Meta-Analysis in R

Lin Gu, Cooperative Studies Program Epidemiology Center, Veterans Affairs Health Care, Durham, NC; Duke University, Durham, NC;

Alyssa Jasmine Bullard, Cooperative Studies Program Epidemiology Center, Veterans Affairs Health Care, Durham, NC;

Kevin Su, Cooperative Studies Program Epidemiology Center, Veterans Affairs Health Care, Durham, NC;

Christina D. Williams, Cooperative Studies Program Epidemiology Center, Veterans Affairs Health Care, Durham, NC; Duke University, Durham, NC

ABSTRACT

With the accumulation of a large body of information and data in clinical trial studies, questions are often raised to estimate the pooled effect size for similarly designed studies for a particular research topic. Meta-analysis is such a statistical technique employed for this purpose and has become a universally accepted research tool. In drug development, the U.S. Food and Drug Administration (FDA) has published draft guidance on how to conduct a meta-analysis of randomized clinical trials to evaluate safety risks associated with the use of human drugs or biological products within the framework of regulatory decision making (FDA, 2018).

SAS® has been a widely used tool to perform a wide range of statistical analyses, but SAS® procedures for meta-analysis have been lacking. On the other hand, R is a powerful and increasingly popular software not only for statistical analysis but also for visualization. In this paper, R is the tool used to demonstrate end-to-end how to perform meta-analysis.

Key words: Meta-analysis, Meta, R, Hazard Ratio, Forest Plot

INTRODUCTION

Meta-analysis is a statistical methodology that leverages published studies by integrating the outcome estimates from multiple studies into a single estimate. This method is frequently used in medical and social sciences. Because combining multiple studies potentially results in increased sample size and precision, meta-analytic results are considered the most trustworthy source of evidence by the evidence-based medicine literature.

Meta-analysis is especially useful for subgroup analysis. In an individual randomized controlled study, adequate sample size and power are preserved for the primary objective analysis whereas analyses on population subgroups such as age and race do not have adequate power and sample size. Pooling multiple studies increases the power and sample size, thus improving the validity of the statistical analysis.

Over the past several decades, concurrent with development of statistical methodology for meta-analysis, numerous software programs have been developed for meta-analysis. R is an example of such software and is considered to be the most powerful and versatile language. R is open-source, and its packages are often developed with participation of statisticians. This makes R as particularly suitable for meta-analysis. Incorporating the knowledge of statisticians allows the most current statistical methodology to be integrated into R packages in a timely manner, which is less likely for proprietary software. It is worth pointing out that over the last decades, R has evolved to include web-based technology, which makes interactive visualization in statistical analysis possible.

There are numerous R packages written for meta-analysis, including Metafor (Viechtbauer, 2010), Meta (Schwarzer et al, 2015), and Rmeta (Lumley, 2018). Meta package, developed by Guido Schwarzer et al., includes comprehensive collection of functions with a plethora of options that allow meta-analysis on desired statistical methods. This paper will use Meta packages to illustrate the process of conducting meta-analysis of hazard ratios (HR).

STEPS OF META-ANALYSIS

To start meta-analysis, the first step is to propose a focused research question. For example, suppose we have observed conflicting reports about whether a drug X would increase risk of internal bleeding and we would like to investigate what a unified conclusion will be if we combine all the studies related to this question. A focused question can be asked: does a drug X increase internal bleeding? If it does, are there any differences in risk of internal bleeding among different age groups?

Once a research question is determined, the next step is to define eligibility criteria for selecting studies that will be included in the meta-analysis. The widely accepted principle to follow is called PICO, an abbreviation for Population, Intervention, Control, and Outcome. For examples, what population do we want to include? What age group? For intervention, how long of an exposure duration do we want to include? Is there any placebo control or active comparator? What type of studies to include (e.g., randomized, open label, observational)? For outcomes, is it hazards ratio (HR), odds ratio, or mean differences?

After you choose the research topic and decide the eligibility criteria, the next step is to write an analysis plan. This is an essential step to make your analysis organized, intentional, and reproducible. In the analysis plan, the number of studies, type of effect sizes, statistical analysis method(i.e., fixed-effect model or random-effect model), and power analysis are defined. Specification of analysis plan before the actual data are extracted is critical to ensure methods are appropriate for the research question and to minimize the tendency to modify methods for favored results.

The next step is to search literature databases for the studies related to your research question. The objective at this point is to find all available evidence on the research topic in order to get an unbiased and comprehensive view of the findings. For medical research, PubMed is a commonly used database (PubMed (nih.gov)), and the Cochrane Central Register of Control Trials is another option for identifying studies

The database searches may yield numerous references that are related to your research question. At this stage, follow the **Preferred Reporting Items for Systematic Reviews and Meta-Analyses (**PRISMA, Moher et al. 2009) framework to select studies.

The final step is to perform the meta-analysis based on methodology established in the analysis plan.

OBJECTIVE

The objective of this analysis is to synthesize results of the multiple published studies to estimate the pooled HR between figurative investigational drug X and standard care.

DATA

After extensive database searching and evaluation of literature according to PICO and PRISMA guidance, HR and its Confidence Intervals (CI) for each study are extracted (Table 1). Normally, in scientific reports related to HR, CI are expected to be reported along with HR. If CI are not reported, other parameters could also be used. For example, median, interquartile range and range (arguments median, Q1, Q3, min, and max) can also be used (Schwarzer et al, 2015).

Author	HR	lower.HR	upper.HR
Author A	0.5529	0.2801	1.0914
Author B	0.9275	0.7934	1.0842
Author C	0.7910	0.5951	1.0515
Author D	1.1796	0.6405	2.1724
Author E	0.1926	0.0503	0.7374
Author F	1.3434	0.5812	3.1051
Author G	0.7554	0.3724	1.5323

Author H	0.5664	0.2731	1.1745
Author I	1.6134	0.7763	3.3533
Author J	2.6738	0.8405	8.5056

Table 1. Hazard Ratio (HR) and Confidence Interval (Lower and Upper of HR) of Select Studies

METHOD

R package meta is used to perform the meta-analysis for this pooled study. The input data used in this report is the pre-calculated effect size, which is the hazard ratio of treatment over placebo associated with risk of getting internal bleeding and its CI for each study. For this type of data, meta recommends using function metagen, which stands for **gen**eric inverse variance meta-analysis. This method uses inverse-**variance pooling method** to calculate the pooled effect size for combined studies. The required parameters for this function to run are TE (treatment effects) and seTE (standard errors of treatment effects). The standard errors are not typically reported in literature. CIs can be used to estimate seTE. The estimated TEs and CIs must be log transformed.

If study effect sizes do not come from a homogenous population, a random-effects model will be used to estimate the pooled effect size. Meta incorporated many methods to estimate the between-study variance Tau^2. Below is a list of the available methods in meta-analysis functions.

- method.tau = "REML" Restricted maximum-likelihood estimator

- method.tau = "PM" Paule-Mandel estimator

- method.tau = "DL" DerSimonian-Laird estimator

- method.tau = "ML" Maximum-likelihood estimator

- method.tau = "HS" Hunter-Schmidt estimator

- method.tau = "SJ" Sidik-Jonkman estimator

- method.tau = "HE" Hedges estimator

- method.tau = "EB" Empirical Bayes estimator

For illustrative purposes, the Tau^2 is estimated using DerSimonian-Laird method, the simplest and most commonly used method when fitting the random effect model for meta-analysis.

CI for random effects estimate: The following methods are available in all meta-analysis functions to calculate a CI for the random effect estimates.

- method.random.ci = "classic" based on standard normal quantile

- method.random.ci = "HK" method by Hartung and Knapp

- method.random.ci = "KR" method by Kenward-Roger

To estimate the CI for the pooled effect size, use method.random.ci = "classic" based on the standard normal quantile (DerSimonian and Laird, 1986).

RESULTS

R code

Below are the R codes for using generic inverse variance method in meta-analysis.

```
install.packages("meta")
library(meta)

## function settings.meta print and change default settings to conduct and print or plot meta-
analyses in R package meta.
## The following general settings are available: Review Manager 5, Journal of the American
Medical Association.
settings.meta('RevMan5')
#settings.meta('JAMA')

## input data
n=10
study  <- paste(rep('Author', n), LETTERS[1:n])

HR       <- c(0.5529,0.9275,0.7910,1.1796,0.1926,1.3434,0.7554,0.5664,1.6134,2.6738)
lower.HR <- c(0.2801,0.7934,0.5951,0.6405,0.0503,0.5812,0.3724,0.2731,0.7763,0.8405)
upper.HR <- c(1.0914,1.0842,1.0515,2.1724,0.7374,3.1051,1.5323,1.1745,3.3533,8.5056)



## run function metagen to get estimates using generic inverse variance method
m<-metagen(
    HR = log(HR), lower = log(lower.HR), upper = log(upper.HR),
    sm = "HR", fixed=F, random=T,
    studlab = study,
    method.tau = "DL",   ## method to calculate Tau
    method.random.ci = "classic",  ## method to calculate estimator's CI
)

summary(m)

## Forest plot of the estimates
forest(m, xlim = c(0.3,4.5))

## Funnel plot of the meta-analysis
funnel.meta(m,  studlab = TRUE)
```

At minimum, the function **metagen** takes log of HR, log of lower bound, log of upper bound, summary measure sm, and studlab as arguments. The estimation method of between-study heterogeneity Tau^2 is default as DL and estimation of CI for estimator of pooled effect is 'classic'.

Summary results

Below is the output after running the code m*etagen*.

```
           HR         95%-CI %        W(random)

Author A 0.5529 [0.2801, 1.0914]      8.6

Author B 0.9275 [0.7934, 1.0842]     24.6

Author C 0.7910 [0.5951, 1.0515]     19.9

Author D 1.1796 [0.6405, 2.1724]     10.0

Author E 0.1926 [0.0503, 0.7374]      2.9

Author F 1.3434 [0.5812, 3.1051]      6.4

Author G 0.7554 [0.3724, 1.5323]      8.2

Author H 0.5664 [0.2731, 1.1745]      7.8

Author I 1.6134 [0.7763, 3.3533]      7.8

Author J 2.6738 [0.8405, 8.5056]      3.8


Number of studies combined: k = 10


                HR        95%-CI    z p-value
Random effects model 0.8851 [0.6952, 1.1269] -0.99    .322


Quantifying heterogeneity:
 tau^2 = 0.055 [0.013, 0.844]; tau = 0.2350 [0.1162, 0.9187]
 I^2 = 48% [0%, 75%]; H = 1.38 [1.00, 1.99]


Test of heterogeneity:
   Q d.f. p-value
 17.23   9   .0453


Details on meta-analytical method:
- Inverse variance method
- DerSimonian-Laird estimator for tau^2
- Jackson method for confidence interval of tau^2 and tau
```

The first section of the outputs shows the input data you have extracted from the literature, including Authors, HRs (the standard care is the reference), and CIs. The `%W(random)` column contains the weights (in percent) that the random-effects model attributed to each study. The Author B has the greatest weight because its CI is narrowest.

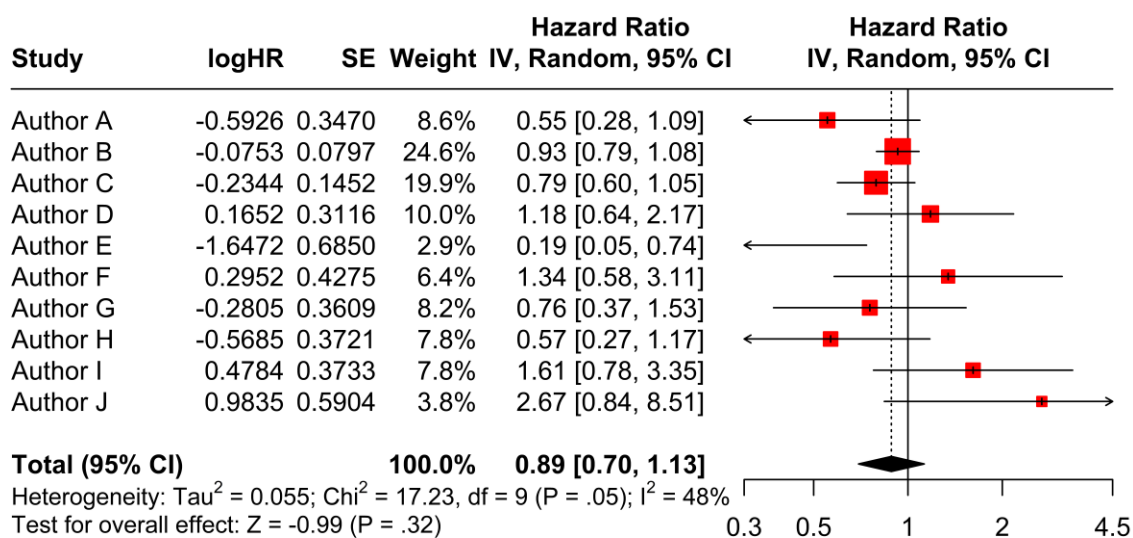The outputs show that there are 10 studies combined.

The third section displays the core results. The pooled effect is 0.8851 with 95% CI from 0.6952 to 1.1269. The effect is not significant with p = 0.322 at a significant level of 0.05. We can conclude that the meta-analysis did not find statistically significant different risk of internal bleeding among patients treated with Drug X and patients received the standard care.

The Tau^2 is found to be 0.055 with CI from 0.013 and 0.844, which does not cover 0, suggesting that study heterogeneity exists among studies. The I^2 statistic (J. P. Higgins and Thompson 2002) is another way to quantify between-study heterogeneity, and directly based on Cochran's Q. The I^2 is estimated to be 48% indicating moderate heterogeneity.

The test of heterogeneity section shows the **Cochran's** Q **test statistics, which follow a** Chi-square distribution. The p value is 0.0435, indicting the significant heterogeneity between studies and so use of the random-effect mode is a good choice.

Forest plot

To graphically depict the output of individual studies and pooled estimate, a forest plot is created below.

Study	logHR	SE	Weight	Hazard Ratio IV, Random, 95% CI	Hazard Ratio IV, Random, 95% CI
Author A	-0.5926	0.3470	8.6%	0.55 [0.28, 1.09]	
Author B	-0.0753	0.0797	24.6%	0.93 [0.79, 1.08]	
Author C	-0.2344	0.1452	19.9%	0.79 [0.60, 1.05]	
Author D	0.1652	0.3116	10.0%	1.18 [0.64, 2.17]	
Author E	-1.6472	0.6850	2.9%	0.19 [0.05, 0.74]	
Author F	0.2952	0.4275	6.4%	1.34 [0.58, 3.11]	
Author G	-0.2805	0.3609	8.2%	0.76 [0.37, 1.53]	
Author H	-0.5685	0.3721	7.8%	0.57 [0.27, 1.17]	
Author I	0.4784	0.3733	7.8%	1.61 [0.78, 3.35]	
Author J	0.9835	0.5904	3.8%	2.67 [0.84, 8.51]	
Total (95% CI)			**100.0%**	**0.89 [0.70, 1.13]**	

Heterogeneity: Tau^2 = 0.055; Chi^2 = 17.23, df = 9 (P = .05); I^2 = 48%
Test for overall effect: Z = -0.99 (P = .32)

Error! Reference source not found.

In Figure 1, each row of the first section represents the individual studies. The second and third columns represent effect sizes (TE, the logged HR) and its standard errors (SE). The Weight column represents the weight (the inverse of squared SE, usually but not always represents the sample size of the study) of each study attributed. The weight is calculated by inversing the squared SE. A study with bigger sample size usually has a smaller SE, so it carries more weight in the meta-analysis. This is not necessarily the case though – a larger study with few numbers of events may have a larger SE. The fifth column represent HRs and their 95% CIs.

In the second section, the pooled effect size estimator and its CI are in the 4th column (HR 0.89, 95% CI: 0.70-1.13). The model statistics including between-study heterogeneity and test scores are displayed in the second row. The Tau^2 is 0.055 and chi-square is 17.23 corresponding the outputs in the summary Results section.

In the forest plot on the right, the square box represents point estimator of HR for each study. The size of the square box is proportional to the weight. The CIs are represented by lines.

The diamond in the bottom represents the pooled estimate of the effect size. For HR, if the value crosses 1, it suggests that no statistically significant differences in hazard ratio between treatment and placebo were found from the meta-analysis.

Publication Bias

Publication bias exists when negative and not significant results did not get an equal chance to be published. The consequence of publication bias has potential risk of invalidating summarized results.

There are many ways to assess the publication bias. One commonly used method is to use a funnel plot to examine the relationship between precision errors and effect sizes of studies. In essence, a funnel plot is a scatter plot of the studies' observed effect sizes on the x-axis versus its standard errors on the y-axis. The rationale behind the funnel plot is that small studies with greater standard errors should get published with symmetrical distribution of effect sizes along the x-axis.
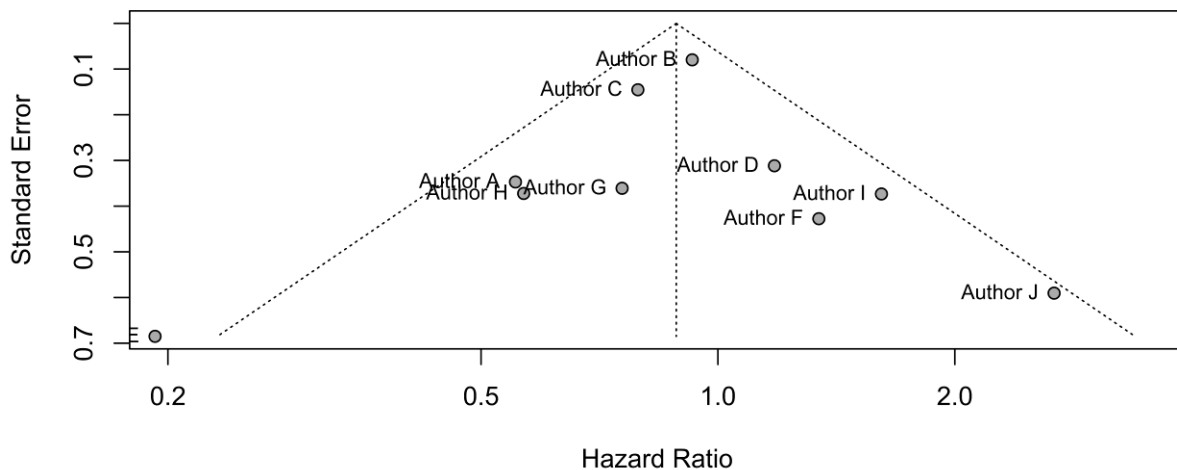

Figure 1. Funnel plot of Standard Error against Hazard Ratio

Figure 2 is a funnel plot of standard errors against hazard ratios. The vertical line in the middle represents the pooled effect size. Without publication bias, the studies should be scattered symmetrically around the vertical line. If there is a publication bias, the distribution of studies will be skewed to one side. It can be seen in Figure 2 that the studies are approximately distributed symmetrically, suggesting that publication bias is minimal.

It should be noted that using a funnel plot to assess the publication bias is based on visual judgment on symmetry and thus is subjective. However, the funnel symmetry can be tested using Egger's regression test (Egger et al. 1997). Due to paper restrictions, h, this test will not be discussed here.

CONCLUSION

This paper briefly introduced the meta-analysis, demonstrated the process of how to use R package Meta to conduct meta-analysis on HR with CI, and explained in detail how to interpret the output of the analysis.

It should be noted that this paper is a brief overview. Readers who are interested in the details of statistical theory should read the references provided for topics including but not limited to methods for estimating pooled effect size, estimation of Tau^2, and the full use of Meta package.

REFERENCES

DerSimonian, Rebecca, and Nan Laird. 1986. "Meta-Analysis in Clinical Trials." Controlled Clinical Trials 7 (3): 177–88.

Egger, Matthias, George Davey Smith, Martin Schneider, and Christoph Minder. 1997. "Bias in Meta-Analysis Detected by a Simple, Graphical Test." *BMJ* 315 (7109): 629–34.

Meta-Analyses of Randomized Controlled Clinical Trials to Evaluate the Safety of Human Drugs or Biological Products. FDA 2018.

Higgins, Julian PT, and Simon G Thompson. 2002. "Quantifying Heterogeneity in a Meta-Analysis." *Statistics in Medicine* 21 (11): 1539–58.

Moher, David, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, Prisma Group, et al. 2009. "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement." *PLoS Medicine* 6 (7).

Schwarzer et al. (2015), https://link.springer.com/book/10.1007/978-3-319-21416-0.

Thomas Lumley (2018), "Rmeta." Journal Article. CRAN, Version 3.0.

Viechtbauer W (2005): Bias and efficiency of meta-analytic variance estimators in the random effects model. Journal of Educational and Behavioral Statistics, 30, 261–93

Viechtbauer W (2010): Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software, 36, 1–48.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Lin Gu
Enterprise: Cooperative Studies Program Epidemiology Center, Durham, Veterans Affairs Health Care System
Phone: 1-919-268-0411
E-mail: lin.gu@va.gov