How to perform meta-analysis in R: a simple yet comprehensive guide

Muhammad Aaqib Shamim[1]🔵, Aravind P Gandhi[2]🔵, Pradeep Dwivedi[3]🔵, Bijaya Kumar Padhi[4]*🔵

[1] Department of Pharmacology, All India Institute of Medical Sciences, Jodhpur, Rajasthan, India.
[2] Department of Community Medicine, All India Institute of Medical Sciences, Nagpur, Maharashtra, India.
[3] Centre of Excellence for Tribal Health, All India Institute of Medical Sciences, Jodhpur, Rajasthan, India.
[4] Department of Community Medicine and School of Public Health, Postgraduate Institute of Medical Education and Research, Chandigarh, India.

*Correspondence: bkpadhi@gmail.com

Cite this Article

Evidence in Context

Prior Evidence: In evidence synthesis research, there is a noticeable scarcity of comprehensive tutorials on conducting meta-analysis using R, a robust and open-source software platform. Furthermore, existing literature does not adequately address the integration of various study designs with the latest statistical methodologies in meta-analysis.

Evidence added by this study: We start with a concise elucidation of the core statistical principles inherent to meta-analysis. This is followed by an in-depth exploration of their application in R. Our article methodically covers an array of study designs, integrates illustrative figures crafted with academic rigor, provides comprehensive statistical reporting templates, and elucidates the nuances of modern statistical techniques. We follow up with pointers for further reading.

To view Article

Abstract

Systematic review and meta-analyses (SRMA) form the cornerstone of evidence-based medicine and evidence-based public health. SRMA occupies the highest position in the evidence pyramid. Thus, researchers, clinicians, public health professionals, policymakers, and other stakeholders place much trust in the findings of an SRMA. So, these interpretations must be methodologically robust and statistically sound. No comprehensive and sound tutorial is available for meta-analysis of different effect sizes condensed in a single article. This is an easy yet comprehensive guide to fill this gap. It also incorporates some advanced but essential meta-analytical techniques. Everything has been explained simply in open-source and free software, i.e., R. This allows greater access by researchers from lower- and middle-income countries, ensuring inclusivity and equitable access to health research. Although the primary target readers are beginners of R, seasoned R users can also benefit from the advanced analyses introduced in the meta-analysis, thereby addressing the needs of both groups. It explains statistical concepts briefly as necessary and demonstrates their implementation. The code file for the demonstrated examples and the relevant data extraction sheets have been attached. Then, we have provided examples for reporting the results of a systematic review and meta-analysis, troubleshooting common mistakes, customising the output, changing the finer settings, and exploring other important aspects of meta-analysis. We have finally discussed other sources for further reading for those interested in delving into more advanced and complicated techniques.

Keywords: *Systematic review, Tutorial, Course, Step-by-step guide, Statistical analysis, R Studio, R Software, Troubleshooting, Funnel plot, Egger's regression, Egger's test, Doi plot, LFK index, Prediction interval, Heterogeneity, Bubble plot, Meta-regression*

Introduction

High-quality evidence generation and synthesis of evidence from research studies form the cornerstone of Evidence Based Medicine (EBM)[1–3] and Evidence Based Public Health (EBPH) [4–8]. Meta-analysis is one of the powerful techniques for evidence synthesis[9–11]. It estimates the pooled (summary) effect size from multiple studies answering the same research question. This is done by applying appropriate statistical methods. Apart from the pooled effect sizes, it also assists in assessing and exploring the heterogeneity by means of sub-group analysis, meta-regression, outlier assessment, sensitivity analysis and evaluating the publication bias.

Statistical software and programs are available to enable the effective and accurate conduct of the meta-analysis. RevMan, STATA, SPSS, R and Comprehensive Meta-Analysis (CMA) are some of the most well-known tools. Among these RevMan and R software are free of cost for researchers worldwide. While RevMan operates by graphic user interface (GUI)/Menu-driven user interface, R software is a command line interface (CLI). Though the CLI makes R cumbersome for beginners, its customisability expands the scope of the analysis that can be undertaken with R. For instance, meta-regression and network meta-analysis are not possible with RevMan, while it is feasible in R. There are several other advantages, including but not limited to a better assessment of publication bias, and computation of prediction intervals. Thus, R has a significant advantage over other software for meta-analysis. However, the cumbersome and CLI nature of the R can make it difficult for the researchers, especially the clinical and medical researchers, to accept, adopt and implement the R software in their meta-analysis.

An easy-to-use, stepwise guide to assist the clinical and medical researchers involved in EBM and EBPH would improve the quality and quantity of the evidence synthesis work done. Moreover, R has a very rewarding learning curve. Once comfortable with the software, one can use it for various uses, including SRMAs, health economics and outcomes research, bioinformatics, and other analyses in primary research.

Studies in the past have reported stepwise guides by including the codes and applications for conducting the meta-analysis in R software[12–16]. However, they were limited to the two-group meta-analysis with binary outcomes[14,16] or continuous outcomes only in pre-post designs[15]. Several important aspects like heterogeneity assessment and exploration, publication bias assessment and sensitivity analysis have been skipped or explained inadequately[12–15]. With the rise in the single group meta-analysis (proportion and continuous outcome), it is imperative to disseminate the stepwise R guide to enable such meta-analysis. Newer methods to detect and report publication bias (Doi plots and LFK index)[17], steps for the outlier detections, and prediction intervention calculations[18] were also not reported in the previous studies. Hence, the current article has been framed to address the above gaps and provide a stepwise guide from installing R, along with R codes for the standard and newer analyses for single-group and two-group meta-analyses (proportion/binary and continuous outcome).

Learning with an example

This article focuses on assisting researchers who have never used R or have no knowledge about R, but are eager to follow the steps closely to learn and perform their first meta-analysis using R. We have summarised the steps in **Figure 1**. Though this may seem daunting, it is explained in simple words in the subsequent sections. It should serve as a good recap once the reader has finished reading the article, alongside being a primer to the workflow.
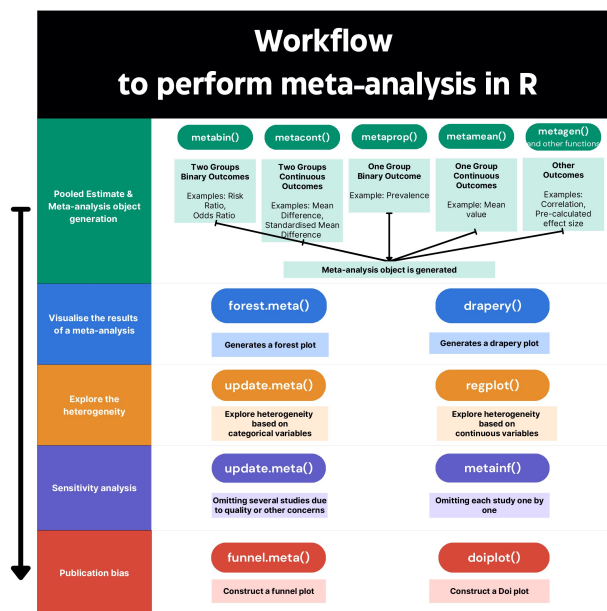

Figure 1: Primer to the workflow for systematic review and meta-analysis in R

To begin with, we will discuss a common scenario. Here, we are performing a meta-analysis of several randomised controlled trials (RCTs) that assess the effect of intervention X on outcome Y. The outcome Y is expressed as continuous data (e.g., weight), and is mentioned in the RCTs as mean ± standard deviation. Once we complete this systematically to produce the primary and ancillary analyses, it will be easier to proceed to other cases. These other cases may differ in having only a single group (instead of a comparison between two groups) or having proportional data (instead of continuous data). The background steps will be the same, with minor modifications only in the R packages and commands. The researchers must know only a few minor adjustments for these other cases to replicate most or all these analyses for meta-analysis of such studies.

Installation and preparation of R

Initially, you need to install R[19], and then RStudio[20]. These steps have been discussed extensively online[12,21–24]. After installing R and RStudio, one should see four panes, as shown in **Figure 2**. If all four panes are not seen, one can click View in the top bar, then click on *Panes* followed by *Show All Panes*. The lines of code can be typed in either the console pane or the source pane (which is more convenient, easier to view, easier to edit, and allows saving the code for reproducibility and updating). To write in the console pane, one can directly start writing there as it is opened. For the source pane, one can write in the empty file that is open. Alternatively, they can click the 'new' button in the top left and open a new 'R Script' **(Figure 2)**. Then, there should be a blank space with line number(s) on the left. We can write the code here. One can save the file by pressing Ctrl + S (Command + S) or clicking File and Save.
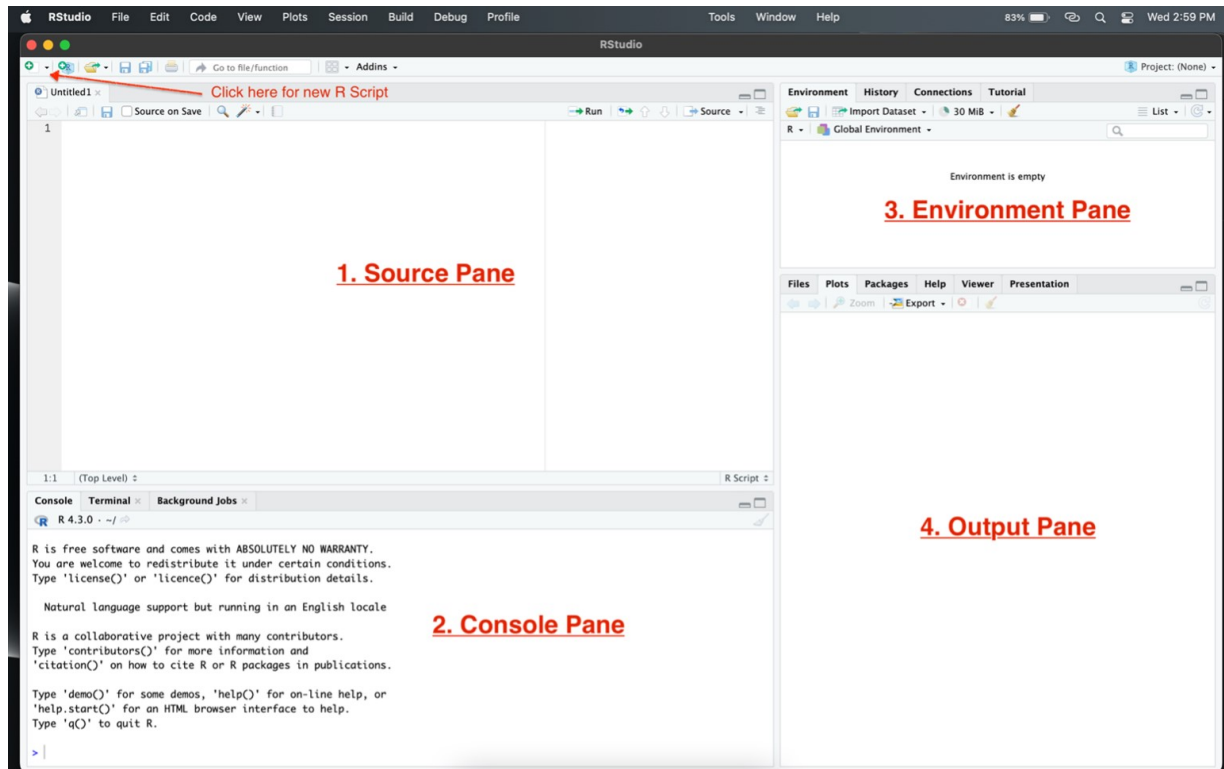
Figure 2: RStudio interface and the four consoles

We will install four packages (meta[16], metafor[25], metasens[26] and readxl[27]) that simplify the whole workflow. Download the sample code file in the supplementary **[Data S1]** and open it in RStudio. Run the first four lines of this file. This can be done by selecting the first four lines completely and pressing Ctrl + Enter (Command + Enter in MacOS) or clicking the 'run' button to run the selected lines. This should install these packages.

> *install.packages("readxl")*
> *install.packages("meta")*
> *install.packages("metafor")*
> *install.packages("metasens")*

This step only needs to be done when one is setting up their R for meta-analysis for the first time. Later, they do not need to keep downloading and reinstalling these packages. Moreover, this is the only step that requires a connection to the internet. All the other steps of meta-analysis in R can be completed offline. Whenever we need to do a meta-analysis, the installed packages must be called through the following four lines in the code.

> *library("readxl")*
> *library("meta")*
> *library("metafor")*
> *library("metasens")*

If this is done correctly, we have successfully prepared R for performing meta-analysis and can now prepare or clean our datasheet.

Preparation of our spreadsheet (or .xls file) and importing it

Spreadsheet preparation

The readers can download a template data extraction spreadsheet from the supplementary **[Data S2]** and open it. It is much simpler and easier than it looks at first glance. We will initially see what the acronyms mean. 'n' stands for the number of participants, 'mean' stands for the mean value, 'sd' stands for standard deviation, 'e' stands for experimental / exposure arm, and 'c' stands for control arm. Thus, the column names (column headings) convey these meanings:

- authors: It is the study identifier. It is a convention to write it as the name of the first author followed by the year of publication. However, this can be written in any format.

- n.e: Number of participants in the experimental / exposure arm

- mean.e: Mean value of the outcome in the experimental / exposure arm

- sd.e: Standard deviation of the outcome in the experimental / exposure arm

- n.c: Number of participants in the control arm

- mean.c: Mean value of the outcome in the control arm

- sd.c: Standard deviation of the outcome in the control arm

This is the minimum requirement. The other columns will be discussed later in the respective sections as required.

Importing the spreadsheet into R

Now, both R and the spreadsheet file are ready. Before integrating the two, we need to prepare our R working directory. The working directory is a location or folder on our desktop where we can store the input (spreadsheet files) and save the output (plots). Thus, all the files relevant to a project are stored in one place, ensuring ease. We have two options to choose this folder or location as our working directory. We can use the keyboard shortcut Ctrl+Shift+H. Alternatively, we can click "Session -> Set Working Directory -> Choose Directory" **[Figure 3]**. Here, select the desired location. Keep the required spreadsheet file containing the data (the datasheets attached in the supplementary in this case) here.
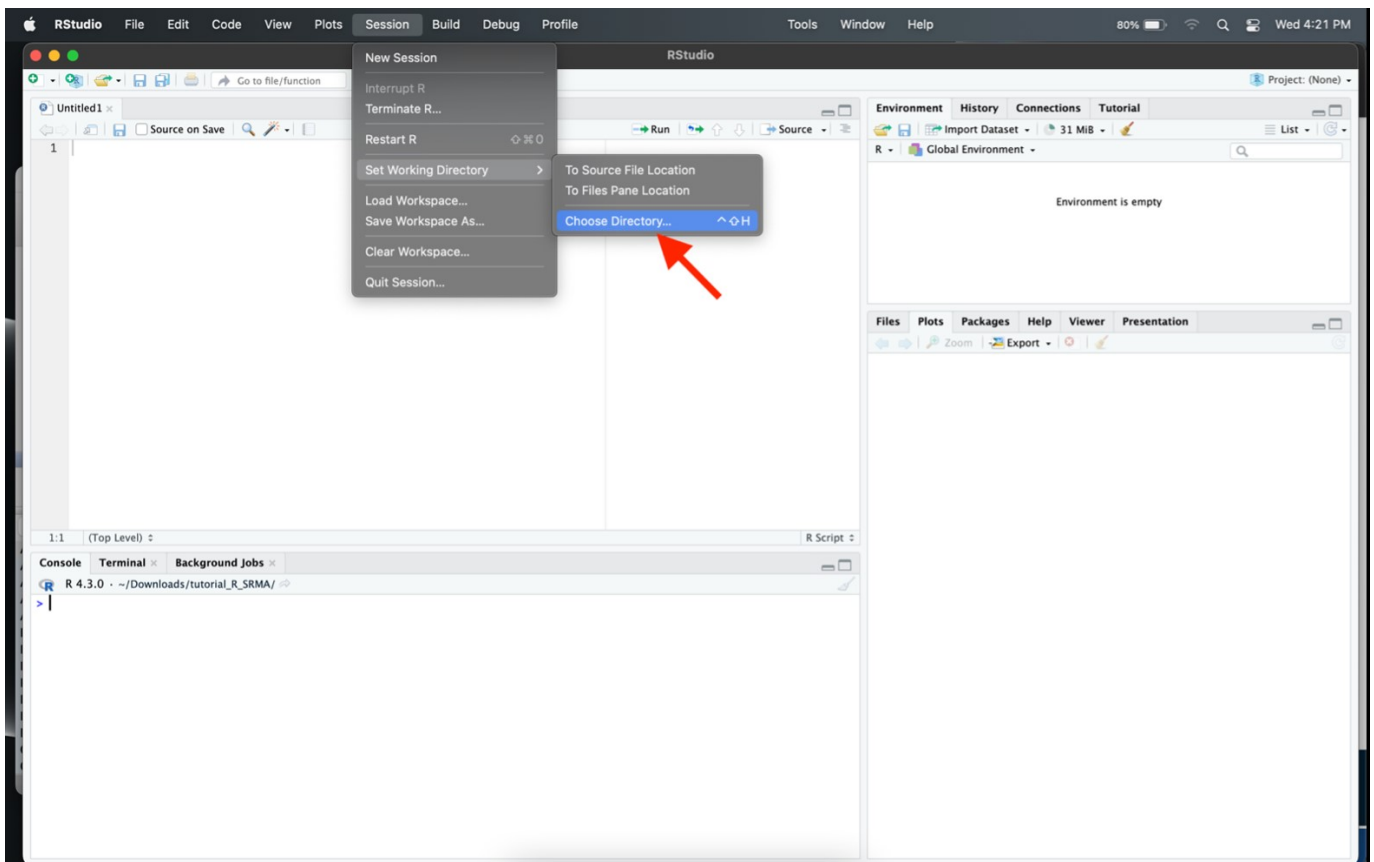
Figure 3: Setting working directory in R

Now, it is time to integrate R and the spreadsheet file. Open the R script file **[Data S1]**, and after calling the packages using *install()* and setting the working directory, run this line.
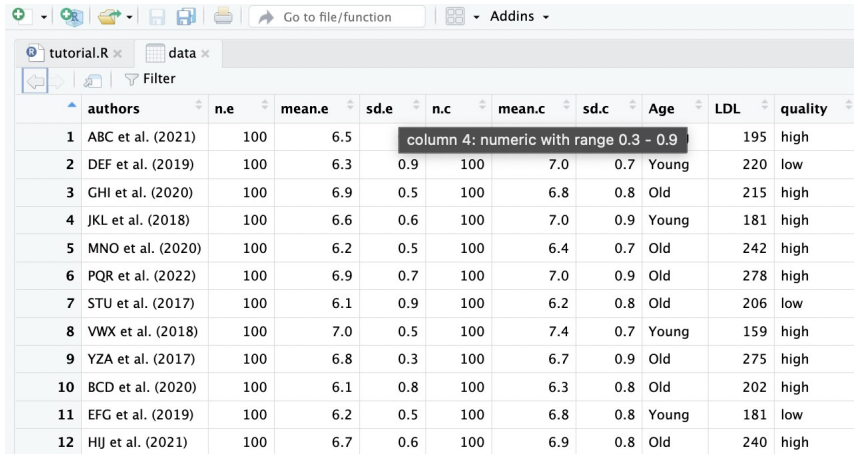
> *data <- read_xlsx("tutorial_md.xlsx")*

This imports the table as an object named *data*. In other words, it imports the spreadsheet file named *tutorial_md.xlsx* present in the assigned working directory. It is imported into the R environment and stored as an object named *data*.

If we wish to break down this code, it says it uses a function named read_xlsx for this work. The assignment operator (<-) assigns this table (the output) to the *data* object. Thus, the output of this function (i.e. the table) is stored as an object named *data*. We can now view this data by running this line **[Figure 4]**.

> *View(data)*

Figure 4: Viewing imported data in R

Now, one can check the type of data in each column. The main thing to verify is that the numerical data (the six columns from n.e to sd.c in this case) should show their structure as *num* instead of *chr*.

> *str(data)*
Now, we can move to the meta-analysis proper.

Meta-analysis: forest plot and pooled estimate

As mentioned earlier, we are performing a meta-analysis comparing HbA1c (a continuous outcome expressed as mean and SD) between two groups (experimental and control). We will use the *metacont* function. We will "feed" the function with the data and several arguments (like the name of our data and column names). This is an example according to the *data* that we have imported earlier.

> *metta <- metacont(n.e = n.e ,*
mean.e = mean.e,
sd.e = sd.e,
n.c = n.c,
mean.c = mean.c,
sd.c = sd.c,
studlab = authors,
data = data,
sm = "MD",
prediction = TRUE)

Here, we are performing the meta-analysis using a function *metacont*. There are several arguments within this function. The words to the left of the equal symbol are argument names, and the words to the right are the values we are passing to those arguments. The commas separate the arguments. We are storing this as an object named *metta*.

The arguments mean the following:
- n.e: Specify the column containing the number of participants in the experimental / exposure arm

- mean.e: Specify the column containing the mean value of the outcome in the experimental / exposure arm

- sd.e: Specify the column containing the standard deviation of the outcome in the experimental / exposure arm

- n.c: Specify the column containing the number of participants in the control arm

- mean.c: Specify the column containing the mean value of the outcome in the control arm

- sd.c: Specify the column containing the standard deviation of the outcome in the control arm

- studlab: Specify the column containing the study identifiers or author names.

- data: Specify the object where the table of data is stored

- sm: Specify the summary measure used for meta-analysis. We are taking it as mean difference (MD) here.

- prediction: Specify whether you wish to compute the prediction interval or not. This is an example of an argument that takes logical inputs, i.e., *TRUE* or *FALSE*.

Once we run the above code line, the meta-analysis object named *metta* should be generated and stored. Now, we have to view this. We can use one or both of the next two lines of code to view the results of this meta-analysis.

> metta
> summary(metta)

The results can be seen more concisely by printing the forest plot. This can be printed this way.

> forest.meta(metta)

This outputs a basic forest plot[28–30]. It can be seen in the *Plots* tab in the bottom-right pane in RStudio **[Figure 5]**.
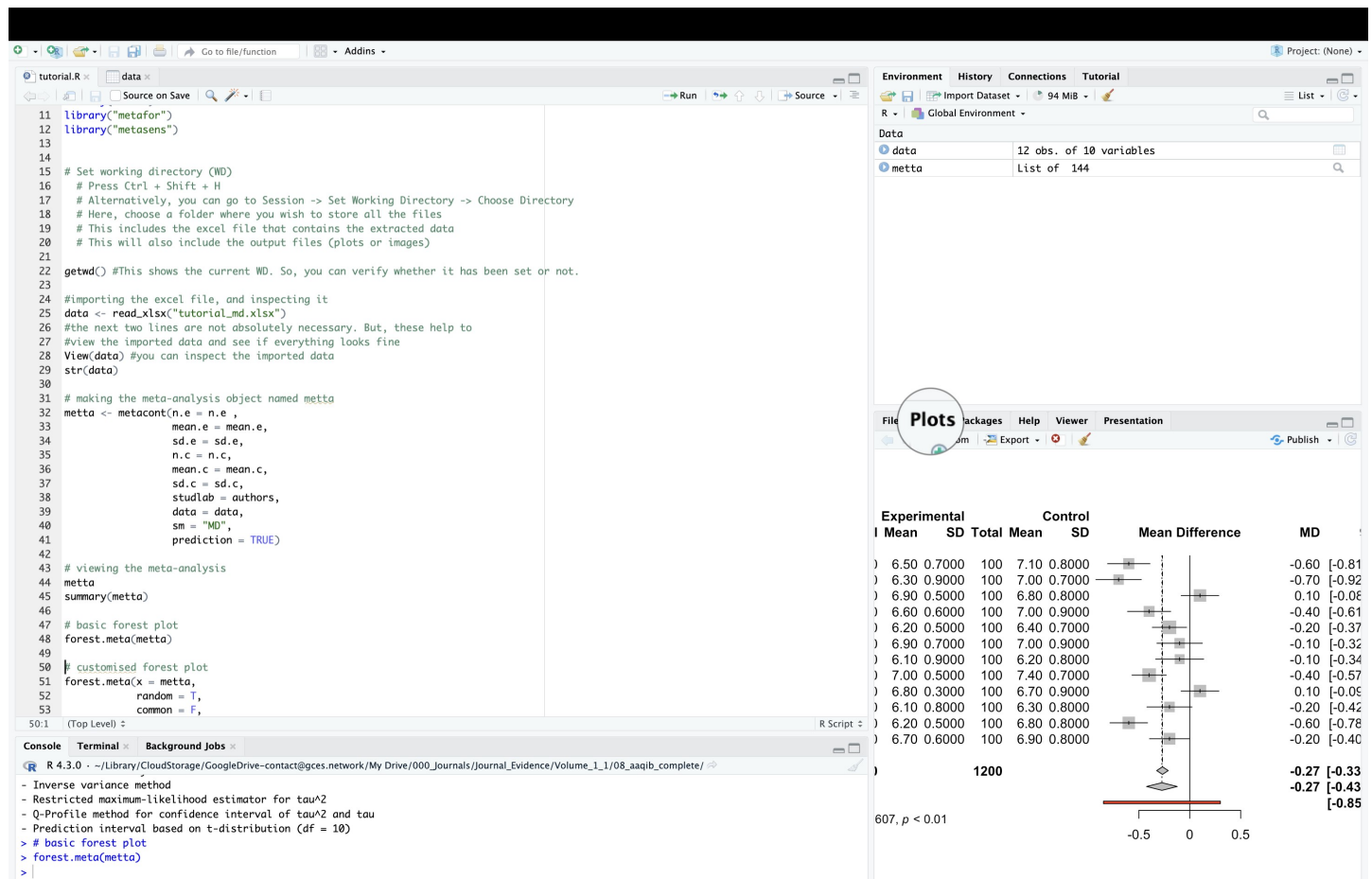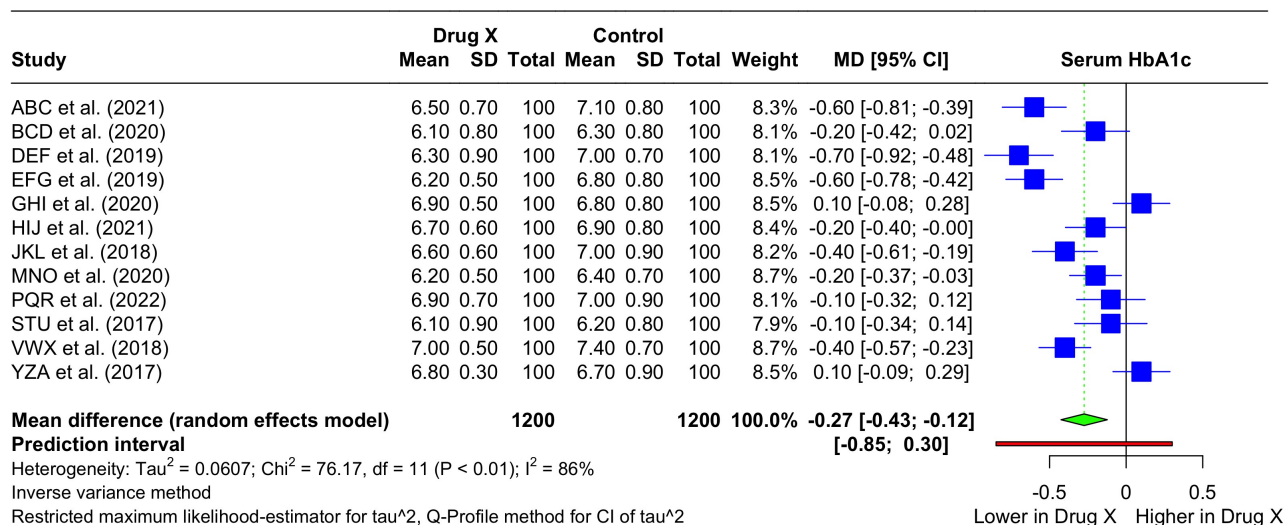

Figure 5: Viewing plots in R

It plots individual study and summary estimates (shown as common effects and random effects). Though this plot is complete, we can draw more visually pleasing and more comprehensive forest plots **[Figure 6]**. The code for the more comprehensive forest plots is in the code file and attached here, too

Study	Drug X Mean	SD	Total	Control Mean	SD	Total	Weight	MD [95% CI]	Serum HbA1c
ABC et al. (2021)	6.50	0.70	100	7.10	0.80	100	8.3%	-0.60 [-0.81; -0.39]	
BCD et al. (2020)	6.10	0.80	100	6.30	0.80	100	8.1%	-0.20 [-0.42; 0.02]	
DEF et al. (2019)	6.30	0.90	100	7.00	0.70	100	8.1%	-0.70 [-0.92; -0.48]	
EFG et al. (2019)	6.20	0.50	100	6.80	0.80	100	8.5%	-0.60 [-0.78; -0.42]	
GHI et al. (2020)	6.90	0.50	100	6.80	0.80	100	8.5%	0.10 [-0.08; 0.28]	
HIJ et al. (2021)	6.70	0.60	100	6.90	0.80	100	8.4%	-0.20 [-0.40; -0.00]	
JKL et al. (2018)	6.60	0.60	100	7.00	0.90	100	8.2%	-0.40 [-0.61; -0.19]	
MNO et al. (2020)	6.20	0.50	100	6.40	0.70	100	8.7%	-0.20 [-0.37; -0.03]	
PQR et al. (2022)	6.90	0.70	100	7.00	0.90	100	8.1%	-0.10 [-0.32; 0.12]	
STU et al. (2017)	6.10	0.90	100	6.20	0.80	100	7.9%	-0.10 [-0.34; 0.14]	
VWX et al. (2018)	7.00	0.50	100	7.40	0.70	100	8.7%	-0.40 [-0.57; -0.23]	
YZA et al. (2017)	6.80	0.30	100	6.70	0.90	100	8.5%	0.10 [-0.09; 0.29]	
Mean difference (random effects model)			**1200**			**1200**	**100.0%**	**-0.27 [-0.43; -0.12]**	
Prediction interval								[-0.85; 0.30]	

Heterogeneity: Tau^2 = 0.0607; Chi^2 = 76.17, df = 11 (P < 0.01); I^2 = 86%
Inverse variance method
Restricted maximum likelihood-estimator for tau^2, Q-Profile method for CI of tau^2

-0.5   0   0.5
Lower in Drug X   Higher in Drug X

Figure 6: Forest plot showing the pooled estimate, heterogeneity indicators, prediction interval and individual study estimates

```
> forest.meta(x = metta,
random = T,
common = F,
layout = "RevMan5",
col.diamond.random = "Green",
col.random = "Green",
col.square = "Blue",
col.square.lines = "Blue",
col.study = "Blue",
text.random = "Mean difference (random effects model)",
digits = 2, digits.se = 2,
smlab = "Serum HbA1c"
pooled.events = T,
text.addline1 = "Inverse variance method",
text.addline2 = "Restricted maximum likelihood-estimator for tau^2, Q-Profile method for CI of tau^2",
header.line = "both",
label.e = "Drug X",
label.c = "Control",
label.left = "Lower in Drug X",
label.right = "Higher in Drug X",
sortvar = studlab
)
```

For more information on the two functions used for this analysis, one can refer to the package documentation and help. It is a very detailed document mentioning many fine points, especially regarding customisations. Some of it has been incorporated into the code file.

```
> help(metacont)
> help(forest.meta)
```

The results can be reported in the following manner. "The pooled mean difference (with 95% confidence interval) in serum HbA1c between the interventional arm (Drug X) and control arm (placebo) is -0.27, (-0.43; -0.12). The individual study estimates have been pooled using a random effects model."

Ancillary analyses
Exploring heterogeneity – subgroup analysis and meta-regression

Heterogeneity is the measure used to represent the variation between the studies included in the meta-analysis. Broadly, heterogeneity is of three types: clinical, methodological, and statistical. Statistical heterogeneity is the component which is calculated in the pooled estimates. Statistical heterogeneity can be influenced by clinical and methodological heterogeneity[31].

Between-study heterogeneity is an important factor in meta-analysis. Among other things, it decides our choice of model – Common Effects (CE), also known as Fixed Effects (FE) or Random Effects (RE) - for a meta-analysis. It also helps interpret the results of a quantitative synthesis. Heterogeneity can be assessed by visually looking at the results. It can also be assessed by checking the formal measures of heterogeneity like tau, tau-squared, I^2, and Cochran's Q[32–34]. These are printed in the meta-analysis summaries and forest plots.

The CE model assumes that all the studies are similar or homogenous and thus provides a pooled estimate assuming a common true effect. On the other hand, the RE model assumes the individual study estimates are heterogenous, thus providing a summary estimate of the various true effect sizes. The choice of model is based on the homogeneity or heterogeneity of the studies concerning the research question. It need not be restricted to a single decision rule. Generally, it is relevant to apply RE model when the heterogeneity is substantially high (I^2>50%)[35–38].

Additionally, since some heterogeneity between the studies is bound to occur, presenting it as a 95% prediction interval (PI) is prudent. Conventional estimators of heterogeneity are difficult to interpret[39,40]. PI, instead, is much more intuitive. PI presents the heterogeneity in a more interpretable manner. It is expressed in terms of the outcome of the index studies (in the current example, it will be mean)[41–44].

In case of a high heterogeneity, we need to explore the causes. This is needed to interpret the results correctly. We need to answer why the studies differ between themselves. We try to find out whether it is merely by chance or whether there is a factor that leads to heterogeneous results. If we had an a priori hypothesis about a factor influencing the results of individual studies, and then this factor successfully explores the cause behind the between-study heterogeneity, it reaffirms our hypothesis and helps interpret the findings of an SRMA[45–47].

There are two chief methods to explore this heterogeneity. Firstly, the suspected moderator variable can be categorical. For example, the age group of patients may moderate the outcome (HbA1c). So, we can divide all the studies into two or more categories (studies on young and elderly populations). Each subgroup category is analysed separately. Then, the differences in results (or between-study heterogeneity) are assessed between the two or more subgroups. This is called subgroup analysis.

The second method is for continuous variables. We suspect a continuous variable to moderate the outcome instead of a categorical variable (as in subgroup analysis). For example, an investigator may suspect that differences in baseline LDL-cholesterol (low density lipoprotein) levels may moderate the differences in HbA1c. This is called meta-regression.

We specify the subgrouping variable(s) and the meta-regression variable(s) in the datasheet. We also had these two columns in the imported spreadsheet we imported and, consequently in the *data* object.
- subgroup: Any subgrouping variable. Suppose if we expect the outcome to differ in studies recruiting young and old participants, then we can use this as a subgrouping variable. We can then mention the subgroup of each study category in this column.

- regression: If we expect any continuous variable to alter the outcome of interest, we can include it here. In this case, the LDL value is given.

Now, running a subgroup analysis is simple. Here, we will introduce another function, *update.meta*. The code is the following:

```
> metta.sg <- update.meta(object = metta,
subgroup = Age,
print.subgroup.name = FALSE)
> metta.sg
> forest.meta(x = metta.sg)
```

We are feeding the *update.meta* function with our meta-analysis object *metta*. We are asking it to perform a subgroup analysis based on the data in the column named *Age*. The third argument regarding printing is purely cosmetic. This updated meta-analysis does not disturb the current meta-analytic object named *metta*. Instead, this updated meta-analysis is assigned to a new meta-analysis object named *metta.sg*. Then, we see the results by typing the object name and running. Similarly, we can view it in a simple forest plot by calling it within the *forest.meta* function. Alternatively, we can construct a more detailed and visually better forest plot by running the code chunk shown earlier in the section on forest plot and replacing the name of the meta-analysis object being supplied (the value of the *x* argument within the forest.meta function) **[Figure 7]**.

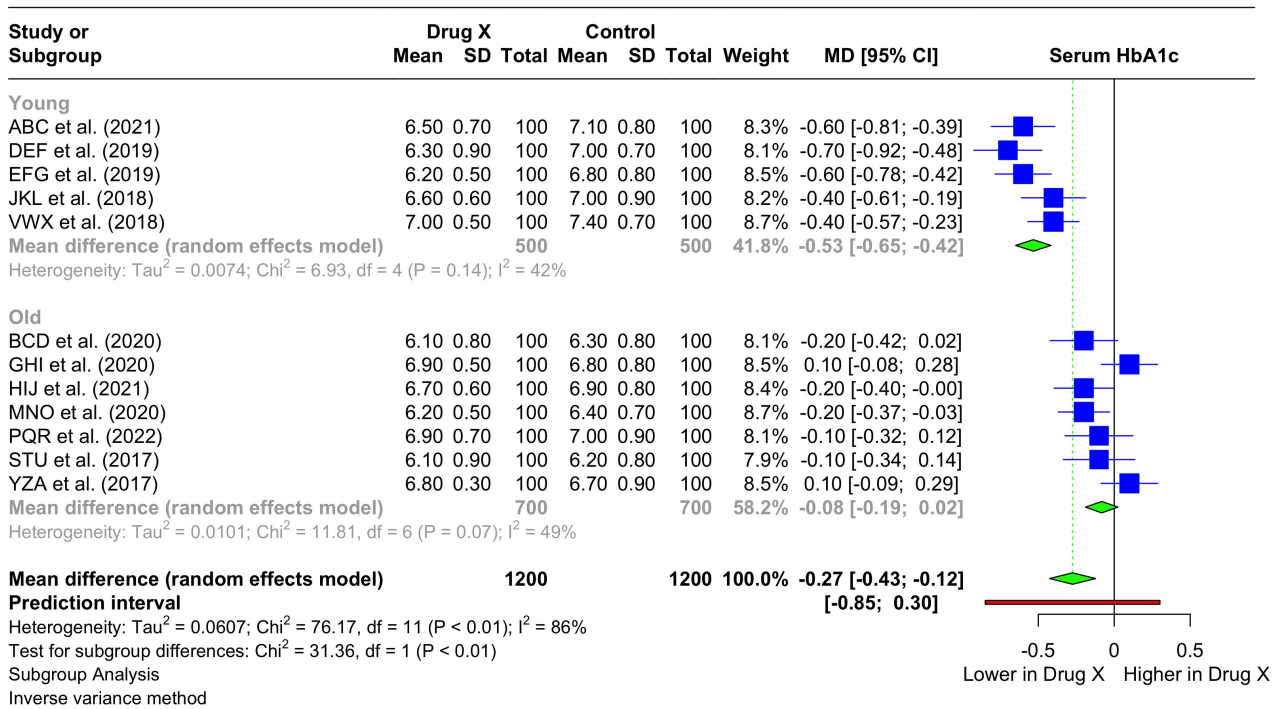| Study or Subgroup	Drug X Mean	SD	Total	Control Mean	SD	Total	Weight	MD [95% CI]	Serum HbA1c
Young									
ABC et al. (2021)	6.50	0.70	100	7.10	0.80	100	8.3%	-0.60 [-0.81; -0.39]	
DEF et al. (2019)	6.30	0.90	100	7.00	0.70	100	8.1%	-0.70 [-0.92; -0.48]	
EFG et al. (2019)	6.20	0.50	100	6.80	0.80	100	8.5%	-0.60 [-0.78; -0.42]	
JKL et al. (2018)	6.60	0.60	100	7.00	0.90	100	8.2%	-0.40 [-0.61; -0.19]	
VWX et al. (2018)	7.00	0.50	100	7.40	0.70	100	8.7%	-0.40 [-0.57; -0.23]	
Mean difference (random effects model)			500			500	41.8%	-0.53 [-0.65; -0.42]	
Heterogeneity: Tau^2 = 0.0074; Chi^2 = 6.93, df = 4 (P = 0.14); I^2 = 42%									
Old									
BCD et al. (2020)	6.10	0.80	100	6.30	0.80	100	8.1%	-0.20 [-0.42; 0.02]	
GHI et al. (2020)	6.90	0.50	100	6.80	0.80	100	8.5%	0.10 [-0.08; 0.28]	
HIJ et al. (2021)	6.70	0.60	100	6.90	0.80	100	8.4%	-0.20 [-0.40; -0.00]	
MNO et al. (2020)	6.20	0.50	100	6.40	0.70	100	8.7%	-0.20 [-0.37; -0.03]	
PQR et al. (2022)	6.90	0.70	100	7.00	0.90	100	8.1%	-0.10 [-0.32; 0.12]	
STU et al. (2017)	6.10	0.90	100	6.20	0.80	100	7.9%	-0.10 [-0.34; 0.14]	
YZA et al. (2017)	6.80	0.30	100	6.70	0.90	100	8.5%	0.10 [-0.09; 0.29]	
Mean difference (random effects model)			700			700	58.2%	-0.08 [-0.19; 0.02]	
Heterogeneity: Tau^2 = 0.0101; Chi^2 = 11.81, df = 6 (P = 0.07); I^2 = 49%									
Mean difference (random effects model)			**1200**			**1200**	**100.0%**	**-0.27 [-0.43; -0.12]**	
Prediction interval								**[-0.85; 0.30]**	
Heterogeneity: Tau^2 = 0.0607; Chi^2 = 76.17, df = 11 (P < 0.01); I^2 = 86%									
Test for subgroup differences: Chi^2 = 31.36, df = 1 (P < 0.01)									
Subgroup Analysis									
Inverse variance method									

Figure 7: Subgroup analysis to explore heterogeneity due to a categorical variable

The text output displayed in R is quite detailed. Breaking it down, we can see two major parts. Results and tests for subgroups are shown for both common effects (CE) and random effects (RE). However, CE and RE do not refer to the overall pooled estimate but mean something else here. It signifies whether the studies within a subgroup (in this case, whether all the studies on a young population) are assumed to follow a common true effect and should be combined as such (CE), or they should be pooled using RE.

Output for subgroups (CE) gives the results for each subgroup category (young and old here). Then, the test for subgroup differences (CE) assess whether the studies within a subgroup category are significantly different from each other (within groups) and whether the two subgroup categories are significantly different from each other (between groups). These are repeated for RE. We can see the key details summarised in the forest plot for this subgroup analysis.

We perform meta-regression next. In brief, it is run using the following code:

```
> metta.mr <- rma(yi = metta$TE, sei = metta$seTE, method = metta$method.tau, mods = data$LDL)
> summary(metta.mr)
# Meta-regression: basic bubble plot
> regplot(metta.mr)
```

In the first line of code, the argument to focus on is *mods*. We specify the column name where the suspected moderating continuous variable (LDL) is stored. Then, we assign this meta-regression to an object named *metta.mr*. Next, we can see the results. After that, we visualise it using the function *regplot* and feeding our meta-regression object named *metta.mr*. An easier-to-comprehend and aesthetically better output [Figure 8] can be obtained using the code in the file [Data S1]. It must be noted that we have used hash (#) in the code here. Anything written in a line after this symbol (#) is taken as plain text and not as code. We use it to annotate our code to understand it better when we revisit it later. When we share our code file, annotations help others interpret the code quicker, irrespective of their knowledge level. It is considered a good practice in the coding community.

A detailed explanation of all the code, including its customisations (and the underlying syntax), along with every minute detail in the generated results, is outside the scope of this manuscript. Hence, we are discussing enough details in the code to make it reproducible for individual cases and are touching on the key points in the results section. These key points are mainly focused on while writing and interpreting the results of a meta-analysis.

The output of this meta-regression can be checked for the test of moderators (to see the significance of the moderating effect of the given variable, i.e., LDL in this case). The *estimate* and *se* (standard error) of *mods* give the strength of moderation. Next, a bubble plot is constructed. This shows the same effect visually. In the customised bubble plot **[Figure 8]**, the outcome (HbA1c) in the Y-axis is plotted against the suspected continuous moderating variable (LDL) in the X-axis. The individual bubbles reflect each study along with its weight. The green dotted lines represent the 95% confidence interval, while the blue dotted lines indicate the 95% prediction interval. Like other cases, *help(regplot)* gives more details on this function.

Meta-regression: Bubble plot with confidence and prediction intervals

Test of Moderators: Beta = 0 (0 - 0.01), p = 0.03
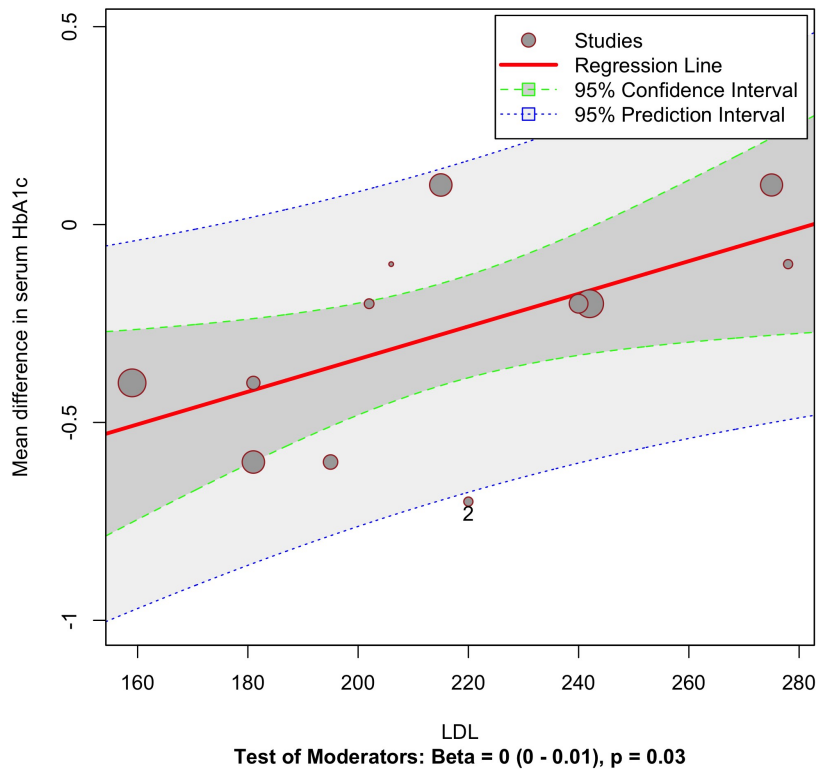
Figure 8: Bubble plot demonstrating meta-regression to explore heterogeneity due to a continuous variable

We can briefly report the heterogeneity assessment and exploration results: "There is considerable between-study heterogeneity (I^2 = 86%, tau^2 = 0.06, Cochran's Q: p < 0.01). The pooled estimate shows a prediction interval of -0.85 – 0.30 i.e., 95% of the similar future studies are expected to give a summary estimate within this range.

To explore this heterogeneity, we conducted subgroup analysis and meta-regression. Subgroup analysis based on the age of the population revealed significant differences between the two age categories (Q = 31.36, p < 0.01). The studies on young participants reported a higher mean difference (-0.53, -0.65 - -0.42) than those on old participants (-0.08, -0.19 - -0.02). Meta-regression exploring the effect of LDL on mean difference in serum HbA1c is visualised as a bubble plot. Higher LDL is associated with a higher mean difference. Here, the test of moderators shows a significant effect of LDL on the pooled estimate (p = 0.03)." The subgroup analysis can be expressed as a figure (like in **Figure 7**) or as a table (like in **Table 1**).

Table 1: Subgroup analysis

	Mean	Lower 95% CI	Upper 95% CI	I2	Test for subgroup differences
Subgrouping variable - Age					
Young	-0.53	-0.65	-0.42	42%	p < 0.01
Old	-0.08	-0.19	0.02	49%	

Publication bias, study asymmetry, and small-study effects

It is well known that significant studies are more likely to be published than their non-significant counterparts[48–50]. There can also be other factors for a systematic difference in publication or availability of studies[51]. Since this can influence the summary estimate, we must adjust for this difference in the publication of studies. Publication bias explores the systematic error in reporting and accessing the studies (non-availability of full text and language restriction) [52–54].

Multiple methods and statistical tests are available to depict and assess the publication bias in the meta-analysis[55–57]. Funnel plot and Egger's test are commonly used measures among them[58–60]. Doi plot and LFK index are relatively newer and better measures to explore the publication bias[17]. Publication bias is predicted by means of the asymmetry in both funnel plot and Doi plot. However, Doi plot gives a relatively better and unambiguous visual asymmetry than the funnel plot. Similarly, LFK index quantitatively estimates the potential publication with better validity than the Egger's test. While funnel plot and Egger's test are valid only when a minimum of ten studies are included in the meta-analysis, LFK index can assess publication bias with a number of studies as low as five. Single-group meta-analysis requires modifications in the variables used in the funnel plot, while no such modification is required in Doi plot[17,41–44].

We can construct a basic funnel plot by running one code line. Next, we can use a trim-and-fill funnel analysis to trim outlying (suspected to be a result of bias) studies and then fill back both these studies and their mirroring (opposite size of effect) studies. Simply put, it is an adjusted meta-analysis considering the effect of potential publication bias or study asymmetry. We can graphically view this trim-and-fill adjusted funnel plot, along with contours coloured with shades of grey, indicating different levels of significance. This **[Figure 9]** can be constructed using the code file **[Data S1]**.
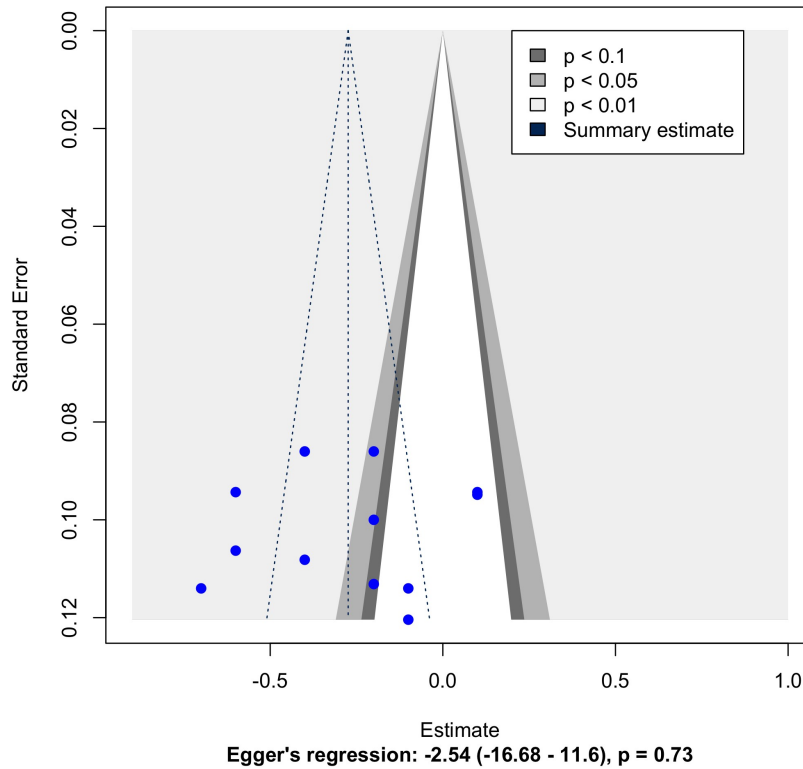
Figure 9: Trim-and-fill adjusted contour-enhanced funnel plot to assess publication bias, and small study effects

 # Publication bias: basic funnel plot
> funnel.meta(metta)
 # Publication bias: basic funnel plot with study labelling
> funnel.meta(metta, studlab = TRUE)
 # Publication bias: Trim-and-fill analysis
> meta.tf <- trimfill(x = metta)
> meta.tf

Egger's regression is a commonly used formal statistical method to test for the significance of small study effects or publication bias. It often accompanies the funnel plot. However, it is not recommended when there are less than ten studies.

 # Publication bias: Egger's regression
> metabias(x = metta, method.bias = "Egger")
 # Publication bias: Egger's regression with plot
> metabias(x = metta, method.bias = "Egger", plotit = TRUE)

A key detail in the output is the p-value for this regression. We should consider that, unlike most other cases, the significance threshold for Egger's regression is conventionally taken at 10% (i.e., $p \leq 0.10$)[61].

The Doi plot is a valuable alternative in many situations, including single-group meta-analysis (no comparator group) and when less than ten studies are present[41–43]. Moreover, it can complement the funnel plot in the other cases where the latter is used. This is because it (along with LFK index) has been reported to be much more sensitive in many circumstances. Like the Egger's regression accompanies the funnel plot, the LFK index is a quantitative interpretation of the Doi plot. Here, we specify a priori (before

The SRMA) the suspected direction of publication bias. Accordingly, a value outside the range of -1 to +1 is considered significant[17]. We can get the Doi plot and the accompanying LFK index with one line of code.

```
# Publication bias: Doi plot and LFK index
> doiplot(metta)
# Publication bias: Reformatted Doi plot with LFK index
> pdf(file = "doi.pdf", height = 10, width = 5)
> doiplot(metta, xlab = "Transformed estimates", main = "Effect of Drug X v/s placebo on HbA1c")
> dev.off()
```

The second part of the above code segment introduces two important points. First, plots and images can be exported directly from the code itself without clicking *export* in the *plots* tab or right-clicking on the image after clicking *zoom* in the *plots* tab. The way is first to write the code required for generating an image. Then, in the preceding line, we can use the *pdf* function and supply the file name and certain other optional attributes. In the line after the main code for the plot, we write *dev.off()*. This is a generic line common to every situation and marks the end of plotting. Thus, we first specify the file name, then give the command for the plot, and then end the code segment. This exports our image [Figure 10] to the working directory.
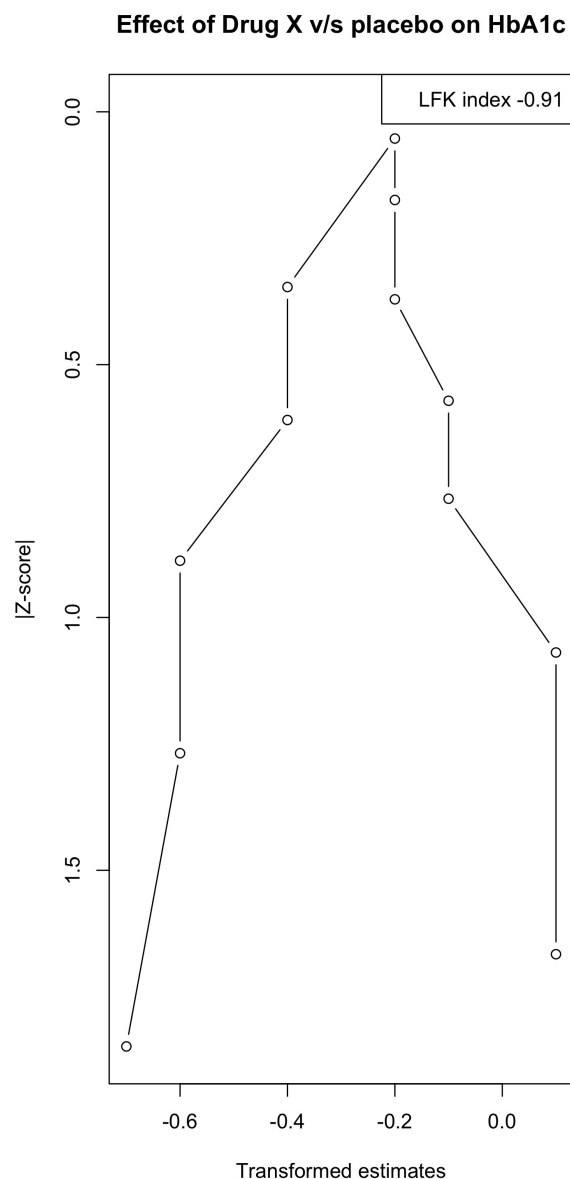
Figure 10: Doi plot with LFK index for publication bias in meta-analyses – useful even with less number of studies or only a single group of data (not comparative)

Another important point is the values for the *height* and *width* arguments. We can often leave it to the default value, while in some cases, we edit it. The Doi plot should be generated such that the Y-axis is double the length of the X-axis[62], and hence, we have used these specific values within the *pdf* function.
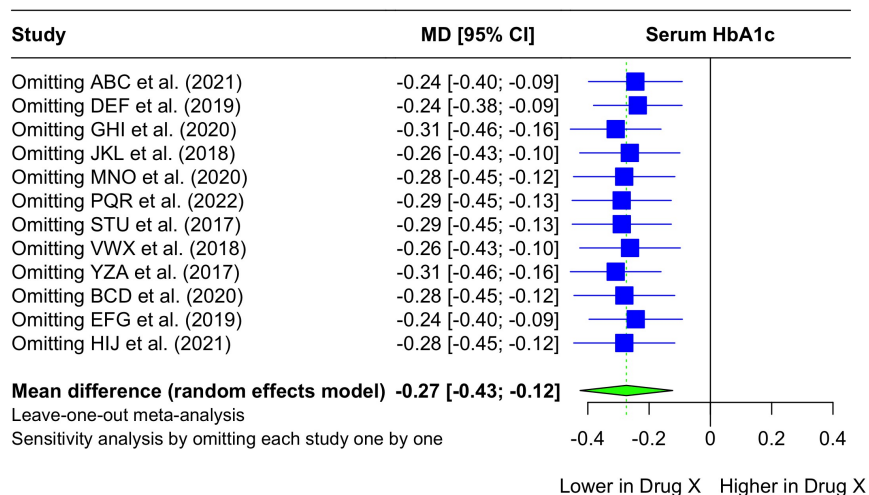
We can report the results for publication bias as follows: "There is no evidence of publication bias or small-study effects according to the trim-and-fill contour-enhanced funnel plot and the accompanying Egger's regression (p = 0.73). Similarly, the Doi plot is visually symmetrical. Its quantitative interpretation, the LFK index (-0.91), also falls within the range of -1 to +1 and does not indicate any publication bias or small-study effects."

Sensitivity analysis

The results of a meta-analysis need to be assessed for robustness. This is required to answer questions like, "Have we obtained these results because some poorly conducted studies have skewed the summary estimate?" or "Did a single study overly influence the pooled results, and would excluding it change the results significantly?". Thus, we need sensitivity analysis[63–65]. We resort to two chief types of sensitivity analyses. First, we rate the studies for evidence pertaining to our research question. Depending upon the quality of evidence obtained from individual studies and depending upon the tool used to do so, we can rate the studies for their quality. Then, we can add a column to store this data. Our datasheet [Data S2] shows this as a column named *quality*. It is present even in the imported data frame named *data*. Another method to perform a sensitivity analysis is a leave-one-out meta-analysis. We can omit each study one by one and observe the change in the pooled estimate, between-study heterogeneity, and other things. This can help identify any study having an overly influential effect on the pooled estimate. We can perform these two sensitivity analyses using the following code.

```
#Sensitivity analysis for study quality
> metta.rob <- update.meta(object = metta, subgroup = quality )
> metta.rob
#Sensitivity analysis for leave-one-out
> metta.l1o <- metainf(metta, pooled = "random")
> forest.meta(metta.l1o)
```

The new objects *metta.rob* and metta.l1o can be read and interpreted using the text output. Alternatively, a forest plot can be constructed by simply feeding it into the *forest.meta()* function. Alternatively, a customised, comprehensive forest plot can be constructed using the code discussed earlier and using these as the meta-analysis objects for the argument *x* **[Figure 11]**.

Study	MD [95% CI]	Serum HbA1c
Omitting ABC et al. (2021)	-0.24 [-0.40; -0.09]	
Omitting DEF et al. (2019)	-0.24 [-0.38; -0.09]	
Omitting GHI et al. (2020)	-0.31 [-0.46; -0.16]	
Omitting JKL et al. (2018)	-0.26 [-0.43; -0.10]	
Omitting MNO et al. (2020)	-0.28 [-0.45; -0.12]	
Omitting PQR et al. (2022)	-0.29 [-0.45; -0.13]	
Omitting STU et al. (2017)	-0.29 [-0.45; -0.13]	
Omitting VWX et al. (2018)	-0.26 [-0.43; -0.10]	
Omitting YZA et al. (2017)	-0.31 [-0.46; -0.16]	
Omitting BCD et al. (2020)	-0.28 [-0.45; -0.12]	
Omitting EFG et al. (2019)	-0.24 [-0.40; -0.09]	
Omitting HIJ et al. (2021)	-0.28 [-0.45; -0.12]	
Mean difference (random effects model)	**-0.27 [-0.43; -0.12]**	

Leave-one-out meta-analysis
Sensitivity analysis by omitting each study one by one

-0.4 -0.2 0 0.2 0.4

Lower in Drug X   Higher in Drug X

Figure 11: Sensitivity analysis using leave-one-out meta-analysis

We can briefly report the leave-one-out analysis results as follows: "The pooled result is stable according to a sensitivity analysis using the leave-one-out meta-analysis. No single study had a disproportionate effect on the summary estimate." The sensitivity analysis results for risk of bias (or quality assessment) can be reported similar to subgroup analysis.

Drapery plot: a complement to the forest plot

This is a newly introduced concept. It serves as a helpful complement to the forest plot and adds value. It allows us to avoid an arbitrary significance threshold of 0.05 and to visualise the estimate and error from individual estimates in a logical sequence (of effect sizes). This allows quicker interpretation visually[66]. This can be obtained using a simple line of code **[Figure 12]**.

Effect against a range of p values
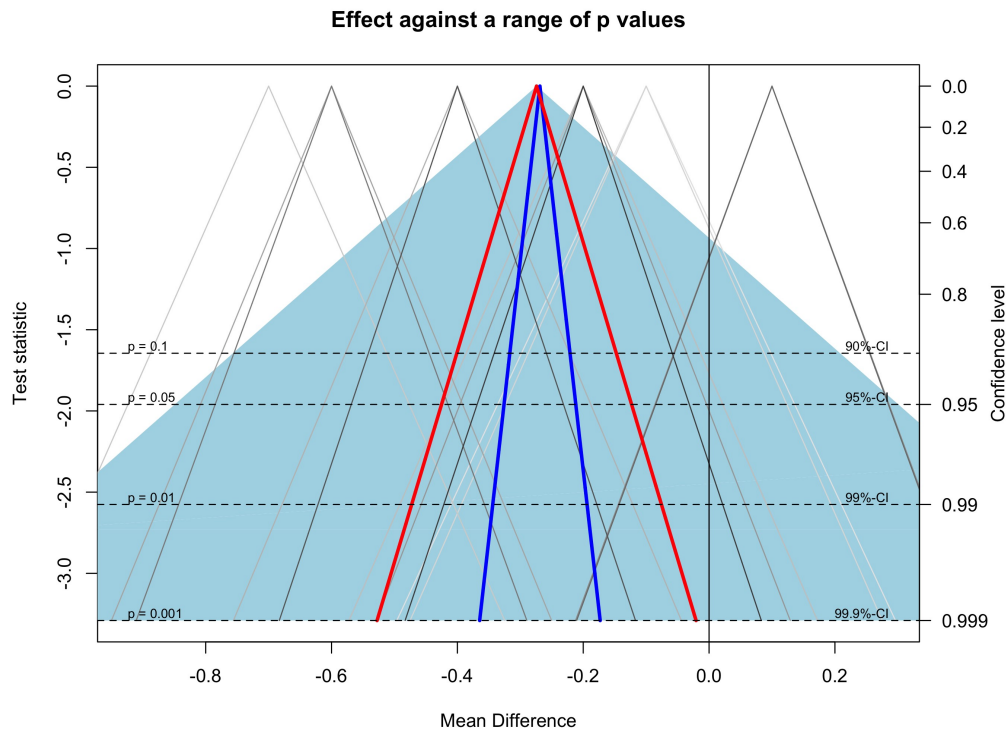
Figure 12: Drapery plot to visualise the results against a p-value function (a wide range of significance threshold levels)

> drapery(metta, legend = FALSE)

Here, the individual study estimates and their uncertainty are plotted opposite a p-value function, i.e., at varying p-values. The pooled estimates, with their uncertainty and the prediction interval, can also be traced against a range of p-values. The dark blue line reflects the CE model, while the red line depicts the RE model. The light blue shaded region indicates the prediction interval. **[Figure 12]**

Other study types in meta-analysis

Now that we have completed a meta-analysis, the remaining part is straightforward. We can proceed for other cases just like we have computed for continuous data comparing two groups. In most cases, the only requirement is restructuring the data collection spreadsheet and redefining the meta-analysis object. Once we have the meta-analysis object (like *metta* in the previous case), we can proceed similarly with the subsequent steps (like forest plot, heterogeneity exploration). So, in this section, we have covered spreadsheet preparation and meta-analysis object creation for other common types of meta-analysis.

Meta-analysis comparing two arms or groups – categorical data with proportions

The data collection spreadsheet should have four columns for data collection instead of the six seen earlier. The two columns for number of participants (*n.e* and *n.c*) persist. For the outcome, we use *event.e* and *event.c* instead. *event.e* specifies the number of participants in the experimental / exposure arm who have experienced the outcome of interest. *event.c* gives the same information for the control arm. For example, there are two groups of diabetic patients, with one arm assigned an experimental drug and the other assigned standard care. The outcome of interest is the hypoglycaemic event. *event.e* and *n.e* tell the number of patients with hypoglycaemic events and the total number of patients, respectively. An example data extraction spreadsheet is attached in the supplementary **[Data S3]**. We will use this for the sample analysis.

The next and final step in learning this meta-analysis is to create an appropriate meta-analysis object. It is pretty straightforward after that. Here, we use the *metabin()* function for binary outcomes instead of *metacont()* in the previous case. The code for this is the following.

 # Two group - proportion meta-analysis

```
> data_2_prop <- read_xlsx("tutorial_rr.xlsx")
# Meta-analysis of binary outcomes
> metta.bin <- metabin(event.e = event.e,
n.e = n.e,
event.c = event.c,
n.c = n.c,
studlab = authors,
data = data_2_prop,
sm = "RR",
prediction = TRUE
# Here, this performs a meta-analysis of risk ratio
# Just change the sm argument, and you can perform other 2-group proportional meta-analyses (like odds ratio or risk difference
instead of risk ratio)
```

This code segment allows us to import a datasheet named *tutorial_rr.xlsx* stored in the working directory folder. We imported it as a data frame named *data_2_prop*. The *metabin()* function includes the arguments for outcomes as discussed earlier and certain other arguments like study identifiers similar to *metacont()*. One key point to note is the *sm* argument. This can be selected to choose odds ratio, risk ratio, and risk difference. The values for *sm* are *"OR"*, *"RR"*, and *"RD"*. This meta-analysis result is assigned to an object named *metta.bin*. We can run the previous analyses using this meta-analysis object *metta.bin* instead of *metta* as done earlier.

Meta-analysis comparing two arms or groups – continuous data with mean and standard deviation

This has been covered in the main example presented in the article.

Meta-analysis describing a single arm or group - categorical data with proportion (prevalence)

Here, we provide a summary estimate of a proportion in a single group. Examples are meta-analysis of the prevalence of a specific disease amongst the population or the pooled prevalence of a specific symptom, a particular disease characteristic, or a feature amongst all the patients of a disease. We perform this type of meta-analysis in every case where we assess the proportion of a binary outcome in a single group.

We can utilise the datasheet that we have used for the meta-analysis of binary outcomes in two groups *tutorial_rr.xlsx* **[Data S3]**. We are importing the table from this spreadsheet and feeding it into the function named metaprop(). In this case, we are only assessing the pooled proportion in one group (ex: the prevalence of hypoglycaemic episodes in diabetic patients). The meta-analysis object obtained here *metta.prop* can be used for the analyses as discussed earlier. The code is attached here.

```
# One group - meta-analysis of proportion
> data_2_prop <- read_xlsx("tutorial_rr.xlsx")
> metta.prop <- metaprop(event = event.e,
n = n.e,
studlab = authors,
data = data_2_prop,
prediction = TRUE)
```

Meta-analysis describing a single arm or group - continuous data with mean and standard deviation

We can summarise the continuous outcome in a particular group. For example, we can continue from the data used earlier regarding HbA1c in two groups **[Data S2]**. We can use that same table but only pool the HbA1c. Thus, we are computing the pooled estimate for HbA1c in the experimental group. We give the required values as arguments to the *metta.mean()* function. We can run this code, and then use the *metta.mean* meta-analysis object for further analysis as discussed earlier.

```
# One group - meta-analysis of mean and SD
> data <- read_xlsx("tutorial_md.xlsx")
# We can try this with the earlier datasheet that we used for mean difference meta-analysis
# We will just take data for the experimental group, ignoring control
> metta.mean <- metamean(n = n.e ,
```

>*mean = mean.e,*
sd = sd.e,
studlab = authors,
data = data,
prediction = TRUE)

Common errors and important considerations

Common errors

Error: `path` does not exist: '…...xlsx

This can appear in the console when the file name or the location of the file has not been set properly. A common mistake is not setting the working directory correctly. One can recheck the current working directory by running *get.wd()*.

Error in ……. : could not find function "……..

Often, an error like this may appear. This is most commonly seen when a function cannot be found because the relevant package has not been loaded. Hence, we should call all the required packages using the *library()* function whenever restarting R.

Another common oversight is that we make changes in the spreadsheet file. Then, we rerun our analysis in R but do not find the changes in the analysis. The same old results come. The mistake is that we may not have re-imported the spreadsheet file. Unless it is re-imported - i.e., the *read.xlsx()* function is rerun - the old data frame stored in R will still be used for the analyses. Changes made outside R do not reflect here. If we re-import the data frame, the analyses should reflect the changes made.

Naming conventions for Excel files and headings in the datasheet

R is a case-sensitive programming language. A convention in coding is to use lowercase wherever possible. However, it is not a compulsion, especially in R. But one mandatory rule is that we should exercise caution with spaces. Whenever a space is required – for example, when writing two words to describe a name like *diabetes data* – an underscore can be used (*diabetes_data*), the space can be omitted (*diabetesdata*), the first letter of the second word can be written in upper case (camel case – *diabetesData*), or any other reformatting can be done avoiding a space.

Another example of caution in naming is when we specify the subgroup category that each study falls into. If it is mentioned *young* in 4 studies, *Young* in 3 studies, *old* in 3 studies, and *Old* in 4 studies, then R takes it as four categories, not two. It is a case-sensitive programming language and is ordinarily unable to recognise that *young* and *Young* are the same or that *old* and *Old* are the same. Hence, we should be consistent in our spelling throughout.

How to add multiple outcomes in the same Excel file?

We have discussed cases where we analysed a single outcome for the particular SRMA. We can similarly analyse multiple outcomes for the same SRMA. For example, individual studies provide HbA1c and fasting blood sugar (FBS) data. Then, we can add both the data in the first spreadsheet file. We can rename the columns containing outcome data for HbA1c as *hba.n.e*, *hba.mean.e*, *hba.sd.e*, *hba.n.c*, *hba.mean.c*, and *hba.sd.c*. Here, *hba* stands for HbA1c, and the other letters mean the same as discussed earlier while preparing the spreadsheet. Similarly, we can add six new columns for FBS and name them as *fbs.n.e*, *fbs.mean.e*, *fbs.sd.e* and so on. We must be careful that no two columns have the same name. Otherwise, it is straightforward.

Change labels and limits of both the X and Y axes in the plot

Readers might have observed that in the customised code for the funnel plot [Ancilliary Analyses], we had two arguments named *xlab* and *ylab*. These arguments specify the labelling for the two axes (X and Y) in the funnel plot. Next, we also used two arguments named *xlim* and *ylim*. These specify the limits for the X and Y axes in the plots. These four arguments can be used for many other plots (including the bubble plot) to customise their appearance. One can experiment with these.

More information on the possibilities

For further help, one can use the *help()* function and supply the name of a particular function within the parentheses. For example, running *help(metacont)* or *help(forest.meta)* gives us detailed information on the possibilities with these functions. This can be done for any function or package.

Another easy method is to press the tab function after typing a function's name and writing the opening bracket. When the cursor is between the parentheses of a function and the tab button is pressed, R prompts the names and brief details of the arguments within that function. One can scroll with their keyboard, press Enter (or Return) at the argument of choice, and pass a value for it as required. Pressing F1 (or Fn + F1) opens a detailed help section. One can thus learn more about the minor tweaks required for different circumstances. For example, suppose we have to perform a meta-analysis of standardised mean difference instead

Of mean difference. In that case, we change the value of the <u>sm</u> argument in *metacont()* from *"MD"* to *"SMD"*. This is explained in the documentation for the *metacont()* function.

Troubleshooting

The best way to tackle other errors that could not be discussed here is a simple web search. Given the vast possibilities with base R, the packages, and the functions therein, various errors may crop up. A tried and tested method is copying the error string and searching it online. Chances are someone else faced the same error, asked their query, and received a resolution for it. If this does not solve the error, one can visit online resources or forums like Stackoverflow[67], StackExchange[68], R-project[69], and R Seek[70]. At the forums, one can mention their query reproducibly. People at forums often answer queries when appropriately explained and in detail (with the data file, the whole code file, or the code segment required to reproduce the error message). A search at these forums and websites helps solve many of the queries.

Reporting the results

We have seen that a robust meta-analysis has a lot of output. Properly drawn plots capture a great deal of this information so one can check them all at a glance. However, since key results are written in text in the results section of a manuscript, we have provided brief examples in each section. These show how the results of an SRMA can be reported in manuscripts. We are also introducing a draft example write-up for the data analysis subsection in the methodology section of a systematic review and meta-analysis.

"We have pooled the mean differences for HbA1c between Drug X and control using the inverse variance method. We used I^2, tau-squared, Cochran's Q, and 95% prediction interval for assessing heterogeneity. We used a random-effects model due to high between-study heterogeneity. We used subgroup analyses and meta-regression (visualised as bubble plots) to explore this heterogeneity for suspected categorical and continuous moderators, respectively. We assessed the robustness of results by carrying out sensitivity analysis based on study quality followed by a leave-one-out meta-analysis. We assessed publication bias using funnel plots. Following this workflow (Shamim et al., 2023), the analysis has been performed using meta, metafor, and metasens packages in R version 4.3.0 following this workflow."

For further information, readers can refer to published systematic reviews and meta-analyses for ideas. However, they should still be careful as other published articles also have great scope for improvement[41–44].

Validation

In order to assess the comprehension, completeness and validity of the steps given in the index article, it was shared with two researchers: one medical student and one medical faculty, who had none to little experience with R and asked to undertake the analysis from scratch. Feedback was taken from these researchers under ease, comprehension, and completeness domains of the article, enabling them to undertake the meta-analysis in R. Based on their feedback, we revised the draft to improve the comprehensibility of the article at the beginner level.

Further reading

A series of manuscripts are being planned and in the pipeline for an easy yet refined understanding of key concepts in applied statistics for SRMAs using R. These include but are not limited to heterogeneity, publication bias, and visualisation of pooled outcomes. Moreover, we also intend to cover other advanced meta-analytic techniques, including Bayesian analysis, and network meta-analysis, in the future.

Conclusion

In conclusion, statistical analysis is an integral component of evidence synthesis. The present article is supposed to be a bridge by introducing both the basic statistical concepts of meta-analysis and the practical implementation of it in open-source software (R). It assists novice researchers with R to undertake the basic analysis required to complete a meta-analysis of analytical and single-group studies. Since no measure of a step-wise guide can replace hands-on practice, it is impressed upon the readers and researchers to apply the guide and commands to undertake their meta-analysis to realise the utility of the paper and incorporate the experience into their routine practice. This covers the routine steps, some important advanced techniques, and several enhancements they can use to develop a methodologically robust and statistically sound meta-analysis. We have also explained reporting the results of these analyses. Comprehensively covering other advanced techniques in meta-analysis is outside the scope of this paper. Step-wise guides for some analyses, like clustering to detect outliers, a graphical depiction of heterogeneity for the same, and influence assessment of individual studies, need to be explored. These can further improve the quality of meta-analysis and its interpretation.

Supplementary files
Data S1: Code file
Data S2: Spreadsheet for meta-analysis of continuous data
Data S3: Spreadsheet for meta-analysis of categorical data

References

[1] Sauerland S, Seiler CM. Role of systematic reviews and meta-analysis in evidence-based medicine. World J Surg 2005;29:582–7. https://doi.org/10.1007/s00268-005-7917-7.

[2] Gopalakrishnan S, Ganeshkumar P. Systematic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare. J Family Med Prim Care 2013;2:9–14. https://doi.org/10.4103/2249-4863.109934.

[3] Ferreira González I, Urrútia G, Alonso-Coello P. Systematic Reviews and Meta-Analysis: Scientific Rationale and Interpretation. Revista Española de Cardiología (English Edition) 2011;64:688–96. https://doi.org/10.1016/j.rec.2011.03.027.

[4] Jenicek M. Epidemiology, evidenced-based medicine, and evidence-based public health. J Epidemiol 1997;7:187–97. https://doi.org/10.2188/jea.7.187.

[5] Wang S, Moss JR, Hiller JE. Applicability and transferability of interventions in evidence-based public health. Health Promot Int 2006;21:76–83. https://doi.org/10.1093/heapro/dai025.

[6] Fielding JE, Briss PA. Promoting evidence-based public health policy: can we have better evidence and more action? Health Aff (Millwood) 2006;25:969–78. https://doi.org/10.1377/hlthaff.25.4.969.

[7] Howes F, Doyle J, Jackson N, Waters E. Evidence-based public health: the importance of finding "difficult to locate" public health and health promotion intervention studies for systematic reviews. J Public Health (Bangkok) 2004;26:101–4.

[8] Anderson LM, Brownson RC, Fullilove MT, Teutsch SM, Novick LF, Fielding J, et al. Evidence-based public health policy and practice: promises and limits. Am J Prev Med 2005;28:226–30. https://doi.org/10.1016/j.amepre.2005.02.014.

[9] Cornell University Library. A Guide to Evidence Synthesis: What is Evidence Synthesis? 2023. https://guides.library.cornell.edu/evidence-synthesis/intro (accessed June 18, 2023).

[10] Morton SC, Murad MH, O'Connor E, Lee CS, Booth M, Vandermeer BW, et al. Quantitative Synthesis—An Update. 2008.

[11] García-Perdomo HA. Evidence synthesis and meta-analysis: a practical approach. International Journal of Urological Nursing 2016;10:30–6. https://doi.org/10.1111/ijun.12087.

[12] Ariel de Lima D, Helito CP, de Lima LL, Clazzer R, Gonçalves RK, de Camargo OP. HOW TO PERFORM A META-ANALYSIS: A PRACTICAL STEP-BY-STEP GUIDE USING R SOFTWARE AND RSTUDIO. Acta Ortop Bras 2022;30:e248775. https://doi.org/10.1590/1413-785220223003e248775.

[13] Shim SR, Kim S-J. Intervention meta-analysis: application and practice using R software. Epidemiol Health 2019;41:e2019008. https://doi.org/10.4178/epih.e2019008.

[14] Chen D-G. Meta-analysis for psychiatric research using free software R. Shanghai Arch Psychiatry 2015;27:195–9. https://doi.org/10.11919/j.issn.1002-0829.215063.

[15] Li X, Dusseldorp E, Su X, Meulman JJ. Multiple moderator meta-analysis using the R-package Meta-CART. Behav Res Methods 2020;52:2657–73. https://doi.org/10.3758/s13428-020-01360-0.

[16] Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health 2019;22:153–60. https://doi.org/10.1136/ebmental-2019-300117.

[17] Furuya-Kanamori L, Barendregt JJ, Doi SAR. A new improved graphical and quantitative method for detecting bias in meta-analysis. Int J Evid Based Healthc 2018;16:195–203. https://doi.org/10.1097/XEB.0000000000000141.

[18] IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open 2016;6:e010247. https://doi.org/10.1136/bmjopen-2015-010247.

[19] Comprehensive R Archive Network. CRAN Mirrors 2023. https://cran.r-project.org/mirrors.html (accessed June 19, 2023).

[20] Posit. DOWNLOAD RStudio Desktop 2023. https://posit.co/download/rstudio-desktop/ (accessed June 19, 2023).

[21] TC Technology Knowledge Base. Install R and R-Studio for Windows n.d. https://teacherscollege.screenstepslive.com/a/1108074-install-r-and-r-studio-for-windows (accessed June 21, 2023).

[22] TC Technology Knowledge Base. Install R and R-Studio for Mac n.d. https://teacherscollege.screenstepslive.com/a/1135059-install-r-and-r-studio-for-mac (accessed June 21, 2023).

[23] Hands-On Programming with R : Installing R and RStudio n.d. https://rstudio-education.github.io/hopr/starting.html (accessed June 21, 2023).

[24] DataFlair. How to Install R, RStudio and R Packages – The Easiest Process Ever! n.d. https://data-flair.training/blogs/how-to-install-r/ (accessed June 21, 2023).

[25] Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. J Stat Softw 2010;36. https://doi.org/10.18637/jss.v036.i03.

[26] Schwarzer G, Carpenter J, G R. metasens: Statistical Methods for Sensitivity Analysis in Meta-Analysis. 2023.

[27] Wickham H, Bryan J. readxl: Read Excel Files. 2023. https://doi.org/https://CRAN.R-project.org/package=readxl.

[28] Sedgwick P. How to read a forest plot in a meta-analysis: BMJ 2015:h4028. https://doi.org/10.1136/bmj.h4028.

[29] Dettori JR, Norvell DC, Chapman JR. Seeing the Forest by Looking at the Trees: How to Interpret a Meta-Analysis Forest Plot. Global Spine J 2021;11:614–6. https://doi.org/10.1177/21925682211003889.

[30] Andrade C. Understanding the Basics of Meta-Analysis and How to Read a Forest Plot: As Simple as It Gets. J Clin Psychiatry 2020;81. https://doi.org/10.4088/JCP.20f13698.

[31] Higgins JPT. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. Int J Epidemiol 2008;37:1158–60. https://doi.org/10.1093/ije/dyn204.

[32] Cochran WG. The Combination of Estimates from Different Experiments. Biometrics 1954;10:101. https://doi.org/10.2307/3001666.

[33] Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539–58. https://doi.org/10.1002/sim.1186.

[34] Huedo-Medina TB, Sánchez-Meca J, Marín-Martínez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? Psychol Methods 2006;11:193–206. https://doi.org/10.1037/1082-989X.11.2.193.

[35] Biggerstaff BJ, Tweedie RL. Incorporating variability in estimates of heterogeneity in the random effects model in meta-analysis. Stat Med 1997;16:753–68. https://doi.org/10.1002/(sici)1097-0258(19970415)16:7<753::aid-sim494>3.0.co;2-g.

[36] Hedges L V., Vevea JL. Fixed- and random-effects models in meta-analysis. Psychol Methods 1998;3:486–504. https://doi.org/10.1037/1082-989X.3.4.486.

[37] Poole C, Greenland S. Random-effects meta-analyses are not always conservative. Am J Epidemiol 1999;150:469–75. https://doi.org/10.1093/oxfordjournals.aje.a010035.

[38] Borenstein M, Hedges L V, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010;1:97–111. https://doi.org/10.1002/jrsm.12.

[39] Rücker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. BMC Med Res Methodol 2008;8:79. https://doi.org/10.1186/1471-2288-8-79.

[40] Borenstein M, Higgins JPT, Hedges L V, Rothstein HR. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. Res Synth Methods 2017;8:5–18. https://doi.org/10.1002/jrsm.1230.

[41] Shamim MA. Real-life implications of prevalence meta-analyses? Doi plots and prediction intervals are the answer. Lancet Microbe 2023. https://doi.org/10.1016/S2666-5247(23)00096-4.

[42] Anil A, Shamim MA, Saravanan A, Sandeep M. HPV DNA and p16INK4a positivity in vulvar cancer and vulvar intraepithelial neoplasia. Lancet Oncol 2023;24:e235. https://doi.org/10.1016/S1470-2045(23)00210-3.

[43] Shamim MA, Dwivedi P, Padhi BK. Beyond the Funnel Plot: The Advantages of Doi Plots and Prediction Intervals in Meta-Analyses. Asian J Psychiatr 2023:103550. https://doi.org/10.1016/j.ajp.2023.103550.

[44] Gandhi AP, Satapathy P, Rustagi S, Hermis AH, Sah R, Padhi BK. Comments on "Shigellosis in Southeast Asia: A systematic review and meta-analysis". Travel Med Infect Dis 2023:102593. https://doi.org/10.1016/j.tmaid.2023.102593.

[45] Thompson SG, Smith TC, Sharp SJ. Investigating underlying risk as a source of heterogeneity in meta-analysis. Stat Med 1997;16:2741–58. https://doi.org/10.1002/(sici)1097-0258(19971215)16:23<2741::aid-sim703>3.0.co;2-0.

[46] Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. BMJ 1994;309:1351–5. https://doi.org/10.1136/bmj.309.6965.1351.

[47] Song F. Exploring heterogeneity in meta-analysis: is the L'Abbé plot useful? J Clin Epidemiol 1999;52:725–30. https://doi.org/10.1016/s0895-4356(99)00066-9.

[48] Chan A-W, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, et al. Increasing value and reducing waste: addressing inaccessible research. Lancet 2014;383:257–66. https://doi.org/10.1016/S0140-6736(13)62296-5.

[49] Scherer RW, Meerpohl JJ, Pfeifer N, Schmucker C, Schwarzer G, von Elm E. Full publication of results initially presented in abstracts. Cochrane Database Syst Rev 2018;11:MR000005. https://doi.org/10.1002/14651858.MR000005.pub4.

[50] Schmucker C, Schell LK, Portalupi S, Oeller P, Cabrera L, Bassler D, et al. Extent of non-publication in cohorts of studies approved by research ethics committees or included in trial registries. PLoS One 2014;9:e114023. https://doi.org/10.1371/journal.pone.0114023.

[51] Page MJ, Sterne JAC, Higgins JPT, Egger M. Investigating and dealing with publication bias and other reporting biases in meta-analyses of health research: A review. Res Synth Methods 2021;12:248–59. https://doi.org/10.1002/jrsm.1468.

[52] Viechtbauer W. Publication bias in meta-analysis: Prevention, assessment and adjustments. Psychometrika 2007;72:269–71. https://doi.org/10.1007/s11336-006-1450-y.

[53] Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. J Clin Epidemiol 2000;53:207–16. https://doi.org/10.1016/s0895-4356(99)00161-4.

[54] Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. Stat Med 2001;20:641–54. https://doi.org/10.1002/sim.698.

[55] Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol 2000;53:1119–29. https://doi.org/10.1016/s0895-4356(00)00242-0.

[56] McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? Lancet 2000;356:1228–31. https://doi.org/10.1016/S0140-6736(00)02786-0.

[57] Rust RT, Lehmann DR, Farley JU. Estimating Publication Bias in Meta-Analysis. Journal of Marketing Research 1990;27:220–6. https://doi.org/10.1177/002224379002700209.

[58] Copas J, Shi JQ. Meta-analysis, funnel plots and sensitivity analysis. Biostatistics 2000;1:247–62. https://doi.org/10.1093/biostatistics/1.3.247.

[59] Sterne JAC, Egger M. Regression Methods to Detect Publication and Other Bias in Meta-Analysis. Publication Bias in Meta-Analysis, 2005, p. 99–110.
https://doi.org/https://doi.org/10.1002/0470870168.ch6.

[60] Sterne JAC, Harbord RM. Funnel Plots in Meta-analysis. Stata J 2004;4:127–41. https://doi.org/10.1177/1536867X0400400204.

[61] Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629–34.
https://doi.org/10.1136/bmj.315.7109.629.

[62] Doi SA. Rendering the Doi plot properly in meta-analysis. Int J Evid Based Healthc 2018;16:242–3. https://doi.org/10.1097/XEB.0000000000000158.

[63] Riley RD, Sutton AJ, Abrams KR, Lambert PC. Sensitivity analyses allowed more appropriate and reliable meta-analysis conclusions for multiple outcomes when missing data was present. J Clin Epidemiol 2004;57:911–24. https://doi.org/10.1016/j.jclinepi.2004.01.018.

[64] Elvik R. Can We Trust the Results of Meta-Analyses?: A Systematic Approach to Sensitivity Analysis in Meta-Analyses. Transp Res Rec 2005;1908:221–9. https://doi.org/10.1177/0361198105190800127.

[65] Mathur MB, VanderWeele TJ. Robust Metrics and Sensitivity Analyses for Meta-analyses of Heterogeneous Effects. Epidemiology 2020;31:356–8. https://doi.org/10.1097/EDE.0000000000001180.

[66] Rücker G, Schwarzer G. Beyond the forest plot: The drapery plot. Res Synth Methods 2021;12:13–9. https://doi.org/10.1002/jrsm.1410.

[67] Stack overflow n.d. https://stackoverflow.com (accessed June 5, 2023).

[68] Stack Exchange n.d. https://stackexchange.com (accessed June 5, 2023).

[69] The R Project for Statistical Computing n.d. https://search.r-project.org (accessed June 5, 2023).

[70] R Seek n.d. https://rseek.org (accessed June 5, 2023).