Current Strategies for Updating Systematic Reviews

Presenters:
Dr. Carlton Fong, Professor James Thomas and Dr. Chiara Arienti

Sponsored by AIR's Center on Knowledge Translation
Disability and Rehabilitation Research (KTDRR)
https://ktdrr.org/training/webcasts/webcast78/index.html

Transcript for audio/video file on YouTube:
https://youtu.be/F5Vl4KRT9Iw

JOANN STARKS: Hello, my name is Joann Starks, and I want to welcome you to today's webcast from the Center on KTDRR, which is housed in the Austin, Texas office of American Institutes for Research. I'm the training lead for KTDRR. And let me introduce myself as a white woman with gray hair and glasses, and I use she/her pronouns. I'd like to, now, introduce the Director of the Center on KTDRR, Dr. Kathleen Murphy.

KATHLEEN MURPHY: Hello, everyone. As Joann mentioned, my name is Kathleen Murphy. I'm a white woman with shoulder length blonde hair, and I'm wearing a red collared dress today.

As Joann mentioned, I direct the Center on KTDRR, which is housed at the American Institutes for Research, and I just wanted to say hello and welcome you all and especially thank the KTDRR staff. You just met Joann and Felice, who have worked on this. You may be wondering what the purpose, overall, of the center is and why are we offering this training, and it is to build the knowledge translation capacity of other grantees of our funder, the National Institute on Disability, Independent Living, and Rehabilitation Research, which most people call NIDILRR.

So, one thing, when people hear the phrase "knowledge translation," they don't always think about the importance of being aware of vetting and synthesizing available research before it's even included in any strategy to promote research use. So that's why we do this training, to highlight that part of the process of knowledge translation. And we do often present webcasts like this and archive them. So, later on, I'll put in the chat a link to our website where you can see other similar offerings that we've done in the past.

So, really, that's it. I just wanted to say hello and welcome. And back to you, Joann.

JOANN STARKS: Thank you, Kathleen. I also want to thank several AIR colleagues who have helped prepare this event. Felice Trirogoff is assisting with the technical aspects today, and other AIR staffers including Shoshana Rabinovsky, Tracy Bauman, Brian Litke, and Lee Nethercott have helped get everything together.

At the end of the presentations, today, I will ask you to provide your feedback through a brief evaluation form.

Today's webcast is Current Strategies for Updating Systematic Reviews, and it is a part of KTDRR's Evidence Synthesis series. This afternoon, we bring methodologists representing several international synthesis organizations to share their perspectives on the important process of updating a systematic review. I wanted to let you know that you can download accessible PDF files of today's presentations from the webcast information page. Just scroll down to Presentation Materials.

Next, I want to introduce our presenters. Dr. Carlton Fong, Editor of Campbell's Disability Coordinating Group, discusses the process implemented in an updated systematic review published in 2021. Professor James Thomas, Deputy Director of the EPPI-Centre, University College London, describes the latest tools and processes for updating, including the living systematic map of evidence on COVID-19. And Dr. Chiara Arienti, representing the Cochrane Rehabilitation fields, shares the process used with a rapid, living systematic review on rehabilitation intervention for the management of COVID-19.

We will start, today, by hearing from Carlton Fong, PhD, who is editor of the Campbell Collaboration, Disability Coordinating Group and is a co-author on several Campbell Systematic Reviews. Dr. Fong is an Assistant Professor in the graduate program in Developmental Education within the Department of Curriculum and Instruction at Texas State University. He examines the motivational, psychological, and instructional factors that influence success, achievement, and persistence in post-secondary education, primarily using research synthesis and meta-analytic techniques.

Our second presenter, Professor James Thomas, is Deputy Director at the EPPI-Centre. His research covers substantive disciplinary fields such as public health and education as well as computer and information science. He has written extensively on research synthesis, including methods for combining qualitative and quantitative research and reviews. His activities in computer science include implementing novel techniques and processes such as machine learning and crowdsourcing to improve the efficiency of systematic reviews.

Our final presenter is Dr. Chiara Arienti, the Coordinator for Cochrane Rehabilitation and of the Clinical Trial Unit of IRCCS at Don Carlo Gnocchi Foundation. She has experience in leading methodological work and supporting systematic reviews of rehabilitation interventions. She is also an editorial board member of JBI Evidence Implementation journal, a statistical advisor of European PRM Journal and a member of the GRADE Working Group.

Carlton, I'll ask you to briefly introduce yourself and go ahead with your presentation. And, when he's finished, I'll ask James Thomas to introduce himself before his presentation. Carlton, are you ready to go?

CARLTON FONG: I am. Thank you so much. Thank you for the invitation to present with you all today and for hosting this important event. Once again, my name is Carlton Fong. I will introduce you to me as an Asian male with short black hair, and I'm wearing a gray sweater.

And I'm really excited to share with you a recent update that we did to two systematic reviews and walk through the process that we engaged in to update two prior syntheses, and so it'll kind of be a case example and some of the rationale for why we undertook this systematic review update. On the slide, you'll see that I'm associated with the Campbell Collaboration as the Editor of the Disability Coordinating Group. You can see our Twitter handles as well, @campbellreviews and then my personal handle, @carlton_fong.

So I first want to start with describing what a systematic review update is and why it's important and why our field needs these updates to better understand the evidence that we use. So this isn't really a formal definition, but the way I like to think of an update is taking an existing review, updating that review and its procedures so that what's produced is a more recent snapshot of the available evidence.

Why would we undertake such an update? I think one rationale is reviews grow old, right? Actually, the day, the moment you complete your search strategy for available evidence to be included in your synthesis, it's already old. Right, there might be a study that gets published the next day, and that wouldn't be captured in your current systematic review. So there are continually being, studies are continually being produced, and the reviews of such evidence might become stale or might become outdated unless an update is conducted.

The second rationale is this is how science works. I think we are all experiencing that in the pandemic that we're currently struggling through right now, that science evolves, more evidence emerges, and we take action on that evidence. And so, as we think about making evidence-informed decisions, we really want the most up-to-date evidence in order to make those informed choices.

Rationale number three, in our case, the two reviews that we wanted to update, both of those were empty reviews. They were reviews with inclusion criteria. But when we searched for studies to meet those inclusion criteria, none of those studies met the criteria, so they were two empty reviews. And so we were aware that new evidence was being generated, and so we wanted to particularly update these reviews because they didn't have any evidence to really talk about. So we thought it was pertinent to go about doing an update, in our case.

So the update was recently published by myself and some other colleagues in 2021. The title is "Interventions for Improving Employment Outcomes for Persons with Autism Spectrum Disorders A Systematic Review Update." On the right, you'll see the citations for the two older reviews, Westbrook et al, 2013, and Westbrook et al, 2012.

Once again, our case was a little unique in that we decided to update two reviews and actually combine the scopes of those two reviews together. And so the review published in 2012, that one focused on adult populations, and the one published in 2013 focused on individuals transitioning from school, school-based programs, right, so kind of that secondary school transition. So the reviews were more or less the same in terms of the criteria for types of interventions, types of outcomes, types of disability, but the age range was kind of the distinguishing factor of these two reviews. So, for the update, we thought we could update both together, not have an age restriction that these two prior reviews did, and combine them in a larger review/update of the two reviews combined. Hopefully that made sense.

So how do we go about our process of updating these systematic reviews? Well, once again, this is just kind of a case example of how we approached it. We first reached out to the Campbell Collaboration, who had published the prior two reviews that we were updating. We talked to Coordinating Group Editors in the education field as well as the Editor-In-Chief to discuss our idea. Not only did we want to update these two reviews, but we wanted to do it on two reviews and combine them. So it was kind of a novel idea that we wanted to run by the leadership of the Campbell Collaboration.

We also confirmed with the editors that we would bypass the title and the protocol stages. So if you don't know how Campbell works, and a lot of these other organizations, there are various phases to the systematic review process, registering a title which describes the general purpose and scope of the synthesis, then a more detailed protocol phase, and then a systematic review phase with the entire review completed. And so since the title and protocol were more or less the same from the prior ones that were already done, we went ahead and bypassed those stages and agreed with the editors that we were going to combine the titles and protocols of these two prior reviews together in order to engage in this systematic review update.

We went ahead and, once again, it's kind of unique. We had two reviews, so we made sure we looked at the inclusion criteria and the search terms for both reviews and made the most inclusive joint search strategy possible for both of these reviews so they would encompass the scope of the two prior reviews. And, thankfully, with some financial support, which we'll describe in the Acknowledgment section at the end, and a new review team, a team of coders and synthesists, we began our update.

So the next step was that we updated the electronic database searches with, once again, the complete inclusivity from the two existing reviews. And then we reviewed all the titles and abstracts that we received after the search based on the following six criteria. It had to include some sort of experimental research design, participants of any age but with a diagnosis of autism, intervention focused on employment outcomes. The outcomes had to be the attainment of employment. And then publication status, we had no specific criteria. It could be unpublished or published, and there were no language restrictions as well.

So when we conducted our search, we took the date of the last year of the earliest review. So the earliest review was in 2012, and that search was done, I believe, in 2008. So we searched for all the evidence from 2008 to when we began this update, which I believe was 2018. So we searched for that new 10-year period that we hadn't looked at before and, from that, we received 20,053 potentially relevant studies identified and screened for retrieval.

We removed all those prior to 2008, like I mentioned, and that took off about 7,000 studies. And then, with the remaining 13,000 or so, we screened those abstracts, and we ruled out about 12,700 because they weren't empirical studies or they weren't about autism spectrum disorders and employment. We narrowed it down to close to 300 studies that were potentially relevant, and then we screened all of those full texts with greater scrutiny to see if they met all of our inclusion criteria.

We removed the lion's share of those studies, and we remained with three eligible studies for our systematic review update. Remember, the first two reviews that we were updating, those were both empty. Right, there were no eligible studies from those prior reviews, and so finding three, even though it sounds like a pretty small number, it actually was kind of exciting for us to see that there were studies being produced within the last 10 years from once we did the search.

Like any other systematic review, we took those three studies and we put that through the same coding procedures that we had planned for those first two reviews, where we had two research team members independently code studies for various characteristics and effect size information. Some of those characteristics you see on our slide on the right, where we coded for study characteristics, whether it was published or unpublished, the method and study design, where the study took place or the setting, aspects of the participants, characteristics about their age and the gender breakdown, details about the intervention-- right, what the intervention entailed for the individuals in these studies, and the outcomes being measured, effect size information, mostly odds ratios because they measured whether they were employed or not, and then risk of bias characteristics as well.

So what did we find? We found three studies, once again. They were all RCTs. The first was a study conducted in 2014, and they used the Project SEARCH approach with individuals with autism, and this involved a nine-month internship model. And their overall odds ratio for employment outcomes was 4.65, which is pretty large considering the effects of the study.

The second-- once again, this is kind of a quick overview-- but the second one was conducted by Smith et al in 2015, also using an RCT. And they employed a virtual reality job interview training with, also, a pretty large odds ratio as well, 3.43 on employment outcomes.

The third study that we located was another study by Wehman and colleagues, published in 2017, an RCT incorporating the Project SEARCH model that was examined before, but also introducing some Applied

Behavioral Analysis, or ABA techniques, as well. And that had a similar magnitude of odds ratio for employment, which is 3.89.

So, overall, we updated the search. We located three new studies that were newly conducted and disseminated since those two prior reviews, and we found that there were, overall, promising results from three studies. But evaluating the risk of bias, it was likely that there was risk of bias for all three of these studies as well.

So what can we take away from doing this systematic review update? Updating the review yielded new understanding. Once again, the two reviews were empty before. We had, really, little evidence to go off of about how can we intervene and support individuals with autism to have gainful employment. There were really no rigorous studies out there back in 2012 and 2013. But, by doing this update, we see that there is a growing interest and a growing trend in examining this issue with more rigor and with examining outcomes that focus on employment.

However, they were still a few studies with relatively small sample sizes and likely risk of bias. These studies, though, do suggest large effects, and I think it's promising to continue this line of work and to see potentially another update to see where the field goes.

There are still relatively few experimental studies focusing on this topic. Therefore, I think there's a greater need for greater funding in this area for high quality research and technical assistance to support the employment for individuals with autism.

I want to thank the generous support that we received to conduct this review update, the Center on Knowledge Translation for Employment Research at the American Institutes for Research. We also want to acknowledge two of our Search Specialists that helped us gather the thousands of titles and abstracts that we screened, Liz Scalia and David Pickup as well as other research assistants provided by AIR staff. So I want to thank you so much for your attention, and I'll go ahead and turn it over to James Thomas.

JOANN STARKS: Thank you very much, Carlton. Let's see if we do have any questions from the audience before we move ahead, and I think I did see one question come through in the chat.

KATHLEEN MURPHY: Sure thing, Joann. I believe you're looking for a question from Peggy King Sears, who is asking Carlton, "Is it better to exclude all except randomized controlled trials, even if it reduces the yield?"

CARLTON FONG: Yeah, that's a great question. And hi, Peggy. Thanks for joining.

Yeah, that's an important question. I think that this is, I think, still a debated topic, especially in education and education-related fields like disability, where randomized controlled trials are hard to come by and

tricky to employ. And so I think there is debate about whether QEDs, Quasi-Experimental Designs, should be included and how so.

I believe, in our case, we were open to quasi-experimental designs as well, but none met all of the other criteria as well. And so, while we didn't intentionally exclude everything else, we were open to the quasi-experimental designs and RCTs, but it just kind of ended up being that we have three RCTs ultimately be included in our update.

JOANN STARKS: OK, great. Thank you. Do we have any other questions? And you can always ask a question later, too, but if you have a question right now, feel free to bring it up.

I'm not seeing any, so I'm going to go ahead and move on, now, to our next presentation. So, James, will you please briefly introduce yourself and then go ahead with your presentation? Thanks.

JAMES THOMAS: Thank you very much, and thank you for the invitation to join you, for me, this evening, for others with you, this afternoon. I'm James Thomas, and I'm based in London, as you might be able to tell from my accent. I'm middle-aged white man. Hair's not yet all gray. Pronouns, he/him.

And it's really nice to see some familiar names or faces in the participant list, so nice to see you. And I'm going to be talking about some of the approaches we've been taking to updating reviews and in our recent work. OK, so I'm going to move on to the next slide.

And I'm based at UCL when I'm there in London, which has been a while since I've seen that building. And I've been doing systematic reviews to inform policy and practice for many years now, based at the EPPI-Centre And, really, what I've found over the years is a sort of frustration about how long it takes to find the records.

Carlton mentioned the number of, you know, sort of the phase of having to look through and find lots and lots of records and filter them down to find the ones that actually that you're interested in, and that awareness that we waste a lot of time looking at records that actually are not relevant really has been the prompt for some of the work that we've done to find new methods and new tools and new technologies to make that process more efficient. So that's really what has underpinned this work which, I should start off by saying, is under construction. It is work that we're still actively developing, and so this is by no means the finished product that I'm talking about this evening.

So what I wanted to sort of raise is really a question that I've had in my mind for a little while now, which is, really, are we starting to see opportunities for a new paradigm in searching, a new way of locating relevant research for systematic reviews, other types of evidence synthesis, and for maps? And I'm going to give you a little bit of background in terms of to explain why I might be thinking this, and then I'm going to illustrate what we've been doing using a case study, and that's our COVID-19 map of research

that we've been maintaining at the EPPI-Centre in collaboration with the University of York and certain reviews of dissemination there and also the London School of Hygiene and Tropical Medicine. So it's been a collaboration, and we've been doing this now for nearly two years.

So this question I've got around a new search paradigm, I think there are two what I'm calling enablers for this, and they're operating together to open newer possibilities here. So there's increasingly open bibliographic data. You've all, no doubt, heard about moves towards open science and open data and open research data, and open bibliographic data is part of that same picture. There was the launch of the open ALEKS database recently, at the beginning of the year, which is very much in line with that thinking, a database which is aiming to-- somewhat emulates the Library at Alexandria all those years ago and try and contain all the world's research.

So there's questions around increasing the open bibliographic data and the opportunities that that might provide and then, in conjunction with that, new automation technologies, machine learning. Some people call it artificial intelligence, though there's not always very much intelligence there. So I tend to call it machine learning at the moment, but certainly this idea that we've got these two possibilities which are opening up for us, more bibliographic data than we've, up until now, had the opportunity to work with, but then also new automation technologies to help us make sense of it.

So I'm sure those of you who have been involved in systematic reviews are very familiar with this stage of the review. It's the finding the eligible studies. What we typically do is we find that the research we're interested in is scattered or stored or scattered and stored in multiple databases and websites, and what we have to do when we're doing our review is go to all these different databases, websites, and retrieve everything that might be relevant.

And we use Boolean searches to retrieve the most likely relevant into our local database. And, because we've come from lots and lots of different websites, we end up with lots of duplicates, so we have to deduplicate the records, and then we sit down and we screen through them, often thousands and thousands of records, most of which are completely irrelevant. But because we don't want to miss the relevant records, we have to look through them, and this is very time-consuming and it is just the way it is.

So why is it like this? Well, it's like this partly because of, for historic reasons, really. The various organizations have indexed the research in their area. And it costs money to sit there and index research, and so various business models have developed which have really been around putting a certain field's content into a database and then charging people to access the database.

And so what has now emerged are these different subject's domain-specific databases which are driven by commercial imperatives to guard their content. It sounds slightly bizarre, but to sort of almost keep the research they contain secret unless you're paying for it. So that's why we end up looking at lots and lots of different places. It's kind of historic, and it's the way that the databases have been set up.

But, as I mentioned in the intro, there's been this movement towards more open data, and what that's enabled to emerge are more comprehensive sources. So I've mentioned open ALEKS. CrossRef is part of this. There's a database called Dimensions. Google Scholar to some extent, though it's not open data. But what they are all doing is publishing bibliographic data at scale. And Microsoft deserves some of the credit for this because it developed this thing called Microsoft Academic, which was really the biggest open source of bibliographic material that we had until it closed at the end of last year.

And thinking with a sort of economic hat on, we can think about bibliographic data now as becoming commoditized as in that you can get the same content from multiple sources now. You don't have to go to just, for example, Web of Science to get particular content from there. You can get some of the same content from elsewhere.

So what this is starting to mean, essentially, is that we can ingest all the world's published research into a local database. If you've got a big enough hard drive on your computer, you can actually download the lot. You can download the whole of open ALEKS onto your local computer.

However, what does that mean? It means, OK, so we can maybe have all of the world's research stored in these comprehensive repositories, and anyone can do it, but anybody who's needed to work with large databases knows that that's only part of the story. So I can download 200 to 250 million records onto my computer, just about. How on Earth, then, do I go about finding what I'm looking for in amongst all of that noise? And that's, of course, where the machine learning comes in.

And so, within our EPPI-Reviewer system, we've been developing this workflow which is aiming to keep a constant surveillance of the literature as it's published. So, using cloud compute resources, we've got a copy of the open ALEKS database. Before this year, we had the Microsoft Academic database, all of the records, almost all of the world's research, sitting there in a database and new papers arrive every couple of weeks or so.

So we've got 200 million, 250 million records, depending on which database we're thinking about, and it increases in size. 600,000, a million new records arrive every couple of weeks to a month. So great in that we've got all of this research and potentially we've got the ability to keep on top of the research as new research is published, but we've obviously got a very big signal-to-noise problem. There's an awful lot of research out there which is not relevant to most of the reviews that we're doing.

And so, within the system, we've got individual reviews. And what we do is we build machine learning models for each review, for the machine essentially to learn the scope of the review. And then, as new papers arrive in the system, they're fed through these machine learning models. And we're using something called a distance algorithm which basically gives us a number which tells us how similar or different every new paper is to every review that we're looking at, and new papers which have got a

sufficiently high relevance score are then automatically puts into the reviews that we're updating in this way.

And then people, reviewers, people can come along and judge for themselves as to whether or not those new papers are indeed relevant or actually whether they're not relevant. And we obviously, being systematic reviewers, we like to be comprehensive, and so we'd rather see more irrelevant records than not, if that means that we're going to include everything that we should be including.

The big picture, the fact that you can now get all of the world's research and you can use machine learning to, up to a point anyway, to be able to say this is relevant for this review, this is relevant for this review. And what I think is quite important about the context that we're thinking of, at the moment, this evening, in this meeting, is around updating. So when you've got a systematic review, a set of studies, a systematic map already in place, you've got lots of information about the scope of what you're looking for, which can be used for machine learning. And that's why this process is particularly suitable for updating reviews, updating maps.

So we undertook some research on this, and we did it within a map of COVID-19 research. And our overarching research questions was, well, OK, so how can we use these new tools to maintain this map?

And there were two research questions, really, that we asked. One was, does this big data set actually contain the records that we're interested in? If it doesn't, then there's a problem there and we need to go and either look at other databases and we're kind of edging ourselves back into the old paradigm of looking in lots of different places. But, if it does, then, OK, that's good.

But then, of course, we've got all of these millions of records. Can we efficiently identify the records that we need for our systematic reviews in this big, big data set? So two questions-- all the records are there? If they are there, can we find them?

And this map of research started at the end of February, beginning of March 2020, and it's been updated almost every week. And, to start with, it was populated using normal searches of Embase and PubMed, comprehensive data sets. But, as a large team, we were doing it for the Department of Health in this country and putting it online for other people to use as well. So we had the resource to be able to work on this large, large project.

And we conducted something which we've called the Octopus Study because it's an eight-arm cost-effectiveness study looking at different options for maintaining this map. I'm not going to look at every single arm in detail. I'm going to give you the top-level picture, but what we did was we collected lots of data about how long it was taking us to do things manually, all of the different tasks, and we collected data for four weeks back in June/July 2020. And we compared Embase and PubMed as our main search

sources against the Microsoft Academic stroke, the open ALEKS as a single search source, and then we also compared using different machine learning options within that updating workflow.

I think I might just skip on to the results because I don't want to sort of bog us down in the detail. Essentially, what we found was-- we've got various different arms here, and I've just highlighted a couple of arms for your interest here. Really, the first arm is our main comparison, our completely manual work flow which we did for weeks and weeks and weeks.

And what we found here was we searched MEDLINE and Embase and, because of the fact-- and this was a little bit surprising to me-- that MEDLINE and Embase actually didn't contain as much of the relevant research as we thought, our so-called gold standard sources only had 83% recall. So Microsoft Academic, at the time, had more relevant research units than the standard sources, which was really interesting to us. And what we were finding, also, was there was more non-English language literature coming through from the Microsoft Academic route, too. So that was an important finding straightaway for us.

Microsoft Academic potentially had more relevant information. When we started using the machine learning systems as well-- remember, that big signal-to-noise ratio that I talked about-- we were able to come up with a workflow which gave us a higher recall and a higher precision than we were getting using the manual processes. So that's the big-picture findings. And if you want to look at this in more detail, there's a paper on Wellcome Open Research platform, where we've got all the details for all the arms.

But the team carried on, still classifying the research into lots of different categories, and we have over 100,000 records in there now, categorized into diagnosis, health impacts, treatments, that kind of thing. So we had lots more data for machine learning, and we use a BERT model, now, which is something which a lot of computer scientists have been paying a lot of attention to recently, in order to automatically categorize things into those categories. So we've got what I was talking about before, that workflow of all of the research in the world, machine learning to go into the map-- as in, is it relevant, is it not relevant-- and then we have another lot of machine learning to decide which category something belongs to.

So there's an awful lot of technology here, a big comprehensive data set goes through one machine learning classifier to decide whether something is actually about COVID-19 or not, goes through to another robot, another classifier which decides which category something belongs to. And where the robots, where the machine can't decide, where it's less-- where the text is more ambiguous, the record then goes through to a human decision to decide on relevance and category. But what we've been able to do with this workflow is essentially reduce the number of people and the amount of time that it's taking us to maintain this and still actually, now, maintain a higher throughput of references through the workflow.

So, in conclusion, we conducted this study. We didn't know what we were going to find. We found some things which were surprising to us, and our traditional search sources had lower recall than we expected. And what's really critical, here, is it works so well because we had so much training data, high-quality

data, accurate data from humans. Machines can't do this kind of thing without all of that input from people.

And we're still working on it, but I think one of the things, the lessons that we've learned from this, also, is the importance of the technology people working with the reviewers, the systematic reviewers, to both in sort of dialogue to determine what appropriate levels of accuracy might be, but also to determine what, which direction the tool development should go in. So it's been a lot of technical work but, actually, what it's enabled us to do is maintain this surveillance of COVID-19, now approaching two years of it. And potentially we can just keep this going, now, with relatively little manual input compared with the beginning because of the automation that has been evolving over the last couple of years.

So I think I'm going to stop there and say thank you. And are there any questions?

JOANN STARKS: Thank you very much, James. We do have a couple of questions that have come in. Excuse me. This was a question from Joshua, early on, and he asked, "When you say study, do you mean an abstract or the full study details?"

JAMES THOMAS: That's a good question. I mean the former. So it's a report of a study, and what we're looking at mostly are titles and abstracts, though the manual work involves going in, looking at the full text reports when we can't make a decision based on the titles and the abstracts.

But, yeah, the big difference between our map and the Cochrane COVID-19 Study Register is that Cochrane spent a lot of work in what they call studifying the records. So they're using the trials registry databases, and they're linking all of the publications where there is a trials registry record to that trial's registry record or linking linked publications to one another. So it's sort of like the unit of analysis there is the study, whereas, strictly speaking, the unit of analysis for us is the bibliographic record.

JOANN STARKS: OK, thank you. Excuse me, we have another question from Jessica. "Are there any AI programs keeping track of when search results for research articles do not yield any results so that it can learn different ways of finding reference materials?"

JAMES THOMAS: That's an interesting one because it's actually something that we're encountering in another project, which is where machine learning is trying to find results and is actually having to recognize the difference between papers which do contain results and those which don't, but also the difference between papers which should contain results and those that shouldn't. So protocols are an obvious example of a report. You see lots and lots of protocols out there which, of course, you wouldn't expect to contain results.

Results are really hard for machine learning to accurately determine, at the moment. There are these tools called sentiment analysis, which some people have used, looking at the abstracts and looking broadly

about whether or not the abstract's got an upbeat, positive vibe about it or whether it's a bit sort of gloomy about the finding of the research. But I think it's fair to say, at the moment, that there's no really good tool that can just look at a research report and accurately extract the research results.

But, having said that, it's quite a hard task for people, too. There's lots and lots of errors made in the extraction of results in systematic reviews, too. I hope that's answered your question.

Yeah, and I think the other thing on that is that it sort of talks to the value of the maps and the value of the mapping process. And we've been talking, and the NIHR and other funding organizations have been looking at them, too, of course, because you can use them to identify gaps in research and places where, actually, you need to commission new research in order to generate findings and to fill uncertainties.

JOANN STARKS: Thank you. We had a comment. "That's really cool, James." Thought you'd like that.

JAMES THOMAS: Oh, thanks, Mark.

JOANN STARKS: And then, also, we have another question. "This is excellent work. How do you also capture gray or unpublished literature? What percent of papers is ML unsure of that requires human intervention?"

JAMES THOMAS: That's a good question. Yeah. No, gray and unpublished literature, we were finding that the comprehensive database was doing better at that and also doing better at addressing some geographical bias that you see in PubMed, for example, and other big databases.

I did a very-- I mean, this isn't a proper research project. I didn't write protocols. I just did it one afternoon, just for fun, but I did an analysis of the longitude of the institutions of the first author of publications in the map, and I found that the PubMed records had a larger longitude than the records that we got from the wider Microsoft Academic feed. So basically the average longitude of the PubMed records was just to the west of Spain, I think. And for the non-PubMed, for the Microsoft Academic, it was somewhere towards Russia.

So it's not much, but it sort of indicates that there are all sorts of biases in how we find and how we store research. Whether or not something has a DOI and whether it's classed as published or not is one of them. One of the challenges, actually, through this period has been what do you do with preprints and all of that kind of thing.

Oh, now why was Microsoft Academic discontinued? I think a lot of people have asked that question, probably including the team itself that put it together. I don't think it was a good decision.

Having said that, the people that open ALEKS have just jumped in with huge enthusiasm and energy to pick up where they've left off. So the Microsoft Academic team did a great job, and hopefully the signs

are good that open ALEKS will just carry on. And the whole, the way that they're doing it in sort of like the spirit of open data, but also in terms of linked data, they're linking the concepts in their data set through into Wikipedia and Wikidata, and so I think there are all sorts of opportunities that will open up over the next few years as we get data sets linked in increasingly conceptually useful ways. And I think that's one of the things that we could maybe participate in, as systematic reviewers, is in terms of just helping that movement to organize the literature in a way that ultimate users of the research would find useful.

JOANN STARKS: OK. Well, thank you very much. If anyone else has a question for either Carlton or James, please feel free to go ahead and bring it up right now in the chat box.

OK. Thanks again, James, and now let's go to our final presentation. Chiara, can you please introduce yourself and then begin your presentation?

CHIARA ARIENTI: Yes. Thank you, and good morning, good evening, whatever you are. I'm Chiara Arienti, a white woman with brown hairs, and it's fine for me the she/her pronouns. I'm the Coordinator of Cochrane Rehabilitation, and today I'm going to show you a very big and important project of Cochrane Rehabilitation on the management of COVID-19 disease and how and what is the role of Cochrane Rehabilitation in this topic.

As we know, in March 2020, the World Health Organization declared the health emergency relating to the COVID-19 outbreak. During this period, it was very important to manage the situation, the emergency, and so a lot of people produced a lot of evidence. So, also, Cochrane rehabilitation wanted to give its support to give a timely and updated answer and information about evidence in this important situation.

So we set, we started with a very important project, important action, the Rehabilitation-COVID-19 Evidence-Based Response action, and it included first projects. The first was the Rapid Living Systematic Reviews on Rehabilitation COVID-19, an interactive living evidence map on rehabilitation and COVID-19. And in collaboration with the WHO about the definition, the Rehabilitation Research Framework for Patients with COVID-19. And that, again, together to Cochrane Library, we performed a special collection dedicated to the coronavirus and rehabilitation.

Today, I will show you deeply the first project and the normative methodology that we used for the conduct of the Rapid Living Systematic Reviews. And, at the end of my presentation, I will show you rapidly the last three projects.

So the Rapid Living Systematic Reviews on rehabilitation and COVID-19, up to now, included three main editions, three main papers. And, for each papers, we included a series of updates. At the beginning, during the 2020 year, we updated the rapid review each month. Later, during the '20-'21st year, we updated the review every two months.

Now, I will show you this specific methodology and why we can call it innovative. First of all, we combined two specific methodologies, the rapid systematic review together to the living systematic review. Rapid means a form of knowledge synthesis that accelerates the process of conducting traditional systematic reviews. And, during pandemic, it was very, very important to be rapid, to give rapid answer. The second step was to keep this rapid review living. That means systematic review which is continually updated. Indeed, at the end, up to now, we published 14 updates of this rapid and living systematic review.

Now, I will show you the evolution of the methodology. Before March 2020, we started with a very simple systematic review, just a general collection of the main papers on COVID-19 and rehabilitation, and we published it just to give our rapid support. Then we checked the production of the evidence in COVID-19, and we decided to improve the methodology because, at that time, we needed to have a more strong and rigorous methodology to give a more rigorous information about the decision-making on the management of COVID-19.

So we had the first step. The first step was to establish an international multi-professional steering committee. We involved any information specially to start with a rigorous search strategy for the collection of the evidence, and then we strengthened the internal process. Indeed, we divided our internal team in two teams, the clinical working group and the methodological working group that, together, the international Multi-professional Steering Committee, lead the rapid and living systematic reviews.

After that, so this first improvement was related to a more rigorous process about the screening and selection of the papers. Then we started the collaboration with the WHO Rehabilitation Program to define the specific research questions related to the COVID-19 because it was very important to focus the problem related to the COVID-19 for the rehabilitation, so what is the role of the rehabilitation inside COVID-19.

So we created-- yes, we set the list of research questions and we prioritized these research questions inside our systematic reviews. So we, after a Delphi Round process, we define that the epidemiology of limitation of functioning and disability one of our priorities. The second one was to set the evidence on rehabilitation for COVID-19 at individual level. So a micro-level means at for the population, for the patients. Then we set to identify the evidence at the meso-level, and so means inside the organization of the hospitals, Rehabilitation Center, and so on. And then to identify evidence at the system level, at a macro level, and this means for the health system level for each country.

So, with these research questions, we identify the epidemiology of the COVID-19 to identify the problems related to this disease. And then we wanted to identify three specific evidence for three specific, three different levels-- patient level, hospital level, and systematic level. System level, sorry.

Then, after that, we started with declassification of the papers selected using these priorities and for specific study design for each priorities. So the research questions was the prevalence, to identify the prevalence and/or the characteristic of emerging disability after COVID-19, the rehabilitation approaches dedicated to the COVID-19 patients and eventually-- but it was very early for that time-- to identify the efficacy and effectiveness of rehabilitation intervention for the management of COVID-19. Then we identified the organization of rehabilitation service after COVID-19, the impact of COVID-19 on disease of rehabilitative interest, and the late complications that may be of rehabilitation interest. So we, using the priorities defined with the WHO, we identified these five research questions, and we selected the papers using these specific research questions.

Then, at the beginning, since we notified a lot of production of evidence with a huge heterogeneity of the study design, we wanted to exclude the expert opinion and all papers not reporting patients' data because, at the beginning, the most of the papers was just expert opinion because there was not the time to produce more rigorous primary studies. So, in this context, there were a lot of papers that describe the expert opinion. So, in this moment, in that moment, we decided to exclude the expert opinions and to include the observational studies, intervention studies, experimental studies-- but there were very few experimental studies, and some systematic reviews, but there were also a few systematic reviews at that time.

So, after that, we decided to use a specific sort of evaluation of the methodological quality of these studies. Since the level of evidence was very low, we decided to use the levels of evidence performed by the Oxford Center for Evidence-Based Medicine table, OK, and this table provides just a classification following the level of evidence perform the classification of the study design. And so this was the first step to improve the methodology of our rapid review.

Then we performed a data extraction. So this step was a more specific selection, applying the priorities and the research question. We introduced the critical price at using the level of evidence table. And then we used a specific data extraction, and we extracted the data about the type of study design; the type of the research question; the COVID-19 phase; the limitation of the functioning, disability, or rehabilitation interest and the type of rehabilitation service; the PICO elements; and the main findings.

Obviously, it was not still possible to perform a statistical analysis, and I mean a meta-analysis, for this specific systematic reviews because, up to now, we didn't-- we don't still have experimental studies, so we cannot put together the data of other types of study design. So, in this moment, we just can describe the main findings.

At the end of the day, during the third edition, we introduce this important step. That is, we passed from the classification of level of evidence to a critical appraisal. And for this, we used two specific tools. The first is the methodological checklist of Joanna Briggs Institute, and this checklist gave just a methodological quality of the study design for the observational studies.

For the randomized controlled trials, and up to now, we included just three randomized controlled trials. We used the Cochrane Risk of Bias tool for the evaluation of the risk of bias.

And, since there were a very heterogeneity of the study design and how the authors call the study design, we needed to define which kind of study design classification to use. And so we decided to use this report, the Methods Research Report. That is a document that defines a classification of study design in systematic reviews of intervention and exposure. This means that this risk classification, we are able to classify observational study and experimental studies. This support us to identify the correct study design and to use the correct tools, checklist, or risk of bias tool for the evaluation of methodology, the methodological quality.

Now, here, I just show you an example of our work and how we have managed all this information inside our rapid living systematic review. And, as you can see, for each study design, we identify the specific research question. And, for example, here, we can see that we have, for the epidemiology prevalence research question, we have 22 cross-sectional studies, a lot of studies, observational studies of prevalence. This is the most information that we found inside our rapid and living the systematic reviews.

And then you can see that, for example, for the micro-level, at the patients level, we found five randomized controlled trials at 2 August 2021. Not many, but is just an important information, I think. At the end, you can see at the end, the last row of the table, that, in the third edition, we found 8 randomized controlled trials, 35 cross-sectional study designs, 22 cohort studies, 3 non-randomized controlled trials, for a total of 68 studies.

Results, yes, up to now, not 13 but 14 papers have been published in European Journal, and all the results have been also published in our website in different formats. The first is a mapping that summarizes all the information according to the COVID-19 research topics agreed upon with the WHO Rehabilitation Program. Then we synthesize these results in dynamic tables-- you can see this dynamic tables, always, on our website-- and a geographic map that shows where the evidence is produced. All these three tables and map, you can see that on our website.

And now I show you just what you can see on our website. The first is the evidence map, and the second one is the dynamic table. With the evidence map, you can see a sort of overview of how many studies have been published for each research question. And when you have identified a specific research question and the specific studies, you can see the dynamic table, all details of each studies included in evidence map and in our rapid and systematic reviews.

What we can say about all this work? That the quality of evidence is still low. We need of more rigorous primary studies to give more appropriate information on how to manage this disease, the COVID-19 disease. The body of evidence is heterogeneous and rapidly growing.

Overall, at the beginning of the outbreak, at the beginning, the most of this evidence was performed by case reports, case series, and expert opinions. And now it's changing because, as I showed you before, we are a lot of observational studies. OK.

So far, we have only eight randomized controlled trials. And, in this moment, it's quite difficult, if not impossible, to draw any efficacy conclusion. In this moment, we cannot say if the rehabilitation intervention is efficacy on the management of COVID-19 disease, if with the rehabilitation intervention we can manage the COVID-19 consequences.

And so we continue our work because we believe that it is very important and we still need to know how manage better the consequences of COVID-19. And we also think that the management of COVID-19 consequences is a priority of the rehabilitation field more than other biomedical fields.

Now, very rapidly, I show you the additional REH-COVER actions. How I mentioned before, we collaborate with the Cochrane Library to perform a special collection. And you can see, always, this collection on our website, and obviously we keep, also, this updated.

Then we performed two data projects, a series of rapid reviews in collaboration with the WHO that are a sort of proxy of how we can use the rehabilitation intervention to treat the COVID-19 consequences or the different phases of the COVID-19. And, how I already said, we collaborated with the WHO for the publication of the priorities for the researcher in COVID-19. Now, we are working with the WHO for important guide on the post-COVID-19 condition and what is the role of the rehabilitation inside this specific context.

Here, I left you some references of our work. And if you have any questions, doubts, or just information, you can write me to my email or eventually to the Cochrane Rehabilitation email. All this work is online on our website. Thank you so much for giving me the opportunity to show you this important project of Cochrane Rehabilitation.

JOANN STARKS: Well, thank you so very much, Chiara, for your presentation today. It's really quite a project to address the important topic of COVID-19.

I want to thank all our presenters Carlton Fong, James Thomas, and Chiara Arienti, for sharing their experiences and perspectives in updating systematic reviews. If you have any questions for these researchers, their contact information is included in the presentation slides, which you can download from the KTDRR website. Go to ktdrr.org/training/webcasts and look for Webcast 78.

We would very much appreciate your feedback using the brief evaluation survey form found at the link on this slide. And please visit our website at www.ktdrr.org. We have many online resources related to

systematic reviews and evidence synthesis, including the archive of past webcasts as well as materials on knowledge translation.

So, to end today, I want to thank our presenters and everyone for participating this afternoon. Once again, we thank NIDILRR for their support for these webcasts and other activities of the Center on KTDRR. We look forward to seeing you at our next event.