Finding systematic evidence: search and scan principles

Zhengting (Johnathan) He, MBBS.

2021/4/6

# 1. Why do we need to search literature systematically?

## 1.1 Why do we need to search literature systematically?

* I think it is similar to the question, “why do we need to conduct a systematic review?”, since searching and scanning are the very first steps after you select a topic for a systematic review.
* Why do we need to conduct a systematic review?
  + A simple answer is, **evidence is overwhelming but inconsistent**
* We are not conducting a systematic review, why do we need to systematically review literature in our studies?
  + Ensure not to duplicate the work of others
  + Gain insights from other’s work

## 1.2 What is a Systematic Review?

* A systematic review:
  + Focuses on a specific question
  + Uses explicit, pre-planned scientific methods to **identify, select, appraise, and summarize** similar but separate studies
* Cochrane handbook define systematic review as:
  + Systematic reviews seek to **collate evidence that fits pre-specified eligibility criteria** in order to answer **a specific research question**. They aim to **minimize bias** by using explicit, systematic methods documented in advance with a protocol.

## 1.3 Types of reviews

* All reviews
  + Reviews that are not systematic (traditional, narrative reviews)
  + Systematic reviews
    - **Qualitative** systematic reviews
    - **Quantitative** systematic reviews: systematic reviews with **meta-analyses**
* Definition of **systematic review**
  + Broad scope definition: the type of studies that focuses on a specific question and uses explicit, pre-planned scientific methods to identify, select, appraise, and summarize similar but separate studies
  + Narrow scope definition: those systematic reviews that only use **qualitative** synthesis method
* Definition of **meta-analysis**
  + Broad scope definition: those systematic reviews that use **quantitative** synthesis method
  + Narrow scope definition: the **quantitative synthesis method/process** used in conducting a systematic review

## If you would like to know more about systematic review:

* Cochrane Handbook for Systematic Reviews of Interventions: <https://training.cochrane.org/handbook/current>
* Introduction to Systematic Review and Meta-Analysis. Coursera. Johns Hopkins Bloomberg School of Public Health: <https://www.coursera.org/learn/systematic-review/home/welcome>
* Chinese textbook by Huaxi School of Public Health: 系统评价、meta-分析设计与实施方法

# 2. How to conduct a systematic electrical searching

## 2.1 Starting point for developing a search strategy: break your research question into concepts (PICOS)

* consider the main concepts being examined in a review: **PICOS**
  + The **review PICO** (planned at the protocol stage) is the **PICO** on which eligibility of studies is based (what will be included and what excluded from the review).
  + The **PICO of the included studies** (determined at the review stage) is what was actually investigated in the included studies.
* PICOS
  + **P**opulation
  + **I**ntervention(s)
  + **C**omparison(s)
  + **O**utcome
  + **S**tudy design

### 2.1.1 Components of research questions: “PICOS/PECOS” principle

* **P**: Patients or populations
  + Define condition or disease, including explicit diagnostic criteria
  + Population and setting of interest (age, race, sex, community, hospital, outpatient)
* **I/E**: Intervention/exposure
  + Timing of exposures
  + Route of administration
  + Dose intensity
  + Duration of exposure or therapy
* **C**: Comparison group(s)
  + For trials: placebo, standard therapy, no treatment
  + For epi: no exposure, non-cases (hospital, neighborhood, etc.)
* **O**: Outcome
  + Criteria for defining
  + Important to cunsumers, providers
  + Unpublished data?
* **S**: Study design
  + Experimental studies or observational studies (Levels of evidence: <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009>)
  + Details of study design, eg. placebo in control group, endpoint blinding, follow-up time

### 2.2.2 Examples of breaking down the research questions

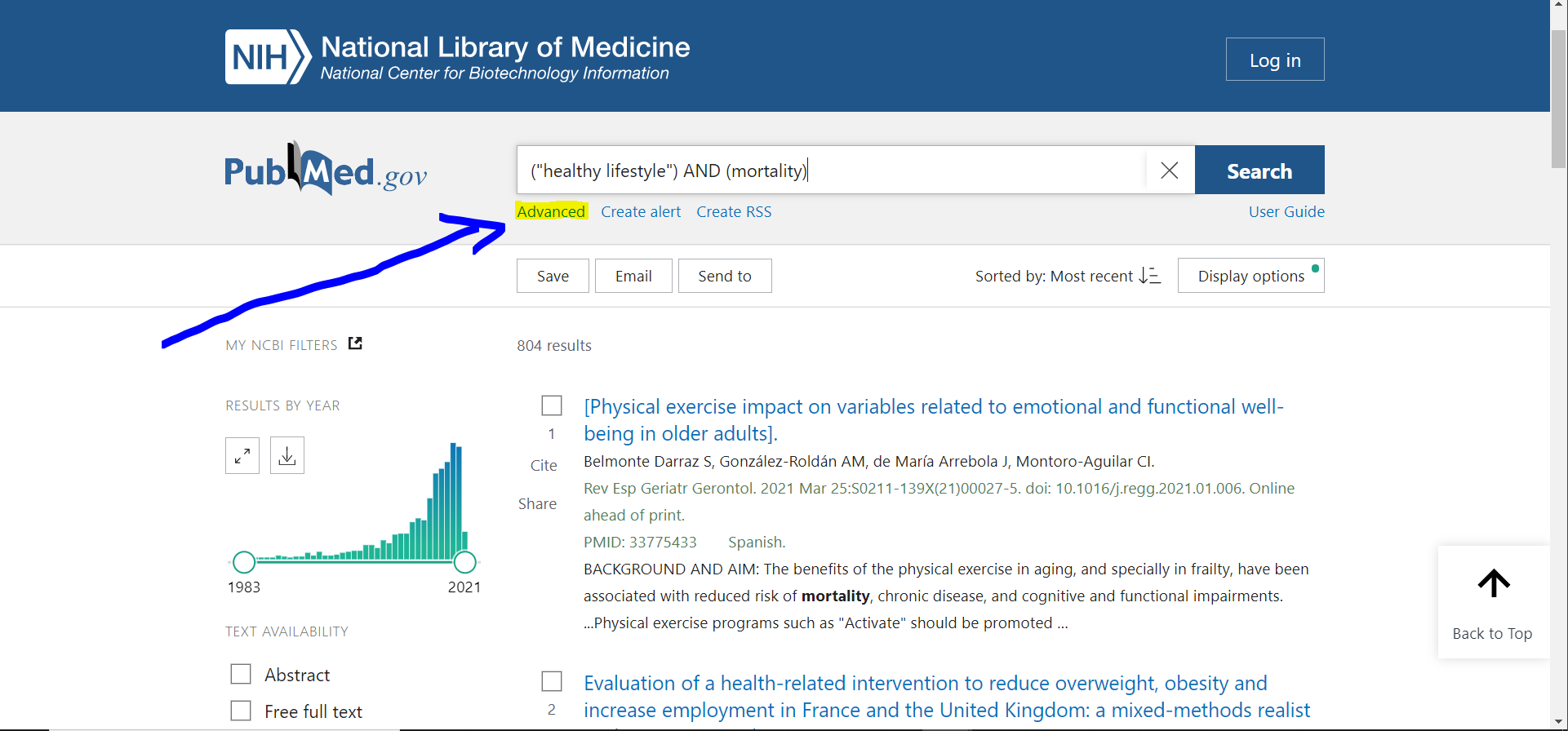
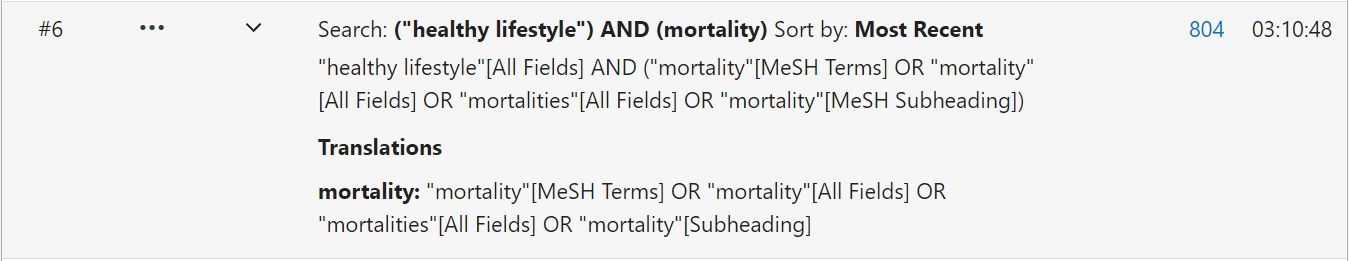
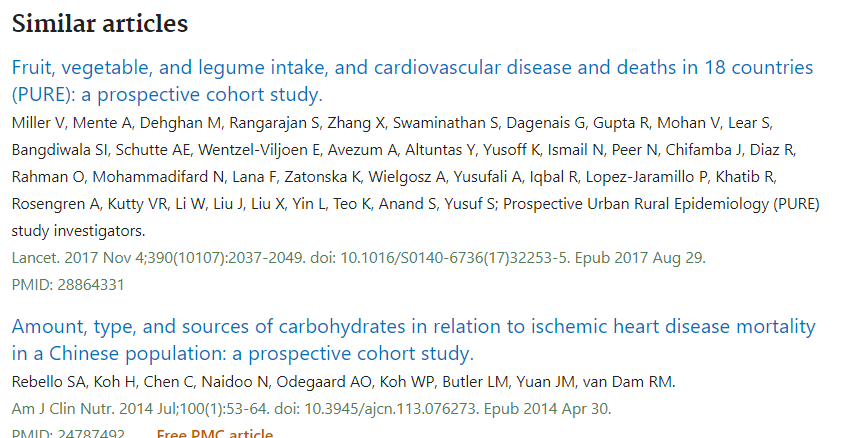
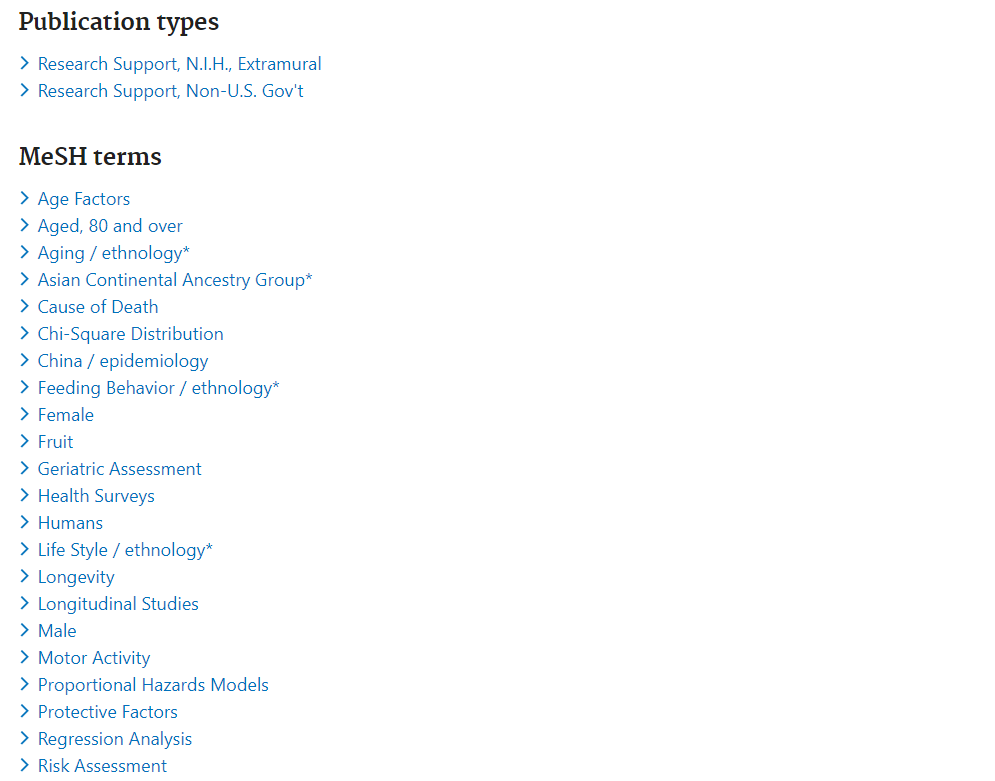
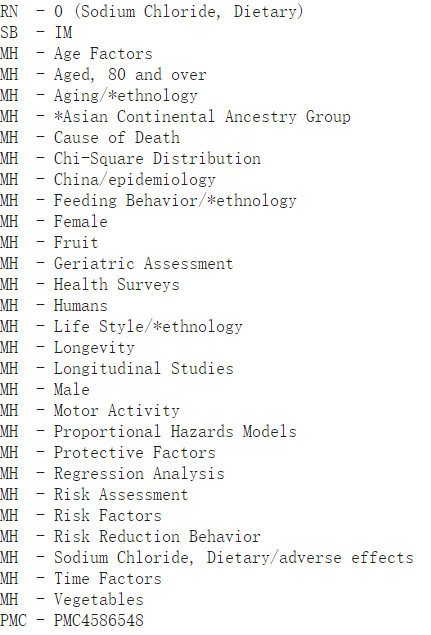
* Combined lifestyle factors, all-cause mortality and cardiovascular disease: a systematic review and meta-analysis of prospective cohort studies
  + P: N/A (without the outcome of interests at the start of the study)
  + E: combined lifestyle factors (healthy?)
  + C: combined lifestyle factors (non-healthy?)
  + O: all-cause mortality and cardiovascular disease
  + S: a systematic review and meta-analysis of prospective cohort studies
* Lifestyle factors, cardiovascular disease and all-cause mortality in middle-aged and elderly women: a systematic review and meta-analysis
  + P: middle-aged and elderly women (without the outcome of interests at the start of the study)
  + E: lifestyle factors (healthy?)
  + C: lifestyle factors (non-healthy?)
  + O: cardiovascular disease and all-cause mortality
  + S: a systematic review and meta-analysis (of cohort studies)

## 2.2 Structure of a search strategy

* **Comparators/outcomes** typically may not be well described in the title or abstract of an article and are often not well indexed with controlled vocabulary terms
* A search strategy will typically has:
  + terms to search for the health condition of interest, i.e. the **population**
  + terms to search for the **intervention(s)** evaluated
  + terms to search for the types of **study design** to be included
* Structure:
  + **a broad set of search terms** gathered for each concept
  + combined with the **OR** Boolean operator to achieve sensitivity within concepts
  + combined with the **AND** Boolean operator for results of each concept
* **Structure from my point of view:**
  + **word[field] + logic**

### 2.2.1 Word

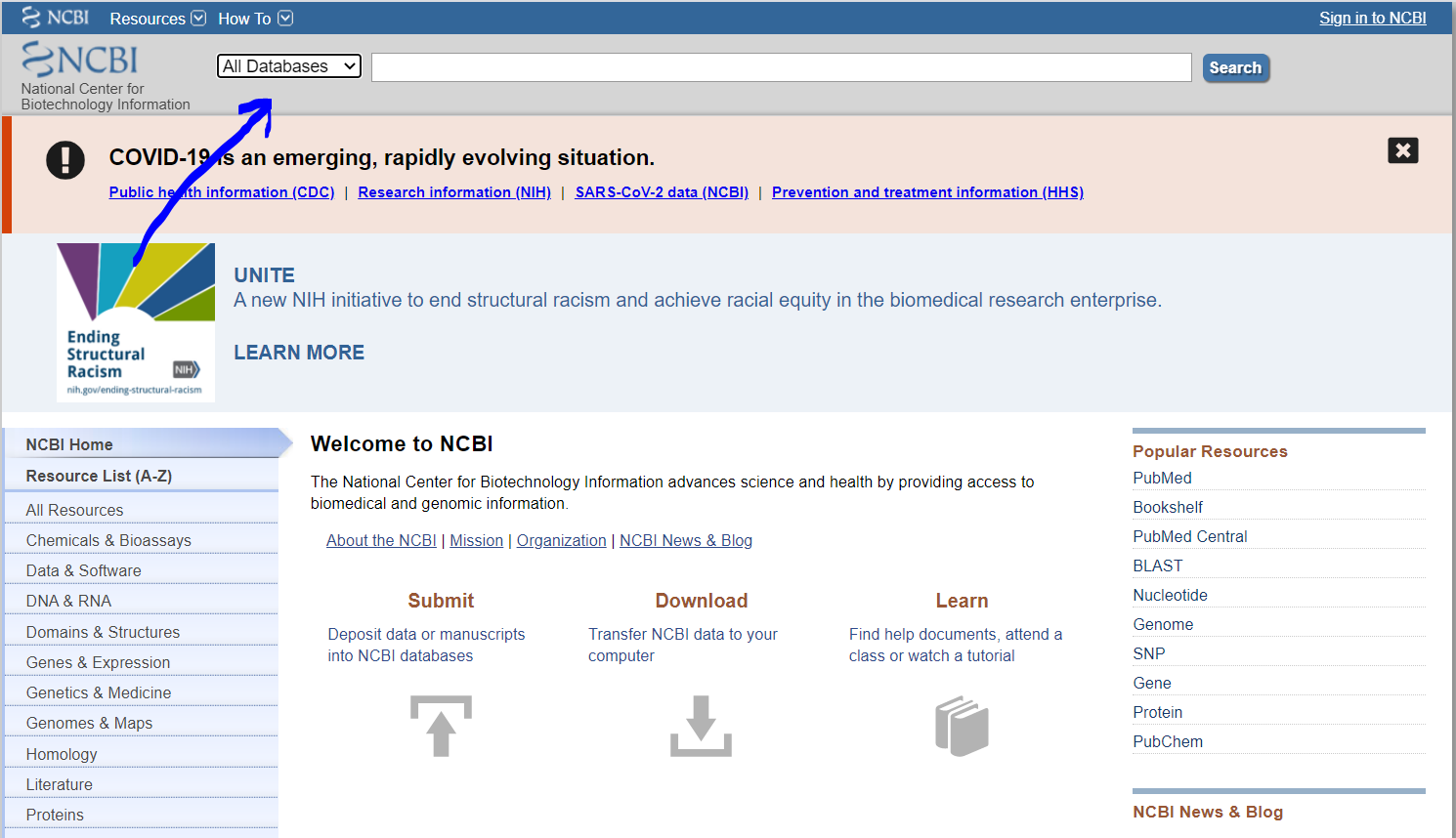
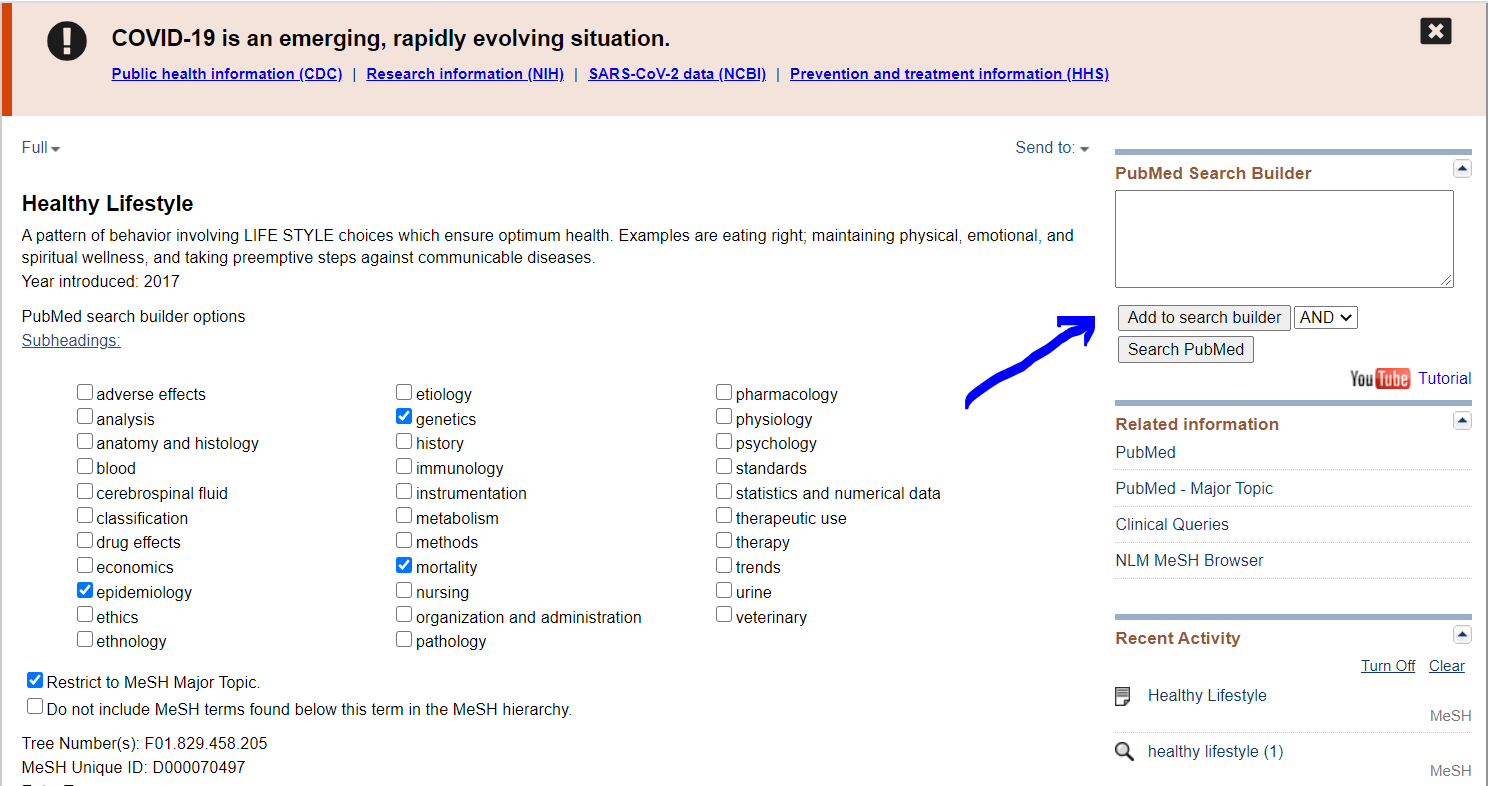
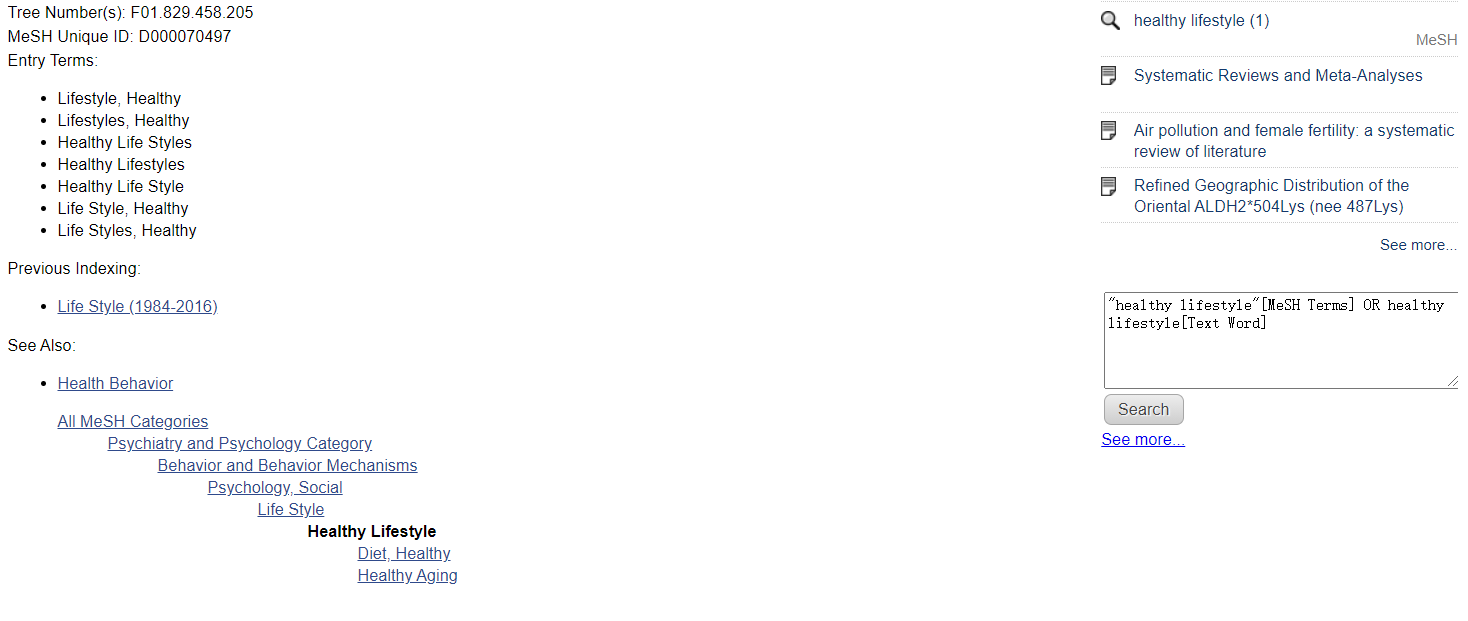
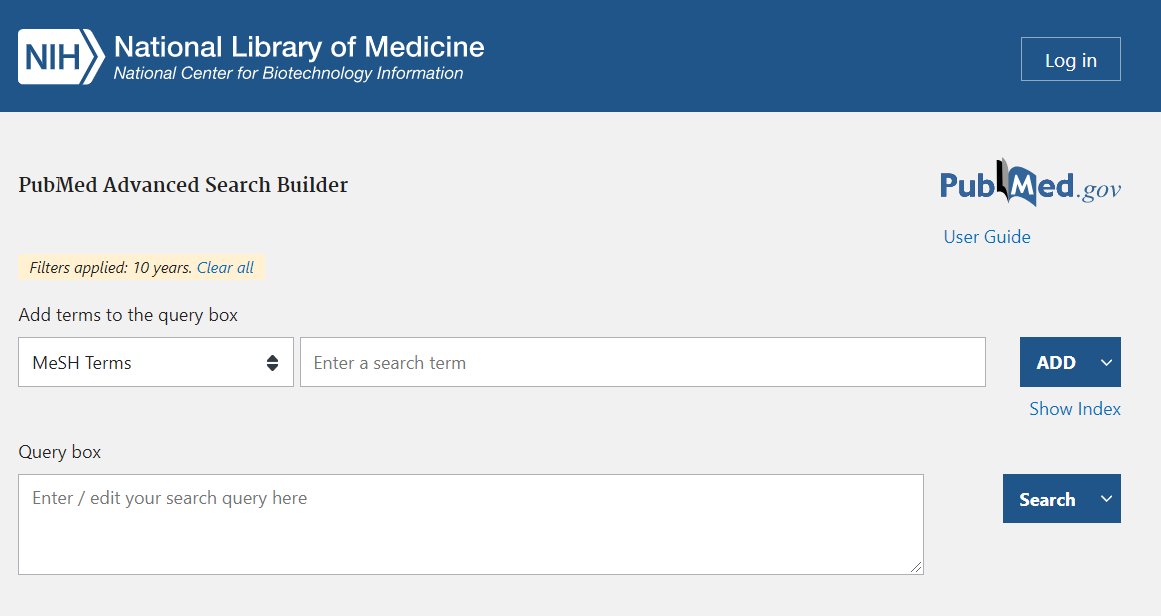
#### 2.2.1.1 Find synonyms for text words

* You need to start with **simple search strategy**
* Start with the general format:
  + (Population OR synonym #1 OR synonym #2) AND
  + (Intervention OR synonym #1 OR synonym #2) AND
  + [if applicable] (Comparator OR synonym #1 OR synonym #2) AND
  + [if applicable] (Outcome OR synonym #1 OR synonym #2)
  + Add study type filter terms
* Take “healthy lifestyle and mortality” as an example
  + You can simply go to PubMed database at <https://pubmed.ncbi.nlm.nih.gov/>, type **(“healthy lifestyle”) AND (mortality)** in the search box: 
  + Take a look at the “Advanced” buttom, where you can see what is actually searched in the search history. You can use the “Translations” to find synonyms. 
  + You may find some relavent articles in the previous searching. Click on these articles. Roll down to see the **keywords** and **Similar articles**. You can expand synonyms through these places.  
  + Further roll down to see the **Publication types** and **MeSH terms**. You can further expand synonyms through these places. 
  + Another way is to click on the **Display options**, choose **PubMed**, and you can see the indexing code behind the article. 
  + Roll down the code to see **Mesh terms** 
* Use **key words** and **mesh terms** in known articles to expand synonyms and build search strategy

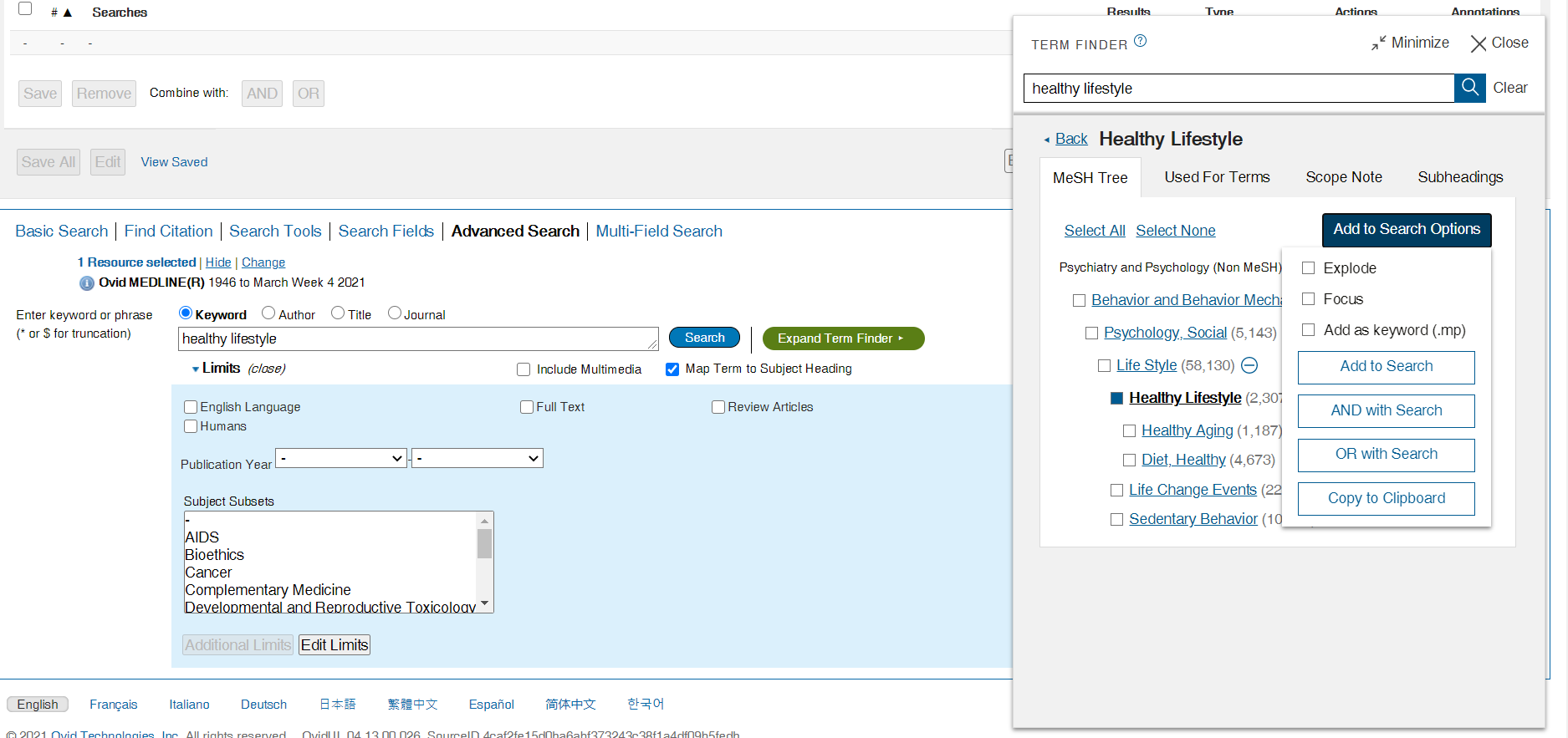
#### 2.2.1.2 Use **controlled vocabulary** when available

* **text words** and **controlled vocabulary**
  + **text word**: the terms occurring in the title, abstract, or other fields
  + **controlled vocabulary**: standardized subject terms assigned to the references by indexers
* Use both **text words** and **controlled vocabulary**
  + If you use only keywords, you could miss articles that don’t use your precise terms
  + If you use only controlled vocabulary, you could miss articles that have not been indexed yet or that have older indexing

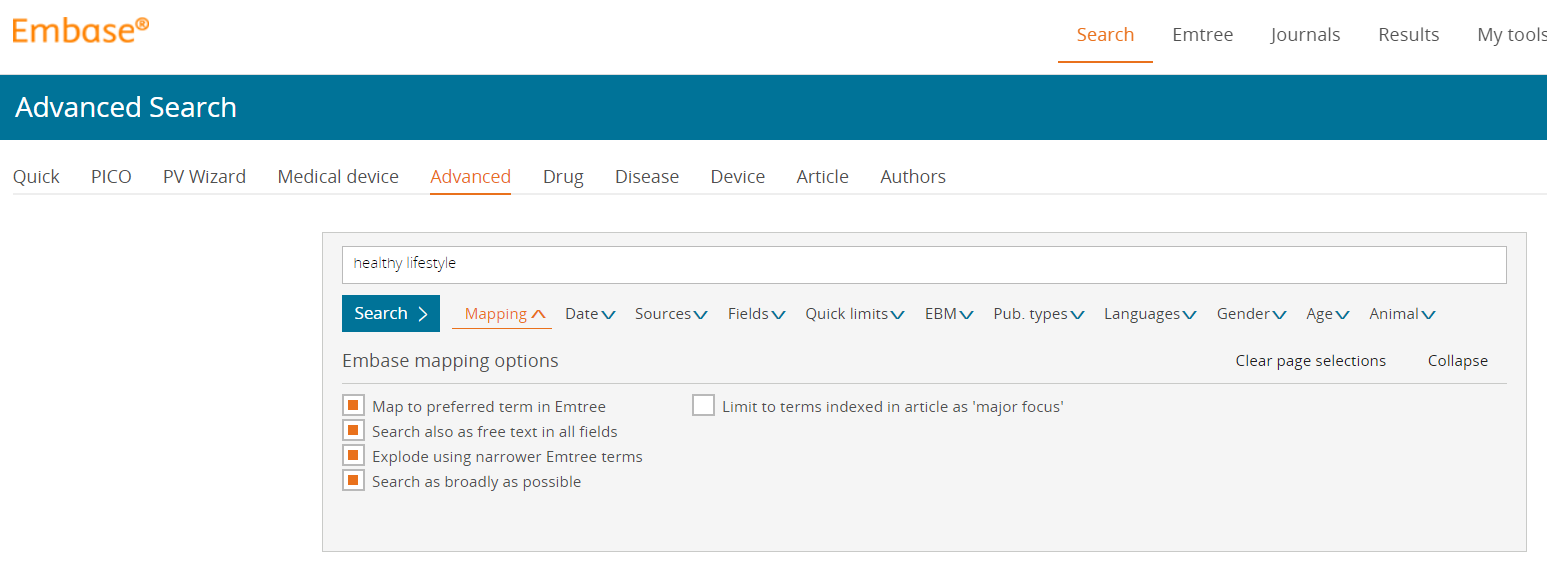
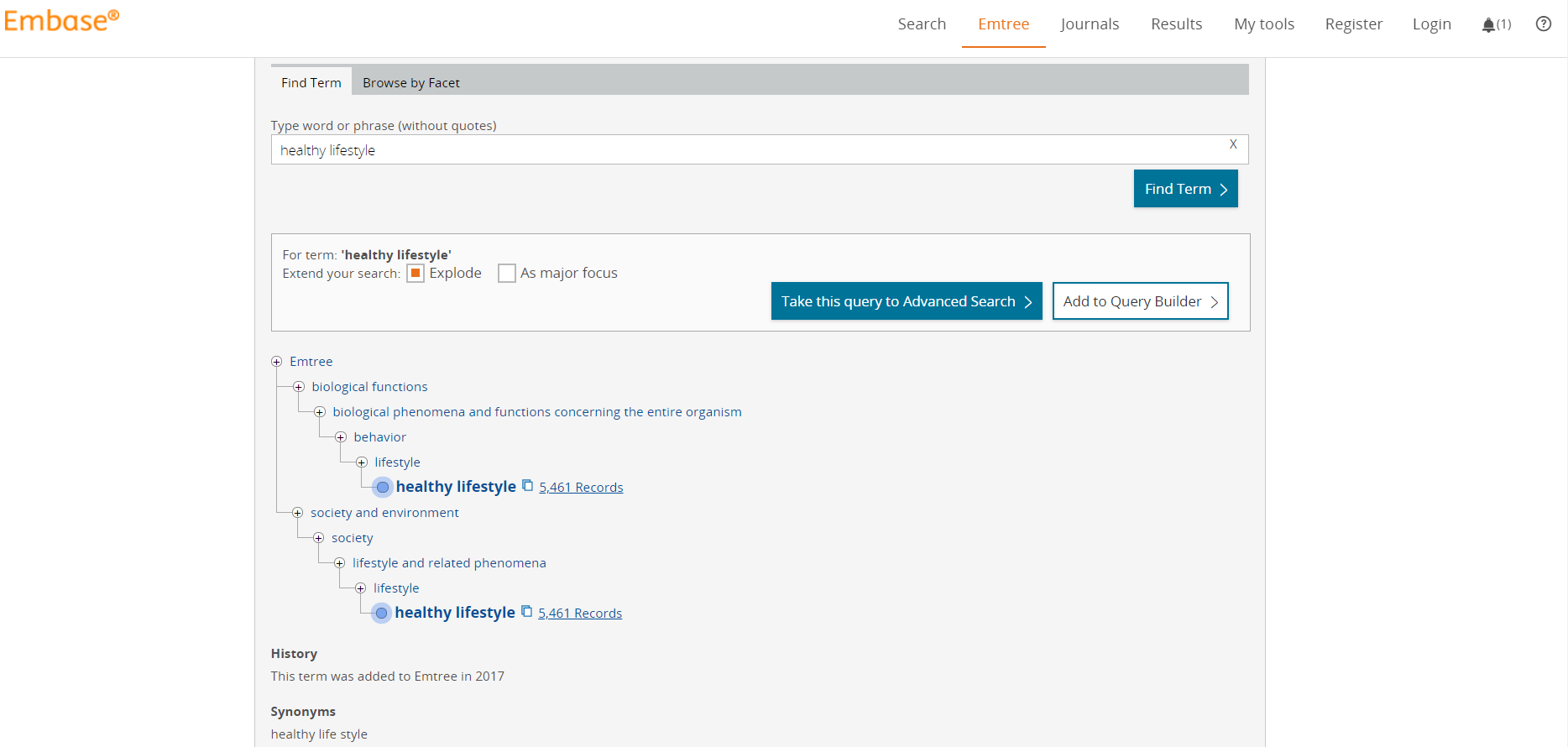
##### 2.2.1.2.1 PubMed: **MeSH Database**

* Go to NCBI’s website <https://www.ncbi.nlm.nih.gov/>, choose “Mesh” Database, and click “Search” 
* Type **healthy lifestyle** in the search box, and it will retrieve the Mesh term **Healthy Lifestyle**. It is not common to retrieve just a simple Mesh term like Healthy Lifestyle, normally it will retrieve multiple Mesh terms, and you need to select which term fits with your scope.
* You can see multiple **subheadings** here. You can click on the subheadings that is correlated to your research topic, or simply do not click any, which will retrieve all subheadings for this Mesh term.
* Another important point is the 2 options below **subheading**, i.e., **Restrict to Mesh Major Topic** and **Do not include MeSH terms found below this term in the Mesh hierarchy**. Select these options may help you to retrieve less articles that are more focus on your Mesh term, but may omit articles that you need. The best way to decide whether to select these options or not is to conduct pilot searching, and is case by case based on your research question.
* After selecting these options, you can simply click “Add to search builder” and “Search PubMed”. 
* You can further roll down to see the **Entry Terms** and Mesh **hierarchy tree**. Please note that although it is allowed to search terms and terms under its hierarchy, eg., **Healthy Lifestyle[Mesh] AND Healthy Aging[Mesh]**, actually **Healthy Aging[Mesh]** is not meaningful, since all articles it will retrieved are covered by **Healthy Lifestyle[Mesh]**, unless you click **Restrict to Mesh Major Topic**. This is simple when you are using one Mesh term, however, when you are using multiple synonyming Mesh terms, you do need to have this info in mind and always check the hierarchy tree. 
* Another way to access Mesh Database is to directly go to PubMed advance searching <https://pubmed.ncbi.nlm.nih.gov/advanced/> and choose **MeSH Terms** as the field. However, this may be misleading, since it also allow you to input a non-Mesh word or phrase as a Mesh Term and allow you to search. So I strongly disencourage you to use this way. 

#### 2.2.1.2.2 MEDLINE (Ovid interface): **MeSH Database**

* To access Medline (Ovid interface), first get on Duke Library, click on **Research Database**, and search for **medline** <https://guides.library.duke.edu/az.php?q=medline>
* Since PubMed is also called “Old Medline”, Medline has nearly identical **Mesh Database** and function with PubMed. You can directly enter **healthy lifestyle**, select **Map Term to SubJect Heading**, and click **Search**; or you can use **Expand Term Finder** 
* The **Explode/Focus** is similar to the **Restrict to Mesh Major Topic** option in PubMed. The **Subheadings** is similiar to the **Subheadings** in PubMed.

#### 2.2.1.2.3 Embase: **Emtree**

* To access Embase, first get on Duke Library, click on **Research Database**, and search for **Embase** <https://guides.library.duke.edu/az.php?q=embase>
* Embase has an **Emtree** database, which is similar to **Mesh** database in PubMed/Medline. Simply fo to **Advanced** searching in Embase, select **Mapping**, and select the preferred options at the left side. You can enter **healthy lifestyle** in the search box, and click **Search**. 
* Another way is to select **Emtree** in Embase, and to search for the **Emtree term**. You can see the hierarchy tree of the searched term. Sometimes, your search term appeared as a **candidate term**, without hierarchy tree retrieved. This is because Embase has a database of **candidate Emtree Terms**, which is database of terms that are currently not included Emtree, but may be included in further indexing. You can choose whether to include these terms or not. 

#### 2.2.1.2.4 Web of Science: no controlled vocabulary available

### 2.2.1.3 Other issues on Consistency

* Plurals: Acuity, acuities
* Abbreviations: CVD for cardiovascular diseases
* Synonyms: variation, mutation, polymorphism
* Spelling variations: Randomized/randomised
* Scope: cardiovascular disease, heart disease

### 2.2.2 Field

* We do not have time to discuss about details here. Generally speaking, we mostly use “[Title/abstract]” as the search field.
* Here are the common field tags for PubMed. Other databases have similar fields.
  + MeSH: mh
  + Text word: tw
  + Title and abstract: tiab
  + Title: ti
  + Date:dp
  + Author: au
  + Publication type: pt
  + All fields: all

### 2.2.3 Logic

#### 2.2.3.1 Boolean logic

* similiar to coding: AND, OR, NOT
  + “healthy lifestyle” AND mortality
  + “healthy lifestyle” OR “healthy diet habit”
  + “healthy lifestyle” NOT “animal”

#### 2.2.3.2 Proximity searching

* NEAR/ADJ
  + healthy adj1 lifestyle
  + **PubMed and Web of Science do not have proximit searching function**

#### 2.2.3.3 Truncate searching

* similiar to coding: \* or $
  + varian\*: variant, variance

### 2.2.4 Specific for study type

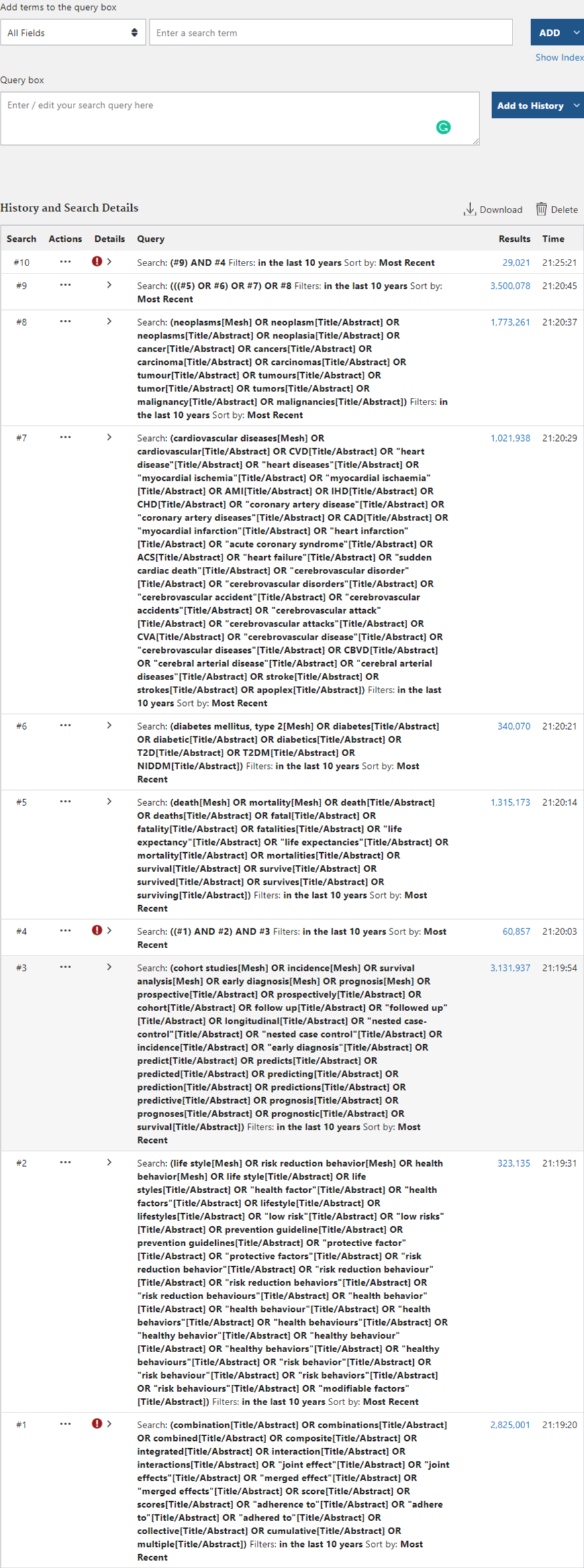
* There are existing sensitivie search strategy for different study type
* eg. Highly Sensitive Search Strategy for identifying randomized trials in PubMed:
  + #1 randomized controlled trial [pt]
  + #2 controlled clinical trial [pt]
  + #3 randomized [tiab]
  + #4 placebo [tiab]
  + #5 drug therapy [sh]
  + #6 randomly [tiab]
  + #7 trial [tiab]
  + #8 groups [tiab]
  + #9 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8
  + #10 animals [mh] not (humans [mh] and animals [mh])
  + #11 #9 not #10
* eg. Highly Sensitive Search Strategy for identifying randomized trials in Ovid:
  + #1 randomized controlled trial.pt.
  + #2 controlled clinical trial.pt.
  + #3 randomized.ab.
  + #4 placebo.ab.
  + #5 drug therapy.fs.
  + #6 randomly.ab.
  + #7 trial.ab
  + #8 groups.ab.
  + #9 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
  + #10 animals.sh. not(humans.sh. and animals.sh)
  + #11 9 not 10

## 2.3 Fit search strategy to each database

### 2.3.1 Common databases

* PubMed
* MEDLINE (Ovid interface)
* EMBASE
* Cochrane Central Register of Controlled Trials (CENTRAL) & Cochrane Database of Systematic Review
* Citation databases: Web of Science
* Other databases:
  + Citation databases: Scopus
    - Dissertations, thesis databases: ProQuest
  + National and regional databases (often local language): LILACS
  + Subject-specific databases:
    - CINAHL
    - PsychINFO
    - OTSeeker
  + Gray literature databases: Opengrey

### 2.3.2 Example

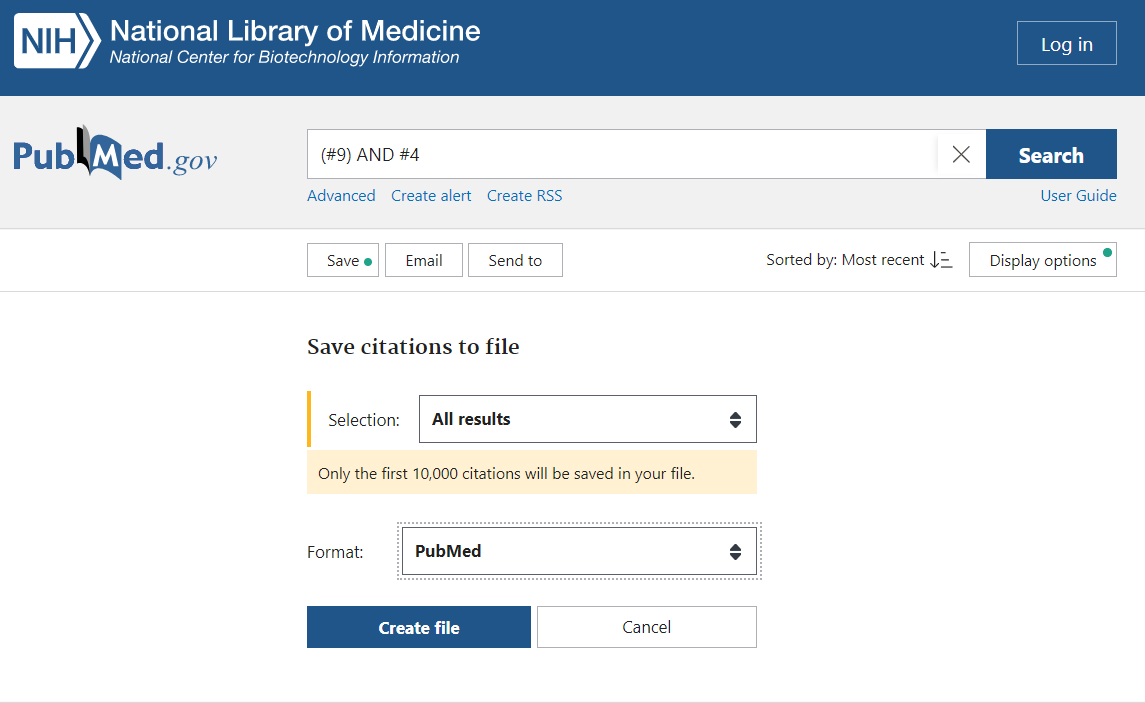
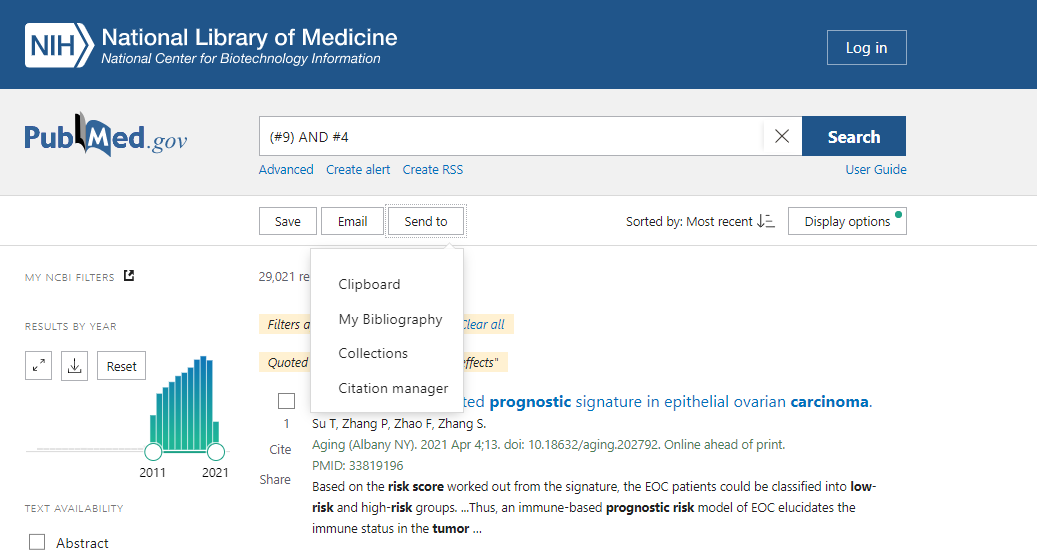
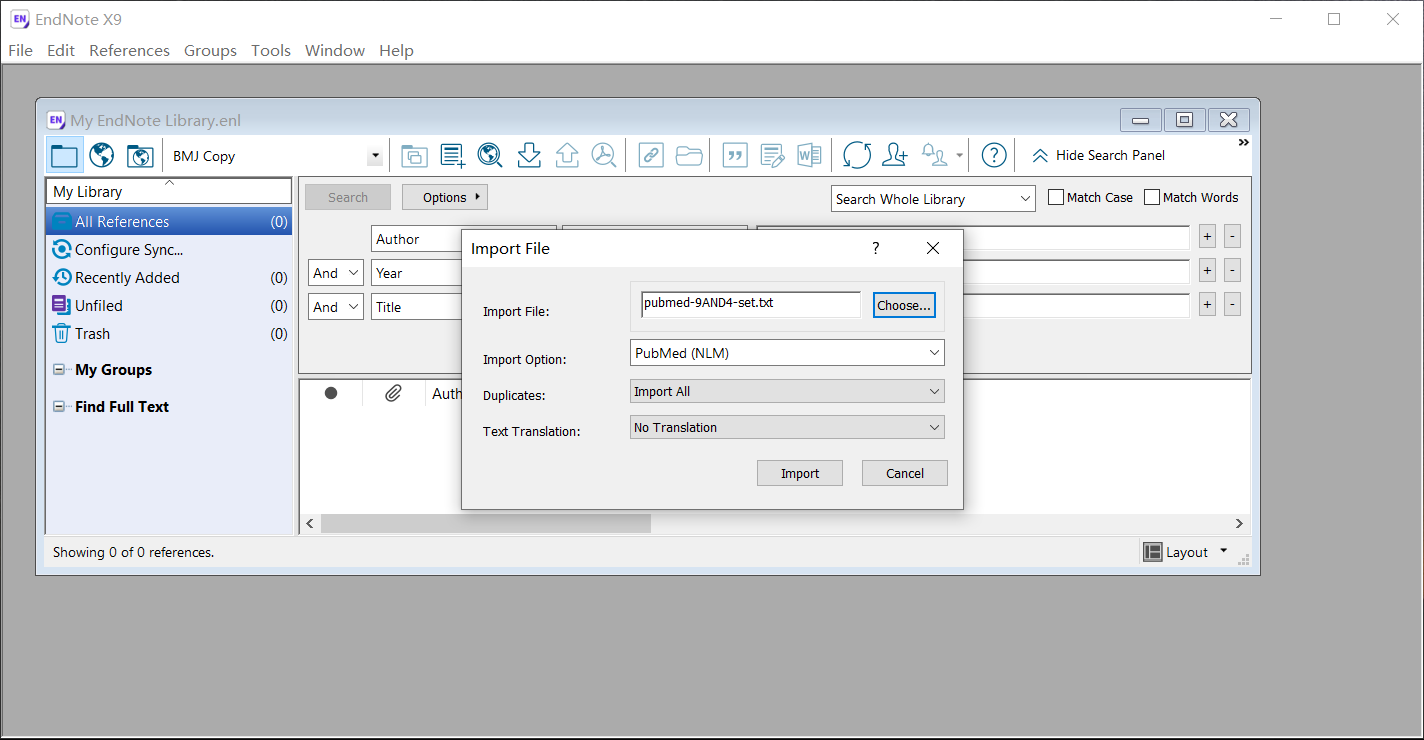
* Take this article as an example [[1]](#footnote-1): “In brief, the search terms included the Medical Subject Heading terms and related exploded versions as well as keywords in titles or abstracts related to the following themes: ‘diabetes’, ‘cardiovascular disease’, ‘cancer’, ‘mortality’, ‘combined’, ‘lifestyle’ and ‘cohort studies’. The search themes were then combined using the Boolean operator ‘or’ for the four health outcomes (diabetes, CVD, cancer and mortality) and then combined with other themes using ‘and’.”
* You can try to find synonyms based on the brief search strategy the author provided here, and to try to search on PubMed and MEDLINE, and to see how many articles your search strategy will retrieve, and how many articles the author’s search strategy will retrieve.
* You can also try to modify the search strategy the author provided, eg. to use combin\*[Title/Abstract] instaed of “combination[Title/Abstract] OR combinations[Title/Abstract] OR combined[Title/Abstract]”, and to see how the searching result changes.
  + PubMed
    - #1 (combination[Title/Abstract] OR combinations[Title/Abstract] OR combined[Title/Abstract] OR composite[Title/Abstract] OR integrated[Title/Abstract] OR interaction[Title/Abstract] OR interactions[Title/Abstract] OR “joint effect”[Title/Abstract] OR “joint effects”[Title/Abstract] OR “merged effect”[Title/Abstract] OR “merged effects”[Title/Abstract] OR score[Title/Abstract] OR scores[Title/Abstract] OR “adherence to”[Title/Abstract] OR “adhere to”[Title/Abstract] OR “adhered to”[Title/Abstract] OR collective[Title/Abstract] OR cumulative[Title/Abstract] OR multiple[Title/Abstract])
    - #2 (life style[Mesh] OR risk reduction behavior[Mesh] OR health behavior[Mesh] OR life style[Title/Abstract] OR life styles[Title/Abstract] OR “health factor”[Title/Abstract] OR “health factors”[Title/Abstract] OR lifestyle[Title/Abstract] OR lifestyles[Title/Abstract] OR “low risk”[Title/Abstract] OR “low risks”[Title/Abstract] OR prevention guideline[Title/Abstract] OR prevention guidelines[Title/Abstract] OR “protective factor”[Title/Abstract] OR “protective factors”[Title/Abstract] OR “risk reduction behavior”[Title/Abstract] OR “risk reduction behaviour”[Title/Abstract] OR “risk reduction behaviors”[Title/Abstract] OR “risk reduction behaviours”[Title/Abstract] OR “health behavior”[Title/Abstract] OR “health behaviour”[Title/Abstract] OR “health behaviors”[Title/Abstract] OR “health behaviours”[Title/Abstract] OR “healthy behavior”[Title/Abstract] OR “healthy behaviour”[Title/Abstract] OR “healthy behaviors”[Title/Abstract] OR “healthy behaviours”[Title/Abstract] OR “risk behavior”[Title/Abstract] OR “risk behaviour”[Title/Abstract] OR “risk behaviors”[Title/Abstract] OR “risk behaviours”[Title/Abstract] OR “modifiable factors”[Title/Abstract])
    - #3 (cohort studies[Mesh] OR incidence[Mesh] OR survival analysis[Mesh] OR early diagnosis[Mesh] OR prognosis[Mesh] OR prospective[Title/Abstract] OR prospectively[Title/Abstract] OR cohort[Title/Abstract] OR follow up[Title/Abstract] OR “followed up”[Title/Abstract] OR longitudinal[Title/Abstract] OR “nested case-control”[Title/Abstract] OR “nested case control”[Title/Abstract] OR incidence[Title/Abstract] OR “early diagnosis”[Title/Abstract] OR predict[Title/Abstract] OR predicts[Title/Abstract] OR predicted[Title/Abstract] OR predicting[Title/Abstract] OR prediction[Title/Abstract] OR predictions[Title/Abstract] OR predictive[Title/Abstract] OR prognosis[Title/Abstract] OR prognoses[Title/Abstract] OR prognostic[Title/Abstract] OR survival[Title/Abstract])
    - #4 ((#1) AND #2) AND #3
    - #5 (death[Mesh] OR mortality[Mesh] OR death[Title/Abstract] OR deaths[Title/Abstract] OR fatal[Title/Abstract] OR fatality[Title/Abstract] OR fatalities[Title/Abstract] OR “life expectancy”[Title/Abstract] OR “life expectancies”[Title/Abstract] OR mortality[Title/Abstract] OR mortalities[Title/Abstract] OR survival[Title/Abstract] OR survive[Title/Abstract] OR survived[Title/Abstract] OR survives[Title/Abstract] OR surviving[Title/Abstract])
    - #6 (diabetes mellitus, type 2[Mesh] OR diabetes[Title/Abstract] OR diabetic[Title/Abstract] OR diabetics[Title/Abstract] OR T2D[Title/Abstract] OR T2DM[Title/Abstract] OR NIDDM[Title/Abstract])
    - #7 (cardiovascular diseases[Mesh] OR cardiovascular[Title/Abstract] OR CVD[Title/Abstract] OR “heart disease”[Title/Abstract] OR “heart diseases”[Title/Abstract] OR “myocardial ischemia”[Title/Abstract] OR “myocardial ischaemia”[Title/Abstract] OR AMI[Title/Abstract] OR IHD[Title/Abstract] OR CHD[Title/Abstract] OR “coronary artery disease”[Title/Abstract] OR “coronary artery diseases”[Title/Abstract] OR CAD[Title/Abstract] OR “myocardial infarction”[Title/Abstract] OR “heart infarction”[Title/Abstract] OR “acute coronary syndrome”[Title/Abstract] OR ACS[Title/Abstract] OR “heart failure”[Title/Abstract] OR “sudden cardiac death”[Title/Abstract] OR “cerebrovascular disorder”[Title/Abstract] OR “cerebrovascular disorders”[Title/Abstract] OR “cerebrovascular accident”[Title/Abstract] OR “cerebrovascular accidents”[Title/Abstract] OR “cerebrovascular attack”[Title/Abstract] OR “cerebrovascular attacks”[Title/Abstract] OR CVA[Title/Abstract] OR “cerebrovascular disease”[Title/Abstract] OR “cerebrovascular diseases”[Title/Abstract] OR CBVD[Title/Abstract] OR “cerebral arterial disease”[Title/Abstract] OR “cerebral arterial diseases”[Title/Abstract] OR stroke[Title/Abstract] OR strokes[Title/Abstract] OR apoplex[Title/Abstract])
    - #8 (neoplasms[Mesh] OR neoplasm[Title/Abstract] OR neoplasms[Title/Abstract] OR neoplasia[Title/Abstract] OR cancer[Title/Abstract] OR cancers[Title/Abstract] OR carcinoma[Title/Abstract] OR carcinomas[Title/Abstract] OR tumour[Title/Abstract] OR tumours[Title/Abstract] OR tumor[Title/Abstract] OR tumors[Title/Abstract] OR malignancy[Title/Abstract] OR malignancies[Title/Abstract])
    - #9 (((#5) OR #6) OR #7) OR #8
    - #10 (#9) AND #4 
  + EMBASE
    - 1 exp lifestyle/
    - 2 exp health behavior/
    - 3 (life style or life styles or health factor or health factors or lifestyle or lifestyles or low risk or low risks or prevention guideline or prevention guidelines or protective factor or protective factors or risk reduction behavior or risk reduction behaviour or risk reduction behaviors or risk reduction behaviours or health behavior or health behaviour or health behaviors or health behaviours or healthy behavior or healthy behaviour or healthy behaviors or healthy behaviours or risk behavior or risk behaviour or risk behaviors or risk behaviours or modifiable factors).ti. or (life style or life styles or health factor or health factors or lifestyle or lifestyles or low risk or low risks or prevention guideline or prevention guidelines or protective factor or protective factors or risk reduction behavior or risk reduction behaviour or risk reduction behaviors or risk reduction behaviours or health behavior or health behaviour or health behaviors or health behaviours or healthy behavior or healthy behaviour or healthy behaviors or healthy behaviours or risk behavior or risk behaviour or risk behaviors or risk behaviours or modifiable factors).ab.
    - 4 1 or 2 or 3
    - 5 (combination or combinations or combined or composite or integrated or interaction or interactions or joint effect or joint effects or merged effect or merged effects or score or scores or adherence to or adhere to or adhered to or collective or cumulative or multiple).ti. or (combination or combinations or combined or composite or integrated or interaction or interactions or joint effect or joint effects or merged effect or merged effects or score or scores or adherence to or adhere to or adhered to or collective or cumulative or multiple).ab.
    - 6 exp cohort analysis/
    - 7 exp incidence/
    - 8 exp survival analysis/
    - 9 exp early diagnosis/
    - 10 (prospective or prospectively or cohort or follow up or followed up or longitudinal or nested case-control or nested case control or incidence or early diagnosis or predict or predicts or predicted or predicting or prediction or predictions or predictive or prognosis or prognoses or prognostic or survival).ti. or (prospective or prospectively or cohort or follow up or followed up or longitudinal or nested case-control or nested case control or incidence or early diagnosis or predict or predicts or predicted or predicting or prediction or predictions or predictive or prognosis or prognoses or prognostic or survival).ab.
    - 11 exp prognosis/
    - 12 6 or 7 or 8 or 9 or 10 or 11
    - 13 exp death/
    - 14 exp mortality/
    - 15 exp survival/
    - 16 (death or deaths or fatal or fatality or fatalities or life expectancy or life expectancies or mortality or mortalities or survival or survive or survived or survives or surviving).ti. or (death or deaths or fatal or fatality or fatalities or life expectancy or life expectancies or mortality or mortalities or survival or survive or survived or survives or surviving).ab.
    - 17 13 or 14 or 15 or 16
    - 18 exp non insulin dependent diabetes mellitus/
    - 19 (diabetes or diabetic or diabetics or T2D or T2DM or NIDDM).ti. or (diabetes or diabetic or diabetics or T2D or T2DM or NIDDM).ab.
    - 20 18 or 19
    - 21 exp cardiovascular disease/
    - 22 (cardiovascular or CVD or heart disease or heart diseases or myocardial ischemia or myocardial ischaemia or AMI or IHD or CHD or coronary artery disease or coronary artery diseases or CAD or myocardial infarction or heart infarction or acute coronary syndrome or ACS or heart failure or sudden cardiac death or cerebrovascular disorder or cerebrovascular disorders or cerebrovascular accident or cerebrovascular accidents or cerebrovascular attack or cerebrovascular attacks or CVA or cerebrovascular disease or cerebrovascular diseases or CBVD or cerebral arterial disease or cerebral arterial diseases or stroke or strokes or apoplex).ti. or (cardiovascular or CVD or heart disease or heart diseases or myocardial ischemia or myocardial ischaemia or AMI or IHD or CHD or coronary artery disease or coronary artery diseases or CAD or myocardial infarction or heart infarction or acute coronary syndrome or ACS or heart failure or sudden cardiac death or cerebrovascular disorder or cerebrovascular disorders or cerebrovascular accident or cerebrovascular accidents or cerebrovascular attack or cerebrovascular attacks or CVA or cerebrovascular disease or cerebrovascular diseases or CBVD or cerebral arterial disease or cerebral arterial diseases or stroke or strokes or apoplex).ab.
    - 23 21 or 22
    - 24 exp neoplasm/
    - 25 (neoplasm or neoplasms or neoplasia or cancer or cancers or carcinoma or carcinomas or tumour or tumours or tumor or tumors or malignancy or malignancies).ti. or (neoplasm or neoplasms or neoplasia or cancer or cancers or carcinoma or carcinomas or tumour or tumours or tumor or tumors or malignancy or malignancies).ab.
    - 26 24 or 25
    - 27 17 or 20 or 23 or 26
    - 28 4 and 5 and 12 and 27

## 2.4 Document your searching

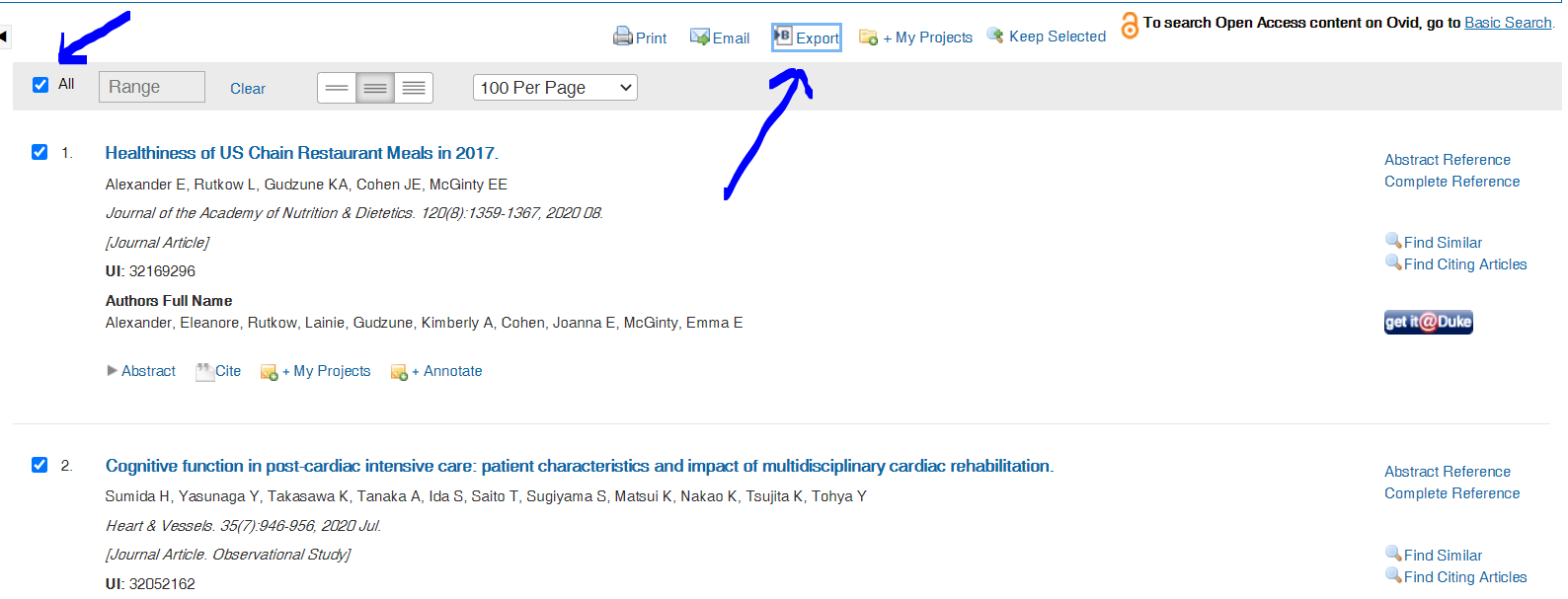
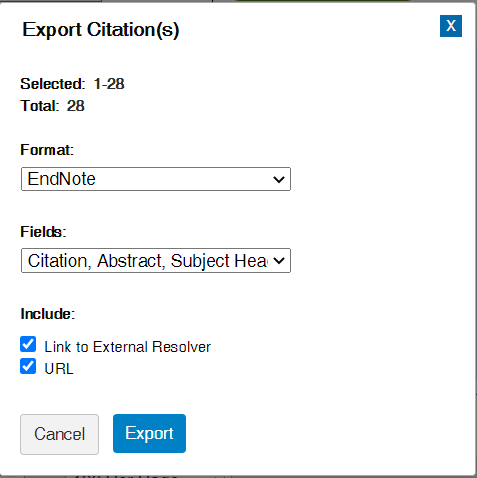
### 2.4.1 Bibliographic Management Software

* Download and store results of electronic searches in EndNote
  + EndNote is most popularly used, and is also purchased by Duke and DKU for faculty, staff, and students
  + EndNote 20 version is not quite perfect, so I still recommend you to use EndNote X9
  + Win and Mac version can be downloaded here: <https://duke.box.com/s/yp0k9jro6t4247we6f0xv2xn2ndkqiz1>
* Other similar softwares
  + ProCite
  + Reference Manager
  + RefWorks
  + QUOSA
  + Mendeley: good to be used in reading fulltext <https://www.mendeley.com/?interaction_required=true>
  + Sciwheel: also Duke support, online interface <https://duke.box.com/s/7k9ec7vf4rnr3rh1c0f8qoiiprdf3yr6>

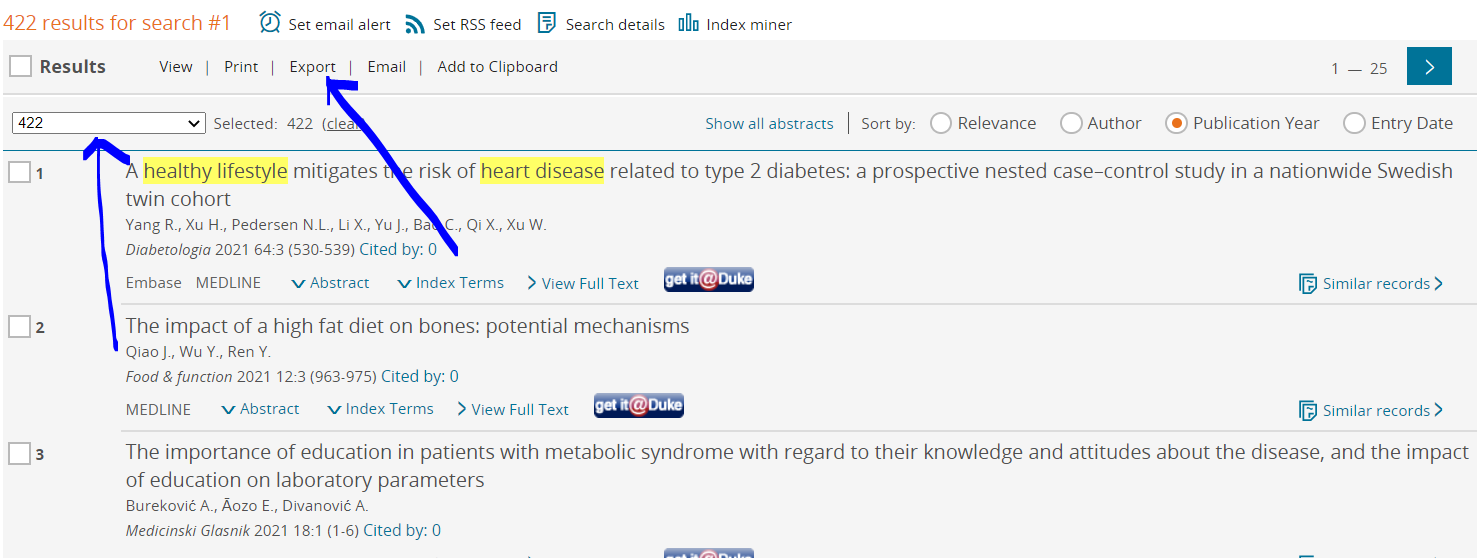
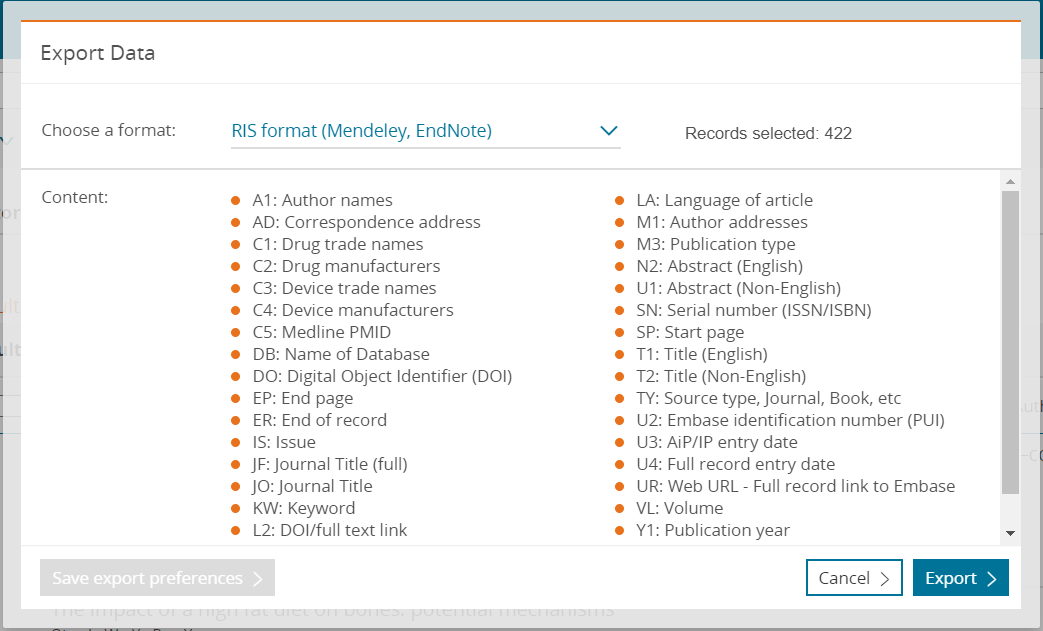
### 2.4.2 Import your searching result from PubMed

* Download citations
  + One way is to select **Save**, **All results**, and **Summary (text)**. A txt document will be created shortly. 
  + Another way is to select **Send to**, **Citation manager**, and **All results**. A nbib. document will be created shortly. 
* Import citations
  + For the first option, you need to select **PubMed (NLM)** in the **Import Option**. 
  + For the second option, you can directly double-click the nbib. document, or select **PubMed (NLM)** in the **Import Option**.

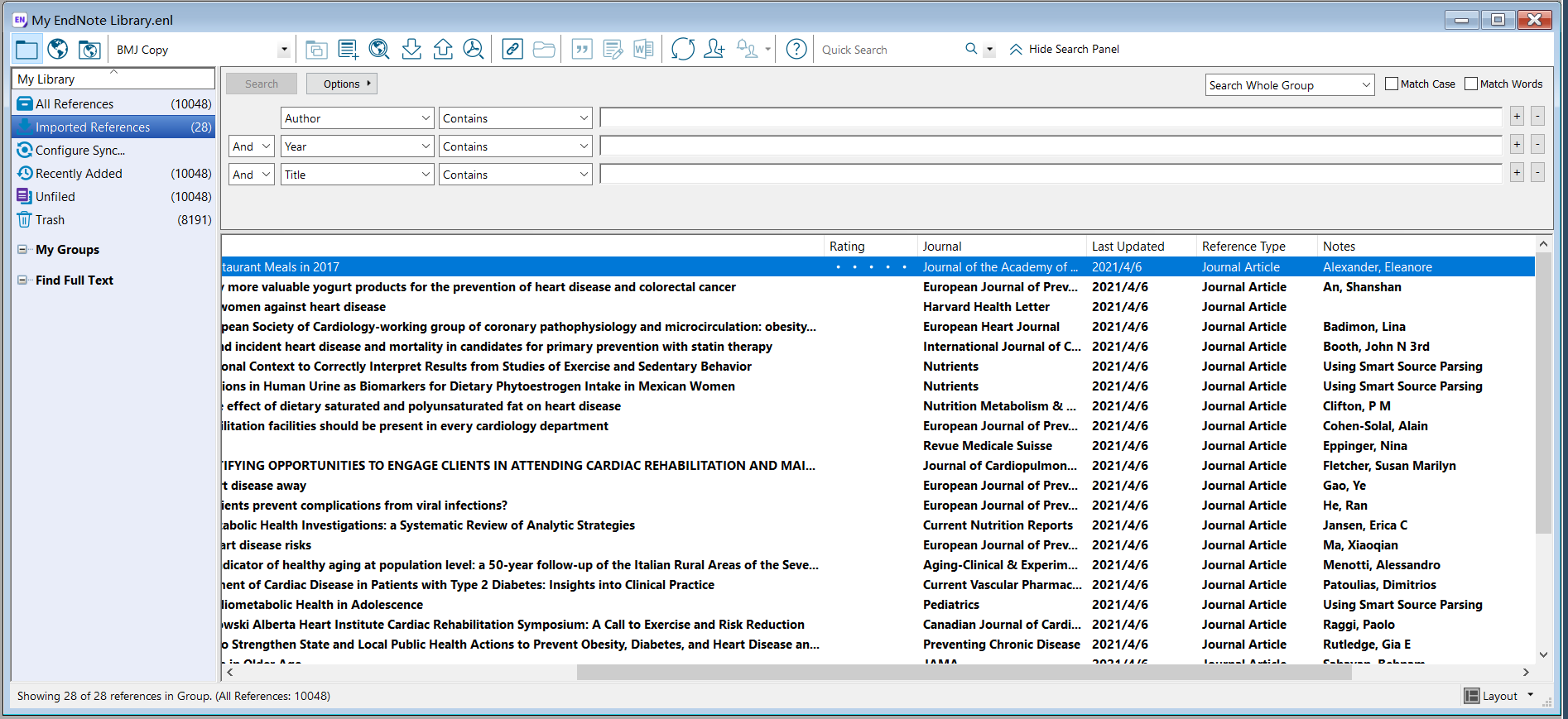
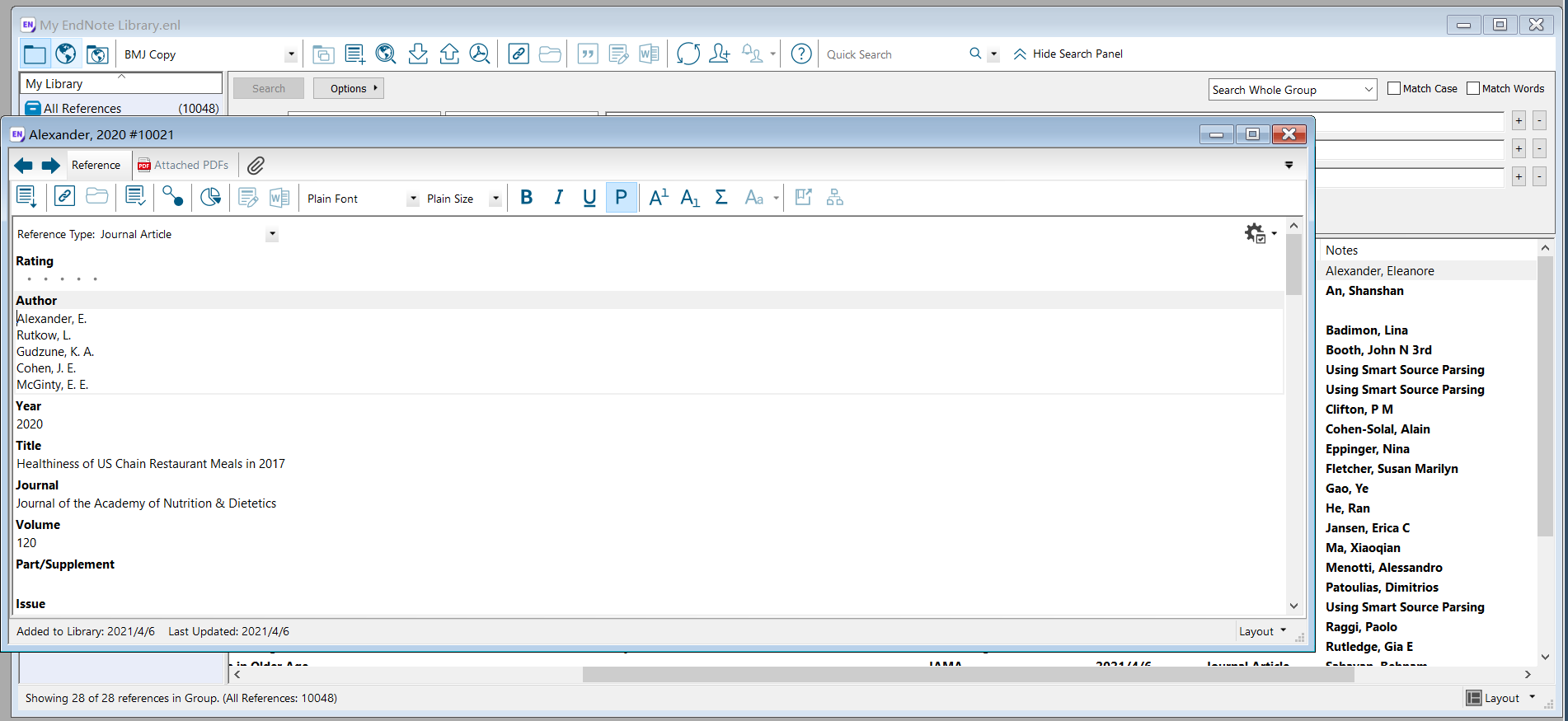
### 2.4.3 Import your searching result from MEDLINE

* Download citations
  + You need to select **All**, and then click **Export** 
  + Then choose **Format** as **EndNote**, and select appropriate **Fields**, and click **Export**. A document will be created shortly. 
* Import citations
  + Simply double-click the document to import it to EndNote.

### 2.4.4 Import your searching result from EMBASE

* Download citations
  + You need to select the **number** to all articles retrieved, and then click **Export** 
  + Then you can use the default format **RIS format (Mendeley, EndNote)** to **Export** 
* Import citations
  + Simply double-click the document to import it to EndNote.

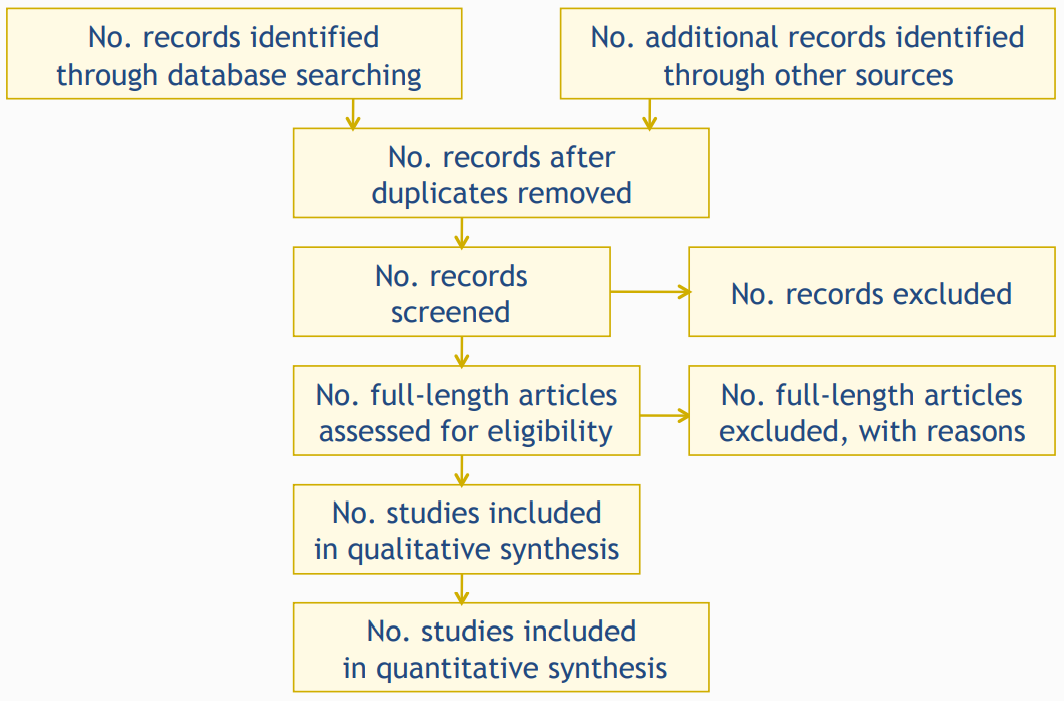
### 2.4.5 Use EndNote to manage literatures

* Use **Notes** to store key information, eg. how record is identified, reason for exclusion
* Use **Rating** to rate the literature to different levels, eg. “highly relevant”, “medium”, “low”
* Click on **Notes**/**Rating** will help you to sort literature based on this field
* Double lick on the literature, and you can change details of these fields, and to attach PDF of full text  

## 2.5 Systematic Snowballing

* Use included articles as a source to identify missing studies:
  + Reference lists: Check **cited** articles = backward searching
  + Web of Science, SCOPUS: Check **citing** articles = forward searching
  + “Related articles”/ “Find Similar” in MEDLINE and Embase: Also could use to find new terms to expand your search strategy

## 2.6 PRISMA Flow of Information

* To keep record of complicated searching and scanning, I strongly encourage you to document infromation in PRISMA flow diagram. Details can be found here: <http://www.prisma-statement.org/PRISMAStatement/FlowDiagram> 

## 2.7 Summary on searching

* Create a comprehensive, sensitive searching (comprehensive vs. sensitive)
* Starting with your question/PICOS, build the search strategy on the structure of word, field, and logic
* Use all electronic resources
* Document when, where, and how you searched and what you found
* We are not experts on searching, if your team cannot solve a problem in searching, it is time to **Ask a Librarian**. It will be also helpful to consult a content expert.

# 3. How to conduct study selection?

## 3.1 The process of study selection

* Import all citation records retrieved from all databases, remove duplicates (both electronic and manually), set up a reference library
  + Electronic: In EndNote, select **Reference**, **Find duplicates**
  + Manual: recommend to perform at least 2 rounds, first sort literature by **Author** and remove duplicates, and sort literature by **Title** and remove duplicates
* Scan literature by **Title/Abstract**, remove articles that are clearly not fit with your research topic
* Find full text of all articles left
* Group the left articles by outcome of interests or different outcomes
* Scan literature by **Full text**, contact the author if needed for missing information
* Final decision on included articles

## 3.2 Study selection critera: build based on PICOS

* **P**opulation
  + diagnostic criteria of diseases
  + main characteristics of participants
  + demographic characteristics
  + research institutions
  + who make the diagnosis
  + participants needed to be excluded from systematic review
  + how to deal with studies only contained a subgroup of the study population
* **I**ntervention/**E**xposure/**C**omparison
  + what is the intervention/exposure/comparison of interests
  + how to deal with heterogeneity in intervention, eg. doses, treatment period
  + how to deal with studies only contained part of the intervention of interests
  + how to deal with studies not only included the intervention of interests
* **O**utcome
  + whether the outcome is of significance to patients and decision-makers in clinical work
  + primary outcomes are 2 or 3 of the target outcomes
  + secondary outcomes are the rest target outcomes, or outcomes to evaluate the effects of the interventions
  + outcomes should contain potential or certified adverse-effect
  + consider all outcomes that are related to decision-making, including economic factors
  + consider the scope and time for evluating the outcomes
* **S**tudy design
  + Experimental studies or observational studies (Levels of evidence: <https://www.cebm.ox.ac.uk/resources/levels-of-evidence/oxford-centre-for-evidence-based-medicine-levels-of-evidence-march-2009>)

## 3.3 Critical points in study selection

* The **sequence** of study selection criteria
* Study selection is a **subjective** process
* **Blinding**
* **Pilot studies**
* Follow your **protocol**
* Document the reasons for **exclusion**

# Thank you

* Similar to programming, it is hardly possible to cover all details in a workshop
* The fastest way to progress is to **try it**, and to find solutions when meeting problems
* Hopefully you will find this process interesting and good luck

# References

* 刘鸣. 系统评价、Meta-分析设计与实施方法.人民卫生出版社
* Kay Dickersin, Tianjing Li. Introduction to Systematic Review and Meta-Analysis. Coursera
* Julian Higgins, James Thomas, et al. Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Training
* Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

1. Zhang Y, Pan X F, Chen J, et al. Combined lifestyle factors and risk of incident type 2 diabetes and prognosis among individuals with type 2 diabetes: a systematic review and meta-analysis of prospective cohort studies[J]. Diabetologia, 2020, 63(1): 21-33. [↑](#footnote-ref-1)