

Introduction to Systematic reviews and evidence syntheses - Video 3: Developing a search – PubMed - transcript

I've drafted my protocol, so now I can start developing the search. These are some of the issues I want to think about.

The key words that authors might have used in the titles and abstracts of papers, so key words to describe each of the concepts within our PICO, so things like children, different synonyms for taxation, sugar sweetened beverages and so on. We'll also consider the database we're searching and whether it has a controlled vocabulary, sometimes called index terms. Here I'm referring to them as subject headings where the database has attached tags to each article record to describe the subject of that article. We'll often find that each article has, after a short period of time, between 10 and 15 subject headings assigned to it. For a broad search, it is essential that we combine both author keywords and subject headings within our search strategy.

We might also want to think whether we're going to limit our search by study type, randomised controlled trial, observational studies, qualitative research. There are pre-prepared and validated search filters, sometimes called hedges, that we can apply to our search to restrict to that type of study. In fact, there are hedges available for all kinds of subjects and one that's useful for this search is a hedge potentially for age groups. We'll see whether there is a hedge available for child studies. You'll often find them also for things like geographic areas.

Once we have our key words, we need to know how to combine them correctly and we'd use AND / OR generally to do this to make sure we have a sensitive search. We'll think about whether introducing limits might bias our search. When I talk of limits, things like language and date of publication. This is an iterative process. So you would test out your initial search and think - does it capture my key papers? Is there anything missing? If there is, how do I broaden it out to make sure it captures those papers? If I make it too broad, what percentage of irrelevant papers are coming up with a sensitive search? It's inevitable that we are going to have a large number of irrelevant papers, but we don't want that to be, too big that we couldn't manage the screening process. We want to introduce an element of precision into

our search. One way we can do this is by sharing your search with colleagues or a librarian who'd be able to advise you on the coverage of the search, how you might adapt it to either improve sensitivity or precision to increase the likelihood that you're capturing all the studies, but without having to screen lots of papers unnecessarily. They could also provide comments regarding accuracy, typing errors, that kind of thing as well. We'll review and finalise before applying it to multiple databases.

I'm just going to demonstrate now using PubMed, how I might go about developing a search.

You would do this on a database that you're familiar with, that you use frequently so that you're comfortable with the interface before you start testing out your search. I'll go into PubMed and start thinking about my search. If I've used PubMed previously that day, I just want to double check that there's nothing in my search history. I'm going to delete everything I've been doing all day. I start with a clean slate and I'll go back to the PubMed home page. The first thing that I'm going to do is to use the MeSH database. I've already mentioned subject headings and the subject headings within PubMed are called MeSH or Medical Subject Headings, and these are tags that are assigned to articles to describe their contents.

I'm going to start off with my first concept, which is sugar sweetened beverages. We won't search for papers at this stage, we will search for the index terms or subject headings. I see that it has got a subject heading that's useful for me. This is what it covers, a long list here of the kinds of papers where this tag will be used. I also notice it's only been used since 2020, so to get earlier papers, it's really important for me to also do a title and abstract or textword search within the advance search option to maximise my search retrieval. At this stage, I'll add to search builder. Sometimes if you find a search heading, you might want to see whether there are useful associated headings, you can do that with PubMed because the search terms are hyperlinked. So I can go up a level to see whether there are other beverages that I might want to add to my search box. I don't want to add alcohol or artificially sweetened beverages. However, energy drinks are quite often high in sugar. I might consider adding energy drinks to my search box. I click on the term, I'm going to add it to my search folder, but because I want papers indexed with this sugar sweetened

beverages or energy drinks, I need to change the operator here so that I get a sensitive search, I add to search builder with OR and search that.

I get 1100+ papers. In order to maximise retrieval, I am going to do a textword search, so I'll go in to my Advanced search. From that scroll down menu, I will select Text Word, so these are generally for keywords, title and abstract keywords and the like. We'll look at sweetened beverages, I'll use the asterisk for the plural.

Sometimes in the US, they're called soda, in the UK, we call them soft drinks, and I'll add energy drinks here. Add as many synonyms as you think appropriate. I'll add to my search box and click on Search. And I have 13,000+ in this set.

If I go back to my Advanced search, what I'll see is that I now have two searches on the same topic, but done in a slightly different way. I want to consolidate the results, using the Actions option here. I want to add query one OR query two. I want anything that's captured in either of these two searches to make it a really broad master set for sweetened beverages. I click on Search and I get it a few more by consolidating them. You'll see the difference in numbers here. Most would have been captured by number 2, but I did get a few additional papers by adding 1. I'm particularly looking at child consumers of sugar sweetened beverages. I mentioned Hedges before, ISSG is a good source of hedges, the entire INTERTASC sub-group for search filters. The URL will be in the handbook.

If we scroll down, you'll see that on the left hand side we have search filters with more of a design focus. But under other filters, we do have age filters. If you were to do this, you might have a look at some of these to see what would be most appropriate. As I'm using PubMed, I see that there is one here by Leclerc that's specific to PubMed. I link through to that article, I've read the article and I'm happy that this filter is sensitive enough for my purpose. So I'll copy the filter and I will paste directly into PubMed. I'm not gonna do anything more to it just as it is, I have my master set for children. I'm not even going to look at the MeSH term for children because this is a validated filter, it's been designed with maximising sensitivity in mind.

However, when I go to my next stage, which is looking at taxation, which is our intervention, I will do the subject heading search again. Back into the MeSH

database and look at taxation. Click on Search. I search for taxes here, which is the nearest match to taxation. It would also capture income tax, tax equity and tax exemption. If I'm happy to include that, I just add it to search builder. If I want to make it more specific, I would tick this box and say do not include MeSH terms found below this term in the MeSH hierarchy. I'm going to leave it as it is and add to Search Builder and click on Search PubMed.

Again, repeat the process with Text Word in my Advanced search. From the drop down, I look at **tax OR taxes OR taxation?** Now, previously, I did use the asterisk when I had a common word stem. **Beverage*** for beverage OR beverages. I could in theory, put **tax*** rather than typing all the key words out. But I'm concerned that **tax** starts other keywords. It would appear at the start of taxonomy, for example. In this case, I'm going much more specific. And the same with levy, **lev*** will get things like level which are outside the scope of my search.

I'll add it to my query box and click on Search. Again, back to my Advanced search to consolidate those two lines as a master search for taxation again, Actions add 6 OR 5. I'm building it up step by step.

Quite often for a systematic review, we would check our search at this stage to see how many results we get with our population terms and our intervention terms. Often we don't add the outcome terms unless we have one very specific outcome or if we have too many results. In this case, if I was looking at this systematic review, I'm looking at a reduction in consumption or purchasing, so multiple outcomes. I'm going to see at this stage, if I look at combining **children** and **sugar sweetened beverages** and **taxes**, does it give me a reasonable enough number of results that I'd be happy to screen? I'm going to **AND** the search terms together. So 3 is my master set for sugar sweetened beverages, 4 is the master set for children. Add with **AND** so when we have more than one different concept, we always combine with **AND** to narrow down, finally adding taxation. I click on Search and see what number of results I get.

I would say that in terms of a systematic review process where we sometimes screen several thousands or tens of thousands of papers, that 206 is actually a reasonable number. I wouldn't at this stage want to restrict further by adding a

separate search concept for consumption or reduction. I would be happy that I'm maximising the sensitivity of my search. I would be ready now to export my results in to reference management software before I went forward to duplicate the search on a different database.

Before I did that, I might actually just have a look at what the retrieval is like, check some of the titles and abstracts just to make sure that I'm not missing anything. I might want to look at some of these to see what else I want to add in. The way that it's highlighting key words can be really useful for me, for example, this one is talking about fiscal policy, not necessarily taxation. So I might want to add that into my group of synonyms. To go back, adapt and edit my search if I think that I need to either add terms for increased sensitivity or add another search concept to improve precision.

Please contact your local librarian if you have access to one to help you with this process. You'll see, my finalised search looks something like this. Next thing to do is to think about which other databases I'm going to look at to make sure that I'm capturing research from different clinical areas, from different disciplines. I'll choose my database, I might adapt the search for the screen or interface that we're going to use, and then I'll export the results into some kind of software so I can then deal with any duplicates. Here are some examples of databases that we might use within the health care environments if we were looking at systematic reviews. Alongside Medline, we almost always use Embase as another general health and medicine database. In this case, because it's around food nutrition, we'd add CINAHL because it's the main nursing allied health database. We're looking at an intervention, so Cochrane Library would be useful because it will retrieve anything related to trials or systematic reviews. Because it's a behavioural intervention, we might add in PsycINFO psychiatry and psychology databases that might have journals related to health behaviour within them.

Ask again, if you have access to a librarian, which are the most suitable databases for your subject area, because it will change from question to question. Once we've done this, we then move onto identifying grey literature. We'll think about additional search methods that might capture papers that our search will have missed.

Sometimes even when we try our best to do a sensitive search, we will still miss some papers.