**Online Workshop: Qualitative Research Synthesis**

**Additional Materials for Session 1:**

**Steps for Preparing a Qualitative Systematic Review**

Presenter:

Karin Hannes, PhD

A webinar sponsored by SEDL’s Center on Knowledge Translation for

Disability and Rehabilitation Research (KTDRR)

Edited Transcript for YouTube Video- http://youtu.be/u5eSx1XKRUY

Karin Hannes For the final part of this presentation which is probably the most boring part as well – I’m sorry for that – I’d like to run you through a couple of slides outlining how to protocolize a qualitative evidence synthesis. I’m borrowing some slides from many of my colleagues working in the methods group, including slides from Jenny Popay, Andrew Booth, and Janet Harris who have inspired me a lot in the last years.

Just defining what a protocol is, let me just pick out the most comprehensive definition of a protocol portrayed by the Cochrane Collaboration. It’s a plan or sets of steps to be followed in a study. A protocol for a systematic review should, first of all, describe the rationale for the review, the objectives, and the specific methods that will be used to locate, select, and critically appraise studies. It also outlines details about how to collect and how you plan to analyse data from the included studies.

So, most protocols follow some sort of linear outline that starts with the scoping of the literature; then outlines rationale for the review. It describes the condition, situation you are interested in. It describes the intervention or your phenomenon of interest; and above all, it’s required to actually state why it is important to do a particular review, and that can be theoretically inspired or effectively inspired. It then sets out to formulate a question and it sets objectives. It works with predefined inclusion criteria that evaluate the relevance of a particular study that has been retrieved through search for the objectives of the review and the criteria mainly linking to population you’re interested in, again your phenomenon, and the sort of studies you intend to include. You then outline your search strategy in terms of which databases you want to use and which key terms you think of using to retrieve relevant articles; and it stops with actually outlining methods for data collection and analysis, and with anticipating of how your results will visually be presented.

Now, the difference between quantitative and qualitative protocols is that quantitative protocols are always linear. Clearly, protocols in qualitative evidence synthesis are seldom as linear as they are presented in the previous outline, or even in the reports people write, because many of these reviews opt actually for non-comprehensive or maybe even purposeful samples of literature based on initial set of key papers from which the rest of the relevant papers is then generated for some sort of snowball sample. So the questions of a qualitative evidence synthesis may even be adapted based on initial review findings and inclusion criteria. So what we actually do is we reconstruct, or rather iterative logic into some sort of linear structure that is required in the most researched reports. This is not to say that qualitative review have never used a linear approach; only that we are convinced that it is not the best way for every sort of review, and that’s an important thing to know, I guess. The iterative process is actually the cause of trying to develop theory, and theory cannot be developed from some sort of fixed perspective; theory is built on components that are proven in different phases of the review; hence, that explains the iterative process that we often go to.

So let us start with this initial phase of scoping the literature; that’s pretty much the same concept for effectiveness reviews and qualitative evidence synthesis. We kind of search using a first set of key terms that we think is relevant to generate the right articles for our review, and it kind of gives us an overview of the amount of literature; the diversity in the population, interventions, and outcomes we come across; and it gives us an indication on how we can define our review question. So for effectiveness review, it kind of sets out priori limits for the inclusion criteria; and the thing that effectiveness review has done is that they don’t deviate anymore once these criteria or elements of population, intervention, and outcome are defined. So they stick to the same definitions of these concepts through route screening of the articles. Qualitative evidence, on the other side, actually uses the scoping literature already as a starting point to start analyzing data, and then the information from the scoping review becomes part of the key papers that are extracted.

Providing a rationale in an effectiveness review really relates to describing the condition of a situation, describing the intervention you’re interested in, and justifying the importance of compiling evidence of effectiveness. The same counts for qualitative evidence synthesis. We rather describe experiences and attitudes, behaviors related to a certain condition or a situation, or even the reactions of authority through interventions that have already been tested out and need to be improved. So that’s a way of trying to justify why it’s important to compile qualitative evidence that may explain heterogeneity, for example, in effectiveness reviews.

So the Lewin and Glenton reviews, the ones on lay health workers, I’ve incorporated an example of how they actually positioned their rationale; and for the effectiveness part, it was situated in chronic shortage of health workers, increasing the need for treatment; and into task shifting to alleviate the demand from doctors and nurses who are probably heavily loaded. So that’s why this review needed to be done to actually explore the potential of these lay health workers. The qualitative evidence synthesis started from the same situation, but the rationale was different in the sense that they bring to view, in trying to explain the effects identified in the effectiveness review that was already conducted at the moment the qualitative evidence synthesis was done.

Formulating a question for effectiveness reviews is often based on the PICO acronym, which actually requires review authors to define their population, their intervention, comparison of more interventions like the control groups, and the outcomes they are interested in, while qualitative evidence synthesis would more go into questions of how, why a particular intervention works for some people and doesn’t work for others. So it’s very much oriented towards exploring particular phenomenon that is relevant to the population intervention or outcome described in the effectiveness review.

Again, this is the way it has been done in the lay health worker review. So they’re questioning objectives related to effectiveness of lay health worker interventions in primary and community health care on maternal and child health; and then if you compare that with the qualitative evidence bit, you see that they were actually interested in exploring factors affecting the implementation of lay health worker programs for maternal and child health. So they actually identify barriers and facilitators to the implementation process and try to integrate these findings with the results of the effectiveness review, and that actually helped to enhance and extend the understanding of how these complex interventions may work and how context actually impacted on the implementation.

In setting criteria for inclusion, there’s some sort of standard procedure there for effectiveness reviews. What we should describe is what type of studies you will include. You would provide detailed definitions and parameters for both the population and the intervention; you would set out parameters of primary and secondary outcomes where secondary outcomes might be relevant; and you sort of take a position into what sort of quality measurement that you would include on the studies that you have to treat; and in effectiveness reviews that’s very important because they sort of understand that if you put crap in, then the overall effect will be crap as well. So they’re very sensitive to the methodological quality of the trials and other quantitative studies that they include. The qualitative evidence synthesis, on the other hand, they often include all sorts of qualitative methodologies. They may limit it to particular methodologies of types of papers. I’m aware of many reviews that exclude, for example, opinion papers because they don’t consider that empirical qualitative research. Some would only address ethnographic studies and would neglect, for example, action research; so there are a lot of decision points to make there. It’s best that the population that you search, research matches somehow that of the effectiveness review, although you can include probably - or other populations as well if you like. Outcomes are not really the phenomenon of interest here. We describe outcome maybe in terms of experiences; and quality is not as important as relevance in qualitative reviews. So if we judge further an article should be in or out, then we basically judge how relevant the insight of the article or how insightful, how illuminative they are to help us formulate or generate theory, and we do adopt articles that may have certain methodological flaws as long as they are minor enough not to affect the credibility of the overall theory.

So for the lay health workers review, they opted for the inclusion of RCTs only in the quantitative strength but they employ their broad definition of qualitative studies, included all sorts of studies that use qualitative methods for the qualitative evidence strength.

If you looked at the criteria for inclusion that have been chosen by the review authors, then you can see that they considered any lay health workers, paid or voluntary, including community health workers, village health workers, birth attendants, peer counselors, et cetera. They also defined the term “lay health worker” as any health worker who performs functions related to health care delivery, was trained, and had received no formal professional certificate. So as you can see, it’s very much operationalized and that really helps your second screener to actually evaluate whether a particular article should be in, based on the population, or not. For the qualitative bit, they use the same criteria, but they expand that to stakeholders to include family members, patients themselves, policymakers, and so on; and they include their perspectives and opinions as well on the topics. So here, this is an example of program you’re targeting.

Considering the intervention, the effectiveness review actually considered any intervention delivered by lay health workers intended to improve maternal and child health care. So child health was operationalized as children aged less than five years, and maternal health was health care aimed at improving reproductive health, ensuring safe motherhood for women in their role as carers for children aged less than five years. So that really sets the boundaries over a few, because the authors could not include many more outcome measures on the level of maternal or child health. On the qualitative side, they took a broader intake like programs that intend to improve maternal or child health and that use any type of lay health worker, including community health workers and the whole batch that have been considered in the effectiveness review as well.

Here you can see the division between primary outcomes on the next slide through going to secondary outcomes of the effectiveness review, like for example, health behaviors, type of care plan agreed, adherence to care plans, et cetera. Health care outcomes as assessed by a variety of measures and these may include mortality, physiological measures, also participants’ self-reports of symptom resolution, quality of life, and the classical ones like patient self-esteem. What is worthwhile to mention is that this review also included outcome measures related to harms or adverse effects, and sometimes that’s what the qualitative evidence synthesis may take out as well. So if you look at the outcomes defined for the qualitative review, you can see that includes studies where the phenomenon of interest is a description and interpretation of what I already mentioned, experiences and attitudes of stakeholders toward these lay health worker programs.

Now, we move into the plan for searching. What you need to define in your search strategy is, first of all, a list of sources that you will be consulting, and most of the time this is major electronic databases that have been identified within your specific scientific field, topped on with grey literature and that may be true for example, most participations, abstracts of articles that you could chase the article out of from, but also author contacts. A really useful thing to do is if you have papers that are really spot on, what I often do is contact the author to see whether he has any other studies related to the topic lying in his desk that may not have been published, and sometimes you get positive responses from that. You also define search term and if you want, you can use filters that sift out particular designs, such as RCTs only or qualitative evidence only. You then set date and language limits, and it’s very important that these limits are actually preferably theoretically inspired, because most of the people take a cut out of 10 years based on arguments that are really too pragmatic, for example, like of time, sometimes you find that samples of review authors that state that, “Well, I start in 2010, because that was when this particular intervention has been published on the first time.” So they actually define some sort of theoretical criteria for a couple of the dates. For qualitative evidence synthesis, the sources may differ. There’s a famous study from Trisha Greenhalgh in which she tested how much evidence that was retrieved via major databases and it was only 30% of the relevant studies that were retrieved through major databases. So a lot more work needs to be done to actually change the grey literature on qualitative studies, and actually that’s sort of unfolds, because we tend to use these very creative titles to position our qualitative basic research papers; and that’s very interesting. But the titles more the index terms give us any indication of whether or not it is qualitative, so that leads to the fact that these studies are often not indexed as qualitative studies and that is so hard to retrieve. So you may need to contact more people to generate evidence particularly when what you found is really thin in level of description.

So here is an example of a search strategy developed for the lay health worker review. As you can see, they actually captured a lot of databases here. The qualitative evidence synthesis used the same sort of databases, but the filter is different. So for retrieving the RCTs, a filter incorporated in Cochrane or MEDLINE was actually used; and for the qualitative review, they also filter that, only qualitative studies using pre-specified filters in these particular databases.

Then there’s the selection of studies. You actually screen preferably with two independent screeners. You screen all the retrieved studies through your search strategy for relevance for inclusion, and that is done based on titles and abstracts; and you actually share your inclusion and exclusion criteria you defined with your second reviewer, so that he can clearly see, “Okay, I need to evaluate this type of abstract based on these particularly in an exclusion criteria.” As such, if it’s not at all clear whether it should be in or out, you then retrieve a full text, and you look into the full text, and then you make an overall judgment of whether it should be included. Often you have conflicts between screeners on whether something should be in or out. So you can ask a senior researcher to actually be a judge on that, and look at the article with a fresh perspective to actually solve these, sort of, in or out conflicts. The same strategy is used in qualitative evidence synthesis. Relevance though may be interpreted a bit more broadly, and in situations where relevance is a bit uncertain, then maybe we would place these studies in some sort of holding pile and return to them later, because they might be interesting to actually create a comprehensive theory.

Now, the next phase is then that you should describe how you’re going to appraise the study quality. In an effectiveness review, that is produced putting Cochrane. They actually always use the similar quality framework that is outlined in the Cochrane handbook on reviews of effectiveness, and it’s based on some sort of quality measurements related to sequence generation, concealment of allocation, blinding of data, and so on. We already discussed that bias is not always an issue in qualitative research. Relevance counts a bit more in that sort of evidence strength. So the criteria that qualitative reviews use may also differ. There’s no such thing as expanded instruments that you could turn to. There are lots and lots of variance on the market and I think you need to decide what type of criteria you are comfortable with and how they actually fit your own personal review project.

In the lay health worker review, for the effectiveness bit, they use the Cochrane instrument to assess risk of bias. For the qualitative bit, they use the qualitative assessment instrument from the Critical Appraisal Skills Program. I’ve published an article in which different online appraisal instruments are assessed based on their ability to assess validity in qualitative research and it kind of outlines the pros and cons of, for example, the CASP instrument compared to the Joanna Briggs instrument of quality appraisal that is also often used by qualitative researchers. So you may want to look into that article to make a final decision.

Both stands engage in the assessment of heterogeneity. In effectiveness reviews, we actually try to look at the effect of – or to look at what happens when we remove low quality articles from our set or from our pile, and we evaluate whether it changes the effectiveness measure or the overall outcome at all. We adopt the same logic in qualitative evidence synthesis in the sense that for a sensitivity analysis, we actually remove all the findings generated from low methodologically quality articles and we try to evaluate whether that changes our theory or whether that changes our line of argument at all. It’s some sort of a test that we incorporate in our reviews to see what the effect of low versus high quality studies sets.

Now data synthesis, that’s a complex issue, because for effectiveness reviews, we usually group according to similarities in population, intervention, and outcomes; and we compile an overall measure of effectiveness. But for the qualitative bit, you have this whole range of qualitative methods that you need to choose from. I’ve given you the choice points already and my colleagues in the upcoming talks, they will go into the details of all these different options for you. But in the sense, you need to be able to justify your choice of method and that justification should be in your protocol.

Again, the work example from the lay health worker group, they didn’t define grouping of studies a priori. The studies were grouped within the review by type of health issue and so that was their major focus in that protocol. For the qualitative bit, the protocol mentioned that a thematic analysis approach would be used, because they sort of sense that the articles that were retrieved were really thin in description, so they wouldn’t really allow for an interpretive account that would meet them in a way of theory.

To sum up this whole bit, I think when I’m looking at the quantity researchers using information from and the quality researchers using information from – in my use on the meta level, I’ve always called the quant researchers my “masters of the window,” because they measure what they can see. They hypothesize and they’re very linear in their approach, and so that helps me in my synthesis. It makes the whole process very transparent. But I call my quality researchers, which I include studies, my “masters of the lantern,” because they search for light and place this where no one else has gone before. They sort of wander around in the difficult issues, the things that we can’t express in numbers in sort of an iterative fashion with the sole purpose, actually, to illuminate, try to better understand. My goal as a reviewer is to try and do justice to both of them by creating a few things that have enough methodological expertise from mixing these apples with oranges.

I think I’m going to close my thought here. I think I’ve talked way too much. Thank you for having me and I’m really happy to respond to any of the questions that I may have provoked through my talk. Thank you very much.