical.html

Google

U.S. Department of Health & Human Services

www.hhs.gov

Chemical & Drug Information

SPECIALIZED INFORMATION SERVICES

National Library of Medicine
National Institutes of Health NLM

SIS Home | About Us | A-Z Index | SIS News | Contact Us

Search This Site

SIS Home > Chemical & Drug Information

ChemIDplus

Information on 400,000+ chemicals including synonyms and structures

ChemIDplus Lite

Search by chemical name or Registry Number:

Search

ChemIDplus Advanced



Search by registry number, name, molecular formula, structure, physical and toxicological properties.

Searching ChemIDplus

- ChemIDplus Locators
- ChemIDplus Lite Help
- ChemIDplus Advanced Help
- FAQ
- TOXNET and Beyond Training Class Schedule and Workbook

Connect with Us






SIS Resources

AIDSInfo Drug Database

Information on FDA-approved and investigational HIV/AIDS-related drugs

Dietary Supplement Label Database

Full label contents from a sample of dietary supplement products sold in the U.S.

Drug Information Portal

Information on drugs from clinical trials through entry in the market place

Hazardous Substances Data Bank

Comprehensive, peer-reviewed toxicological data on chemicals

LactMed - Drugs and Lactation Database

Database of drugs to which breastfeeding mothers may be exposed

LiverTox

Information on liver injury caused by drugs, herbals and dietary supplements

Pillbox

Rapid identification of unknown tablets/capsules using physical features and images

TOXNET

Databases on hazardous chemicals, environmental health, and toxic releases

Additional Resources

Toxicology Information from the National Library of Medicine

Enviro-Health Links

Evaluated links to Internet resources on selected toxicology and environmental health issues

DailyMed

Manufacturers' drug package inserts. From the U. S. Food and Drug Administration

FDA Substance Registration System

Provides Unique Ingredient Identifiers (UNII) for substances in drugs, biologics, foods, and devices

PubChem

Information on the biological activities of small molecules

RxNorm

A standardized nomenclature for clinical drugs

TOXMAP

Uses maps of the U.S. to explore data from the US EPA's Toxics Release Inventory (TRI) and Superfund Program

Web Tools and Services for Chemistry Professionals (PDF, 500 KB)

The PubChem Project x

← → ↻ <https://pubchem.ncbi.nlm.nih.gov> ☆ ☰

Databases > Upload Services > Help more >



 BioAssay ?  Compound ?  Substance ?

[Advanced Search](#)

Try the new [PubChem Search](#)

New PubChem celebrates 10 years of service to the chemical biology community! [Read more...](#)

[more ...](#) 

[Write to Helpdesk](#) | [Disclaimer](#) | [Privacy Statement](#) | [Accessibility](#) | [Data Citation Guidelines](#)
National Center for Biotechnology Information
NLM | NIH | HHS

[BioActivity Summary](#) 
[BioActivity Datable](#) 
[BioActivity SAR](#) 
[BioActivity DataDicer](#) 
[Structure Search](#) 
[3D Conformer Tools](#) 
[Structure Clustering](#) 
[Classification](#) 
[Upload](#) 
[Download](#) 
[PubChem FTP](#) 

<https://pubchem.ncbi.nlm.nih.gov/#>

PubChem provides information on the biological activities of small molecules.

Pcsubstance contains more than 140 million records.

Pccompound contains more than 51 million unique structures.

PCBioAssay contains more than 1 million BioAssays. Each BioAssay contains a various number of data points.

RxNav Home Page
rxnav.nlm.nih.gov


[Home](#)
[APIs](#)
[Applications](#)
[Documentation](#)
[Data Sources](#)
[Terms Of Service](#)
[Contact Us](#)

RxNav

RxNav is a browser for several drug information sources, including RxNorm, RxTerms and NDF-RT. RxNav finds drugs in RxNorm from the names and codes in its constituent vocabularies.

[Launch RxNav](#)



RxNav

APIs

RxMix

RxClass

News

Statistics

RxNav brand names

Recent brand names added

Dermakleen	Eye Stream
IsoSol	Periglene
Peritussin Medicated Syrup	Peritussin

Drug Interaction API



A [drug interaction API](#) is now available. The API serves as a replacement for the drug interactions removed from the NDF-RT data set in September.

The RxClass API is here!

The [RxClass API](#) is a web service for accessing drug classes and associated drug members. The [RxClass Browser](#) uses the API for exploring and navigating through the class hierarchies to find the RxNorm drug members.

INFORMATION FOR CONSUMERS



- Information **for patients** from the National Institutes of Health and other trusted sources on over 900 diseases and conditions, drugs and wellness issues.
- A medical encyclopedia and a medical dictionary, easy-to-understand tutorials.
- Health information in Spanish and other languages.
- Health information from the media updated daily.
- No advertisement and no endorsement for any company or product.





- Linking Patient Portals and EHRs to Consumer Health Information
- MedlinePlus Connect accepts requests for information on diagnoses (problem codes), medications, and lab tests, and returns related MedlinePlus information. It is available as a Web application or a Web service.
 - For problem code requests, MedlinePlus Connect supports: ICD-9-CM (International Classification of Diseases, 9th edition, Clinical Modification) ; ICD-10-CM , SNOMED CT[®] (Systematized Nomenclature of Medicine, Clinical Terms)
 - For medication requests, MedlinePlus Connect supports RXCUI (RxNorm Concept Unique Identifier); NDC (National Drug Code)
 - For lab test requests, MedlinePlus Connect supports: LOINC[®] (Logical Observation Identifiers Names and Codes)
- May help an EHR achieve one of the criteria for Meaningful Use of Health Information Technology.



Genetics Home Reference ... x +

ghr.nlm.nih.gov

About Site Map Contact Us

Genetics Home Reference
Your Guide to Understanding Genetic Conditions

A service of the U.S. National Library of Medicine®

What's New

- nephronophthisis
- laryngo-oncho-cutaneous syndrome
- Legg-Calvé-Perthes disease
- More...

Newborn Screening

Detecting genetic disorders for early treatment

In the Spotlight

- Back to School with Genetics Home Reference
- GHR results now available from MedlinePlus Connect
- What is direct-to-consumer genetic testing?

Genetic Disorders A to Z
and related genes and chromosomes

Conditions

The genetics of more than 1,000 health conditions, diseases, and syndromes.

Genes

More than 1,200 genes, health effects of genetic differences, and gene families.

Chromosomes

Chromosomes, mitochondrial DNA, and associated health conditions.

Concepts & Tools
for understanding human genetics

Handbook

Learn about mutations, inheritance, genetic counseling, genetic testing, genomic research, and more.

Glossary

Medical and genetics definitions.

Resources

Links to other genetics information and organizations.

Genetics Home Reference provides consumer-friendly information about the effects of genetic variations on human health. The resources on this site should not be used as a substitute for professional medical care or advice. Users seeking information about a personal genetic disease, syndrome, or condition should consult with a qualified healthcare professional. See [How can I find a genetics professional in my area?](#) in the Handbook.

Published: September 29, 2014

[External link disclaimer](#)

[Lister Hill National Center for Biomedical Communications](#) ⇨
[U.S. National Library of Medicine](#) ⇨, [National Institutes of Health](#) ⇨
[Department of Health & Human Services](#) ⇨, [USA.gov](#) ⇨
[Freedom of Information Act](#) ⇨ [Copyright](#) ⇨ [Privacy](#) ⇨ [Accessibility](#) ⇨

⇨ Indicates a page outside Genetics Home Reference.
 Links to web sites outside the Federal Government do not constitute an endorsement.
 See [Selection Criteria for Web Links](#).
 PDF resources may require a [PDF reader](#).

This site complies with the [HONcode standard](#) for trustworthy health information: [verify here](#).

- **Consumer** information about genetic conditions and the genes or chromosomes associated with those conditions.
 - Genetic Condition Summaries
 - Gene Summaries
 - Gene Family Summaries
 - Chromosome Summaries
 - Handbook
 - explanations of how genes work and how mutations cause disorders; information about genetic testing, and gene therapy.
 - Glossary

Clinical Trials Registry

ClinicalTrials.gov

A service of the U.S. National Institutes of Health

ClinicalTrials.gov is a registry and results database of publicly and privately supported clinical studies of human participants conducted around the world. Learn more [about clinical studies](#) and [about this site](#), including relevant [history](#), [policies](#), and [laws](#).

[Find Studies](#) ▾ [About Clinical Studies](#) ▾ [Submit Studies](#) ▾ [Resources](#) ▾ [About This Site](#) ▾

ClinicalTrials.gov currently lists **175,684 studies** with locations in all 50 states and in **187 countries**.

[Text Size](#) ▾

Search for Studies

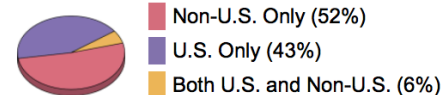
Example: "Heart attack" AND "Los Angeles"

[Advanced Search](#) | [See Studies by Topic](#)
[See Studies on a Map](#)

Search Help

- [How to search](#)
- [How to find results of studies](#)
- [How to read a study record](#)

Locations of Recruiting Studies



Total N = 33,967 studies
Data as of September 30, 2014

- [See more trends, charts, and maps](#)

Learn More

- [ClinicalTrials.gov Online Training](#)
- [Glossary of common site terms](#)

[For the Press](#)

[Using our RSS Feeds](#)

For Patients & Families

- [How to find studies](#)
- [See studies by topic](#)
- [Learn about clinical studies](#)
- [Learn more...](#)

For Researchers

- [How to submit studies](#)
- [Download content for analysis](#)
- [About the results database](#)
- [Learn more...](#)

For Study Record Managers

- [Why register?](#)
- [How to register study records](#)
- [FDAAA 801 Requirements](#)
- [Learn more...](#)

[HOME](#)

[RSS FEEDS](#)

[SITE MAP](#)

[TERMS AND CONDITIONS](#)

[DISCLAIMER](#)

[CONTACT NLM HELP DESK](#)

Clinical Trials Results

ClinicalTrials.gov

A service of the U.S. National Institutes of Health

Example: "Heart attack" AND "Los Angeles"

Search for studies:

Search

[Advanced Search](#) | [Help](#) | [Studies by Topic](#) | [Glossary](#)

[Find Studies](#) ▾

[About Clinical Studies](#) ▾

[Submit Studies](#) ▾

[Resources](#) ▾

[About This Site](#) ▾

[Home](#) > [Find Studies](#) > [Search Results](#)

[Text Size](#) ▾

14563 studies found for: (NOT NOTEXT) [FIRST-RECEIVED-RESULTS-DATE]

[Modify this search](#) | [How to Use Search Results](#)

List

By Topic

On a Map

Search Details

+ [Show Display Options](#)

 [Download](#)

 [Subscribe to RSS](#)

☐ Include only open studies ☐ Exclude studies with unknown status

Rank	Status	Study
------	--------	-------

1	Terminated	Oxandrolone to Heal Pressure Ulcers
---	------------	---

Has
Results

Condition: Pressure Ulcer

Interventions: Drug: Oxandrolone; Drug: Placebo

2	Terminated	Study of Pegylated Interferon-Alfa 2b in Combination With PUVA Therapy In CTCL
---	------------	--

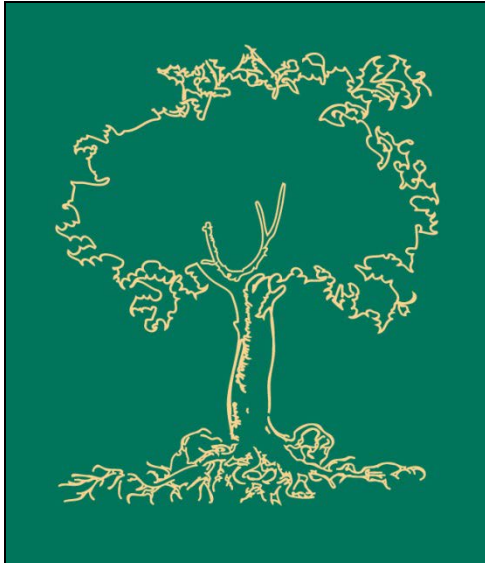
Has
Results

Condition: Lymphoma

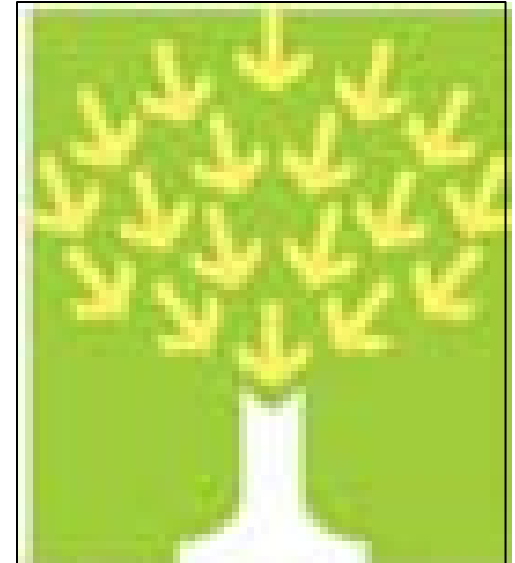
Interventions: Biological: Pegylated interferon α -2b; Other: Psoralens with ultraviolet light A;
Other: Narrowband-ultraviolet light B

INFORMATICS RESOURCES

The UMLS was initiated in 1986 as a long-term research & development project of the National Library of Medicine to reduce system development problems caused by differences in medical terminologies



- Unified
- Medical
- Language
- System



UMLS[®]
Unified Medical Language System[®]

UMLS Knowledge Sources & Tools

- **Metathesaurus**
 - Clusters terms into concepts with unique identifier
 - Over 140 source vocabularies (82 in Active Release) in 21 languages
 - 2,973,458 million concepts (8,646,656 distinct normalized concept names)
- **Semantic Network**
 - Defines relationships between concepts, organizes concepts into categories
 - 133 Semantic types (broad categories)
 - 54 Semantic relations (between categories)
- **Lexical resources**
 - SPECIALIST Lexicon (over 476,000 records)
 - morphology(inflection, derivation); spelling variants; part of speech....
 - Lexical tools (programs and databases)
 - Variant generation; spell check; POS tagger...
- **MetamorphoSys** (installation and customization)
- **UMLS Terminology Services (UTS)**
 - provide web interfaces and web services to search and retrieve UMLS data

UMLS Customization

- It is too broad for a specific purpose
- You are interested in only one of the 21 languages
- UMLS preserves the meanings, hierarchical connections, and other relationships between terms present in its source vocabularies
 - some of these make no sense in your context
 - some are generally suppressible
- You need to add your terms or relations
- You don't have licenses for copyrighted sources

UMLS-Based Resources & Use Examples

- Named Entity Recognition
 - **MetaMap**
- Automated indexing assistance
 - **MTI**
- Information retrieval
 - **ClinicalTrials.gov**
 - **Open-i**

MetaMap

- Purpose: Concept Identification (Named-Entity Recognition)
 - Identify UMLS Metathesaurus concepts in text

Examples:

Text gap, lexical variation: assay/assays

sister-chromatid exchange genotoxicity assays

C1519335:Sister Chromatid Exchange Assay

Composite phrases, word order, text gaps

pain on the left side of the chest

C0541828:chest pain left side(Left sided chest pain)

The Algorithm

- Parsing
 - SPECIALIST minimal-commitment parser
 - SPECIALIST lexicon
 - MedPost part-of-speech tagger
- Variant generation
 - SPECIALIST lexicon
 - Lexical Variant Generation (LVG)
- Candidate retrieval from the Metathesaurus
- Candidate evaluation
- Mapping construction

Use example

- With the help of MetaMap concept positional information XML tags we can now extract more than 130 different variations of the “kidney stone” string that are being documented within our sample corpus (Past Medical History section of 640,000 notes).
- These 137 trigger strings are matched by 10 UMLS synonyms.

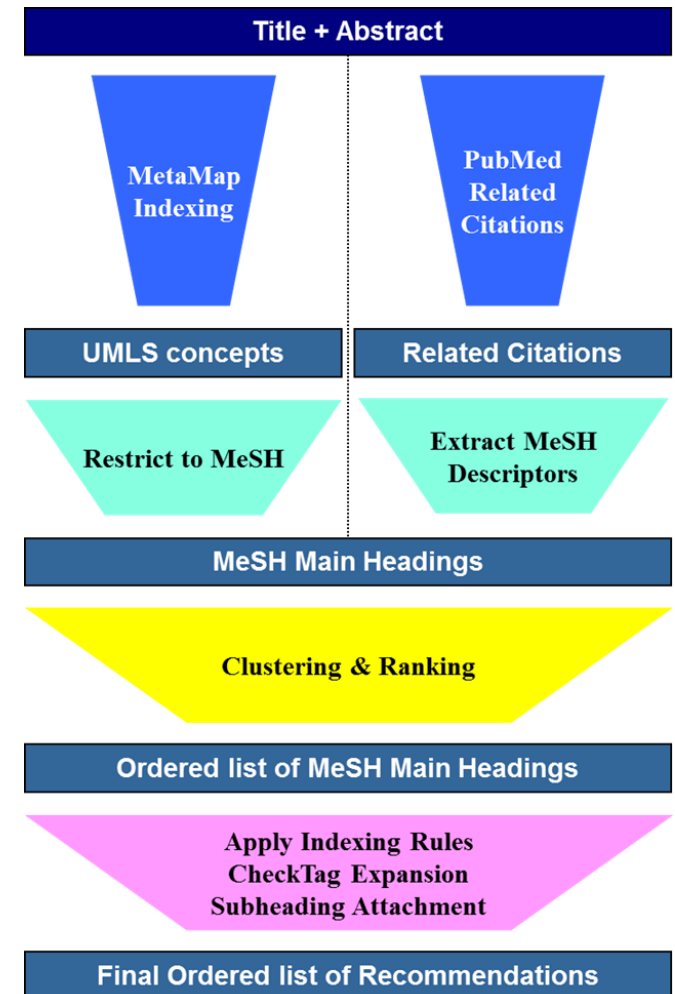
Example:

– C0022650: Kidney Calculi:

- calculi in bilateral kidneys; calculi in both kidneys; calculi in the left renal; calculi within the calices of the left kidney....

The NLM Medical Text Indexer (MTI)

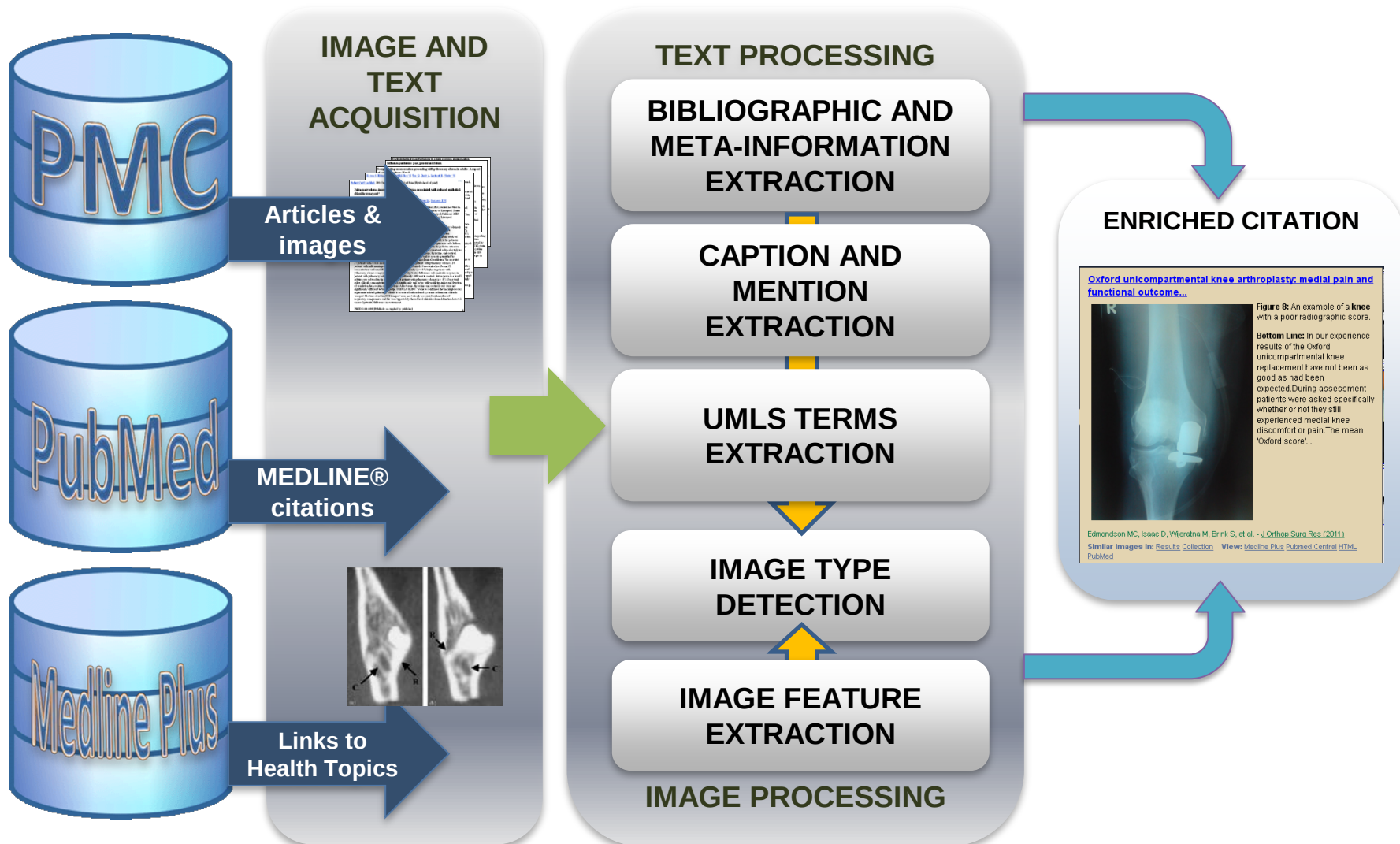
- Product of Indexing Initiative
- Principal developer James Mork
- Assists NLM Indexers
- In use since mid-2002
- Uses article Title and Abstract
- Semi-Automatic MeSH Indexing Recommendations
- Performs initial indexing for small subset of journals (MTIFL)





- **Combines UMLS-based**
 - Text processing
 - Information retrieval
- With image processing
- **To:**
 - Prepare citation
 - Index with search engine
- **Provides:**
 - User interface for searching biomedical literature
 - By text and/or example images
 - Search filters (image type, limits, etc.)

Preparing citations for indexing



Enriched citations

- XML for indexing
- <EnrichedCitation>
 - <Title>....
(other MEDLINE citation fields)
 - <ImageDocument>
 - <Modality>
 - <Caption>
 - <Mention>
 - <Panel>
 - <VisualWords>
 - <Modality>
 - <ROI>

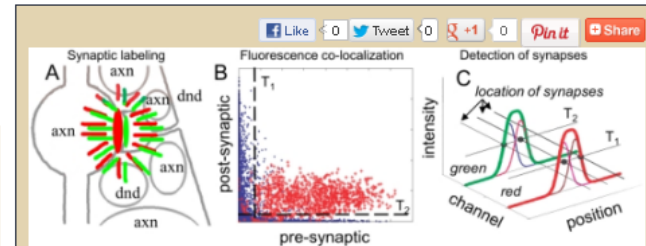
On Optical Detection of Densely Labeled Synapses in Neuropil and Mapping Connectivity with Combinatorially Multiplexed Fluorescent Synaptic Markers

Mishchenko Y · PLoS ONE (2010)

Bottom Line: We simulate fluorescence from a population of densely labeled synapses in a block of hippocampal neuropil, completely reconstructed from electron microscopy data, and show that high-end LM is able to detect such patterns with over 95% accuracy. We conclude, therefore, that with the described approach neural connectivity in macroscopically large neural circuits can be mapped with great accuracy, in scalable manner, using fast optical tools, and straightforward image processing. Relying on an electron microscopy dataset, we also derive and explicitly enumerate the conditions that should be met to allow synaptic connectivity studies with high-resolution optical tools.

Affiliation: Department of Statistics and Center for Theoretical Neuroscience, Columbia University, New York, New York, USA. yuriy.mishchenko@gmail.com

Abstract: We propose a new method for mapping neural connectivity optically, by utilizing Cre/Lox system Brainbow to tag synapses of different neurons with random mixtures of different fluorophores, such as GFP, YFP, etc., and then detecting patterns of fluorophores at different synapses using light microscopy (LM). Such patterns will immediately report the pre- and post-synaptic cells at each synaptic connection, without tracing neural projections from individual synapses to corresponding cell bodies. We simulate fluorescence from a population of densely labeled synapses in a block of hippocampal neuropil, completely reconstructed from electron microscopy data, and show that high-end LM is able to detect such patterns with over 95% accuracy. We conclude, therefore, that with the described approach neural connectivity in macroscopically large neural circuits can be mapped with great accuracy, in scalable manner, using fast optical tools, and straightforward image processing. Relying on an electron microscopy dataset, we also derive and explicitly enumerate the conditions that should be met to allow synaptic connectivity studies with high-resolution optical tools.



pone-0008853-g002: Schematic explanation of synapse detection using co-localization of fluorescence from different pre- and post-synaptic markers. A) Schematic diagram of the synaptic Brainbow, with a red fluorophore on the pre-synaptic side and a green fluorophore on the post-synaptic side of a synaptic cleft. Spatial correlation of the fluorescence from the pre- and post-synaptic fluorophores, occurring due to their proximity across synaptic cleft, allows detecting synapses optically without explicitly resolving them. Due to absence of the fluorophores in the bulk of the axonal and dendritic cytoplasm, nearby processes do not interfere with the detection process even when all neurons are labeled, unlike in regular Brainbow. B) Due to close spatial co-localization of the pre- and post-synaptic fluorophores across the synaptic cleft, their fluorescence intensity is closely correlated near labeled synapses. In this figure we show a simulated scatter plot of the fluorescence intensity in IDLM. Blue dots represent voxels far away from one labeled synapse (further than ≈ 200 nm), and red dots represent voxels closer than ≈ 200 nm. One can threshold this diagram with certain thresholds, T1 for the pre-synaptic marker and T2 for the post-synaptic marker (dashed lines), in order to separate the proximal (red) from distant (blue) voxels, and thus detect presence of a synapse. C) Using correlations in the fluorescence from the pre- and post-synaptic markers, synapses may be detected even when they cannot be explicitly resolved into isolated puncta. Illustrated here are three "synapses", fluorescence from which individually is shown with thin blue, magenta and brown lines. These are observed using two fluorescent markers, green and red. First synapse is tagged only with "green" marker, second synapse is tagged with "green" and "red" markers, and third synapse is tagged with "red" marker. Combined fluorescence from these synapses is shown with thick red and green lines, for the two markers respectively. Even though none of synapses can be seen separately in either green or red channels, by thresholding fluorescence with appropriate thresholds, T1 and T2, three different supra-threshold fluorescence patterns (black dots) indicate presence of three synapses.

Mentions: While one can detect synapses with LM by looking for explicitly isolated fluorescent puncta, one can also use a more powerful, yet simpler, prescription for detecting synapses implicitly. Specifically, consider a synapse labeled with two fluorophores, a fluorophore AFP on the pre-synaptic side and a fluorophore BFP on the post-synaptic side. Because of the spatial proximity of these two fluorophores across the synaptic cleft (i.e., ~ 10 – 50 nm apart), the fluorescence from these fluorophores will be tightly correlated in the region near labeled synapse, Figure 2A and 2B. This correlation may be quantified and used to detect the synapse even when it cannot be resolved as an isolated punctum, Figure 2C.

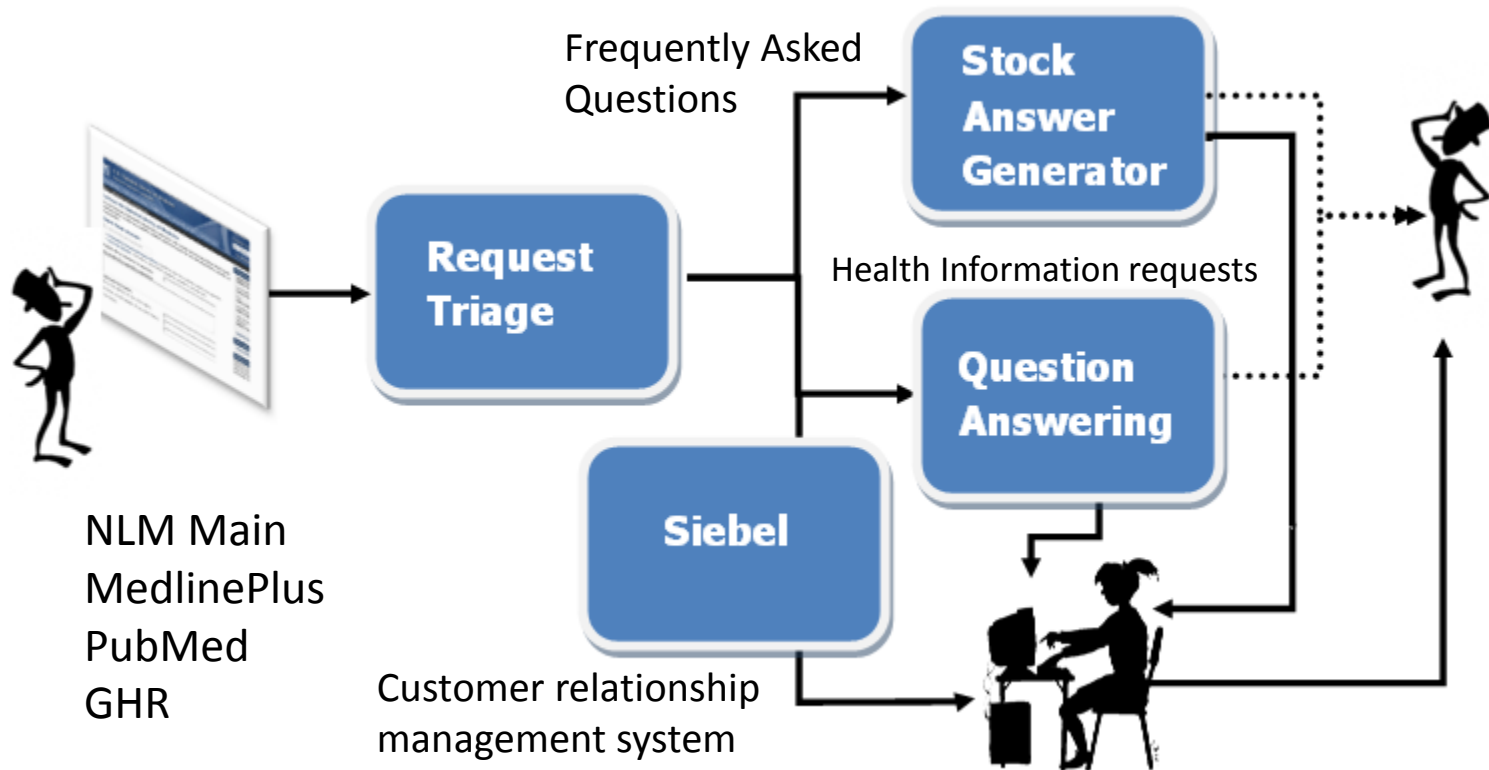
Open-i User Interface Functions

- Images as search input
- LinkOut resources
- Multiple views
- Filtering by image modality, journal type, clinical specialties
- Query expansion via UMLS synonymy *
- Re-rank according to specific task *
- Phrase search and complex queries *

* Essie indexing supports these functions. Essie is a search engine developed by Lister Hill Center for ClinicalTrials.gov

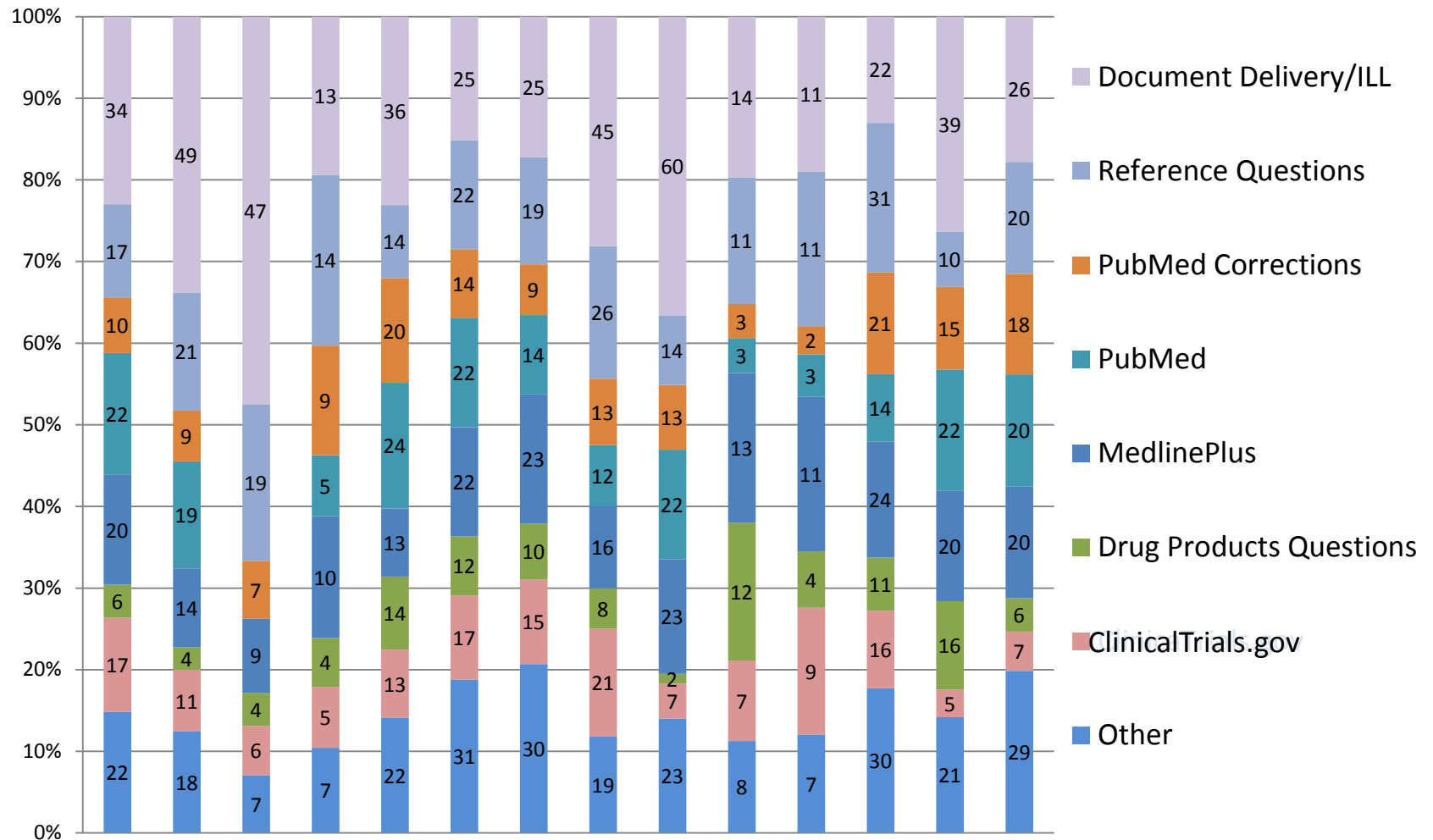
CASE STUDY: SUPPORTING NLM CUSTOMER SERVICES

Request Processing System



Incoming customers requests

Average 131.5 per day



Frequent requests

- PubMed corrections (about 20-25 a day)
 - In this pubmed reference:
<http://www.ncbi.nlm.nih.gov/pubmed/23653210> there is a typo in the title from the pubmed database. The title should be "N6-formyladenosine" instead of "N4-formyladenosine". Thanks!
- Consumer health questions (about 15-20 a day)
 - My sister is diabetic since age 8, now age 20. she is not coping well with it, i've been to the nutrition DR.s with her and we have since change our diet at home just for her, but it's still too havy on her. i want to help her, maybe if i can get her a couselector or support group that understands her situation/diabits

Request triage

- Classify incoming requests as PubMed corrections, Consumer health questions (limited to disorders), other
 - MaxEnt, SVM trained on 100K 2011 requests, 80 - 85% F-score
 - Ongoing / Future work
 - More classes: Junk; Other types of CHQ (e.g., medications); UMLS
 - Subclasses: Find a Doctor or Expert for a given disorder
 - Classes and distributions are changing
- PubMed Corrections Assistant:
 - Prepares stock replies for PubMed correction requests
- Question Answering module:
 - Prepares answers for consumer health questions

PubMed Corrections Assistant

- Searches the request for PMID(s) or reference(s)
- Retrieves PubMed citations using E-Utilities
- Extracts the citation status field:
 - PubMed
 - PubMed - in process (now undergoing changes)
 - PubMed - as supplied by Publisher
 - PubMed - indexed for MEDLINE
- Using the status, finds an appropriate stock reply
- Adds the stock reply to the request
- Sends the augmented request to customer service agent for approval/editing

CRC answers in Siebel

File Edit View Navigate Query Tools Help

SR #:

[Home](#) [Service Request](#) [Search SR](#) [Activities](#) [Change Requests](#)

[Service Requests List](#) | [Service Request Charts](#)

Service Requests | [Menu](#) [New](#) [Query](#) [Reply SR](#) [Forward SR](#)

SR #	Priority	SR Opened	Status	Summary	Description
> 1-56143145	CRC	4/1/2014 01:48:00 PM	Open	medical ear debriefment	SUBJECT: medical ear deb
1-56141205	Low	4/1/2014 01:45:53 PM	Open	set up a let	Hello I need your great as
1-56143125	CRC	4/1/2014 01:44:53 PM	Open	Incorrect year on citation	PMID 22375642 has the y
1-56141185	Low	4/1/2014 01:40:54 PM	Open	dbGaP Best Practices Requirements	Hello, I am not trying to b
1-56142475	Low	4/1/2014 01:36:20 PM	Open	downtime	SUBJECT: downtime EMAI
1-56143105	Low	4/1/2014 01:29:35 PM	Open	bowl movements question.	SUBJECT: bowl movement
1-56141162	Low	4/1/2014 01:24:13 PM	Open	urgente, 8 dias sin evacuar heces	SUBJECT: urgente, 8 dias :
1-56136085	Low	4/1/2014 01:19:33 PM	Open	duplicate MyNCBI accounts	Somehow I have a second
1-56141142	Low	4/1/2014 01:04:13 PM	Open	MedlinePlus Spanish Service Request	SUBJECT: EMAIL: NAME:
1-56136065	CRC	4/1/2014 01:01:37 PM	Open	hypothyroidism	SUBJECT: hypothyroidism

Unexpected consequences:

New PubMed corrections request form

Menu item	Look up	Citation field value	Form Action(s) Before sending to Siebel	Siebel XML tag AREA	Siebel XML tag SUBAREA	Siebel XML tag Group
Misspelling or error in PubMed	PubMed Record Status	MEDLINE	Add to SRDescriptionNotes: -current text/-correction text -PMID -Summary (text) citation	Quality Control of NLM DB	MEDLINE	Indexing
		Lookup CitationSubset For CitationSubset == OM	Same action as above	Quality Control of NLM DB	MEDLINE	MMS
Misspelling or error in PubMed	PubMed Record Status	In-Process OR In-data-review	Add to SRDescriptionNotes: -current text/-correction text -PMID -Summary (text) citation	Quality Control of NLM DB	In Process	Indexing
Misspelling or error in PubMed	PubMed Record Status	Publisher	Add to SRDescriptionNotes: -PMID -Summary (text) citation	Quality Control of NLM DB	Publisher Supplied	Customer Service
Misspelling or error in PubMed	PubMed Record Status	OLDMEDLINE	Add to SRDescriptionNotes: -current text/-correction text -PMID -Summary (text) citation	Quality Control of NLM DB	MEDLINE	MMS
Misspelling or error in PubMed	PubMed Record Status	PubMed-not-MEDLINE	Add to SRDescriptionNotes: -current text/-correction text -PMID -Summary (text) citation	Quality Control of NLM DB	PubMed Only	Indexing
Question about MeSH subject indexing	N/A	N/A	Add to SRDescriptionNotes: -PMID -Summary (text) citation	Quality Control of NLM DB	Assigned MeSH Headings	Indexing
Copy of Full Text	N/A	N/A	Add to SRDescriptionNotes: -PMID -Summary (text) citation	PubMed	Full-Text/Obtaining Articles	Customer Service
Forgot MyNCBI Password	N/A	N/A	Send question to Siebel Forward for action to info@ncbi.nlm.nih.gov	PubMed	MyNCBI	Customer Service

Question Answering Assistant

- Represents question in a structured form (frame) – question understanding
- Uses the frame to generate queries to search engines
- Retrieves articles
- Extracts answers
- Ranks answers

Understanding consumer questions is hard

- Different styles
 - Education level
 - Native language
- Mostly informal language
 - Ungrammatical sentences
 - Inconsistent punctuation and capitalization
 - Abbreviations
 - Misspellings
 - Extraneous information
 - Abundance of anaphora and ellipses

Question understanding: question type

Question type	Question template(s)	All questions could specify context
Information	Information about <i>Disease</i>	
Person/Organization	Geo / internet location of organization / specialist for (testing treatment support) for <i>Disease</i>	
Anatomy	Body part/function affected by <i>Disease</i>	
Management	[Effects of <i>Drug/Procedure</i> as] treatment/prevention for <i>Disease</i>	
Diagnosis	[<i>Procedure</i> as] diagnosis for <i>Disease</i>	
Cause	Cause of <i>Disease</i>	
Susceptibility	Population affected by <i>Disease</i> ; Inheritance patterns	
Complications	Problems caused by <i>Disease</i>	
Prognosis	Outcome/life expectancy/quality of life [in patient] with <i>Disease</i>	
Other Effects	effects of a <i>Disease</i> that are not explicitly Complications or Manifestations	
Manifestations	Signs and symptoms of <i>Disease</i>	

Rule-based question understanding

- Find terms corresponding to entities of interest:
 - UMLS terms corresponding to Diseases, Anatomy, Drugs, Genes
 - Gazetteers for question cues & question type indicators
- Use syntactic patterns that establish relations between terms
- Translate relations to structured form (frames)
- Use methods for anaphora and ellipsis resolution

Example question frames

Are there treatments for trisomy 13?

What is the prognosis?

What is life like for those who have this condition?

- Frame 1

- Question type: Management
 - Indicated by: treatments
- Theme (problem): trisomy 13
- Question cue: are there

- Frame 2

- Question type: Prognosis
 - Indicated by: prognosis
- Theme (problem): trisomy 13
 - Ellipsis resolution
- Question cue: what

- ❖ Frame 3

- ❖ Question type: Manifestations

- ❖ Indicated by: life like

- ❖ Theme (problem): trisomy 13

- ❖ Anaphora resolution

- ❖ this condition → trisomy 13

- ❖ Question cue: what

Frame construction

- Type and Theme argument are mandatory
- Stanford dependencies to identify theme and question cue arguments
 - direct object, passive nominal subject, noun compound modifier, relative clause modifier, etc.
- Transitivity on Information triggers
 - Is there **information** regarding **prognosis** of . . .
- Transitivity on population group terms
 - What is the **life expectancy** for a **child** with **trisomy 13**? . . .

Answer sources

The image displays two side-by-side screenshots of web browsers, both running Firefox. The left browser window shows the MedlinePlus website, and the right browser window shows the Genetics Home Reference website.

Left Browser Window (MedlinePlus):

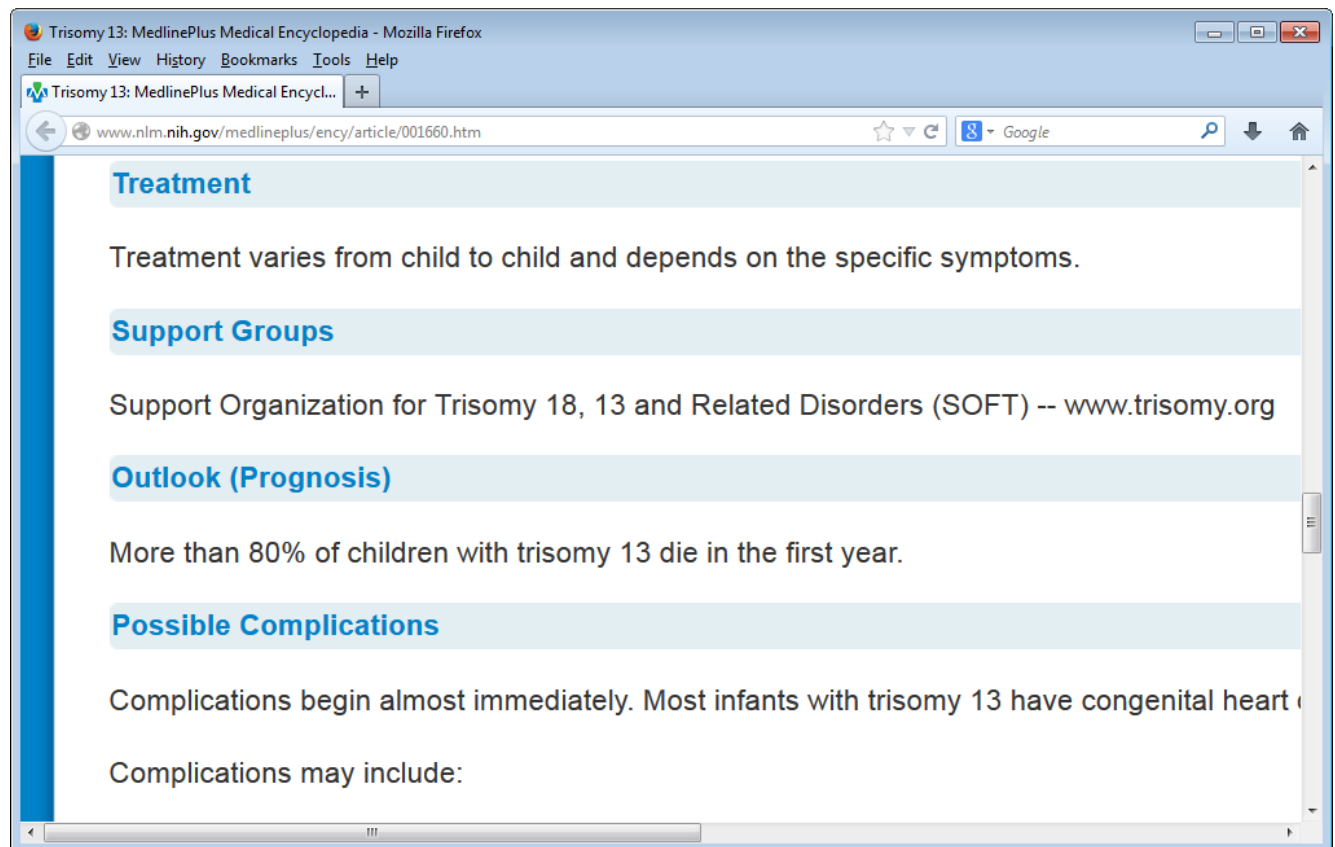
- Address bar: Trisomy 13: MedlinePlus Medical Encycl...
- Logo: MedlinePlus Trusted Health Information for You
- Navigation links: Home, About MedlinePlus, Site Map, FAQs, Contact Us
- Buttons: Health Topics, Drugs & Diseases
- Section: **Trisomy 13**
- Text: Trisomy 13 (also called Patau syndrome) is a genetic disorder... material from [chromosome](#) 13, instead of the usual two copies... another chromosome (translocation).
- Section: **Causes**
- Text: Trisomy 13 occurs when extra DNA from chromosome 13 ap...
- List:
 - Trisomy 13 -- the presence of an extra (third) chromos
 - Trisomy 13 mosaicism -- the presence of an extra chro
 - Partial trisomy -- the presence of a part of an extra chr
- Text: The extra material interferes with normal development.
- Text: Trisomy 13 occurs in about 1 out of every 10,000 newborns. (inherited). Instead, the events that lead to trisomy 13 occur i
- Section: **Symptoms**
- List:
 - [Cleft lip or palate](#)
 - Clenched hands (with outer fingers on top of the inner
 - Close-set eyes -- eyes may actually fuse together into
 - Decreased muscle tone
 - Extra fingers or toes ([polydactyly](#))
 - Hernias: [umbilical hernia](#), [inguinal hernia](#)
 - Hole, split, or cleft in the iris ([coloboma](#))

Right Browser Window (Genetics Home Reference):

- Address bar: Trisomy 13 - Genetics Home Reference
- Logo: Genetics Home Reference Your Guide to Understanding Genetic Conditions
- Navigation links: About, Site Map, Contact Us
- Search bar: Search
- Text: A service of the U.S. National Library of Medicine®
- Navigation links: Home, Conditions, Genes, Chromosomes, Handbook, Glossary, Resources
- Section: **Trisomy 13**
- Text: On this page: [Description](#), [Genetic changes](#), [Inheritance](#), [Diagnosis](#), [Additional information](#), [Other names](#), [Glossary definitions](#)
- Text: Reviewed January 2009
- Section: **What is trisomy 13?**
- Text: Trisomy 13, also called Patau syndrome, is a chromosomal condition associated with severe intellectual disability and physical abnormalities in many parts of the body. Individuals with trisomy 13 often have heart defects, brain or spinal cord abnormalities, very small or poorly developed eyes (microphthalmia), extra fingers and/or toes, an opening in the lip (a cleft lip) with or without an opening in the roof of the mouth (a cleft palate), and weak muscle tone (hypotonia). Due to the presence of several life-threatening medical problems, many infants with trisomy 13 die within their first days or weeks of life. Only five percent to 10 percent of children with this condition live past their first year.
- Section: **How common is trisomy 13?**
- Text: Trisomy 13 occurs in about 1 in 16,000 newborns. Although women of any age can have a child with trisomy 13, the chance of having a child with this condition increases as a woman gets older.
- Section: **What are the genetic changes related to trisomy 13?**
- Text: Most cases of trisomy 13 result from having three copies of chromosome 13 in each cell in the body instead of the usual two copies. The extra genetic material disrupts the normal course of development, causing the characteristic features of trisomy 13.
- Text: Trisomy 13 can also occur when part of chromosome 13 becomes attached (translocated) to another chromosome during the formation of reproductive cells (eggs and sperm) or very early in fetal development. Affected people have two normal copies of chromosome 13, plus an extra copy of chromosome 13 attached to another chromosome. In rare cases, only part of chromosome 13 is present in three copies. The physical signs and symptoms in these cases may be different than those found in full trisomy 13.
- Text: A small percentage of people with trisomy 13 have an extra copy of chromosome 13 in only some of the body's cells. In these people, the condition is called mosaic trisomy 13. The severity of mosaic trisomy 13 depends on the type and number of cells that have the extra chromosome. The physical features of mosaic trisomy 13 are often milder than those of full trisomy 13.

Translating structured questions to search queries

- Search for the theme (problem) in titles
 - If found, return an appropriate section for question type



Selecting the best answer

- Rank articles and sections
- From the best sections, select paragraphs containing more answer type indicators
 - trim long paragraphs
- Due to the presence of several **life-threatening** medical problems, many infants with trisomy 13 **die** within their first days or weeks of life. Only five percent to 10 percent of children with this condition **live past** their first year.
- More than 80% of children with trisomy 13 **die** in the first year.

Unexpected consequences:

Quality assurance for web resources

- **GHR**

- Added a new Handbook page on genetic disorder prognosis:

- <http://ghr.nlm.nih.gov/handbook/consult/prognosis>

- Added AddThis links

External link disclaimer



- **MedlinePlus**

- Revised topics:

- Tuberculosis testing
 - Shingles
 - Tubal ligation ...

Anaphora, Coreference, Bridging Inferences, Ellipsis...

- my question is this: I was born w/a esophagus atresia w/dextrocardia. While the heart hasn't caused problems, the other has [caused problems]. I get food caught all the time. My question is... is there anything that can fix it cause I can't eat anything lately without getting it caught. I need help or will starve!
- Pronominal anaphora
 - it → esophagus atresia
 - it → food
- Other-anaphora (Comparative anaphora)
 - the other → esophagus atresia
- Ellipsis
 - Omission of caused problems: VP ellipsis

Anaphora resolution for frame extraction

	# of frames	Recall	Precision	F1-score
Baseline	14	0.32	0.93	0.48
Anaphora	26	0.54	0.85	0.66

- ❖ Dependency-based frame extraction with simple dictionary lookup provides good precision but low recall
- ❖ Anaphora/ellipsis resolution helps significantly in recovering relevant frames with little negative effect on precision
- ❖ Better term identification → better anaphora/ellipsis resolution → better question understanding

Ongoing & future work

- Expand knowledge base
 - Link to relevant ClinicalTrials.gov
 - Mine NIH ICs, CDC, FDA consumer Web sites
 - Identify PubMed and PubMed Central articles for consumers
- Answer generation
- Understanding questions beyond “known disease”
- Addressing remaining request types