Heterogeneity- Subgroup Analysis and Meta-Regression

Michail Belias

12 February 2019

# Introduction

Bassler et.al (Bassler et al. 2004) conducted a Cochrane Review to evaluate the effects of Ketotifen alone or in combination with other co-interventions in children with asthma and/or wheezing. The primary outcome was the use of rescue bronchodilators. In the systematic review, a random effects model with the risk ratio as measure of treatment effect was used throughout. The meta-analysis of the clinical judgement data contains 10 trials.

Let’s import the data in to start our analysis. In order to do that we have to call the *readxl* package using the library(“readxl”) command. If a package is not installed in your computer you can install by using the *install.packages(‘name of the package’)* command

library("readxl")  
Ketotifen = read\_xlsx("Data/Ketotifen.xlsx")

We can see the first rows of our data with the *head()* command.

head(Ketotifen,   
 n =10 # number of row we want to print  
 )

# A tibble: 10 x 6  
 study Ee Ne Ec Nc blind   
 <chr> <dbl> <dbl> <dbl> <dbl> <chr>   
 1 Chay 1 10 6 10 Blinded  
 2 Rackam 31 68 38 65 Blinded  
 3 Van Asperen 16 52 19 51 Blinded  
 4 Croce 19 39 17 36 unclear  
 5 de Benedictis 7 34 35 41 unclear  
 6 Longo 10 18 15 18 unclear  
 7 Montoya 6 20 14 20 unclear  
 8 Mulhern 6 16 8 15 unclear  
 9 Salmon 8 28 16 34 unclear  
10 Spicak 9 25 20 25 unclear

Our assumption is that the treatment effects come from a distribution rather than from a common effect. We want to generalise our conclusions to broader populations, therefore we will perform a random-effect model. We will use the *meta* package. The primary function for dichotomous outcomes is *metabin()*

library("meta")

**Reminder: An easy way to check the help file of any package and/or functions is to use the *?* or *??* and then the name of the function or package, for instance *??metabin*.** *meta* package has an excellent help file, take advantage of it.

In your computer please fill the … , in the metabin command with the appropriate variables from the data. We want to perform a random-effects meta-analysis with risk ratio as an effect size measure, using the empirical Bayes (EB) as a estimator.

res.RE = metabin(event.e = .. , ## Events of treated  
 n.e = .. , ## Total number of treated  
 event.c = .. , ## Events of control  
 n.c = .., ## Total number of treated  
 sm = .., ## Effect size  
 method = .., ## weight calculation method  
 data = .., ## the data-set  
 studlab = .., ## The trial names  
 method.tau=.., ## tau estimator method  
 comb.fixed =..., ## A logical (TRUE/FALSE) indicating whether a fixed or random   
 comb.random=... ## effect meta-analysis should be conducted.  
 )

Print the output and make a forest plot out of them. Also show the p-value of the overall treatment effect (**Hint: test.overall.random = TRUE**)

**What is the Inverse-Variance method? Are there any other choices? If yes which?**

**Interpret the results**

**Does the Ketotifen help the patients?** **How much is the increase/decrease of the use of rescue bronchodilators ?**

**Describe what Q , , and H are and then report their estimated values**

**Do our data agree with the random-effects model assumption?**

**Is the heterogeneity large?**

## Subgroup analysis

*Until now we 1) did a systematic review, 2) gathered the Ketotifen data, 3) conducted a meta-analysis, in order to find out whether the use of Ketotifen is beneficial or not and 4) we pooled our treatment effects using a random-effects model, because we assumed that the treatment effects of the trials come from a distribution.*

This is one aim of meta-analysis. Another is to investigate our data in order to understand where do these between-study differences come from. We observed, that some trials had a common design. Particularly Chay et al., Rackham et al., and Van Asperen et al. used blinding. We wonder if we can explain some heterogeneity, by splitting the trials into two meta-analyses. One for the blinded and one for not-blinded trials.

**Blinding is a trial characteristic, not a patient characteristic, therefore we can perform subgroup analysis. If the categorical variable was a patient characteristic, for instance gender then we should NOT split the trial and then perform a subgroup analysis. We will discuss this case later on**

TIn order to perform a subgroup analysis we use again the function **metabin()**, but adding also the option **byvar = “name of the subgrouping variable”**.

res.SA = metabin(event.e = .. , ## Events of treated  
 n.e = .. , ## Total number of treated  
 event.c = .. , ## Events of control  
 n.c = .., ## Total number of treated  
 sm = .., ## Effect size  
 method = .., ## weight calculation method  
 data = .., ## the data-set  
 studlab = .., ## The trial names  
 method.tau=.., ## tau estimator method  
 comb.fixed =..., ## A logical (TRUE/FALSE) indicating whether a fixed or random   
 comb.random=..., ## effect meta-analysis should be conducted.  
 byvar = ... ## The splitting variable   
 )

Print the output and make a forest plot out of them. Also show the p-value of the overall treatment effect (**Hint: test.overall.random = TRUE**)

**Note: for better looking forest-plots see the help file ( ??forest.meta )**

The subgroup analysis shows that method of blinding does not explain adequately the statistical heterogeneity between trials. *Why?*

**Report the results for subgroups (random effects model) table (copy-paste)**

**What are the values of , , H and now? What do they represent?**

**Report the Test for subgroup differences (copy-paste). How should you interpret it? Discuss Q, d.f. and p-value.**

Subgroup analysis has a lot in common with a meta-regression with a categorical independent variable. We can also perform a subgroup analysis in a “linear” meta-regression framework. Particularly, we can use the *metareg()* function to fit a meta-regression, using the log risk ratios as a dependent variable and the blinding as an independent one. The options need a *meta* object (the random effects meta-analysis we performed earlier is one) and the categorical variable we wish to use (in our case blinding).

**Fill the ….**

res.MR.SA =metareg(x = .... , ## an object of class meta  
 formula = ....,  
 hakn = T)

Compare the subgroup analysis with the meta-regression output. There some minor differences, but this is due the fact that the within the blinded and unclear groups are different in subgroup analysis, while in meta-regression they are the same. If in the subgroup analysis we “force” the s for both subgroups to be equal, then the results of the meta-regression are exactly the same as Subgroup analysis.

res.SA.2 = metabin(event.e = Ee , ## Events of treated  
 n.e = Ne , ## Total number of treated  
 event.c = Ec , ## Events of control  
 n.c = Nc, ## Total number of treated  
 sm = "RR", ## Effect size  
 method = "Inverse", ## weight calculation method  
 data = Ketotifen, ## the data-set  
 studlab = study, ## The trial names  
 method.tau="EB", ## tau estimator method  
 byvar = blind, ## The splitting variable   
 comb.fixed = F, ## A logical variable (True/False) indicating   
 ## whether a fixed effect meta-analysis should be conducted.  
 tau.common = T  
 )  
  
# forest(res.SA)

## Meta-regression with a continuous covariate

In this section we will try to explain again part of the heterogeneity, but instead of using a categorical variable we will use a continuous.

We will use the meta-analysis of Colditz et al. (Colditz 1994), where he evaluated the overall effectiveness of the Bacillus Calmette-Guerin vaccine against tuberculosis. In addition, covariates that may potentially influence the effect of vaccination were examined.

**Note: we using AGAIN a variable that is a trial characteristic, not a patient characteristic**

library(metafor)  
dat <- dat.colditz1994  
head(dat)

trial author year tpos tneg cpos cneg ablat alloc  
1 1 Aronson 1948 4 119 11 128 44 random  
2 2 Ferguson & Simes 1949 6 300 29 274 55 random  
3 3 Rosenthal et al 1960 3 228 11 209 42 random  
4 4 Hart & Sutherland 1977 62 13536 248 12619 52 random  
5 5 Frimodt-Moller et al 1973 33 5036 47 5761 13 alternate  
6 6 Stein & Aronson 1953 180 1361 372 1079 44 alternate

We believe that part of our data heterogeneity can be due to the country’s place on the world map. That is a resonable thing in epidemic, since the climate may affect the disease. Therefore, we will fit a meta-regression using the logRR as dependent variable and he absolute geographical latitude.

First, perform a random-effects meta-analysis and then fit a meta-regression using the absolute geographical latitude (“ablat”) as a moderator.

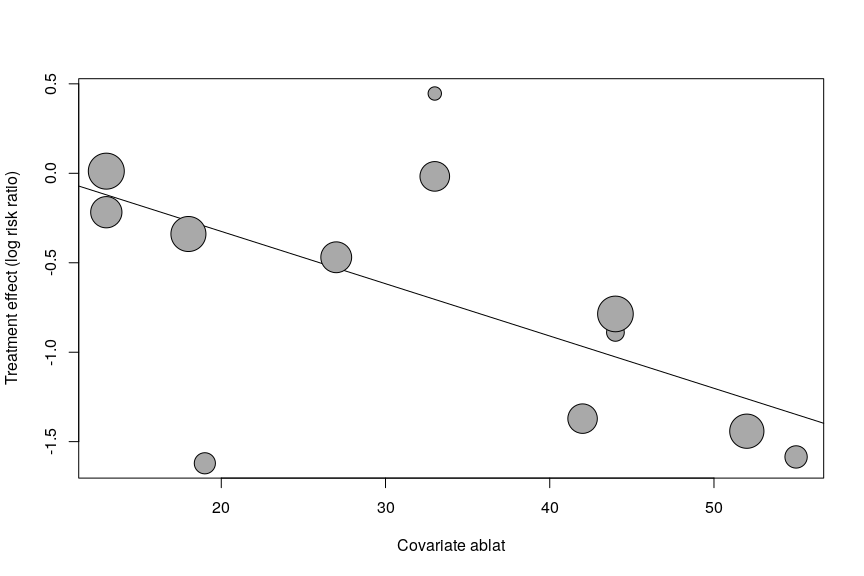
res.RE.MR =metareg(x = .... , ## an object of class meta  
 formula = ....,  
 hakn = T)

**Report the Test for Residual Heterogeneity and the Test for Moderators**

**What are our conclusions?**

We can also plot the effect estimates over the range of the absolute geographical latitude, so that we may have a visual represantion, using the *bubble()* command. See the help file (*??bubble.metareg*).

bubble(res.RE.MR)



**Based on our model calculate the predicted logRR (and then RR) for Nijmegen (or any other city you like)**

**What are the absolute geographical latitude ranges we can safely predict?**

# References

Bassler, Dirk, Andrew AD Mitra, Francine M Ducharme, Johannes Forster, and Guido Schwarzer. 2004. “Ketotifen Alone or as Additional Medication for Long-Term Control of Asthma and Wheeze in Children.” *Cochrane Database of Systematic Reviews*, January. <https://doi.org/10.1002/14651858.cd001384.pub2>.

Colditz, Graham A. 1994. “Efficacy of BCG Vaccine in the Prevention of Tuberculosis.” *JAMA* 271 (9): 698. <https://doi.org/10.1001/jama.1994.03510330076038>.