

## Transcript

>>All right. Welcome, everybody. Welcome to NLM Office Hours. The goal behind NLM Office Hours is to give you a chance to learn more about NLM products and services, and importantly, to get your questions answered. We do record and post these sessions for broader and continuing access to the information provided, and all registrants will receive a link to the recording when it's available. NLM Office Hours is cohosted by the Network of the National Library of Medicine, the education and outreach arm of the NLM. We work to advance the health and well-being of everyone through improved health information literacy and information access. We work regionally through funding and programming, helping organizations within communities serve the public they know best. And we work nationally, like with today's program offering online educational programs for people who serve people. Visit [nlnm.gov](http://nlnm.gov) to learn more.

Today for NLM Office Hours we have Amanda Sawyer from the PubMed team at the National Center for Biotechnology Information at the National Library of Medicine. Amanda will be providing a demonstration of the new proximity search feature of PubMed, and then she will be joined by Jessica Chan and Kathi Canese, also from NCBI, and Alex Sticco, from our indexing section, to help answer your questions about PubMed. Assisting me in managing today's session are Molly Knapp, from the Network of the National Library of Medicine Training Office, and Mike Davidson and Michael Tahmasian, from the Office of Engagement and Training at NLM. Thank you all for being here.

So you have three ways to communicate during today's session. You may use the chat feature. Molly will be relaying questions verbally to our guests. You may raise your hand if you would like to unmute and ask questions verbally. And you may use the Reaction icons to give nonverbal feedback throughout the session. Also, just to note that Zoom does allow you to adjust your screen to increase the size of the presentation. Look for the View menu in the upper right. So without further ado, let me pass this on to Amanda to tell us about proximity searching.

>>Alright, thank you Kate. Let me pull up my slides here. Hi everyone. Like Kate said, my name is Amanda Sawyer and I work on the PubMed team at NCBI. I'm really excited to be talking about proximity today. We know this is something that has been requested for a long time in PubMed, so we're excited that it's offered now and to share that with you.

So just a quick overview. I'm going to go over what is proximity searching, specifically how do you do it in PubMed? We'll look at a few examples, but I want to leave most of the time today for your questions. So we're going to do just a few short examples and then move to your questions.

So proximity is a searching tool that allows you to search for terms that appear near each other in a PubMed record. Proximity searching can be particularly helpful when you're looking for concepts that might be represented in multiple ways, or to look for concepts in the literature that may not appear in a controlled vocabulary or PubMed's phrase index just yet.

So, for example, if we wanted to look for articles and citations about patient-physician relationships, a free search might not capture all of the relevant information because this concept can appear in multiple ways. For example, patient-physician relationship, the physician-patient relationship, the patient-parent-physician relationship, physician and patient relationship, and I'm sure that there are many more ways that this concept can be represented. So proximity is offering us a way to account for these variations without having to add a whole lot of phrases to our search or broadening by just doing keywords. So we'll take a look at a few examples in just a minute, but first let's take a look at the proximity syntax for PubMed.

Proximity searching in PubMed allows you to look for multiple terms appearing in any order within a specified distance of one another in the Title or the Title/Abstract fields. There are three main components to crafting a proximity search in PubMed: your search terms, the PubMed field you want to search in, and a number for the maximum number of words that you want to appear between your search terms. So first enclose your search terms in double quotes, and remember that you have to use two or more terms because proximity searches are based off of the distance between those terms. The rest of your proximity search syntax will after your search terms will be enclosed in brackets. Then you can enter the field you want to search in PubMed. Proximity searching is only available in the Title and Title/Abstract fields. After the field, you need a colon and a tilde [~]. Finally, you'll choose your N value or the maximum number of words you want to appear between your search terms. A higher N value gives you broader results, and a lower end value can help you narrow those results down.

So let's take a look at a few examples. Here I have a proximity search for citations where "hip" and "pain" appear in the Title/Abstract field with no more than two words between them. My search terms "hip pain" are enclosed in double quotes. I've chosen the Title/Abstract field and the N value in this search is 2. Our results might include citations where "hip pain" appears as an exact phrase, but this proximity search will also help us find other variations like "hip related pain" and "hip and groin pain." Remember that proximity returns citations where your search terms appear near each other in any order. So our "hip pain" search also returned this citation where pain appears before hip in the title, "Pain, Hip." And in the Abstract, "pain and hip."

Let's look at one more example, and this is one that's a little bit more specific to PubMed, an example where it's particularly helpful to use proximity on PubMed I think. As many of you know, PubMed uses a phrase index when you're searching for exact phrases, and this means that in cases where you're searching for a phrase that is not frequently appearing in the PubMed database yet, you may get the error message saying, "Quoted phrase not found in phrase index." And then you'll notice that automatic term mapping has translated your query. For example, in the search that I have here on this slide, the large phrase "cognitive impairment in multiple sclerosis patients" is not found in PubMed's phrase index, and I have the yellow warning at the top of my search results saying that it was not found. Automatic term mapping broke up my phrase and translated each of these keywords, and we ended up with 2642

citations. Maybe this was too broad for my purposes, so we can use a proximity search to narrow down those results.

When you set the N value in a proximity search to 0, PubMed will retrieve citations where your search terms appear directly next to each other. Remember, these terms can appear in any order, so it's not the same thing as forcing a phrase search, but it will narrow down our results pretty significantly. A proximity search for "cognitive impairment and multiple sclerosis patients" with an N value of 0 in the Title/Abstract field narrowed us down to just eight citations, where exact search terms appear directly next to each other. But I want to conclude this example with a reminder that proximity searching is another tool you can add to your searching toolkit, but it isn't intended to completely replace other searching techniques.

You may need to try searching for the same terms using a variety of techniques and compare those results to help you decide which option to use. So for example, if we take another look at our last searching scenario, we could run a similar search by combining two phrases that are in PubMed phrase index using the Boolean AND. So our search becomes "cognitive impairment AND multiple sclerosis patients," both in the Title/Abstract field. With this search, we retrieved 181 citations. It's narrower than the search we started with that was translated by automatic term mapping, but broader than the proximity search with an N value of 0. And we can see that from just these first few titles here that if we had stuck with that proximity search, we might have missed out on relevant content. And this is just going to depend on your specific scenario and why you're searching the literature. When you're deciding if a proximity search is the appropriate tool for your searching scenario, keep in mind that proximity searches can be combined with other search terms and other search techniques like phrase searching by using Boolean operators in PubMed.

And so we're going to get to your questions here in just a second, but I always want to point out that we do have a lot of documentation to help you. If you are trying to build a proximity search, we always recommend starting with the user guide. We've updated that with details about proximity searching for you. The Office of Engagement and Training at NLM has also produced a quick and interactive tutorial to help introduce PubMed's proximity searching. And you can find examples and FAQs about proximity searching in our Technical Bulletin announcement from when this feature was released last year. And we will add those links to the chat for you, and these slides will be made available later as well.

And I always have to mention that you are always welcome to write to the PubMed Help Desk with any PubMed questions you have, whether they're about proximity or any other feature. And we welcome your feedback and your suggestions for improving PubMed as well. When you open a case at [support.nlm.nih.gov/](https://support.nlm.nih.gov/) and write about PubMed, your questions are sent directly to a team that is trained to handle public questions and any time that you provide feedback or suggestions about PubMed that is sent to the NCBI PubMed team and all of our user suggestions are taken into consideration when we are planning for the development of PubMed.

And one final thing to mention. If you are trying to implement proximity or you're looking for help deciding what type of search techniques to use in PubMed, we also encourage you to reach out to your local Health Sciences librarian or to the expert searcher in your community for help deciding what is the appropriate technique to use for your scenario. And with that, I'm going to pass the screen back to Kate and to Molly, and I look forward to answering some of your questions.

>>Thank you so much, Amanda. Fantastic. We already have quite a few questions in chat about the details of proximity searching. So Molly, I pass this to you.

>>Yes, we have so many great questions and keep those questions coming because we have till 15 minutes before the top of the hour. So your first question is for Amanda and **Allison wants to know what is the maximum N value?**

>>Yeah, that's a great question. There is no maximum N value you can have as many-- You can have it as high as you would like. It is going to make your search really broad if you make it really high, and so that's maybe not always the best choice, but go ahead and play around with it, see what you get.

>>Exactly. OK. We have another question. And this is from a chat participant who asks **when can we expect proximity search capability with truncated keywords?**

>>Another good question. This is not something that we are able to offer at this time. It's based off of the technology that PubMed is built on. So that's why it's not a part of proximity searching. But again, write to the Help Desk. If this is something that would be useful to you, make sure that we have that down in our records so we can keep it in mind for the future. But as of right now, this is what we're able to offer with the technology that we have.

>>Great. OK, well, here's another question. **John wants to know which words are not included in the search count of N, maybe shorter than a certain length or stop words?**

>>Another good question, all words, and I'm going to double check with my colleagues on this and I'll come back to you, but all words are included in the N value, including stop words. And that's true whether it's the stop words that are between your double quotes for the search terms or in the N value. The stop words are counted as well. They're treated as normal terms when counting.

>>Do any of our other NLM colleagues want to chime in on this one?

>>Amanda's got it.

>>All right. Thanks, Kathi.

>>Alright, great. I want to pause a moment to read a nice comment that we got from Leslie in the chat that the Help Desk has been great whenever she has questions or notices issues. And thank you for your help and responsiveness. So yay. Here's another question from Galen. **Do**

**you need to write a Title/Abstract or can you use TIAB?** So do you want-- I'm sorry about that. Yeah, go ahead.

>>You can use the abbreviations if you'd like. You can write it out like I did in my slides. Or you can use those field abbreviations for TI or TIAB.

>>Oh, here's another good question from Tamara. **How are dashes handled in a proximity search?**

>>In general, the way that dashes-- words with dashes are stored in the PubMed database, the dash is converted to a space, so it's essentially ignored. So, and in general, when you're searching in PubMed, this is just kind of true whether it's a dash or any other special character. We recommend to replace those with a space just to make sure you're not messing with your search syntax too much.

>>OK, so I'm hearing replace a dash with a space might be a best practice there. OK, great. And here's a question from Suzanne. **What is a stop word?**

>>Stop words are words like "of," "the," PubMed will ignore these. If you're doing like a keyword search, and we have a section in the user guide about stop words, I think someone actually linked it in the chat already, so I would recommend going there to look at those, but in proximity or in a phrase search when you're putting it between those double quotes, those are treated as normal terms. There's also a list of all of PubMed stop words in that user guide section if you're curious about it.

>>Great. Yes, I see the stop words in the chat. The chat is flying, but we'll try to get that to you again. **Here's a question from Kerry, can you set a minimum N value? And I know you showed us zero what they mean is as in search results won't return the phrase when words are next to each other, only if there are words between the search terms. So can you set a minimum N value?**

>>No, the N value is the maximum number. That's an interesting question. I hadn't thought of it that way. Yeah, thank you for asking that.

>>And another question from Allison, **is proximity searching only available with field searching? And I think you covered this about. It's either the Title or the Title/Abstract field, right?**

>>Exactly, yes. It has to be in the Title or the Title/Abstract field. And I think if you were trying to do a pretty general proximity search looking through government records, I'd recommend using the Title/Abstract field. That's going to search as much of the record as possible.

>>Here's a question from John. **Can one word of the proximity search be in the title and the second word be in the abstract?** I think you just kind of touched upon this.

>>Another interesting question. If you're using the Title/Abstract field, I think that's going to depend on your N value. But let me check with someone and I will-- We'll come back to that at the end, answer some more questions.

>>OK, here's an easier question from Darlene, **is the user guide for PubMed download as a PDF?**

>>Another good question. I don't think so, you can always print to PDF. That would be the easiest way to do that.

>>OK, we'll check on that for y'all. And here's a question from Teresa. **Can you combine a proximity search statement with other Boolean operators?**

>>Yes, yes. And if you look in the user guide as well as that Technical Bulletin that we released, there are a few examples in both of those where we're combining proximity searches with other types of searches, and you can also combine multiple proximity searches together using Booleans.

>>Cool. I can't wait to try that. There's so many great options for finding different types of-- different ways to find the information that you seek in PubMed. It's almost a little overwhelming sometimes, so I just wanted to ask Kate if there was any hands up. Anything from the audience?

>>I am looking for any hands. And I do not see any. None yet. So, yeah, we'll just-- we may just need a minute to process the questions that we have so far. Be patient. Thank you.

>>I'm seeing a few comments in chat about the PDF version of a PubMed user guide. And first and foremost, we want to provide the most up-to-date information on PubMed. **So the most up to date version is always going to be online because it's kind of a live document, right?** But Rebecca Brown knows that the help guide, if you print it to PDF, is 56 pages long. So, there you go. Let's see. Here's a great question from Ann. **Any other new search options coming down the pipeline?**

>>This is the major change to search that we've had coming recently, but we can take a break and I can show you a few of the updates that we made to PubMed that are kind of adjacent to searching that I think that this audience in particular might be really interested in. So I'm sharing my PubMed search results. The first update that we made in the last month is that we have done a usability improvement to the display of search results when citations have really, really long lists of authors. So this is when a list of authors is hundreds to thousands long. We know that that was a pretty big pain point when you are trying to scroll through a lot of citations, so we've updated the search results when you're using the summary display format, author lists are now truncated after 1200 characters, followed by an ellipsis [...] that you see here at the bottom of this author list and a link that says see abstract for full author list. And I'll click on that just to show you where that takes you. It takes you to the abstract page. In practice, this change is applying to citations with around 100 or more authors. You'll notice that took a very long time to load. That's because this is an incredibly large list of authors. There's a lot of data. So that

would be a pain to scroll through. So 100 or more authors that makes up about .01% of citations in the PubMed database. And if you are curious about this, about why we did this or about what exactly is going on here, we have a Technical Bulletin announcement and someone can link that in the chat for you.

So that was just in the last month, that update and then the second usability update that we made was to the e-mail feature. So now when you are emailing citations to somebody else, there is an optional from field so that people aren't getting anonymous lists of citations in their inbox coming from NCBI. This field remains optional so that if you are sending them to yourself, you don't have to type your e-mail twice. And if you're signed into MyNCBI, just like in the past, both of these fields will be auto populated with your e-mail. So both of these were just released recently and hopefully those will solve some of those pain points that we know especially librarians have encountered with PubMed.

>>Well, both of those examples got a lot of love in the chat and applause and they have also generated more questions. So let's get going on that. Yeah, this this question is from hmayo and they ask, **"and" is both a stop word and Boolean operator and could you elaborate on that a little bit?**

>>Sure, when AND appears between your search terms, it's a Boolean to help you combine concepts. If you put it into the middle of a term, it's-- Basically, when you're searching in a record, "and" is considered a stop word and how it's stored. If you put AND into a search, it's probably going to be interpreted as a Boolean operator, and that's something that's important to keep in mind. In general, when you're using AND, OR, or NOT, I would recommend capitalizing them, but they're-- It's not always necessary in PubMed to capitalize things. Hopefully that helps a bit, but if that didn't answer your question fully, again, I would recommend writing to the Help Desk because sometimes it's easier to dig into the details of these really specific things that way. And I can get some help from our colleagues as well.

>>Great. OK, so here's a question from Eve. **Are you considering adding proximity search to the Text Word field and not just the Title/Abstract field?**

>>Not at this time, we are not. It's again, this is a feature of the technology that PubMed is built on, and this is what we're able to offer, and if Text Words is something that you're really interested in, we ask that you write to the Help Desk. Let us know so that we're aware that it's something that would be beneficial. But as of now, Title and Title/Abstract are the fields that we are able to offer with proximity.

>>So this is a similar question. **Sue wants to know if NLM's considering adding proximity search to Affiliation or Institution because that would really support scholarly communications work.** So another field that is ripe for that option.

>>That is good to hear. And again, you know, I'm hearing you here now, so I'm aware of it, but definitely put it in writing to us as well so that we have that record for the future.

>>Here's a historical question from Sarah. **Just curious what changed with the technology to allow proximity searching to be possible. For example, a while ago it seemed like it wasn't expected to happen.**

>>Right. So a few years ago, we completely rebuilt PubMed in the cloud. And this is one of the major benefits that came out of that. It is this new technology that PubMed is built on that's allowing this to exist. And as to why it took a few years to get it out after that release, well, it was a major effort just to rebuild PubMed in the cloud and there were other things that had to be taken care of first. So we're very excited that this is something that is now possible. In the past with the platform that PubMed was existing on, it was not possible.

>>So this question shifts gears a little bit. It's going back to talking about those truncated author lists. One of those new features, de Castell wants to know, **will the full list of authors be exported to RIS? So if you are exporting the Medline record, would you get the full list of authors? All hundreds of them?**

>>Sure. Just one clarification. PubMed, when you're exporting to a citation manager, we have .NBIB files. That's the only type of file that we offer for downloading. But yes, the full list of authors will be exported. The only thing that this is changing is that display on the search results list. Hopefully, that's helpful.

>>Kate, do we have any hands raised, any questions from the audience?

>>No hands. I do see a comment. I see a couple of folks requesting that the nursing and dental journal filters be restored to PubMed, so there's a suggestion for the PubMed team.

>>OK, in the meantime, here's another question from Ann. **Are you seeing patterns in the problems people are reporting with proximity searches? What are some, like typical bloopers you might encounter?**

>>I have not actually seen a lot of reports of bloopers yet, but I think that part of that is just that this is a new feature and people are still learning that we have it. So we often get people asking questions about phrases that aren't in the phrase index and we say try out proximity so people are excited to learn about it. Some of the things that I would expect that you might run into our syntax errors for getting the tilde [~] or the colon [:] or forgetting to close your brackets or trying to use a field that is not supported for proximity. And in those cases, we-- part of our usability testing for proximity was trying to make sure that the error messages that you receive are pretty specific. So hopefully if you're having trouble building a proximity search and it seems like you are getting a lot of errors or it's not working the way you want. Check out those yellow error messages first and then again, always write to us. If it's just really not making sense, we're happy to answer them. The other thing, just while we're on the question of proximity errors that I'll mention, is that if you try to use truncation with proximity, which is not supported, you will get an error message and the proximity operators will be completely ignored. So it'll be run

as essentially a phrase search with truncation in it. The proximity will be ignored. You'll also get the error that tells you that those proximity operators were ignored.

>>And I guess it's worth mentioning, **if you do notice a glitch, always write the help desk, right?**

>>Absolutely yes. Yeah. If something in PubMed is not working how you expect it, please let us know. I wanted to acknowledge a question that we that we received in chat from Leslie. She's still struggling with the N maximum and gave a great example. Would it retrieve "hip" and then three words and then "pain", three words, and then "treatment" or would it top out and honestly Leslie, we are going to verify with developer about that and then send it out in the Q&A. So congratulations for stumping the experts and we welcome more questions like that in chat. So please if you have a burning question, we still have a little bit of time. Now, in the in the meantime, I'd like to ask Kate a question while we're waiting for some more user questions to come in and that is **what is next for PubMed training?**

>>I'm so glad you asked that question, Molly, because we're really excited because in the next couple of weeks NLM is going to be debuting a new tutorial on topic searching in PubMed using the Medical Subject Headings. This tutorial is specifically written for expert searchers, so watch the NLM Technical Bulletin for an announcement when we make it live. And it's worth mentioning too, that the Network of the National Library of Medicine always has amazing programming coming up. In fact, we have the Health Information Symposium coming up, I believe next week. We have Black Maternal Health Week coming up a couple of weeks after that. Always tons of free training out there.

>>So we welcome some more questions in the chat box, but maybe we have exhausted them. OK, alright, here we go. Looks like some more people are chiming in here, so. John asked, and this is not a proximity search question, it's more of a PubMed question, a burning question.

**John has trouble easily toggling between MeSH and the Advanced Search page without bringing up a new Advanced Search page by using a second MeSH originated search. So how to mitigate those two tabs, MeSH versus the PubMed Advanced Search screen? Is there any solution right now?**

>>I don't have a particular solution for that. I definitely see that that's a pain point for users. I'm going to ask if Kate or Mike maybe has some advice in this area and if not, we can look into this.

>>Yeah, unfortunately I don't have a specific suggestion for that. I tend to toggle between the windows myself.

>>Yeah, so this is Mike Davidson. I'm also on the Training team and I've done a little bit of looking with that. And I think part of what you may be seeing is the fact that you know, as I think we all know, with the current version of PubMed, it's built on a different technology than the previous versions and then also the other what we call the Entrez databases, right, like MeSH and the Bioinformatics databases, etc. And PMC is now sort of moving in that in that

same direction of the new interface as well. And there are, I think you know, I think we've all experienced some glitches with things not connecting quite as cleanly as they used to because not everything's on the same technology right now. So as we move forward with developments from other databases and other things like that, that's something that we can certainly keep in mind to hopefully rectify.

>>Thanks for that. We do have some tips going on in chat too. Rebecca notes she has both bookmarked in her browser and can jump between the two. So Sue agrees with Rebecca. Now here's a question from Allison. This is about the syntax for the proximity searching. **What happens if you add a space before or after the tilde [~] Does it still work?** OK yeah, go ahead.

>>It should still work. And right after I saw that question I went and tested it with my own proximity search. I put a bunch of random spaces in. So yeah, works just fine. Add as many spaces as you like I suppose.

>>Great. OK. Alright, so there's some more chatter and chat about subject headings. All right.

>>I promised to come back, I think there were two things that I wanted to come back to. The first was-- Oh my goodness, looking through my notes. I apologize. Oh, the first was the question about can one of the terms be in the Title and the other in the Abstract, and since those are separate fields, you need them-- the terms need to be either both in the title or in the abstract. It can't be one of each. So just wanted to clarify the answer on that. They're separate fields, so it's not going to find just one in different fields if that makes sense.

And the second thing that I just want to clarify my answer on was "and" as a stop word and a Boolean and that is that when you put "and" into a search PubMed is always going to use that to group together concepts unless it's between the double quotes. So I just want to make sure that I'm really clear on that, that if you put AND whether it's lowercase or uppercase, but there's no quotes, then it's going to group together your concepts as a Boolean. When it's a phrase search or a proximity search, it's just treated as a regular term between those double quotes.

>>Hey, good to know. So if you-- so double quote, if you have a phrase that has "and" in it and the double quotes, it's going to not be it will not be treated as Boolean and it will be Boolean otherwise. OK, makes sense. And I'm kind of curious, you know, sometimes people search like chemical names. And **we talked about dashes, so other special characters like commas [,] and things like that. How do those work with proximity searching?**

>>Sure, and this is advice that I will say applies to proximity, it applies to any type of searching you're doing in PubMed. Special characters have special meaning in PubMed search syntax, and the most clear example of this are parentheses. So you'll see parentheses and chemical names and compounds and things like that. But if you were to put those into a PubMed search, PubMed uses those to group together, to nest your concepts. So when you are searching for anything that includes special characters in PubMed. Our general advice is to replace those

characters with a space so that you're not messing with your search syntax and ending up with things nested together for broken apart that shouldn't be, and if you're really interested in special characters and what's converted to a space and what has a special meaning in PubMed. There is a section of the user guide about that special character conversions as well.

>>Great. That is great. That is so good to know about. Thank you so much. Alright, let's try another question. This is kind of rhetorical. **Why does PubMed use a phrase index?**

>>A very good rhetorical question. PubMed uses a phrase index because with the technology we have, this is the most efficient solution to phrase searching we can offer. And it comes down to the fact that PubMed is a very large database. We have millions of daily users and this is the way that we're able to maintain system speed, system performance for all of our users. So what I will say is that if you're coming up against that phrase not found error a lot with something that you want to be able to do a phrase search for you can put a request into the Help Desk. We consider those requests whenever they come in. As long as it meets our minimum criteria, we will go ahead and add it to the phrase index. We also are looking through PubMed records automatically for new phrases, and those are added twice a month. So this is an ongoing process. The phrase index is not a static thing. It's always changing, being updated. It's millions of phrases strong at this point. And I know that I personally, and other members of the PubMed team as well, have fielded several requests to be added to the phrase index over the last few months, so keep them coming in. Feel free to write if there's something that you really would like to be added.

>>Thank you. Yes, **please send in new terms because I mean we have to crowdsource, right?**

>>Absolutely.

>>Yeah, well, we're coming up to the end of the session. And so I wanted to hand it over to Kate to kind of do some final thoughts and wrap up.

>>Thanks, Molly. All right, I'm going to share my screen again to bring up that slide that Amanda shared on getting help with proximity searching and I'm also going to put the links in chat again so that you're easily able to find this support material. So it is time to wrap up today's session. Thank you so much to Amanda, to Jessica, Kathi, and Alex for lending your expertise. Thank you to Mike and Michael for your support. And thank you to our audience for amazing questions and contributing to an informative session for everyone.