Session date: 3/23//2016
Series: HERC Cost Effectiveness Analysis
Session title: Evidence Synthesis to Derive Model Transitions (Part 2: Quantitative Pooling)
Presenter(s): Risha Gidwani

*This is an unedited transcript of this session. As such, it may contain omissions or errors due to sound quality or misinterpretation. For clarification or verification of any points in the transcript, please refer to the audio version posted at www.hsrd.research.va.gov/cyberseminars/catalog-archive.cfm.*

Unidentified Female: Good morning or good afternoon everybody. Thanks for attending this session of the HERC cost effectiveness analysis cyber course. This lecture today will talk about evidence synthesis for decision modeling and specifically how to do a meta-analysis. This is actually part two of a two-part lecture. The last part happened two weeks ago where we talked about the steps that one needed to talk in order to prepare to do a meta-analysis. Today we are going to talk about what happens after you have completed those steps and you are ready to the quantitative pooling to characterize this meta-analysis.

So the reason that this is in the HERC cost effective analysis cyber course is because meta-analysis are often times needed in order to drive input for a decision model. So for example, we have a decision model on the screen here where we are comparing two hypothetical strategies, drug A and drug B for treating patients that have some sort of an infection. As we all know from previous courses, transition probabilities are the engine to our decision model and they drive how patients progress throughout our model in order to help us understand which strategy represents the best value. The input for a decision model the transition probability can come from a variety of sources and one of the sources can be a meta-analysis which is what we will largely focus on today.

There is really two main ways to get transition probabilities or data inputs for your decision model and the first way is to transform existing data input from the literature which we talked about I think it was three weeks ago in my estimating transition probabilities lecture. The other main way of driving model input is to synthesize available data from multiple studies. That can be done through meta-analysis including individual patient based meta-analysis, mixed treatment comparison analysis or meta-regression. That is going to be the focus of today’s lecture.

So you do a meta-analysis because you have multiple studies that have evaluated your question of interest and you want to create a single pooled estimate from these multiple estimates that come from each one of the studies that you are looking at. The idea is that when you have a pooled estimate that is based on multiple studies, that pooled estimate is going to be higher quality than an estimate provided by any single individual study. So single studies that you are looking at may be too small and they may not be well powered enough. If you combine a pooled estimate that is created for multiple single studies, this can allow you to determine whether your findings are reliable. You can pool these studies together and you are thus increasing your sample size and so in doing so you can reduce the effective random error and produce more precise measures of effect.

We have talked about this in in the last lecture about the steps in meta-analysis and the next couple of slides are actually just going to be a reiteration of what we spoke about in the last lecture and I hope that this reiteration will really sort of cement this in your mind. I think that even though we are repeating this, I think that it is important enough that reputation will actually create hopefully some continue learning. When we do meta-analysis it is really four main steps that are happening. These are all happening behind the scenes of your software program so you yourself are not doing all of these things. The software is going through all of these steps for you.

So in step one of a meta-analysis a summary statistic is calculated for each one of the studies that you have included after doing the systematic literature reveal. Each one of those studies is then weighted…I am sorry. The summary statistics from each one of those studies is then weighted. You can see here that I note that it is conventionally weighted. You do not have to weigh each one of summary statistics. You can assume that each one of your studies contributes equal weight to your pooled meta-analysis estimate. But for reasons that we will talk about in a few slides it is usually a good idea to weight each one of your individual statistics per study. Once you have done that, you actually average these individual weighted estimate from each study and that gives you your pooled point estimate or let’s say your pooled mean or your pooled odds ratio. We know of course that the pooled point estimate has variation around it just like any individual point estimate would have variation around that. So the last step is to calculate the variation around this pooled point estimate.

The meta-analysis is the computation of a weighted mean estimate along with an estimate of variation around this mean. Now when I say that it is a weighted mean estimate, that could be a weighted mean estimate of means, a weighted mean estimate of odds ratios, a weighted mean estimate of probabilities, really any summary statistic. But the important thing to remember here is that it is a weighted mean of that summary statistic whatever that summary statistic is. Here is an example if we had the summary statistic being a relative risk of how you would create a pooled estimate.

You can see here that we have three studies, study A, B and C. From each one of those studies we pull out an individual relative risk. From that relative risk from study A, we transform that relative risk into a log relative \_\_\_\_\_ [00:05:34]. The relative risk from study B because the log relative risk for study B. The relative for study C because the log relative risk for study C. Then we average the log relative risk from all of those individual studies to create a summary log of relative risk or a summary log risk ratio. Then we exponentiate that in order to get the summary relative risk for the summary risk ratio.

When we have binary outcomes that we are looking at in a meta-analysis, we work in the log scale and there are two reasons we do this. First the log scale makes it more likely that the study outcomes from each one of your individual studies polls a normal distribution. The second reason is that when we are looking at relative risk in particular when the binary outcome is characterized as relative risk, logging that relative risk makes the relative risk inverse of each other. We talked about this in the last lecture and there are some slides that relates to that if you would like to go back and refresh your memory of how that works. If we are working with continuous data and let’s say we are pulling out a mean estimate from each of the studies, then we do to need to work in the log scale and all we do is pull out the actual means summary statistic from each one of those studies in order to create a pulled mean estimate which is the results of our meta-analysis.

This slide just represents what the results from a meta-analysis looks like and so you can see that we have multiple studies and there were three outcomes that were evaluated. There was an outcome of mixed prevention, secondary prevention and then whether somebody had cardiac defibrillator device installed. From each one of these studies we have a relative risk. Each one of these studies has its own weight and you can see that each one of these studies has its own square here. This square represents the point estimate from each study and the bars around the square represents the confidence interval around that point estimate from the study. The size of the box here is proportional to the inverse variance of the study. So larger studies have a smaller variance and therefore they have larger boxes and a larger weight associated with them than the smaller studies do.

Last lecture which was two weeks ago, we talked about the steps that you do to conduct a meta-analysis and I am not going to go over all of these again, but you are welcome to go back to the slides from two weeks ago if you want to get details about each one of these steps. However, I will just reiterate them briefly at a high level. When you are doing a meta-analysis you always want to start off with doing a systematic literature review. Once you have done a systematic literature review which includes looking at the grey literature and looking at clinicaltrials.gov, then you do a really quick title and abstract review of all of the studies that you have gathered and you spend about 60 seconds doing a title and abstract review and seeing whether your studies meet your inclusion and exclusion criteria. That inclusion and exclusion criteria should have been determined a apriori before you even started step number one.

After doing your title and abstract review, you extract data from the studies that you have decided meets your inclusion and exclusion criteria and it is best practice to do this according to a very well specified…extraction template. There is an example of one of these templates in the last set of slides. Once you have done your data extraction, you separate out your observation studies and your randomized controlled trials. You convert all of your outcomes to the same scale. So if you have binary data you want it to be all odds ratio, all probabilities or all relative risk. You cannot have a combination of those three different statistics.

After you have decided on which summaries statistics you want to be using for your meta-analysis and converted all of the outcomes from your individual studies into that summary statistic. You evaluate the heterogeneity of your selected studies which you do through both statistical tests as well as graphical examination of \_\_\_\_\_ [00:09:50] plots. Then once you have done all of those things and you have decided that yes, your studies are homogenous enough to combine in a single meta-analysis. It is only at that point that you actually conduct a meta-analysis and that is what we are going to spend our time discussing today.

However, before we continue I have one audience poll and that is, how do you proceed if you have identified heterogeneity amongst the studies after your systematic literature review? You have three options here. The first option is that you do not continue with doing the meta-analysis. The second option is that you exclude the studies that cause heterogeneity and conduct a meta-analysis on the remaining studies. The third is that your run a meta-regression. So Heidi, I think we can just spend about 20 seconds with this poll.

Heidi: Yes. Responses are coming in. I will give everyone just a few more moments before we close it out and go through the results. We are at about 40 percent right now so I will give everyone just a few more seconds. Try to get a few more responses in before we close it out. Looks like we have slowed down. So what we are seeing is 9 percent saying do not continue, 16 percent saying exclude studies that cause heterogeneity and 75 percent run a meta-regression. Thank you everyone.

Unidentified Female: Great. Thanks Heidi. So I am glad that a few people said number two, exclude studies that cause heterogeneity and conduct a meta-analysis on the remaining studies because that is definitely something that you cannot do. You cannot just sort of systematically pick which studies you want to include and exclude. In doing so you are creating a systemic bias in your meta-analysis. So we definitely cannot use that option. Option number one and option number three are both viable and like many things in statistics and research, the option that you choose really depends on the quality of the data that you have and the questions that you want to answer. You can decide that the heterogeneity is too much and you do not continue and then you summarize your results in a systemic literature review. This is actually a perfect acceptable way to go and it is a way that a lot of folks go and there will be no faults to doing this. Some people run a meta-regression and it really depends on how heterogeneous your studies are.

If you are really looking at studies that have…one study has a follow up time of 24 weeks and another study has a follow up time of 30 weeks and you have discussed with your clinical collaborators and you feel like that is really not so different. The heterogeneity follow up time should not really effect the relationship between treatment and outcome then you could just run a meta-regression and use follow up time as a co-variant in that regression. We will talk more about meta-regressions at the end of this lecture. But one of the things to keep in mind about meta-regressions is that they are doing adjustments for co-variates but they are doing so at the study level and not at the individual level.

If you have a study in which you have…one study has a mean age of 50 per participant and another study has a mean age of 60 per participant, when you are going the meta-regression you are really adjusting for the mean age. Well, the study that has the mean age of 50 for its participants really has a standard deviation of plus or minus 14 years and the study that has the mean age of 60 for its participants has a standard deviation of plus or minus 2 years. Then you notice that the first study has a lot more heterogeneity of age within study than the second study has. So that may be a problem for doing a meta-regression because you are really only adjusting for the mean age but you know that study number one the mean age is not really capturing entirely the true ages of the participants in that study. So in that sort of a situation you would really want to think twice about doing a meta-regression and you may decide no, I do not want to continue with doing this quantitative pooling and I am just going to stop at my systematic literature review. Unfortunately, I cannot give you a hard and fast rule of how you should proceed if you do have heterogeneity amongst your studies except for to tell you that option two is definitely not the right way to go. Let’s move on to actually the conduct of the meta-analysis.

We talked before about the four steps that are involved in doing meta-analysis and we talked about each one of those as implemented in the software and you yourself are not actually doing this by hand step one through four. There are really two decisions that you have to make in conducting a meta-analysis before you can actually even allow the software to implement the four steps. The two decisions are whether your fixed versus random effects and how you are going to pool your study. Let’s talk about fixed versus random effects first. So whether you use fixed versus random effects is driven by how you think about the studies that you have. A fixed assessed analysis assumes that the variants that you have amongst the studies is just due to sampling error and that there is some fixed underlying true effect. A random effects analysis assumes that the variance amongst your studies is due to sampling error but is also due to some sort of variation in true effect from study to study. So that may be a situation where you have different participants in different studies or different ways that the intervention was administered or a different follow up time like we spoke about before.

In a fixed assessed analysis, you are only looking at the within study variant because you think that there is no real between study variance that you need to accommodate. However, in the random effects analysis, you think that there is a real between study variance that there is something beyond sampling error that is driving a difference in the relationships you see between treatment and outcomes and you want to accommodate that difference. In a fixed assessed analysis, if is each study had an infinite sample size the sampling error would be zero. But in practice we do not have incident sample size. We are not really looking at the population when we are evaluating each study and that is why we have variation across studies due to the sampling error.

In the random effects analysis, it is a little different. We think that the studies are similar enough to combine but we do not think that they are identical. We do not think that the true effect size is the same in all the studies. Instead we assume that there is a distribution of effects and we assume that in a frequency approach that this distribution of effect is a normal distribution. So we think that there is both sampling error going on as well as true variation in effect sizes that is occurring. When we do a random effects analysis we assume that the studies that we have represent a random sample of studies in the universe. That is why in the random effects analysis the variance includes both the within study variance as well as the between study variance. Because the random effects analysis is incorporating two types of variance within in study and between study. The constant intervals from a random effects meta-analysis are going to be wider than that of a six effect analysis and the imprint is also different. In a random effects analysis, you think that you are getting the mean of the true effect. In the fixed effects analysis your inference is the true effect of whatever you find from pooling the studies.

Fixed effects and random effects analysis also differ in how they handle small studies. So in a fixed effects analysis, the smaller studies are given less weight and in a random effects analysis the small studies are given more weight. So the fixed effects analysis assumes that the only reason there is variation amongst the studies is because of sampling error. So because of that assumption, the fixed effects analysis approach says we can largely ignore the information from smaller studies because we have better information about that same effect size that is coming from the larger study. But conversely, the random effects analysis assumes that you are estimating the mean of a distribution of effects. So we cannot discount the small studies in a random effects analysis because the small study may have information about the effect that no other study in our group of studies have estimated. We also do not want the larger studies to have too much of an influence on the pooled effect in a random effect analysis because we do not want the mean estimate of effect to be overly influenced by the larger study which we think are estimating something a little bit different than the smaller study.

Often times a random effects analysis for meta-analysis is more suitable because there is almost always difference you are going to find amongst your studies. But the thing to keep in mind is that, just because random effects is often times more suitable, it does not mean that it is more conservative. The reason is such, if the smaller studies are…and maybe that is because even the smaller studies have a sicker population and sicker populations are going to demonstrate great improvement with the treatment that you are studying. In that situation, if the smaller studies are systematically different than larger studies, then a random effects analysis will bias the pool…effect. Because you are increasing the weight of the smaller studies when you do the random effects analysis. So if the example is that the smaller studies have a sicker population and the sicker populations demonstrate greater improvement in that particular situation the random effects analysis will actually overestimate the effect of the treatment. So you really want to keep that mind when you are deciding whether to do fixed effects or random effects meta-analysis.

When you are doing the random effects analysis, the width of your pooled estimate…the variation around your pooled estimate describes the degree of heterogeneity amongst your studies. Something to keep in mind is that in the random effects analysis this distribution of the heterogeneity across studies is usually assumed to be normal. If you do have heterogeneity you do have real between study differences. The confidence interval for the random effects analysis will be greater than that for the fixed assessed analysis. Conversely, if you do not have heterogeneity present, the confidence interval for your random effects analysis is going to yield that exact same result as your fixed effects analysis. Now something to keep in mind is that the random effects versus fixed effects will influence only the degree of variation around your pooled treatment effect. The point estimate your pool treatment effect is going to be the same regardless of whether you do the fixed effects analysis or the random effects analysis.

Another thing to keep in mind is that when you are doing this random effects analysis, because you are assuming that the distribution of the study effects is normal, that means that the random effects analysis will only estimate the average treatment effect if the biases are symmetrically distributed. Once you have settled on whether you are going to do a fixed effects analysis or a random effects analysis, you then need to decide what message you are going to use to actually pool your studies. The method that you use to pool the studies depends on a few different factors…three different factors really. It depends on whether you decided to use fixed effects or random effects analysis, the type of outcome data that you have whether it is continuous or binary and the number of studies that you have and are heterogeneity.

I have noted here what are almost all of the different ways that you could pool studies. The inverse-variance and the DerSimonian and Laird methods are the ones that you will see most commonly in the literature. So I will go over those first and then in the next slide I will talk about some limitations and some things you need to keep in mind with these different pooling methods. If you are using an inverse-variance approach it means that you are using fixed assessed analysis. The DerSimonian and Laird approach is the inverse-variance approach with random effects, it just happens to be called DerSimonian and Laird after the two statisticians who came up with this approach. With both the inverse-variance and the DerSimonian and Laird approach, you can use it when you have binary data only if you have the actual 2x2 tables from each study that you are getting the binary data from and we will talk about why that is in a moment.

You can also use inverse-variance and DerSimonian and Laird when you have continuous data and low heterogeneity. DerSimonian and Laird performs better than the inverse-variance approach when you have multiple studies in your analysis. If you have continuous data and you feel as though the studies that you have are heterogeneous and you have a small number of studies, you can use another approach which is the \_\_\_\_\_ [00:23:03] approach. This is an extension of the \_\_\_\_\_ [00:23:09] approach. Just something to be cognizant of is that when you have very few studies and they have very unequal sample size, this approach is not going to work very well. Profile likelihood approach is another approach with random effects. You use it when you have continuous data, you believe your studies are heterogeneous and you believe that your biases are not symmetrically distributed. So you think that the \_\_\_\_\_ [00:23:37] which is the between study variance is not normally distributed.

You can also use the Bayesian approach for those of you who are familiar with the Bayesian versus frequentist approach. The Bayesian approach is very flexible. You can use it if you have binary or continuous data, if you have heterogeneity, if you have very few studies. If you are working exclusively with odds ratios, then you can use the Mantel-Haenszel or the pedo approach. Some of the problems with using an odds ratio, because we always work in the logs scale when we are looking at an odds ratio, if you have some cells in the 2x2 tables that are underlying your studies that have zero values in those cells you can have a real problem. Because of course, you cannot take the log of zero. So if you do have any cells that have zero values you use the pedo method. If not use the Mantel-Haenszel method.

Now I spoke before that…I just mentioned that you should really only use the inverse-variance and DerSimonian and Laird approach if you have binary data and the 2x2 tables that were used to calculate the summary for the binary data. Now some people actually do use inverse-variance and DerSimonian and Laird when they have binary outcomes but they do not have the 2x2 table. I strongly, strongly caution against this. If you have the 2x2 table yourself, you can calculate the variance around each study’s point estimate yourself. If you do not have the 2x2 table, you have to use a normal approximation for the binary distribution to calculate the variance and that makes everything fall apart and we will touch on this.

There are a lot of references that I have noted here about different ways to pool studies. I recommend anybody that is really interested in learning a lot about meta-analysis take a look at all of these different references. There has been a lot of changes actually recently. You can see a couple of studies from 2014…just some new information…relatively new information that has come out that has really changed the way we think about the way we want to pool studies and so I think these are all great references to take a look at.

I am going to go briefly into the equations just so that you can get a little bit of the sense of how studies are pooled behind the scenes at the meta-analysis. I am going to first talk about an inverse-variance approach and remember inverse-variance is with fixed assessment analysis and I am going to talk about this as it relates to continuous data because continuous data is a little more intuitive to understand than the binary data. So you can see here that we have the pooled treatment effect as denoted by T. bar. The pooled treatment effect is a function of the sum of weight of an individual study times the treatment effect from those individual studies divided by the sum of all the different weights from those studies. The variance around this pooled treatment effect is one over the sum of all of the weights from the individual studies. The weight from an individual study is one over the variance of that study. So if we are looking at…this is the inverse-variance approach and remember that larger studies generally have a smaller variance and so larger studies are also more likely to be given more weight in the meta-analysis. A smaller study is going to have a larger variance so it is going to be more imprecise if we are given less weight in the pooled meta-analysis.

Inverse-variance with random effects takes multiple forms. The most commonly known form is the DerSimonian and Laird approach but we also have the \_\_\_\_\_ [00:27:22] and the profile likelihood and the Bayesian approaches that are all inverse-variance with random effects. Something to keep in mind is that the pooled treatment effect which here is denoted by a T. bar random is calculated in the exact same ways as fixed effects meta-analysis. So this is exactly the same. The variance around this pooled effect is also calculated as one over the sum of the weights from the individual studies. What differs is the way that the weights of the individual studies are calculated in the random effects analysis versus the fixed effects analysis. So you can see here that we have the weight is a function of one over two quantities and the quantity on the left is the within study variance. This is the exact same thing that we used in the meta-analysis…I am sorry. In the fixed assessed analysis. But we know also have another quantity in the denominator and that is \_\_\_\_\_ [00:28:13] which is the between studies variance.

The four approaches DerSimonian and Laird, \_\_\_\_\_ [00:28:20] profile likelihood and Bayesian approach, all they do is differ in their calculation of the between studies variance in their calculation of the \_\_\_\_\_ [00:28:31]. You can see here from this equation that if the \_\_\_\_\_ [00:28:35] is zero, there is no between studies variance. Then the estimates you get from the random effects model are going to be the exact same as the ones that you get from fixed assessed model. Something else to keep in mind is the weight from an individual study. So within studies variance is going to be specific to that study, but the between studies variance is going to be the same value for all of the studies in your meta-analysis.

So \_\_\_\_\_ [00:29:04] is what we use in the random effects meta-analysis to give us the between studies variance. It is considered the variance of the true effect sizes. However, we do not really have the variance of the true effect sizes. We would only get the variance of the true effect sizes if we had an infinitely large sample study and each of the studies within this infinitely large sample study had an infinitely large sample size itself and thus it would give us the true effect for each study. That would allow us to calculate the variance of these effects and give us the \_\_\_\_\_ [00:29:41]. But of course we do not have this…we cannot compute this directly because we do not have an infinitely large sample of studies and the studies that we have do not have the population. So instead of computing \_\_\_\_\_ [00:29:53] directly, we estimate it from the observed effects. All of the random effect approaches differ how they handle estimating the \_\_\_\_\_ [00:30:01] which of course is then going to affect your confidence interval and then your test \_\_\_\_\_ [00:30:05].

I spoke before briefly that there are some problems with the inverse-variance and DerSimonian and Laird approaches and I just want to go over them very briefly. Schuster had a really nice…article showing statistics in medicine in 2010 and he was talking about when the problem that occur when you use the DerSimonian and Laird or the inverse-variance method for binary data and you do not have the 2x2 tables directly from each study. The inverse-variance and the DerSimonian and Laird approaches assume that the point estimate and the variance are independent. But if you do not have the 2x2 tables yourself from each study you cannot calculate the variance yourself and you are going to rely on the normal approximation for the binary distribution. When you use this normal approximation to calculate the variance, the variance is actually not independent of the point estimate.

You see the equation for variance is PxQ-N and so you can see that the mean is directly used to calculate the variance and thus they are not independent. I should say…I am sorry. The probability is used to calculate the variance and thus they are not independent. So that is why we recommend that you only use the DerSimonian and Laird or the inverse-variance approach for binary data when you yourself have the 2x2 table. Now that is not often times going to happen and so in that situation you really want to look at a different approach to calculating the pooled estimate from your binary data. Another critique of the DerSimonian and Laird method is that it assumes that we have estimated the \_\_\_\_\_ [00:31:47], the between study variance exactly. That means that we are really thinking that we have a narrower confidence interval than we truly do unless you are more likely to find significant difference.

DerSimonian and Laird themselves have actually come out and said we came up with this approach because it was a simplistic approach to use at a time where we did not have access to much computing power. In today’s day and age that is not a problem. We have enough computing power and we recommend that you use other approaches as well. However, the DerSimonian and Laird approach has just sort of become a conventional practice so much so that the DerSimonian and Laird and inverse-variance methods are actually the default weighting methods for pooling studies RevMan and RevMan is the software that is used by the Cochrane collaboration. Some of you all may be familiar with the Cochrane collaboration. They actually create really high quality systematic literature reviews and some of those systematic literature reviews are also accompanied by meta-analysis. Those meta-analyses are required to be done in this RevMan software. So if you are doing a meta-analysis in RevMan, really think about whether you want to use DerSimonian and Laird method to pool your studies. I would recommend really if you do not have the 2x2 tables from the studies themselves and you have got binary data you do not use this weighting method. You could get a lot of false positives.

Every statistical approach and every research study has limitations and meta-analysis is not immune from limitations either. One of the biggest limitations that you will face when you are conducting a meta-analysis is that of publication bias which is also \_\_\_\_\_ [00:33:40] called the file drawer syndrome. The problem with publication bias is that the studies that you include in your meta-analysis may be systematically different from all the studies that should have been included and you were not able to include all of the studies that you should have because they were not all published in the literature and therefore identifiable through your systematic literature review. So we know that studies with significant results are much more likely to be published in the literature. If you are only including studies with significant results in your meta-analysis because all of the studies with non-significant results are just in a researchers fill drawer somewhere, then your meta-analysis is going to overestimate the effect of treatment.

We also know that larger studies are more likely to be published than smaller studies. If the results of smaller studies are systematically different than larger studies, then you are going to have a problem if you use a random effect meta-analysis because those random effect meta-analyses are of course going to give greater weight to the smaller studies than a fixed effect analysis will. Now fixed effect will not be effected as much from this issue of large studies being more likely to be published than smaller studies, but strictly speaking is also inappropriate. You really want to make sure that you are capturing the entirety of the literature and sometimes that is just going to be not possible because of the publication bias issue. So all of the prep work you did to get to the point where you can even do a meta-analysis, your systematic literature, your data extraction, your assessment of heterogeneity of your studies, you did that so that you could get all of the studies that exist that are relevant to your research question. However, even the best conducted systematic literature review cannot obviate concerns about publication bias.

The question is what do you do then. First you need to assess the degree of publication bias that you have in your meta-analysis and you do so by looking at funnel plots and looking to see whether a funnel plot is asymmetric. We will look at a funnel plot on the next slide. Something to keep in mind is that, even if the funnel plot looks good that does not mean that you are devoid of problems. They are just like many of the test that we have for statistical issues in meta-analyses. If you see a problem, yes the problem is there. But just because there appears to be a lack of a problem does not mean the problem does not exist. So here is an example of a meta-analysis…I am sorry, of a funnel plot. This is actually a really great article. I recommend everyone read it by Stern \_\_\_\_\_ [00:36:16]. Got some really big names in meta-analyses including Alex Sutton and John \_\_\_\_\_ [00:36:21] in there. Published in the British Medial Journey and it shows you what is symmetric versus asymmetric funnel plot looks like.

Here we have on our X-axis our summary statistic. Here we are working with an odds ratio and of course always working in the log scale for our binary outcome. On the Y-axis you have the standard error. Where the standard error goes from high to low. You are looking for asymmetry here in the funnel plot and so here is where the mean estimate is for all of the studies. Your mean pooled estimate for all of your studies. You can see that your studies are pretty well distributed around this mean pooled estimate…they are pretty symmetrically distributed. On the bottom right hand plot you have your mean estimate from your studies. Your mean pooled estimate from your studies and you can see that there are a lot more studies on the left hand side of your mean pooled estimate than there are on the right hand side of your mean pooled estimate. So that indicates that there is asymmetry in your funnel plot. If you have asymmetry in your funnel plot, you know that you have some publication bias that you need to address.

Now one of the things to be careful about is that even…that if you have…it is possible that your funnel plot can be skewed by something other than publication bias. So if the quality of the studies varies with their size, it may look like it is publication bias but it is not. If you do have an asymmetric funnel plot, do a little bit more investigation and make sure it is not just some sort of issue like sample size issues and quality issues that are correlated. Let’s talk about what you do…I am sorry. Let’s talk a little bit more about funnel plot asymmetry and then I will move into what you do if you do have publication bias.

If you have an asymmetric funnel plot, you either have heterogeneity or the quality of your studies varies with the size of those studies. So if you have heterogeneity that is a problem. If you use the quality of your studies varies with size, then that is something that you may be able to address in meta-regression. Now you should not just look at your funnel plot, you need to evaluate it in context of the other information you have about the studies in order to detect whether it is really heterogeneity or the quality of your studies varying with size. Something to keep in mind is that for your funnel plot to have any use, you need to have studies that have various sizes because those various sizes are going to help plot your results on your Y-axis where your Y-axis is the standard error of course a function of sample size. Of course, failure to find asymmetry does not mean that there is no publication bias.

There are test that you can employ for funnel plot asymmetry and they are denoted in this article here. But they typically have low power and so if you have \_\_\_\_\_ [00:39:27] to studies in your meta-analysis, you should not really do a test for funnel plot asymmetry because it is just too difficult to distinguish what is chance variation from real asymmetry.

Unidentified Female: Reesha, we have one question about funnel plots and how to determine the size of I guess the funnel with the dotted lines.

Unidentified Female: So you yourself will not determine that, you will actually put your studies into a software program and ask that software program to create a funnel plot and it will do so for you. Then what you will really just do is interpret graphically the visual results of this funnel plot. So these lines…these dotted lines were created by the software program not by you yourself. But you can see here that there is a lot more variation on the left hand side of the funnel plot than on the right hand side of this funnel plot and that would tip you off that there is a problem here. So if you do have publication bias, if you have this asymmetric funnel plot and you realize that it is not that the quality of your study varies by size but it really looks as though there is a lot of studies that they have either really negative or really positive results and very few studies that look like they have no results sort of indicating maybe there is a publication bias here. There are a number of different ways that you can proceed to deal with your publication bias.

I have noted a number of different ways up here. I am going to talk about the accumulative meta-analysis and the trim and fill method because those are the most common. However, if you do have publication bias I strongly recommend that you read this book. It has a very nice section on publication bias and how to deal with publication bias and it details a lot of how you employ these different methods. The accumulative meta-analysis is an approach where you run your meta-analysis with one study and then another study and then a third study and a forth study and you add your studies to the meta-analysis in order of how precise they are. So the studies that have the least variation around the point estimate are added first and then the ones that have the most variation around the point estimate are added last. If your pooled point estimate has stabilized and it is not affected by the inclusion of the smaller studies which are less precise, then there is no reason to assume that including studies has injected a bias into your analysis.

The trim and fill method is different. It actually is an iterative procedure and it removes the most extreme small studies from the positive side of the funnel plot and the effect size is recomputed at each iteration until your funnel plot is symmetric around the new effect size. So that is the trimming component. However, the problem with trimming is that it reduces the size of the variance because you have kicked out certain studies and you have kicked out the smaller studies and so that will yield too narrow a confidence interval. So to make sure you are not sort of artificially getting too narrow a confidence interval around your pooled mean estimate, the fill portion of the algorithm fills the original studies back into the analysis and imputes a mirror image for each. That has no impact on the point estimate but it corrects your variance so that your variance is not artificially small.

A lot of these approaches are going to require more understanding than I can impart to you in any one-hour lecture. Meta-analysis is often times a one or two semester course in a graduate school. Anybody that is really interested in doing meta-analysis, I would recommend that you actually do take a course on this and/or read the books and the articles that I have denoted here. The point of this lecture is really to give you an overview of this approach and know that you can synthesize multiple studies into a single pooled point estimate in order to use that as an input in your decision model. But if you really are going to undergo doing a meta-analysis, it certainly requires more familiarity than just listening to this one lecture. So I strongly recommend that you continue to educate yourself through the references denoted here or take a class at your local university.

Once you have your meta-analysis, you have decided fixed effects or random effects meta-analysis, you have decided how you are going to pool your studies, you have assessed publication bias through funnel plots and decided how you are going to deal with this publication bias if you do have that. Then from then on is when you actually run your meta-analysis. The reason that we are running the meta-analysis and even describing this in a cost effective analysis course is because you can then use the point estimate from your meta-analysis as your input into you cost effectiveness analysis. So here we have a cost effective…I am sorry, meta-analysis that is looking at point of care test for detecting \_\_\_\_\_ [00:44:44] and looking at their diagnostic accuracy. This may be an input that you need in your cost effectiveness analysis. So where we have what is the combined estimate from all of these individual studies that we are evaluating this point of care diagnostic test.

This point estimate of .96 is going to be the input into your cost effectiveness analysis. However, we know that this point estimate has variation around it and so we also would include that variation around that point estimate into our cost effectiveness analysis so that we can run sensitivity analyses. That will be the topic of my next lecture which is next week. I hope this gives you a sense of why we are doing meta-analysis and how we can use them in our cost effectiveness analyses. There are a number of different software programs that you can use in order to conduct a meta-analysis. You can do this in Stata or SAS or R. I mentioned before that RevMan is the default software for the Cochrane collaboration and I believe they require you conduct your meta-analysis in RevMan in order for them to accept them for publication.

There is CMA which is Comprehensive Meta-analysis software and for folks that are Bayesian inclined you can do meta-analysis in OpenBUGS or WinBUGS. I will say be very careful with plug and chug software. It can be deceptively easy to do a meta-analysis and just throw a bunch of point estimates and variations around point estimates from individual studies into a software program and calculate a pooled estimate. You want to really stay away from that and make sure you think through all the steps that we talked about in this lecture and in the last lecture before deciding how much you can believe in that pooled point estimate.

We do not have a lot of time left but I do want to briefly touch on some advanced topics. We definitely do not have the time to go into anyone of these and in fact, they really each deserve multiple lectures on their own. But I do want to make sure that you are aware of these topics so that you can explore them more on your own time if they are avenues that you want to pursue. So three advanced topics we are going to briefly touch on today, individual patient data meta-analyses, meta-regression and mixed treatment comparisons which are often called network meta-analyses. Individual patient data meta-analyses are when you are in the enviable position of having individual patient data from each one of the studies you want to include in your meta-analysis. There is a real advantage in doing so. The traditional meta-analyses use the summary statistics from each study as input into the meta-analysis and so that means that if you have eight studies, you have eight data inputs.

The individual patient data meta-analysis actually uses the individual patient data from each study and so if you have eight studies and each one of those studies have the sample size of 50, then you have 400 data inputs. So you have much richer data with an individual patient data meta-analysis than you would from a regular meta-analysis. Because you have all these individual data, you can conduct a whole bunch of different analyses. You can calculate different summary statistics, look at what is happening to follow up times, conduct \_\_\_\_\_ [00:48:04] of event analyses and you can impute missing data. However, there is a real challenge in actually getting that individual patient level data.

Essentially, after you decide what studies you are going to include in your meta-analysis after you are doing your systematic literature, you would have to contact the authors of each one of these studies and they would have to share the individual patient level data from their study with you. So it probably requires a lot of IRP approvals and data use agreements and it is pretty rare to actually get these things. If you are the lucky person that does get all the individual patient data from the studies you decided to include in your meta-analysis, keep in mind that doing a meta-analysis on individual patient data could yield different results than if you just do a traditional meta-analysis. So for example, if you handle missing data by imputation and the study authors only used a complete case analysis in reporting their summary statistics, then your individual patient data meta-analysis is going to yield a different analysis…a different result than a traditional analysis would.

Meta-regression is another topic that can…another statistical approach that can be really useful in this space. So we all know regression where you are adjusting for differences at a patient level within a specific study. Meta-regression is also adjusting for differences but it is adjusting for differences at a study level across studies. So in a regression you are adjusting for co-variates that might be imbalance amongst your treatment and control group. In a meta-regression you are adjusting for co-variates that might be imbalanced amongst your study. Meta-regression thus is not recommended when you only have a small number of studies.

Some of you may be aware of this very general rule of thumb that you should only conduct a regular aggression when you have at least 10 events per co-variate. And as a rule of thumb the best way to decide how many variables you should include in your regression is to do a power analysis. But it is at least sort of a guiding rule of thumb to think about apriori when you are thinking about your \_\_\_\_\_ [00:50:17] model. In a meta-regression there is no established rule, but when you have a small number of studies you are not really going to be doing much adjustment because you will not have power to do so. For that reason, it is often times not possible to do a meta-regression because you may only have six or seven studies after doing your systematic literature review that are even eligible to be in your meta-analysis.

Another caveat with meta-regression is that it is subject to the ecologic fallacies. You are adjusting for differences at the study level but if there is heterogeneity within your study that is something that you are not really capturing. You can run a meta-analysis with fixed or random effects and if you are doing a fixed assessed meta-regression, the fixed assessed meta-regression assumes that all of the variation between your studies outcomes can be accounted for by the co-variate that you have included in your regression model. They are assuming that the studies that have the same values for all of their co-variates share the same population effect.

In a fixed assessment or regression, you are testing the null hypothesis that the effect size is the same for all values of the co-variate. A random effects meta-regression assumes that your co-variates explain only part of the variation between the outcomes of your study. Unlike fixed effect analyses…assumes that the studies that have…of the facts and that you are testing the null hypothesis that the means effect bias is the same for all values of the co-variate. So meta-regression works well when all of your studies are evaluated in the same intervention. So let’s say you were interested in drug A versus placebo and you have four studies…let’s say 14 studies that are evaluating drug A versus placebo and they have some heterogeneity at a study level that you think can be adjusted for in meta-regression, that is great. However, most of the time for a cost effectiveness analysis, your interest in the effect of one active intervention versus another active intervention and a lot of studies may not have directly evaluated these active interventions.

You may have the situation where some studies have evaluated drug A versus placebo and other studies have evaluated drug B versus placebo, but what you are really interested in for your cost effectiveness analysis is the effect of drug A versus drug B because you are trying to inform policy and the policy should not be to give the patient placebo. In that situation you can do a mixed treatment comparison analysis which is also called the network meta-analysis. This is a statistical method for estimating the relative treatment effect using a network of evidence. So we have a situation here where some studies have looked at treatment A versus placebo and from those studies we get an estimate of stata A versus placebo.

We have other studies that have looked at treatment B versus placebo and from those studies we get an estimate of stata B placebo. What we are really interested in is the difference between treatment A and…treatment to get a stata A/B. We can do so because we have this network of evidence. So essentially, a mixed treatment comparison analysis looks…drives the estimate of stata A/B as a function of stata A placebo minus stata B placebo. So we are using this network of evidence in order to drive our estimate of interest. We pay for this derivation of our estimate of interest by our variant. So our variation around stata A/B is a function of the variation of stata A placebo plus the variation of stata B placebo. So while we can get a point estimate of the estimated difference about treatment A and treatment B, our variation around this point estimate is going to be pretty large.

So mixed treatment comparisons are an advanced topic and you should not proceed without consulting a statistician. However, I just wanted you to be aware of the statistical technique. When you do consult your statistician about doing a mixed treatment comparison analysis, you should provide information about your network of evidence. By that I mean something like this. I conducted a mixed treatment comparison analysis looking at randomized controlled trials for a different drug that was used to treat menorrhagia or heavy menstrual bleeding. You can see here that there is a number of different types of techniques ranging from surgical ablation to different types of hormones to intrauterine devices. So what we were really interested in is a multitude of different comparisons.

If I wanted to compare ablation to danazol I could do so by evaluating my network or evidence. You can see here that all of my treatments are connected. They are not necessarily connected to each other so danazol and progesterone at three weeks are not connected to each other but they are connected through each other through this network of evidence. So what I can do is really evaluate the relative difference of progesterone at three weeks versus danazol by evaluating progesterone versus TXA, TXA versus progesterone at two weeks, progesterone at two weeks versus danazol. While I can evaluate this estimated at relative…difference between progesterone at three weeks and danazol I pay for this because I add variation to the point estimate with every note in my network I go through.

If you have a network that is broken and you have let’s say what we call two islands, meaning that you do not have a way to connect one treatment to another treatment even through a diverse path then you cannot conduct a mixed treatment comparison analysis. The other thing to keep in mind is that if you have just a few number of studies connecting each one of your nodes, here I just have one study connecting COC and danazol, that is going to be problematic because it may not allow my models to converge. So when you talk to your statistician about doing a mixed treatment comparison analysis, make sure you present this network of evidence along with the number so studies that are connecting each one of the nodes or each one of the treatments of interest.

In summary, a meta-analysis is a single pooled estimate plus a variation around that pooled estimated from waiting and combining individual effects from multiple studies. There are multiple steps that you have to conduct before you can even get to pool the summary statistics from individual studies into a meta-analysis and they involve systematic literature reviews, title and abstract reviews, data extraction, separating out observational studies from randomized controlled trials, converting outcomes to the same scale and evaluating the heterogeneity of those outcomes before you can even conduct the meta-analysis. When you are doing the meta-analysis, the quantitative pooling that is step seven in the previous slide, you have to decide on two things. First whether to use fixed effects or random effects analysis and second how to pool the study.

You use a fixed effect analysis when you assume that you have the universe of all studies in your systematic literature review in meta-analysis and if you can assume that then the \_\_\_\_\_ [00:57:29] is that you have the true effect. If you were doing a random effects meta-analysis you assume you have a sample of studies from the potential universe of studies and thus your interest is that you are evaluating the mean of the true effect. When you are deciding how to pool studies, if you have binary data such as an odds ratio or relative risk you can use the Mantel-Haenszel or pedo method. I should also mention you could use inverse-variance and DerSimonian and Laird if you have the 2x2 tables underlying the binary data from each study. If you have continuous data, you have a lot more flexibility about how you want to pool your stud.

If you are doing a meta-analysis, you may be able to do an individual patient data and meta-analysis if you are lucky enough to get all of the individual patient data from the studies you are including. You do a meta-analysis whether it is a traditional or individual patient data meta-analysis when all the studies have evaluated the interventions that you are interested in evaluating in your cost effectiveness analysis. You do a mixed treatment comparison analysis when you have interventions that you want to evaluate in your cost effectiveness analysis but those interventions have not been directly evaluated in head-to-head randomized control trials. You can run meta-regressions in order to adjust for co-variates to the study level and you can run those meta-regressions whether you are going a traditional meta-analysis or a mixed treatment comparison analysis.

Something to keep in mind about these approaches is that whether you are doing a meta-analysis or a mixed treatment comparison analysis, even if you have randomized controlled trials that are populating your mixed treatment comparison analysis or your meta-analysis, those meta-analyses and mixed treatment comparison analyses themselves are observational studies. That is because people are not randomly assigned to studies in your meta-analysis, they are randomly assigned to treatments within the study. That is your pooled estimate is really itself coming from an observational study. I strongly recommend reading all three of these books for anybody that wants to conduct a meta-analysis. The Borenstein book is a really nice introduction. It is very high level but it is highly accessible. The Sutton book gets into a lot more details. I would not recommend doing a meta-analysis by only reading one of these books. I strongly recommend you read both before you conduct a meta-analysis. I mentioned before the Cochrane collaboration, they have a very nice handbook for how to conduct systematic reviews of intervention and that is available from the urls noted at the bottom of the screen. With that I will open it up to any questions.

Unidentified Female: We had one earlier question again about the funnel plots and whether that was only related to one subgroup. I think that is what they are getting at.

Unidentified Female: I am sorry. What is the question?

Unidentified Female: Whether funnel plots only include sort of one subgroup.

Unidentified Female: No. They can include your entire…the entire sample from each one of your studies and I think that because on this lower right hand side funnel plot, they have evaluated the studies by subgroup and you do not need to do this. But if you are looking at an asymmetric funnel plot it is a good idea to evaluate how your effects differ by subgroup because if you think that asymmetry is due to real differences in the quality of your studies like one study has sicker patients than another study did and that might be what is driving your asymmetry then you would want to do a subgroup analysis but you do not necessarily have to. The first thing you should start off doing is just something that is on the left hand side which is a non-subgroup analysis and if you see that there is no asymmetry then you are feeling pretty good about your analysis and you can continue. If you see that there is asymmetry, then you want to evaluate why that asymmetry is occurring and you can then evaluate whether it is a subgroup asymmetry. However, you yourself with your collaborators have to decide what the subgroups are that you are looking at.

Unidentified Female: Great. We have another question it just came in. Can a single armed non-inferiority clinical trial be included in the meta-analysis?

Unidentified Female: So the single armed…I am not sure what you mean by single armed trial. So that sounds like it would be in non-inferiority so I guess I am…that does not quite…I am not sure quite how to interpret that because it is single armed, I do not see how you can assess non-inferiority. You would need two groups to assess non-inferiority.

Unidentified Female: Maybe we will get a follow up to explain.

Unidentified Female: I mean, if you do have a multitude of single group studies you could include them in a meta-analysis. Those would be observational studies and you would want to keep them separate from your randomized controlled trial.

Unidentified Female: Okay, great. That is all we have coming in for the moment.

Unidentified Female: Great. Well, I do have my email noted so you guys are welcome to email me at \_\_\_\_\_ [01:02:47].gov. I am getting there. There we go. Feel free to email me if you have any other questions.

Unidentified Female: Wonderful. Thank you so much Reesha. We are just pass the top of the hour here so we will wrap things up quickly. Thank you so much for taking the time to prepare and present for today’s session. For the audience you will be prompted with a feedback form in just a moment here. When I close the session out please take a few moments and fill that out. We really do read through all of your feedback and use that for our upcoming sessions. Thank you everyone for joining us for today’s HSR and D-cyber seminar and we look forward to seeing you at a few session. Thank you.