



Practical Bioinformatics (BCH 463)

Instructional Manual Lab (1)

Prepared by

Nora Aljebrin, Emtenan Alkhudair

NCBI Overview and Searching in PubMed Database

NCBI is the National Center for Biotechnology Information. It is part of the United States National Library of Medicine (NLM), a branch of the National Institutes of Health (NIH).

The NCBI houses a series of databases relevant to biotechnology and biomedicine and is an important resource for bioinformatics tools and services. Major databases are divided as the following:

NCBI databases

<div style="text-align: right; font-size: 2em; color: red; margin-bottom: 10px;">1</div> <div style="border: 1px solid gray; padding: 10px;"> <p>Literature</p> <p>The World's largest repository of medical and scientific abstracts, full-text articles, books and reports</p> <p>Bookshelf Books and reports</p> <p>MeSH Ontology used for PubMed indexing</p> <p>NLM Catalog Books, journals and more in the NLM Collections</p> <p>PubMed Scientific and medical abstracts/citations</p> <p>PubMed Central Full-text journal articles</p> </div>	<div style="text-align: right; font-size: 2em; color: red; margin-bottom: 10px;">2</div> <div style="border: 1px solid gray; padding: 10px;"> <p>Genes</p> <p>Gene sequences and annotations used as references for the study of orthologs structure, expression, and evolution</p> <p>Gene Collected information about gene loci</p> <p>GEO DataSets Functional genomics studies</p> <p>GEO Profiles Gene expression and molecular abundance profiles</p> <p>HomoloGene Homologous genes sets for selected organisms</p> <p>PopSet Sequence sets from phylogenetic and population studies</p> </div>	<div style="text-align: right; font-size: 2em; color: red; margin-bottom: 10px;">3</div> <div style="border: 1px solid gray; padding: 10px;"> <p>Proteins</p> <p>Protein sequences, 3-D structures, and tools for the study of functional protein domains and active sites</p> <p>Conserved Domains Conserved protein domains</p> <p>Identical Protein Groups Protein sequences grouped by identity</p> <p>Protein Protein sequences</p> <p>Protein Clusters Sequence similarity-based protein clusters</p> <p>Sparcle Functional categorization of proteins by domain architecture</p> <p>Structure Experimentally-determined biomolecular structures</p> </div>
<div style="text-align: right; font-size: 2em; color: red; margin-bottom: 10px;">4</div> <div style="border: 1px solid gray; padding: 10px;"> <p>Genomes</p> <p>Genome sequence assemblies, large-scale functional genomics data, and source biological samples</p> <p>Assembly Genome assembly information</p> <p>BioCollections Museum, herbaria, and other biorepository collections</p> <p>BioProject Biological projects providing data to NCBI</p> <p>BioSample Descriptions of biological source materials</p> <p>Genome Genome sequencing projects by organism</p> <p>Nucleotide DNA and RNA sequences</p> <p>SRA High-throughput sequence reads</p> <p>Taxonomy Taxonomic classification and nomenclature</p> </div>	<div style="text-align: right; font-size: 2em; color: red; margin-bottom: 10px;">5</div> <div style="border: 1px solid gray; padding: 10px;"> <p>Genetics</p> <p>Heritable DNA variations, associations with human pathologies, and clinical diagnostics and treatments</p> <p>ClinVar Human variations of clinical significance</p> <p>dbGaP Genotype/phenotype interaction studies</p> <p>dbSNP Short genetic variations</p> <p>dbVar Genome structural variation studies</p> <p>GTR Genetic testing registry</p> <p>MedGen Medical genetics literature and links</p> <p>OMIM Online mendelian inheritance in man</p> </div>	<div style="text-align: right; font-size: 2em; color: red; margin-bottom: 10px;">6</div> <div style="border: 1px solid gray; padding: 10px;"> <p>PubChem</p> <p>Repository of chemical information, molecular pathways, and tools for bioactivity screening</p> <p>BioAssays Bioactivity screening studies</p> <p>Compounds Chemical information with structures, information and links</p> <p>Pathways Molecular pathways with links to genes, proteins and chemicals</p> <p>Substances Deposited substance and chemical information</p> </div>

- ◆ To display how to use and search these databases, examples were provided to help guide you throughout the website.

e.g.1 Retrieve all records of **cancer**. Which database returns the highest number of hits?

Since all records are required, your search should be on **All Databases**. Each database will display its number of hits next to it. The highest number is the answer the question in *e.g.1*.

The screenshot shows the NCBI search results for the query 'cancer'. The search was performed on 'All Databases'. The results are categorized into 'Literature' and 'Genes'.

Category	Database	Number of Hits
Literature	Bookshelf	122,305
	MeSH	395
	NLM Catalog	50,904
	PubMed	4,145,627
	PubMed Central	1,841,540
	Genes	Gene
GEO DataSets		935,273
GEO Profiles		15,791,203
HomoloGene		227
PopSet		1,322

- ◆ How to limit the number of search results and get more accurate hits using **PubMed**:

1) Use the proper keywords.

e.g.2 How many articles published on **Arabian camel**? What happen if you search only for **camel**?

This example shows the decline in the number of hits by being more specific in your search

The first screenshot shows a search for 'arabian camel' on PubMed.gov, resulting in 4,562 hits. The second screenshot shows a search for 'camel' on PubMed.gov, resulting in 6,391 hits. This demonstrates that a more specific search yields fewer results.

2) Use the appropriate combining search terms with Boolean operators (NOT, AND, OR)

e.g.3 Retrieve all records of **cancer** *except* those from **BRCA1 mutations**.

e.g.4 How many articles published about **osteoporosis** related to **vitamin D**?

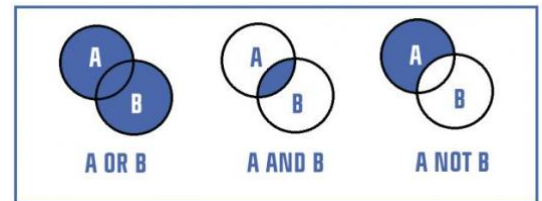
e.g.5 Find citations on the effects of **heat or humidity** on **multiple sclerosis**.

Examples 3- 5 display the importance of using conjunctions to combine or exclude keywords in a search, resulting in more focused and productive results. **So,**

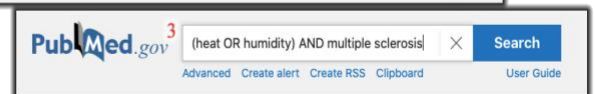
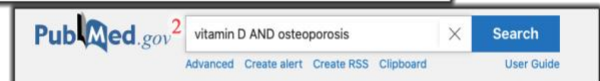
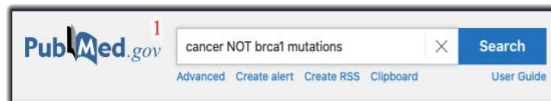
NOT → *narrows* your search by telling the database to *eliminate* all terms that follow it from your search results. *As seen in e.g.3*

AND → *narrows* a search by telling the database that **ALL** keywords used must be found in an article. *As seen in e.g.4*

OR → *broadens* a search by telling the database that any of the words it connects are acceptable. *As seen in e.g.5*



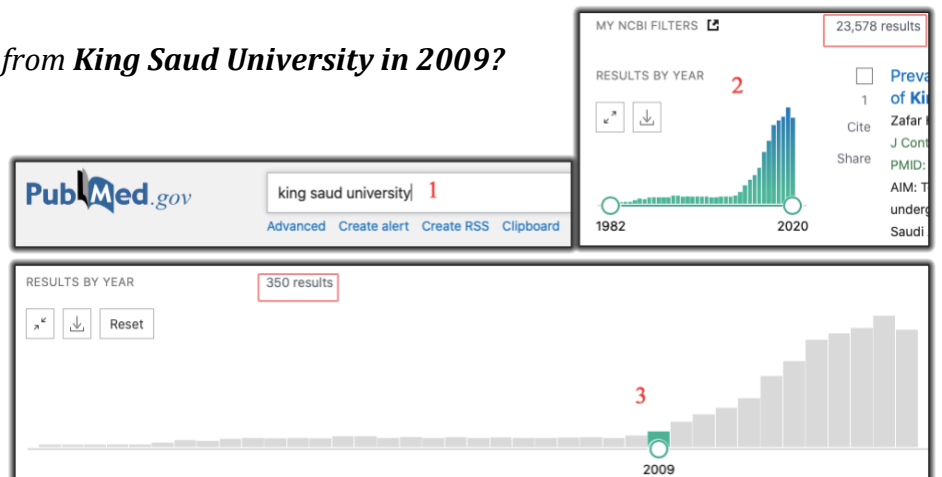
Notice that these terms should be in uppercase characters



3) Use the appropriate filters:

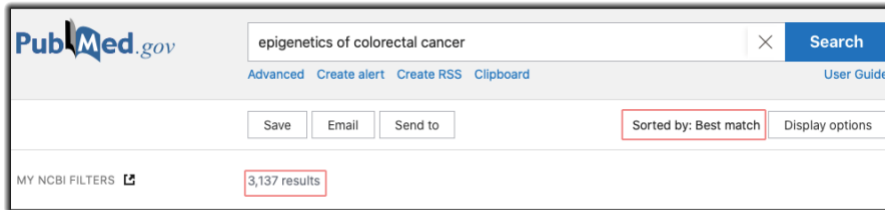
a) Search by years:

e.g.6 How many articles published from **King Saud University** in **2009**?



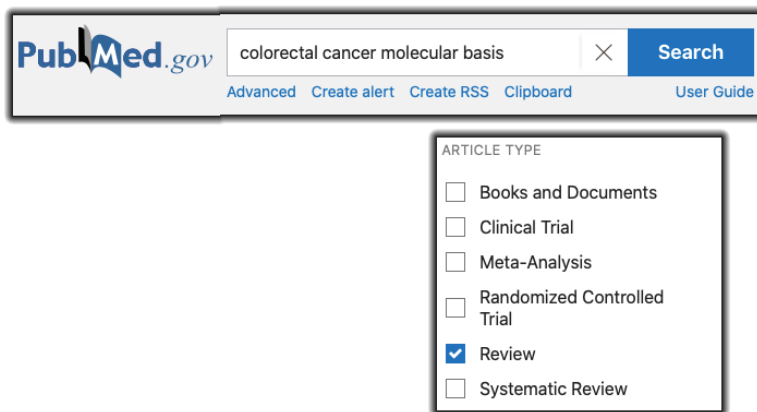
b) Sort your search results in the **Display options**. These include: **Most Recent, Best Match, Publication Date, First Author, and Journal**.

e.g.7 Search for all articles about *epigenetics of colorectal cancer*. Sort the articles by *Best match*.



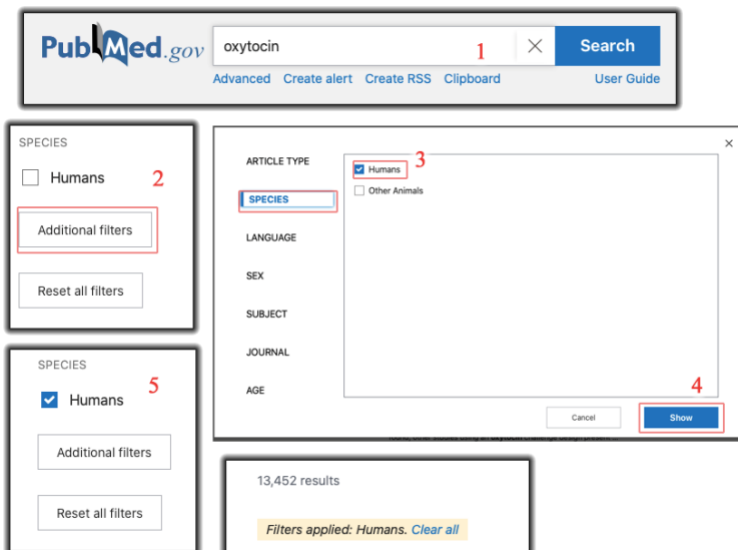
c) Search for a particular article type

e.g.8 Search for a *review* article about *colorectal cancer molecular basis*.



d) Search for a particular species

e.g.9 How many texts published on *human oxytocin*



◆ How to use **Advanced mode** in PubMed to search for:

a) Journal

e.g.10 How many articles published by *the journal of cell biology*

The screenshot shows the PubMed Advanced search interface. At the top, there is a search bar with the text "Search PubMed" and a "Search" button. Below the search bar, the "Advanced" mode is selected. The interface is divided into several sections:

- Left Section:** "Add terms to the query box" with a dropdown menu set to "All Fields" (labeled 2).
- Middle Section:** A list of search filters (labeled 3) including Editor, Filter, Grant Number, ISBN, Investigator, Issue, **Journal** (highlighted), Language, Location ID, MeSH Major Topic, MeSH Subheading, MeSH Terms, and Other Term.
- Right Section:** "Add terms to the query box" (labeled 4) with a dropdown menu set to "Journal" and a text input field containing "The Journal of cell biology".
- Bottom Right Section:** An "ADD" button (labeled 5) and a "Show Index" link.
- Bottom Section:** A "Query box" containing the search string "The Journal of cell biology"[Journal] and a "Search" button (labeled 6).

b) Author

e.g.11 Search for all published articles written by *Arjumand Warsy*.

Notice that the author's last name is written first followed by a (,) then the first name.

The screenshot shows the PubMed Advanced search interface for an author search. At the top, there is a search bar with the text "Search PubMed" and a "Search" button. Below the search bar, the "Advanced" mode is selected. The interface is divided into several sections:

- Left Section:** "Add terms to the query box" with a dropdown menu set to "All Fields" (labeled 2).
- Middle Section:** A list of search filters (labeled 3) including Affiliation, **Author** (highlighted), Author - Corporate, Author - First, Author - Identifier, Author - Last, Book, Conflict of Interest Statements, Date - Completion, Date - Create, and Date - Entry.
- Right Section:** "Add terms to the query box" (labeled 4) with a dropdown menu set to "Author" and a text input field containing "Warsy, Arjumand".
- Bottom Right Section:** An "ADD" button (labeled 5) and a "Show Index" link.
- Bottom Section:** A "Query box" containing the search string "Warsy, Arjumand[Author] and a "Search" button (labeled 6).

c) Combined advanced search (Author, Journal, and Title)

e.g.12 There is an article in *the New England Journal of Medicine* on *Williams-Beuren syndrome* written by *Dr. Barbara Pober*. Construct a **specific search** on PubMed to find this article.

1 Add terms to the query box Author Pober, Barbara 2 ADD Show Index

3 Add terms to the query box Journal The New England journal of medicine 4 AND Show Index

5 Add terms to the query box Title Williams-Beuren syndrome 6 AND Show Index

7 Query box ((Pober, Barbara[Author]) AND ("The New England journal of medicine"[Journal])) AND (Williams-Beuren syndrome[Title]) Search

Important notes regarding the search filters:

- 1) Clear all your filters each time you construct a new search.
- 2) You can find your PubMed search history at the bottom of the advanced page

1 Search PubMed Search User Guide

2 History and Search Details Download Delete

Search	Actions	Details	Query	Results	Time
#1	...	>	Search: ((Pober, Barbara[Author]) AND ("The New England journal of medicine"[Journal])) AND (Williams-Beuren syndrome[Title])	1	13:24:48

◆ How to obtain **Free** copies of some articles, either by:

a) Searching in text availability in the sidebar filter for free full text

e.g.13 How many free texts published on *oxytocin*?

1 Search oxytocin Search User Guide

2 TEXT AVAILABILITY Abstract Free full text Full text

3 6,264 results Filters applied: Free full text. Clear all

1 Oxytocin and social functioning. Jones C, Barrera I, Brothers S, Ring R, Wahlestedt C. Dialogues Clin Neurosci. 2017 Jun;19(2):193-201. PMID: 28867943 Free PMC article. Review.

2 Oxytocin, motivation and the role of dopamine. Love TM. Pharmacol Biochem Behav. 2014 Apr;119:49-60. doi: 10.1016/j.pbb.2014.02.001. PMID: 23850525 Free PMC article. Review.

b) PubMed Central (PMC)

PMC is the U.S. National Institutes of Health (NIH) *free digital archive* of biomedical and life sciences journal literature.

e.g.14 Search for a PMC article about *hepatitis* and **download it as pdf.**

The screenshot shows the PubMed Central search interface. At the top left, the NCBI logo is visible. A dropdown menu labeled 'All Databases' is highlighted with a red box and the number '1'. To its right, another dropdown menu is open, showing 'PMC' selected with a checkmark, also highlighted with a red box and the number '2'. Below these, a search bar contains the text 'hepatitis' and a 'Search' button is highlighted with a red box and the number '3'. The search results section shows 'Items: 1 to 20 of 280430' and a list of results. The first result is highlighted with a red box and the number '4', showing the title 'Hepatitis A virus infection and hepatitis A vaccination in human immunodeficiency virus-positive patients: A review' and the authors 'Kuan-Yin Lin, Guan-Jhou Chen, Yu-Lin Lee, Yi-Chia Huang, Aristine Cheng, Hsin-Yun Sun, Sui-Yuan Chang, Chun-Eng Liu, Chien-Ching Hung'. Below the title, the journal information 'World J Gastroenterol. 2017 May 28; 23(20): 3589–3606. Published online 2017 May 28. doi: 10.3748/wjg.v23.i20.3589' and the PMCID 'PMC5449416' are shown. At the bottom, the 'Formats' section is visible, with 'PDF (1.5M)' highlighted with a red box and the number '4'.

Note: Similar articles in PMC are found on the right side of the article page.

The screenshot shows the full article page for 'Hepatitis A virus infection and hepatitis A vaccination in human immunodeficiency virus-positive patients: A review'. The title and authors are listed at the top. Below the title, there is a section for 'Similar articles in PubMed' which is highlighted with a red box. This section lists several related articles with their titles and publication years, such as 'Serological responses to revaccination with hepatitis A virus (HAV) vaccines among HIV-positive individuals wh [Liver Int. 2018]', 'Comparative effectiveness of two doses versus three doses of hepatitis A vaccine in human immunodeficiency [Hepatology. 2013]', 'Incidence of acute hepatitis A among HIV-positive patients during an outbreak among MSM in Taiwan: Impact of HA [Liver Int. 2018]', 'Hepatitis A and B immunizations of individuals infected with human immunodeficiency virus. [Am J Med. 2005]', and 'Challenges in Management of Hepatitis A Virus Epidemiological Transition in Mexico. [Ann Hepatol. 2019]'. The main article's abstract is also visible on the left side of the page.

◆ **How to build a PubMed search from MeSH:**

MeSH (Medical Subject Headings) is the NLM controlled vocabulary thesaurus used for **indexing** PubMed citations. Skilled subject analysts examine journal articles and assign to each the most specific MeSH term applicable. This ensures that articles are uniformly indexed by subject.

MeSH terms are arranged hierarchically by subject categories with more specific terms arranged beneath broader terms. *PubMed allows you to view this hierarchy and select terms for searching in the MeSH Database.*

e.g.15 Using MeSH term, search for *immunology in cancer*.

- 1) In the search box, write the heading (main topic) → in this example “cancer”.
- 2) Choose the most similar **heading** to your topic → in this case “Neoplasm”
- 3) A page with more specific **subheadings** will appear → choose “immunology”
- 4) Add to search builder then search the PubMed.

NCBI National Center for Biotechnology Information

All Databases

NCBI National Center for Biotechnology Information

MedGen
 MeSH
 NCBI Web Site
 NLM Catalog

MeSH cancer Search

Neoplasms

1. New abnormal growth of tissue. Malignant **neoplasms** show a greater degree of anaplasia and have the properties of invasion and metastasis, compared to benign **neoplasms**.
 Year introduced: /diagnosis was NEOPLASM DIAGNOSIS 1964-1965

antagonists and inhibitors
 biosynthesis
 blood
 blood supply
 cerebrospinal fluid
 chemical synthesis
 chemically induced
 ethnology
 etiology
 genetics
 growth and development
 history
 immunology
 injuries

PubMed Search Builder

Add to search builder AND

Search PubMed

PubMed Search Builder

"Neoplasms/immunology" [Mesh]

Add to search builder AND

Search PubMed

- **How to save and remove results to/from clipboard**

The Clipboard provides a place to collect selected citations from one or several searches. The Clipboard will be lost after 8 hours of inactivity on PubMed or other NCBI databases. The maximum number of items that can be sent to the clipboard is 500. Your web browser must be set to accept cookies to use the clipboard.

e.g.16 Social anxiety is a form of anxiety characterized by continuous fear of one or more social or performance situations. Although multiple treatment modalities exist for social anxiety, they are effective for only 60% to 70% of patients. Could it be possible that oxytocin can be a potential therapeutic option for individuals with social anxiety?? Search **PubMed** for evidence.

Check Oxytocin and social functioning.

*Send one of the articles to the **clipboard**. Remove the article from the **clipboard**.*

Search results

Save Email Send to

Review > Dialogues Clin Neurosci. 2017 Jun;19(2):193-201.

Oxytocin and social functioning

Candace Jones¹, Ingrid Barrera², Shaun Brothers², Robert Ring³, Claes Wahlestedt²

Clipboard
My Bibliography
Collections
Citation manager

1 item added to the Clipboard. Go to the Clipboard page.

oxytocin 1
Advanced Clipboard

Clipboard 1 result x Remove all 1 item selected x Remove selected items from Clipboard 3

Your clipboard expires after eight hours of inactivity.

2

1. PMID: 28867943

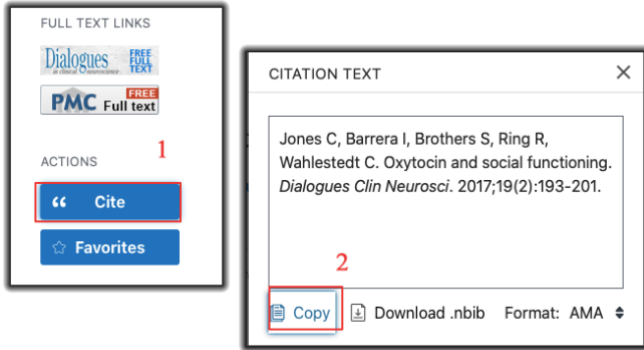
Oxytocin and social functioning.
Jones C, Barrera I, Brothers S, Ring R, Wahlestedt C.
Dialogues Clin Neurosci. 2017 Jun;19(2):193-201.
PMID: 28867943 **Free PMC article.** Review.

Remove from clipboard Cite Share

- **How to manage article citation:**

All the pictures below are applied on the same article of the previous example

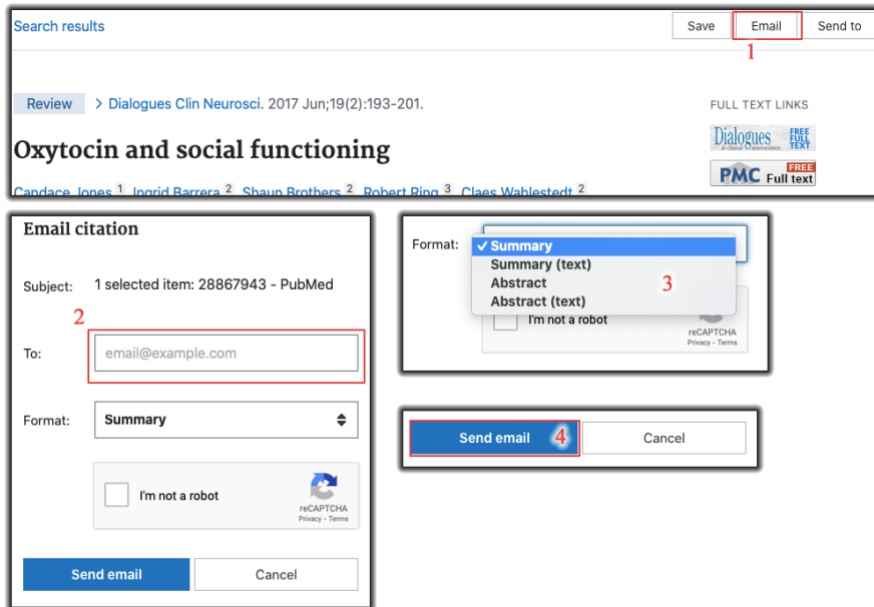
a) By copying citations



b) By exporting citations into citation management software such as Endnote (“.nbib” format)



c) By e-mailing citations



d) By saving citations as a text file in different formats

The screenshot shows a citation tool interface. At the top, there is a 'Formats:' section with a red '1' above it. The options are: Article, PubReader, ePub (beta), PDF (1.5M), and Citation (which is highlighted with a red box). Below this is a larger window with a red '2' above it. It contains the same 'Formats:' section and a table of citation formats for a specific article. The table lists the citation in AMA, MLA, APA, and NLM styles. At the bottom of the window, there is a 'Download as:' section with three options: RIS (highlighted with a red box), NBIB, and JSON.

Formats: 1
Article | PubReader | ePub (beta) | PDF (1.5M) | **Citation**

Formats: 2
Article | PubReader | ePub (beta) | PDF (1.5M) | Citation

Copy and paste a formatted citation from below or use one of the hyperlinks at the bottom to download a file for import into a bibliography manager.

AMA	Lin KY, Chen GJ, Lee YL, et al. Hepatitis A virus infection and hepatitis A vaccination in human immunodeficiency virus-positive patients: A review. <i>World J Gastroenterol.</i> 2017;23(20):3589-3606. doi:10.3748/wjg.v23.i20.3589
MLA	Lin, Kuan-Yin et al. "Hepatitis A virus infection and hepatitis A vaccination in human immunodeficiency virus-positive patients: A review." <i>World journal of gastroenterology</i> vol. 23,20 (2017): 3589-3606. doi:10.3748/wjg.v23.i20.3589
APA	Lin, K. Y., Chen, G. J., Lee, Y. L., Huang, Y. C., Cheng, A., Sun, H. Y., Chang, S. Y., Liu, C. E., & Hung, C. C. (2017). Hepatitis A virus infection and hepatitis A vaccination in human immunodeficiency virus-positive patients: A review. <i>World journal of gastroenterology</i> , 23(20), 3589–3606. https://doi.org/10.3748/wjg.v23.i20.3589
NLM	Lin KY, Chen GJ, Lee YL, Huang YC, Cheng A, Sun HY, Chang SY, Liu CE, Hung CC. Hepatitis A virus infection and hepatitis A vaccination in human immunodeficiency virus-positive patients: A review. <i>World J Gastroenterol.</i> 2017 May 28;23(20):3589-3606. doi: 10.3748/wjg.v23.i20.3589. PMID: 28611512; PMCID: PMC5449416.

Download as: **RIS** | NBIB | JSON

Additional information:

- APA (American Psychological Association) is used by Education, Psychology, and Sciences
- Types of journal article <https://www.springer.com/gp/authors-editors/authorandreviewertutorials/writing-a-journal-manuscript/types-of-journal-articles/10285504>
- Types of scientific evidence <https://www.sciencemediacentre.co.nz/coveringscience/types-of-scientific-evidence/>
- How to Order Author Names and Why That Matters <https://wordvice.com/journal-article-author-order/>