Journal of Research Methodology
Vol. 37 No. 1 (January–April 2024)
ISSN (online) 2697-4835
Academic Article https://doi.org/10.14456/jrm.2024.5

Using R for Meta-analysis in Business Research Context

Gamon Savatsomboon[1], Ong-art Chanprasitchai[2], Anirut Pongklee[3],
Sujin Butdisuwan[4], Piyapun Santaveesuk[5], and Kanokwan Noppan[6]

Received: August 18, 2023 – Revised: October 15, 2023 – Accepted: October 17, 2023

Abstract

R is widely used to perform meta-analysis for international publication. However, R has been rarely used by Thai academics and researchers to conduct meta-analyses for research and publication. This could be viewed as a practice gap. Thus, the goal of this paper is to alleviate this gap by encouraging the Thai academic and research community to use R for meta-analysis in the business research context. This article will demonstrate that R is just as effective as other programs at performing a basic meta-analysis, and beyond. The article is divided broadly into two main parts: systematic review and meta-analysis. The idea of systematic reviews is explained in the first part of the article, which comprises two subsections. The primary goal of Subsection 1 is the application of SRQ (instead of PICO), a framework for precisely articulating the research question that serves as the basis for additional literature inclusion in the meta-analysis at hand. Subsection 2, which defines the criteria for accepting or rejecting studies to be included in a meta-analysis study at hand, deals with the application of PRISMA. The remaining part of the paper focuses on how meta-analysis is carried out in R, as well as related topics, starting with the import of datasets for analysis from an Excel file. Examples of R scripts are provided to conduct meta-analysis using R. Text outputs of all required statistics are generated, for example, pooled effect sizes in the form of common (fixed) and random effects for assessing the heterogeneity of studies included in the meta-analysis at hand. In addition, R can generate both forest and funnel plots. The paper also discusses how to use the funnel plot to determine the publication bias of the study. R is fully capable of doing comprehensive meta-analyses, being on par with or even surpassing other programs in some areas. Thai researchers and academics can utilize R with confidence. Thus, Thai researchers and academics are strongly encouraged to use R for meta-analysis and share their findings, like other international researchers.

Keywords: Meta-analysis, PICO, PRISMA, R, meta Package, dmetar Package

[1] Mahasarakham Business School, Mahasarakham University, Mahasarakham 44150, Thailand.
gamon.s@acc.msu.ac.th
[2] *Corresponding Author,*
Mahasarakham Business School, Mahasarakham University, Mahasarakham 44150, Thailand.
ongart.c@acc.msu.ac.th
[3] Faculty of Agriculture and Technology, Nakhon Phanom University, Nakhon Phanom 48000, Thailand.
anirut.p@me.com
[4] School of Liberal Arts, Shinawatra University, Pathum Thani 12160, Thailand.
sujin.b@mru.ac.th
[5] School of Liberal Arts, Shinawatra University, Pathum Thani 12160, Thailand.
piyapun.s@mru.ac.th
[6] Independent Researcher.
iamnoppan@gmail.com

วารสารวิธีวิทยาการวิจัย
ปีที่ ๓๗ ฉบับที่ ๑ (มกราคม–เมษายน ๒๕๖๗)
ISSN (online) 2697-4835
https://doi.org/10.14456/jrm.2024.5 บทความวิชาการ

การใช้ R สำหรับการวิเคราะห์อภิมานในบริบทการวิจัยธุรกิจ

กมล เสวตสมบูรณ์[1], องอาจ ชาญประสิทธิ์ชัย[2], อนิรุทธิ์ ผงคลี[3], สุจิน บุตรดีสุวรรณ[4], ปิยพันธ์ แสนทวีสุข[5], และ กนกวรรณ นพพันธ์[6]

รับต้นฉบับ : 18 สิงหาคม 2566 – รับแก้ไข : 15 ตุลาคม 2566 – ตอบรับตีพิมพ์ : 17 ตุลาคม 2566

Abstract

R ถูกใช้อย่างแพร่หลายในการวิเคราะห์อภิมานสำหรับการเผยแพร่ในระดับนานาชาติ อย่างไรก็ตามนักวิชาการและนักวิจัยไทยใช้ R ในการวิเคราะห์อภิมานค่อนข้างน้อย ซึ่งอาจมีช่องว่างในทางปฏิบัติ บทความนี้จึงมีจุดมุ่งหมายเพื่อลดช่องว่างนี้โดยการส่งเสริมให้นักวิชาการและนักวิจัยชาวไทยใช้ R สำหรับการวิเคราะห์อภิมานในบริบทของการวิจัยทางธุรกิจ บทความนี้จะสาธิตว่า R สามารถทำการวิเคราะห์อภิมานได้อย่างมีประสิทธิภาพไม่แพ้โปรแกรมอื่น ๆ บทความแบ่งออกเป็น 2 ส่วนหลัก คือ การทบทวนอย่างเป็นระบบและการวิเคราะห์อภิมาน แนวคิดเบื้องต้นเกี่ยวกับการทบทวนอย่างเป็นระบบจะนำเสนอในส่วนแรกของบทความ ซึ่งประกอบด้วยส่วนย่อย 2 ส่วน คือการประยุกต์ใช้ SRQ (แทน PICO) เป็นกรอบการอธิบายคำถามวิจัยอย่างชัดเจน ซึ่งเป็นพื้นฐานสำหรับการรวบรวมวรรณกรรมเพิ่มเติมในการวิเคราะห์อภิมานที่มีอยู่ ส่วนย่อยที่ 2 คือ การกำหนดเกณฑ์การคัดเลือกงานวิจัยที่จะรวมไว้ในการวิเคราะห์อภิมานโดยประยุกต์ใช้ PRISMA ส่วนที่เหลือของบทความเน้นการดำเนินการวิเคราะห์อภิมานใน R รวมถึงหัวข้อที่เกี่ยวข้อง เริ่มต้นด้วยการนำเข้าชุดข้อมูลสำหรับการวิเคราะห์จากไฟล์ Excel ตัวอย่างของสคริปต์ R สำหรับการวิเคราะห์อภิมาน ผลการวิเคราะห์ทางสถิติที่จำเป็นทั้งหมด เช่น ขนาดอิทธิพลรวมในรูปแบบของสมมุติฐานทั่วไป (อิทธิพลคงที่) และขนาดอิทธิพลสุ่มสำหรับการประเมินความหลากหลายของงานวิจัยที่นำมาวิเคราะห์อภิมาน นอกจากนี้ R ยังสามารถสร้างแผนภาพป่าและแผนภาพกรวยได้อีกด้วย บทความนี้ยังอภิปรายถึงวิธีการใช้แผนภาพกรวยเพื่อความลำเอียงจากการตีพิมพ์ R สามารถทำการวิเคราะห์อภิมานได้อย่างครอบคลุมสมบูรณ์เทียบเท่าหรืออาจจะมากกว่าโปรแกรมอื่น ๆ ที่มีอยู่ นักวิจัยและนักวิชาการไทยสามารถใช้ R ได้อย่างมั่นใจ และควรใช้ R สำหรับการทำการวิเคราะห์อภิมานและเผยแพร่ตีพิมพ์ผลงานเหมือนนักวิจัยนานาชาติในวารสารที่มีชื่อเสียงต่อไป

คำสำคัญ: การวิเคราะห์อภิมาน, PICO, PRISMA, R, แพ็กเกจ meta, แพ็กเกจ dmetar

[1] คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม มหาสารคาม 44150
gamon.s@acc.msu.ac.th

[2] *Corresponding Author,*
คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม มหาสารคาม 44150
ongart.c@acc.msu.ac.th

[3] คณะเกษตรและเทคโนโลยี มหาวิทยาลัยนครพนม นครพนม 48000
anirut.p@me.com

[4] คณะศิลปศาสตร์. มหาวิทยาลัยชินวัตร ปทุมธานี 12160
sujin.b@mru.ac.th

[5] คณะศิลปศาสตร์. มหาวิทยาลัยชินวัตร ปทุมธานี 12160
piyapun.s@mru.ac.th

[6] นักวิจัยอิสระ
iamnoppan@gmail.com

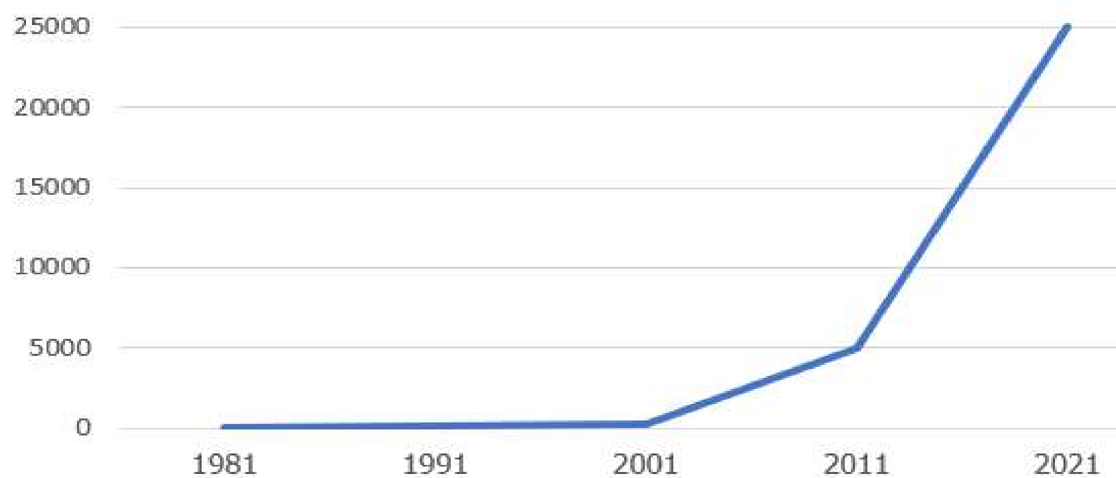
Note: This Thai abstract is generated by JRM.

Introduction

Based on Figure 1, the number of international publications using meta-analysis has increased sequentially during the period of 1991-2021 (Dettori et al., 2022), resulting in the emergence of many meta-analysis software packages, but two packages R and STATA have the most meta-analysis methods (Balduzzi et al., 2019). However, it must be remembered that software packages such as STATA and CMA are costly because they are commercial packages, whereas R is freeware and is widely used internationally. An urge to use R for meta-analysis by Thai scholars could be found in the work of Chananate (2010). In addition, Dechpichai et al. (2017) conducted a meta-analysis using R in the context of health sciences in Thailand. However, the use of R for meta-analysis in the context of business research context in Thailand, based on a literature review, is rare or non-existent. This could be considered a (practice) gap for this paper. For that reason, this paper attempts to fill this gap, by urging Thai academics and researchers to use R to conduct meta-analysis for their business research. R is used widely in practice for meta-analysis in business research in the international context. In short, R is free and fully capable of conducting meta-analysis for business research.

Figure 1

Trends of Articles Using Meta-analysis for International Publications.



Origin of Meta-analysis

O'Rourke (2007) stated that meta-analytic thinking has been seen since 1900. In that year, Karl Pearson studied British soldiers in several areas and compared groups vaccinated against typhoid fever to unvaccinated soldiers. However, Karl Pearson did not coin the term “meta-analysis”. In 1976, Gene Glass was the one who coined the term "meta-analysis". He stated that "the statistical analysis of a large collection of results from individual studies for the purpose of integrating the findings" and added that it was "an analysis of analyses" (Gogtay & Thatte, 2017). The application of meta-analysis in medicines began a few years after the term "meta-analysis" was coined (O'Rourke, 2007). In short, meta-analysis is a research technique used to methodically combine and synthesize data from several quantitative studies in a particular field of study (Paul & Barari, 2022).

Systematic Review in Brief

A systematic review is the required antecedent of meta-analysis. In the systematic review stage, in general, two concepts need to be applied together: PICO and PRISMA. The former leads to the identification of clear research question(s) of the meta-analysis of the study at hand. The latter executes the inclusion and exclusion of studies to be included in the meta-analysis study at hand. In addition, it also spells out the final number of studies to be included in a meta-analysis. One could study systematic reviews in depth from the work, such as Cajal et al. (2020). However, PICO, in the context of business research, may not be applicable. Thus, an alternative to PICO must be sought. This is done in the next section.

A proposed SRQ, as alternative to PICO in business research

It is common to use PICO, P (population), I (intervention), C (control), and O (outcome) as a base for the researchers to ultimately derive a research question. Thus, the overall purpose of PICO is to identify the research question. However, PICO may not work well with meta-analysis studies in the context of business research, because meta-analysis studies in business do not incorporate intervention and control groups in such studies. Thus, this article proposes a new concept, SRQ, to form a basis for a business research meta-analysis study, such as the one used as an example in this article. Let's try to explain SRQ. Here, S stands for sample. What is the target sample of the meta-analysis study? Given the example studies from the work of Ibrahim (2021) as examples for this article, the target samples are customers who are exposed to social media marketing activities of product brands. R stands for relationship. What relationship is being studied? The relationship between social media marketing activities and brand loyalty is being studied. Finally, Q stands for (research) question. What (research) question is being studied? The main research question being addressed by the example case of this article is, is the relationship between social media marketing activities and brand loyalty significantly positive? When the research question becomes clear, the business meta-analysis researcher knows what studies they need to look for during the systematic review phase. The next section describes

PRISMA (see Figure 2), a protocol to include or exclude studies in a meta-analysis at hand.

Using PRISMA to include or exclude studies for meta-analysis

In this article, the research question was obtained through the application of SRQ earlier. The research question is, is the relationship between social media marketing activity (SMMA) and brand loyalty significantly positive? Based on the literature review, this important research question has not been addressed in prior meta-analyses. We borrow this research question from the work of Ibrahim (2021) which we use as a main example in our article. The next step is to employ the concept of PRISMA to include or exclude studies of the meta-analysis based on the criteria set by the researcher beforehand. It is a widely accepted practice that PRISMA is used as a procedure to include or exclude studies for meta-analysis. PRISMA stands for Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Page et al., 2021). Given the example of this study, variables that need attention include the correlations (r) between social media marketing activity and band loyalty and sample sizes (n) of individual studies. Finally, a total of eleven studies are included in the meta-analysis used as an example in this article (see Figure 2).

Assessing the quality of studies to be included in a meta-analysis

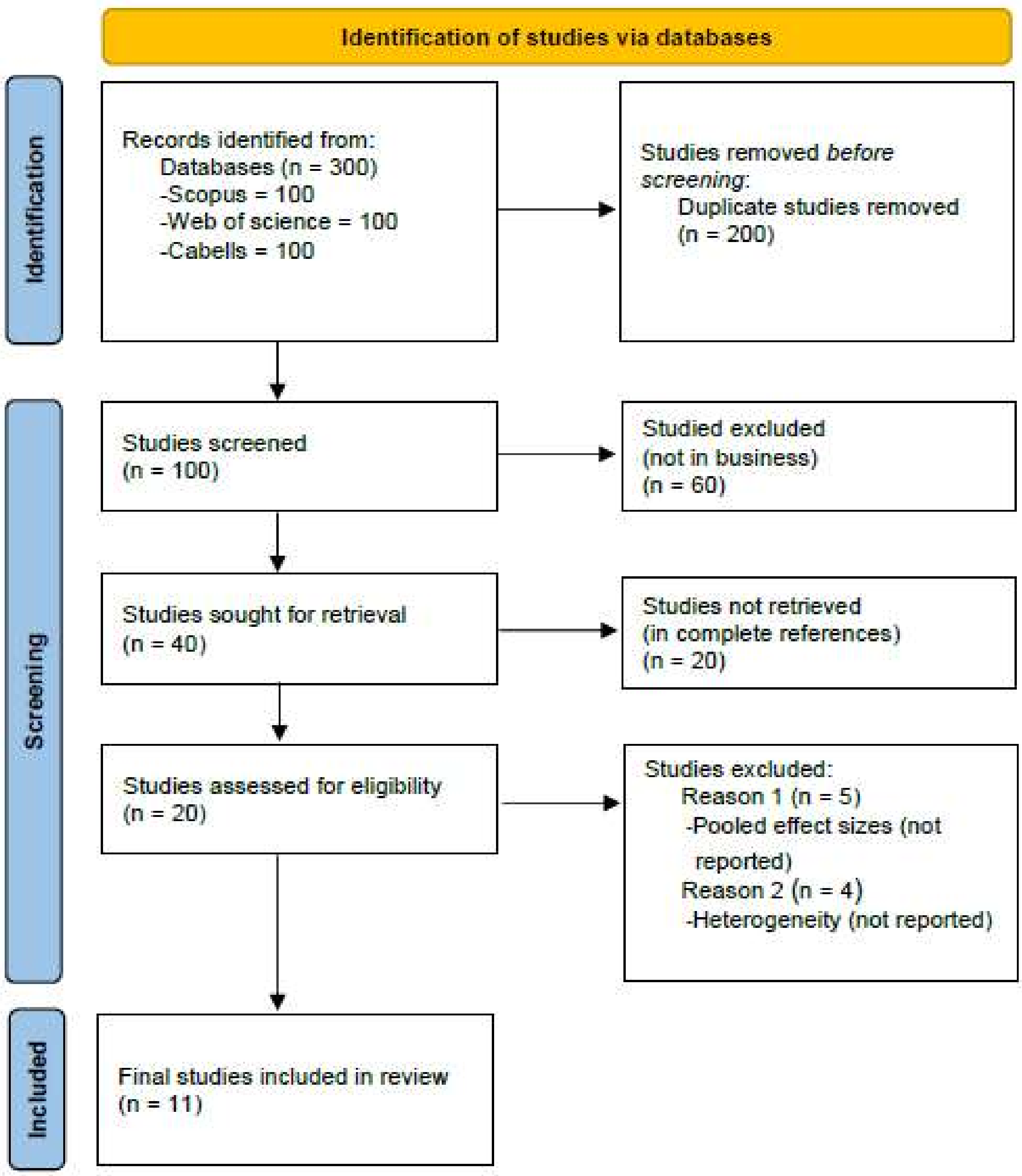
Feeley (2020) discusses the assessment of the quality of studies to be included in a meta-analysis study. Valentine (2009, as cited in Feeley, 2020) outlines four general areas whereby primary studies might be evaluated: (a) internal validity; (b) external validity; (c) statistical conclusion validity; and (d) construct validity. Of course, studies could be assessed using other areas/domains. Finally, Feeley (2020) suggests that there should be a section designated to assess the quality of studies to be included in a meta-analysis study. In addition, the researcher needs to explicitly define the assessment criteria in operational terms.

Conducting Meta-analysis using R

According to R Core Team (2021), R is a language and environment for statistical computing and graphics. To conduct a meta-analysis, an R package is required. This paper mainly uses the meta package. Let's begin by summarizing the capabilities of the meta package to perform meta-analyses. There are, of course, other R packages that could be used to conduct a meta-analysis. Table 1 summarizes the capabilities of the meta package to calculate the pooled effect sizes for meta-analysis. Obviously, the meta package is able to calculate adequate effect sizes for meta-analysis for both single-group and two-group meta-analysis research designs. First, researchers need to learn how to get the data into R for further meta-analysis. There are multiple ways that the data could be brought into R. To be able to do this, researchers wishing to use R need to learn the R syntax to instruct the meta package to perform the desired meta-analyses. The meta package produces both text and graphical outputs. Thus, researchers need to know how to interpret the necessary outputs generated by the meta package. The ability to interpret the outputs generated by R is crucial for reporting and publication.

Figure 2

PRISMA Illustrating the Procedure of Inclusion and Exclusion of Studies for Meta-analysis (Page et al., 2021).



Capabilities of meta package to pool the effect sizes

Here, the capabilities of the meta package to pool the effect sizes are summarized in Table 1. For full information on the meta package, please consult the work of Balduzzi et al. (2019).

Table 1

Statistical Methods, Functions, and Effect Sizes.

Statistical methods	Functions	Effect sizes
Meta-analysis of single means	metameans	MRAW and MLN
Meta-analysis of single proportions	metaprop	Proportion
Meta-analysis of single incidence rates	metarate	Incident rate and HR
Meta-analysis of single correlations	metacor	COR and ZCOR
Meta-analysis of continuous outcome data	metacont	MD, SMD, and ROM
Meta-analysis of binary outcome data	metabin	RR, RD, OR, ASD, DOR, and VE
Meta-analysis of incidence rates	metainc	IRR, IRD, IRSD, and VE
Meta-analysis of genetic inverse variance	metagen	RR, logRR, RD, OR, ASD, DOR, and VE

Remarks: MRAW = mean, MLN = log transformed mean, HR = hazard ratio, COR = correlation, ZCOR = Fisher's z transformed correlation, MD = mean difference, SMD = standardized mean difference, ROM = ratio of mean, RR = relative risk/risk ratio, RD = risk difference, OR = odds ratio, ASD = arcsine difference, DOR = diagnostic odds ratio, VE = vaccine effectiveness, IRR = incidence rate ratio, IRD = incidence rate different, IRSD = square root transformed incidence rate difference

In addition, there are several plots that could be generated by the meta package, including Forest plot, Funnel plot, Galbraith plot/radial plot, L'Abbe plot, Baujat plot, and Bubble plot.

Getting data into R

Again, the research question used as an example in this article is, is the relationship between social media marketing activities and brand loyalty positively significant? The dataset used in this article comes from the work of Ibrahim (2021). Data could be brought into R in two main ways. First, the data could be typed directly into R. Second, the data could be read from an Excel file located, for example, on the D: drive of a computer. Both raw data and effect sizes could be used as data inputs in R. For the sake of simplicity, this article adopted the second way, reading the data from an Excel file located on a D: drive. An example of research data is located on the left side of Figure 3. The data includes these variables, author, year, n, r, and location. R users need to produce an Excel file holding the exact same data, but needs to be done in R fashion (e.g. meta package), so R could read the data from the designated file storage location. We exerted an additional column, called g. This column is used for subgroup analysis in further analysis. In summary, a total of eleven studies are included in the example dataset. Variables included are an author (studlab), year (year), n (n), r (cor), location (location), and column D (g) intentionally added for subgroup analysis intentionally included in the meta-analysis design. In Figure 3 the data on the left, used as an example in this article, comes from the work of Ibrahim (2021) and the data on the right is an Excel data file, created based on the data located on the left.

Figure 3

Excel File Serving as a Data File (Located on Drive D:).

No.	Author	Year	n	r	Location
1	Salem and Salem	2019	240	0.41	Malaysia
2	Khan (2019)	2019	241	0.37	Saudi Arabia
3	Ibrahim and Aljarah	2018	389	0.26	Northern Cyprus
4	Yadav and Rahman	2018	371	0.09	India
5	Laksamana (2018)	2018	286	0.38	Indonesia
6	Bilgin	2018	547	0.63	Turkey
7	Ahmed and Hussain (2018)	2018	250	0.40	Pakistan
8	Ismail	2017	346	0.43	Malaysia
9	Algharabat	2017	400	0.24	Jordan
10	Karamian et al. (2015)	2015	315	0.80	Iran
11	Tariq (2015)	2015	150	0.25	Pakistan
	The final sample size		3535		

	A	B	C	D	E	F
1	cor	n	studlab	g	year	location
2	0.41	240	Salem and Salem	South East Asia	2019	Malaysia
3	0.37	241	Khan	Middle East	2019	Saudi Arabia
4	0.26	389	Ibraham and Aljarah	Middle East	2018	Northern cyprus
5	0.09	371	Yadav and Rahman	South Asia	2018	India
6	0.38	286	Laksamana	South East Asia	2018	Indonesia
7	0.63	547	Bilgin	Middle East	2018	Turkey
8	0.40	250	Ahmed and Hussain	South Asia	2018	Pakistan
9	0.43	346	Ismail	South East Asia	2017	Malaysia
10	0.24	400	Algharabat	Middle East	2017	Jordan
11	0.80	313	Karamian et al.	Middle East	2015	Iran
12	0.25	130	Tariq	South Asia	2015	Pakistan

Instructing R to conduct meta-analysis through syntax

Figure 4 is an example of a set of commands to instruct R to perform a meta-analysis to obtain the desired statistics. Starting with lines 1-2, R is instructed to install packages that the researcher wants, such as the meta package created by Balduzzi et al. (2019) for meta-analysis and readxl package created by Wickham & Brayan (2023) to read data from an Excel file located on drive D: for our example. Line 3, library(meta), is to put the meta package to use. Line 4 is to put the readxl package to use. This package is designed specifically to read the data from Excel files.

Line 4, data <-read_excel ("D:/Brand loyalty_1.xlsx"), reads raw data from the Excel file located on D: drive. Line 5, data, instructs input data to be displayed. Line 7-12 is a set of commands that uses the metacor function to analyze input data consisting of cor, n, and studlab (cor = correlation, n = sample size, and studlab = study). Line 13 provides the results of the analysis in text outputs. Line 14 is a command for generating a forest plot. Line 15 is a command for generating a funnel plot.

Figure 4

An Example of R Syntax to Conduct Meta-analysis.

```
1 install.packages(“meta”)
2 install.packages(“readxl”)
3 library(meta)
4 library("readxl")
5 data <-read_excel("D:/Brand loyalty_1.xlsx")
6 data
7  m1<- metacor(
8      cor,
9      n,
10      studlab,
11      data = data,
12      sm = "COR")
13 m1
14 forest(m1)
15 funnel(m1)
16 b1<- InfluenceAnalysis(m1, random = TRUE)
17 plot(b1, "baujat")
18 s1<-metainf(m1, pooled = "random")
19 print(s1)
20 m.reg <- metareg(m1, ~year)
21 bubble(m.reg, studlab = TRUE, bg = "lightblue")
```

Lines 16-17 are commands for generating a Baujat plot. Lines 18-19 are commands for conducting sensitivity analysis. Line 20 is the command for meta-regression. Finally, Line 21 is the command for bubble plot. We strongly encourage readers to consult the meta and readxl packages further explanations on the commands provided in Figure 4.

By the way, the metacor function is used to analyze single correlations for this paper. This function is applicable to the single-group design. In other words, there is no comparison between two groups here. Common (fixed) and random effects meta-analysis of correlations can be calculated in two ways. First, they can be computed by Fisher’s z transformation of correlations (ZCOR in R command language). Second, they can be computed by untransformed correlations (raw scores) (COR in R command language) (Balduzzi et al., 2019). The example presented in this article uses untransformed correlations (COR) as effect sizes as inputs for conducting the example meta-analysis of this article.

Results of meta-analysis in text outputs

From Figure 5, the results of the analyses can be divided into four major sections from top to bottom.

Figure 5

Results of Meta-analysis in Text Outputs.

```
                            COR              95%-CI    %W(common)   %W(random)
Salem and Salem          0.4100  [0.3045; 0.5155]           4.8          9.0
Khan                     0.3700  [0.2608; 0.4792]           4.4          9.0
Ibraham and Aljarah      0.2600  [0.1672; 0.3528]           6.2          9.2
Yadav and Rahman         0.0900  [-0.0111; 0.1911]          5.2          9.1
Laksamana                0.3800  [0.2807; 0.4793]           5.4          9.1
Bilgin                   0.6300  [0.5794; 0.6806]          20.7          9.5
Ahmed and Hussain        0.4000  [0.2957; 0.5043]           4.9          9.0
Ismail                   0.4300  [0.3440; 0.5160]           7.2          9.2
Algharabat               0.2400  [0.1475; 0.3325]           6.2          9.2
Karamian et al.          0.8000  [0.7601; 0.8399]          33.2          9.6
Tariq                    0.2500  [0.0882; 0.4118]           2.0          8.2

     Number of studies combined: k = 11
     Number of observations: o = 3513
                            COR            95%-CI          z      p-value
     Common effect model    0.4229 [0.5199; 0.5659]    26.42    < 0.0001
     Random effects model   0.3916 [0.2737; 0.5096]     6.51    < 0.0001

     Quantifying heterogeneity:
      tau^2 = 0.0370 [0.0170; .1147]; tau = 0.1932 [0.1304; 0.3386]
      I^2 = 97.3% [96.4%; 89.1%]; H = 6.14 [5.26; 7.16]

     Test of heterogeneity:
             Q  d.f.    p-value
     377.07  10   < 0.0001


    Details on meta-analytical method:
     - Inverse variance method
     - Restricted maximum-likelihood estimator for tau^2
     - Q-Profile method for confidence interval of tau^2 and tau
     - Untransformed correlations
```

Section one summarizes the names of 11 individual studies, correlations, CIs, and common and random weights of individual studies. Section two summarizes the number of studies (k = 11) and the total sample size (n = 3,513 samples). Section three summarizes pooled effect sizes for both common and random effects models. The next two sections summarize heterogeneity values. Finally, the last section summarizes details on the meta-analytical method of the example study. Here, the random effect model may be selected because the heterogeneity is high (97.3%). Another reason that the random effect model is used is that the distribution of effect sizes of the studies is not normal (Spineli & Pandis, 2020). One solution is to identify outliers and remove the necessary outliers (i.e. studies deemed as outliers). In this article, the problem of heterogeneity is solved by using outlier analysis removing studies that may be identified as outliers based on the outlier analysis.

Between-study heterogeneity and other required statistics

Between-study heterogeneity refers to the degree of similarity of the pooled effect sizes (e.g. correlations) of studies that are included in the meta-analysis (Linden & Hönekopp, 2021). If the values of the effect sizes are close together, it could be considered that the degree of heterogeneity is low. This is desirable. On the other hand, if they are not close together, it would mean that the degree of heterogeneity is high. This is certainly not desirable. Heterogeneity levels can be measured by Cochran's Q-statistic, Higgin's & Thompson's I^2, and Tau-squared τ^2. R provides the outputs for all these statistics and their p-values where required (see Figure 5). In general, we do not want heterogeneity to be significant. The degree of heterogeneity could be classified as low, moderate, or high. Moderate to low are desirable levels of heterogeneity. These are expressed as percentages (%), see the levels of heterogeneity at the end of this section, with cut scores at 25% (low heterogeneity), 50% (moderate heterogeneity), and 75% (high heterogeneity). One way to address high heterogeneity is to conduct an outlier analysis (identifying extreme values) as mentioned earlier. If outliers are found, a researcher may consider taking those outliers out, and reconduct the meta-analysis again without outliers.

Forest plot

A forest plot is a graphical representation of the findings of a meta-study (Dettori et al., 2021). In the preceding section, the results of the meta-analysis are shown as text outputs. R also displays graphic outputs, a forest plot, for example. Based on Figure 6 below, eleven studies were included (no outliers excluded) in the forest plot. From left to right, the first column is called Study. This column contains studies included in the meta-analysis at hand. The second column is called Total. This column includes the number of sample sizes of the eleven studies included in the meta-analysis at hand. The third column contains individual studies' effect sizes, and at the end the pooled effect size using visual representations. The next column is COR, symbolizing correlation (r). This column includes the effect sizes of individual studies and the pooled effect size of the study. The remaining columns include CIs and common (fixed) and random weights, both expressed in %. In short, based on the forest plot, the relationship between social media marketing activities (SMMA) and band loyalty (BY) is positive. The bottom left of the diagram showed heterogeneity ($p < .01$) and $I^2 = 93\%$. One solution that could reduce the level of heterogeneity is to exclude studies identified as outliers. This is done later.

Funnel plot

An idea of a funnel plot is to graphically illustrate whether the studies are spread out asymmetrically. Based on Figure 7, it is obvious that the eleven studies included in the example study are asymmetric. This is certainly not desirable. Again, if outliers (studies) are taken out the results may improve. Let's wait until outlier analysis is conducted later to see if the funnel plot is still asymmetric when reanalyzed with the outliers removed. The expectation is the antithesis.

Figure 6

Forest Plot (Outliers Not Removed).

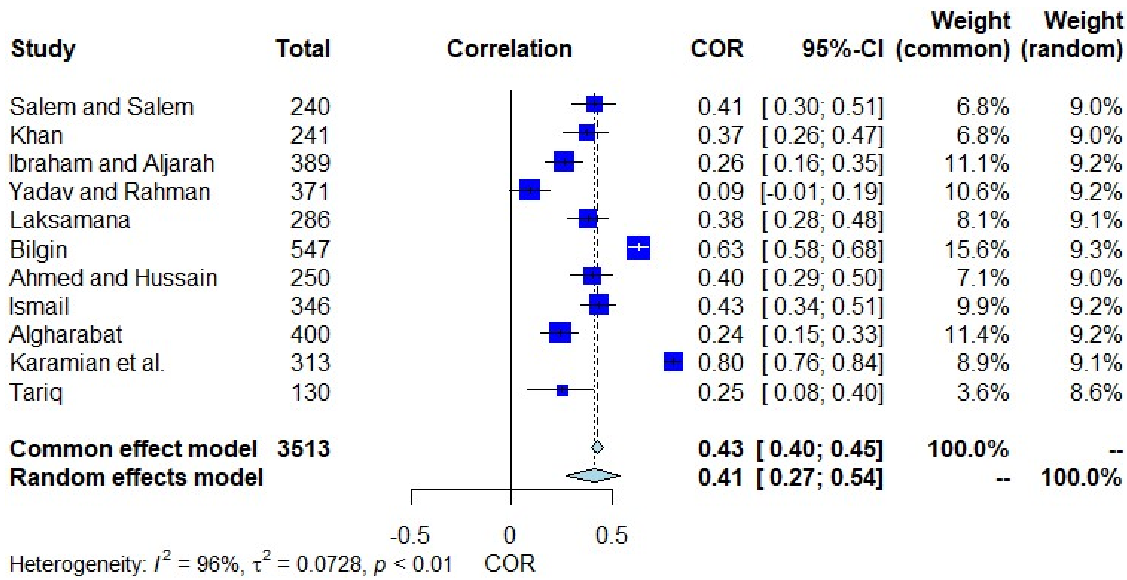
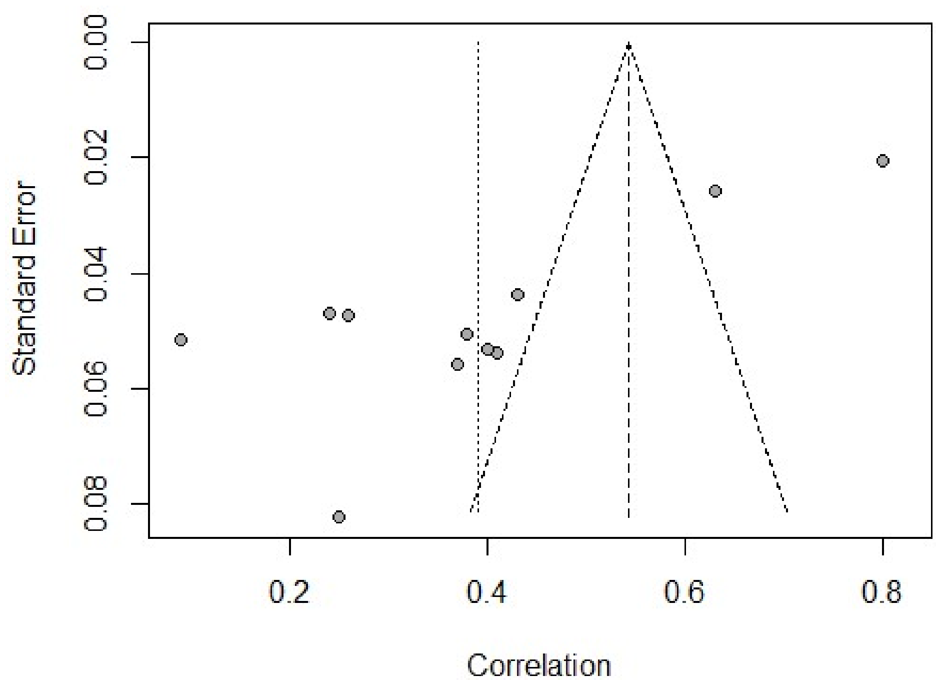


Figure 7

Funnel Plot (Outliers Not Removed).

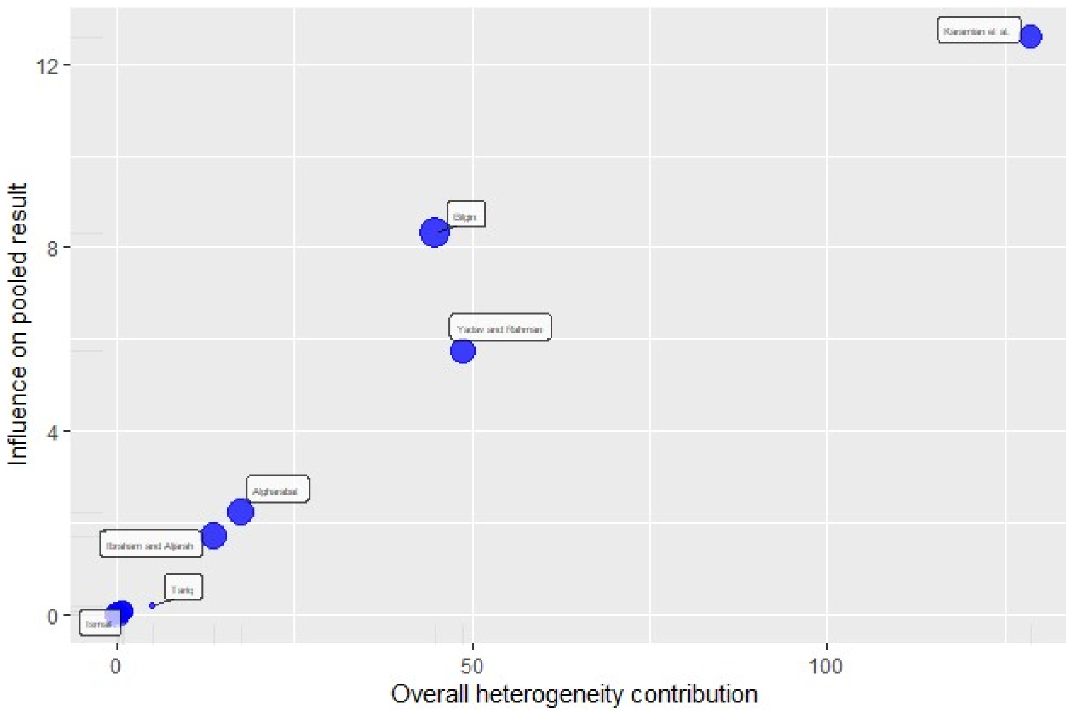


Baujat plot, identifying potential outliers

A Baujat plot can be used to identify potential outliers of the study. Based on the Baujat plot below, potential outliers could be visually examined. However, it is probably challenging to identify the exact number of outliers. Thus, it is recommended that a function of the dmetar package is used to identify outliers of the study. This is done in the next section.

Figure 8

Baujat Plot for Identifying Potential Outliers.

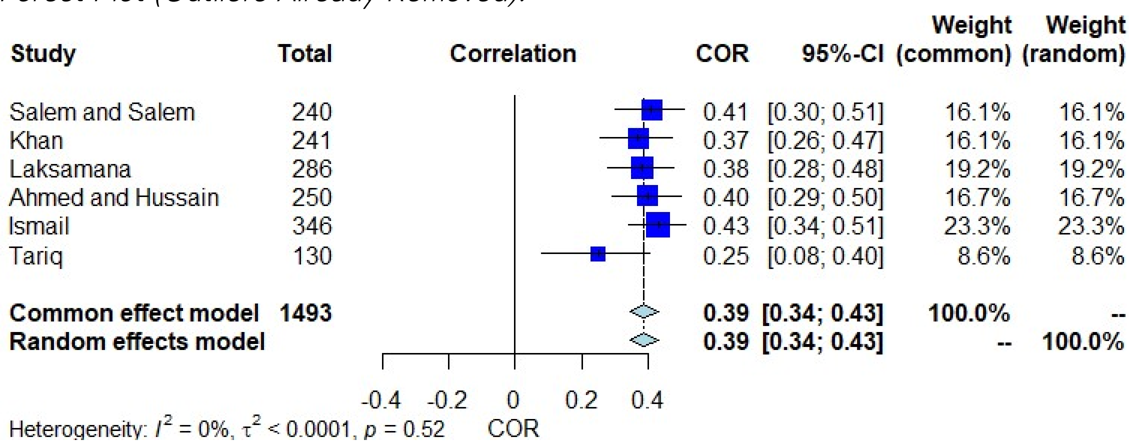


Outlier analysis, identifying and removing outliers

Here, outlier analysis was performed to detect outliers using the dmetar package. Five outliers (out of eleven) were detected; namely, "Ibraham and Aljarah", "Yadav and Rahman", "Bilgin", "Algharabat", and "Karamian et al." If outliers were found, a researcher could take them out from his/her analysis. Doing so could lead to improving the level of heterogeneity of his/her study. Based on Figure 9, it can be seen that both common (fixed) and random effects are significant (based on confidence intervals (CIs) presented in Figure 9). Also important, the heterogeneity tests were conducted, and the I^2 threshold and p =0.52 (p>0.05) passed the criteria, meaning after excluding the outliers, the remaining studies' effect sizes (value) are close. Here, the level of heterogeneity is at a moderate level (acceptable). In short, the heterogeneity level improved drastically (from 96% to 0%) after taking out the identified outliers. In fact, the improvement is excellent.

Figure 9

Forest Plot (Outliers Already Removed).



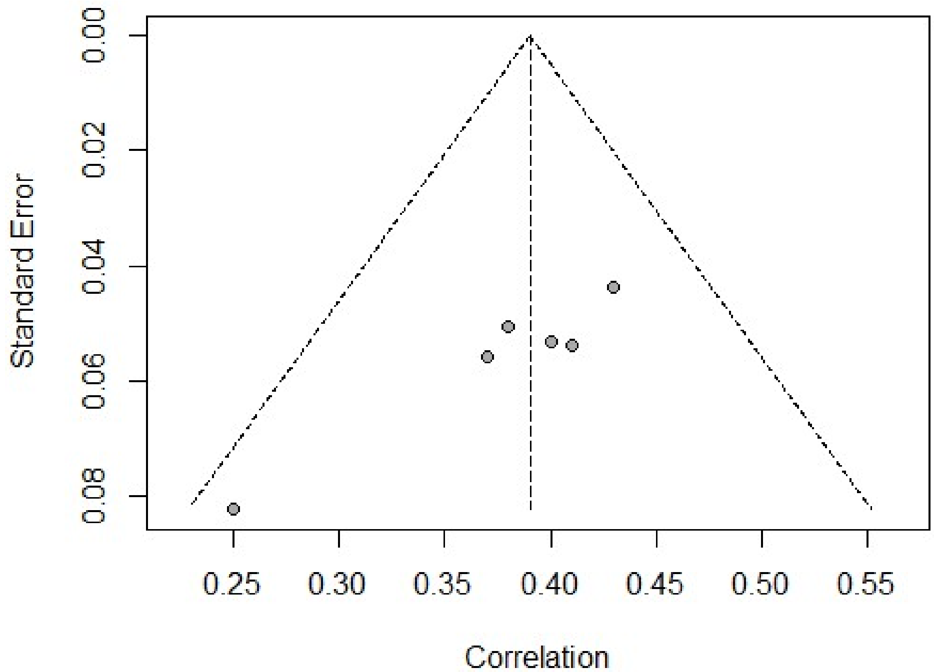
Identifying publication bias through funnel plot

The idea of the funnel plot is to illustrate whether the included studies of a meta-analysis are not asymmetrical. If the funnel plot illustrates that the studies are not asymmetrical, it basically means that there is no publication bias. In other words, a symmetrical funnel plot is desirable because it indicates an absence of both biases and heterogeneity (Simmonds, 2015). Based on Figure 10, it appears that the included studies in this article are not asymmetrical; therefore, no publication bias. Please note that the heterogeneity is at 0%. This is a clear indication that publication bias does not exist. However, this is already confirmed by the text outputs presented earlier in Figure 5. The funnel plot has two major components: correlation (i.e., effect sizes) and standard error.

The former represents the effect sizes of the included studies. The latter indicates the level of asymmetry or the degree of heterogeneity (approaching zero preferable). Based on the text outputs (presented in Figure 5), the p-value of the heterogeneity is nonsignificant. This indicates clearly that there is no publication bias. In summary, both the text outputs and funnel plot confirm that publication bias is not a problem for the study at hand.

Figure 10

Funnel Plot (Outliers Removed) of Effect Sizes (Correlation).



Sensitivity analysis

Sensitivity analysis aims to test whether a single study dominates the pooled effect size of the meta-analysis at hand. We certainly do not want that to happen. Sensitivity analysis uses an analysis called influential analysis. The idea of influential analysis is to omit one study at a time using the leave-one-out method that comes with the meta package. Based on the analysis in Figure 9, the pooled effect is significant ($p<0.05$). Let's take a Salem and Salem study as an example. If a study by Salem and Salem is omitted (taken out), we would expect the pooled effect size to be still significant ($p<0.05$). If not, we can assume that the Salem and Salem study is an influential study. The latter is not desirable. Again, an influential study refers to a study that dominates the pooled effect size of the meta-study at hand. Before analyzing the next leave-one-out round, the last study taken out (omitted) needs to be put back in. Of course, an influential analysis requires a specific set of syntax. With the right syntax, R will perform a desired sensitivity analysis for the researcher. There are two possible sensitivity analysis outcomes. The first outcome, all selected studies pass the sensitivity test, in which case, a researcher does not have to do anything further. The researcher just needs to report the results of the sensitivity analysis without any influential study. If, however, one or more selected studies do not pass the sensitivity test, it would mean that one or more studies need to be taken out to make the meta-analysis at hand free from any influential study.

Figure 11

Sensitivity Analysis (Influential Analysis) Text Outputs.

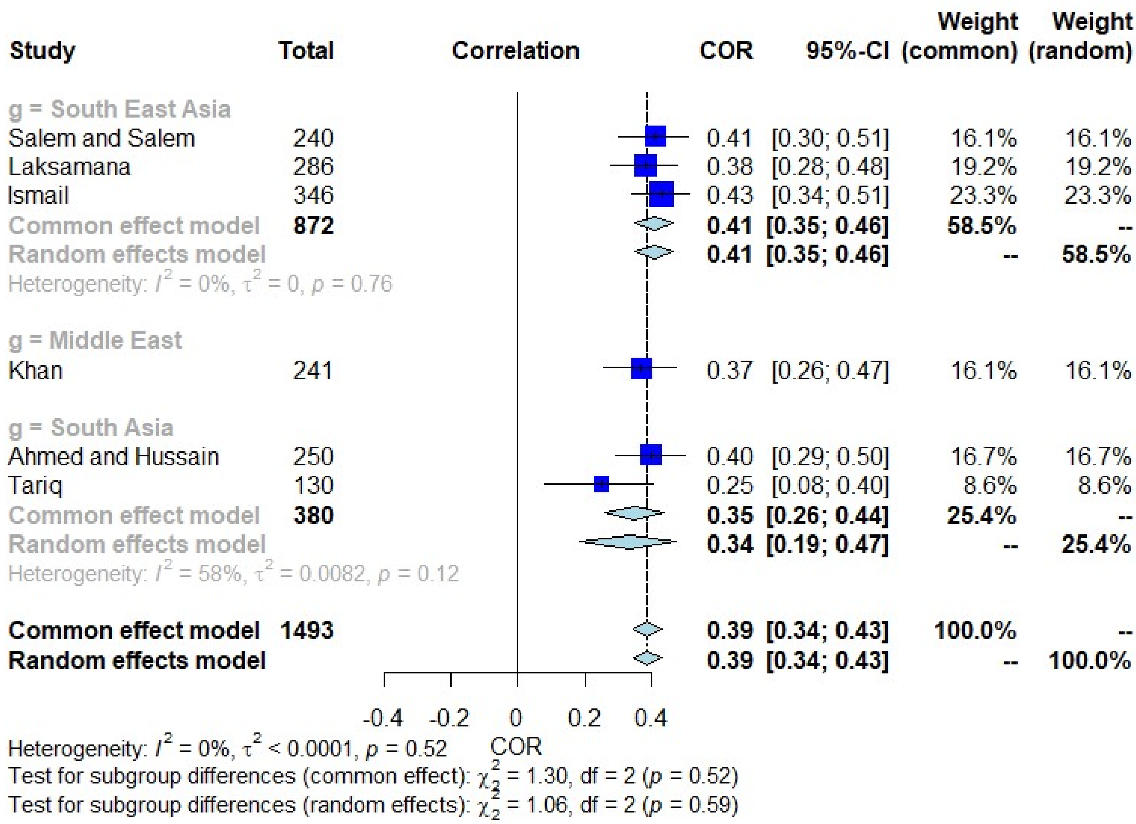
```
                    Influential analysis (random effects model)

                              COR             95%-CI   p-value    tau^2
Omitting Salem and Salem   0.3862 [0.3391; 0.4333]  <0.0001  0.0000
Omitting Khan              0.3938 [0.3470; 0.4407]  <0.0001  0.0000
Omitting Laksamana         0.3925 [0.3448; 0.4402]  <0.0001  0.0000
Omitting Ahmed and         0.3881 [0.3409; 0.4354]  <0.0001  0.0000
Hussain                    0.3768 [0.3271; 0.4266]  <0.0001  0.0000
Omitting Isamail           0.4008 [0.3562; 0.4454]  <0.0001  0.0000
Omitting Tariq
                           0.3901 [0.3471; 0.4332]  <0.0001  0.0000
Pooled estimate            tau      I^2
                           0.0004    0.0%
Omitting Salem and Salem   0.0019    0.0%
Omitting Khan              0.0009    0.0%
Omitting Laksamana         0.0015    0.0%
Omitting Ahmed and         0.0025    0.0%
Hussain                    0.0000    0.0%
Omitting Isamail
Omitting Tariq             0.0020    0.0%

Pooled estimate
Details on meta-analytical method:
- Inverse variance method
- Restricted maximum-likelihood estimator for tau^2
- Untransformed correlations
```

Subgroup analysis

Subgroup group analyses below are divided by geographical region. Based on Figure 12, studies from different countries are subdivided into three regions: South East Asia, Middle East, and South Asia. Please note that the study called Khan would not be counted as a group because there is only one study under the Middle East. From left to right, the forest plot starts out with a study under which subgroups and individual studies are classified. The second column is the total sample sizes of individual studies. The third column presents the effect sizes of individual studies and overall. The fourth column includes actual correlations pulled from individual studies. The two pooled effect sizes are from the common (fixed) effect model and the random effects model. Finally, the remaining columns include CIs and weights from both common and random models. Here, we demonstrated how the meta package can be used to conduct subgroup analysis. For example, the pooled effect sizes of subgroups can be spelled out. However, subgroup analysis is an advanced analysis. Using subgroup analysis is contingent upon the objective of the meta-analysis at hand and the researcher’s ability to carry out subgroup analysis. For in-depth knowledge of subgroup analysis and how to carry out a meta-analysis using R, researchers are strongly recommended to study the work of Harrer et al. (2021).

Figure 12

Forest Plot and Subgroup Analysis.



Meta-regression and Bubble plot

The two topics, meta-regression and bubble plot, are advanced topics. Thus, the decision rests with the researcher whether to conduct meta-regression and bubble plot. It also depends on how sophisticated the researcher is in terms of their ability to conduct meta-regression and bubble plot using R. The results of meta-regression can be illustrated in two ways. However, these two ways go hand in hand. The first way provides the results of the meta-regression analysis in text outputs, and the second provides the results of the meta-regression in a graphic format (e.g. bubble plot). As usual, the researcher needs to instruct R to conduct a meta-regression analysis. No new data is required because meta-regression is performed using the same dataset at hand. If a researcher wants to conduct a meta-regression analysis, he/she needs to identify independent variable(s) carefully. Given, the example data used in this article, a researcher could choose the year as an independent variable. Here, a researcher would like to know if the independent variable (year), for example, has caused any change in the pooled effect size of all studies included over times. Figure 13 is the text output for meta-regression analysis. Based on Figure 13, the independent variable (year) is not the cause of change in the pooled effect sizes of all studies included in the meta-analysis study (p > .05).

Figure 13

Text Output of Meta-regression.

```
Mixed-Effects Model (k = 6; tau^2 estimator: REML)

tau^2 (estimated amount of residual heterogeneity):        0 (SE = 0.0020)
tau (square root of estimated tau^2 value):                        0
I^2 (residual heterogeneity / unaccounted variability):    0.00%
H^2 (unaccounted variability / sampling variability):      1.00
R^2 (amount of heterogeneity accounted for):               100.00%

Test for Residual Heterogeneity:
QE(df = 4) = 3.3545, p-val = 0.5003

Test of Moderators (coefficient 2):
QM(df = 1) = 0.6943, p-val = 0.4047

Model Results:

         estimate      se     zval    pval      ci.lb    ci.ub
intrcpt  -33.4630  40.6290  -0.8236  0.4102  -113.0945  46.1684
year       0.0168   0.0201   0.8332  0.4047    -0.0227   0.0562
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

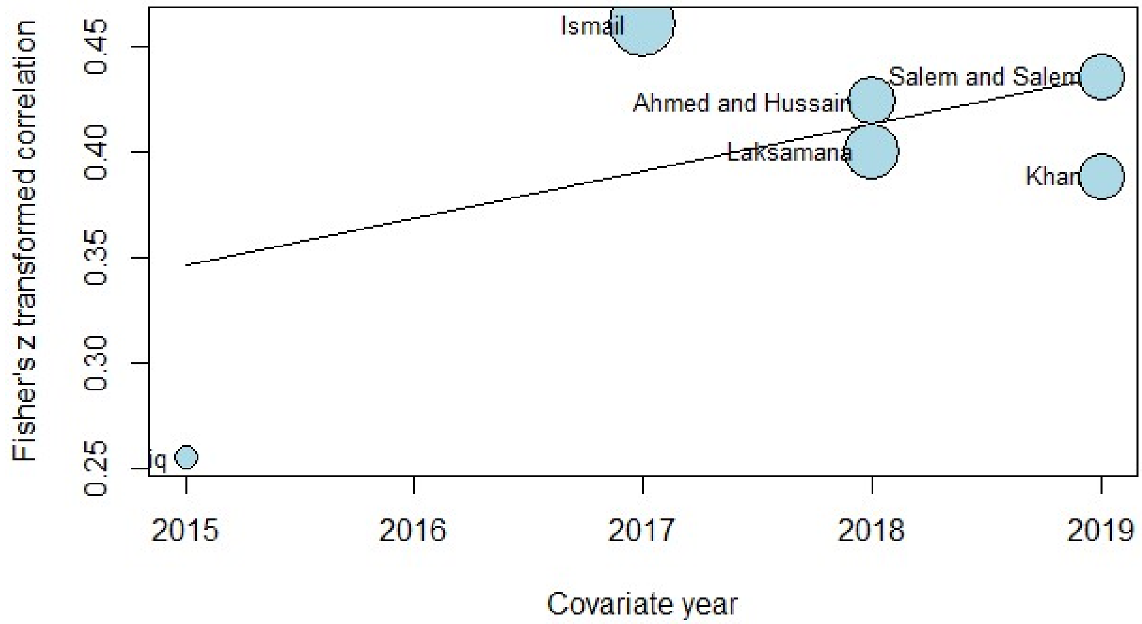
Based on the bubble plot presented in Figure 14, the independent variable (year) is not the cause of change in the pooled effect sizes of all studies included in the meta-analysis study (p > .05). Researchers could study meta-regression further from the work of Tatsioni & Ioannidis (2017), for example.

Advanced Topics

What we have presented so far are the basics of meta-analysis, and some advanced topics including meta-regression and bubble plot. However, there are more advanced topics in meta-analysis. These advanced topics are worthy of learning; thus, researchers are recommended to further study these advanced topics, because these advanced topics would serve as advanced options for analyzing meta-analysis data for their research. The results of the analyses would, of course, be more advanced, leading to more sophisticated and stronger research works. These advanced topics include: Cumulative meta-analysis, Power analysis, Multilevel meta-analysis, Structural equation meta-analysis, Network meta-analysis, and Bayesian meta-analysis.

Figure 14

Bubble Plot and Meta-Regression.



Minimum Requirements for Publishing Using Meta-analysis

This section describes what needs to be included in the researcher's manuscript for publication using meta-analysis at the minimum. First, the researcher needs to describe his/her systematic review by presenting the research problem clearly. It may be easier and more acceptable to use the PICO template to identify the research problem of a meta-analysis study. However, this article proposed and used SRQ instead. The rationale for that was already provided earlier. In addition, it is recommended that PRISMA be adopted as a protocol to execute the inclusion and exclusion criteria of studies to be included in the researcher's meta-analysis study. At the end of PRISMA, how many studies are required to be included in the meta-analysis at hand needs to be spelled out, which was also already discussed in this article earlier. Next, critical text outputs of meta-analysis generated by R need to be included in a report, summarizing and interpreting all the necessary statistics. In addition, the forest plot needs to be presented, followed by the funnel plot. Finally, sensitivity analysis should be conducted after the forest plot to ensure that no outliers are included in the meta-analysis study. If there are outliers, they should be taken out. In the end, sensitivity analysis can improve the level of heterogeneity of the meta-analysis study at hand. The desirable level of heterogeneity would depend on the goodness of the data. The reporting components recommended here in this article include only the basics. The scopes of meta-analysis studies may be different from researcher to researcher. Therefore, reporting may vary. In short, this section provides brief guidelines in terms of what to include in a report/manuscript aiming for publication.

Conclusions and Recommendations

We have demonstrated that R has all the required capabilities to perform basic meta-analysis, and beyond. This article began by briefly introducing a systematic review. Here, two important concepts were introduced: PICO and PRISMA. PICO is a concept that helps researchers to identify their research problem, for those using meta-analysis as a part of the research design. However, SRQ is used instead of PICO, because PICO is not applicable for this article. Thus, SRQ was adopted as a framework to identify the research question of this meta-analysis study. PRISMA is a procedure that executes the inclusion and exclusion of studies to be used in the meta-analysis study at hand. Next, the article illustrated what basic analyses are required to conduct basic meta-analyses.

In addition, advanced topics of meta-analysis were also introduced. However, this article does not provide a demonstration of how to conduct the analyses relating to these advanced topics, because these topics are beyond the scope of this article. Therefore, researchers should study these advanced topics on their own from various sources, such as the meta package 'meta' in a PDF format (https://cran.r-project.org/web/packages/meta/meta.pdf). Another good resource is a book called "Doing Meta-analysis with R: A Hands-on Guide" written by Herrer et al. (2021). There are also other R packages, such as metafor package that researchers should learn. All available R packages designated for meta-analysis could be used in combination under the R environment. In addition, free learning resources are abundantly available on various websites. For those researchers who are less familiar with meta-analysis, they should consider taking short courses or training programs, or both.

Finally, we encourage Thai academics and researchers to use R in their teaching, research, and publications, because R is free, fully capable, and is recognized worldwide. Thai academics and researchers could enjoy using R to publish in world-leading journals, such as PLOS ONE (indexed on Scopus database and Web of Science, H-index of the journal is 404), side by side with those international researchers who use R to publish their researchers in those world-class journals.

Acknowledgements

We want to thank our universities for creating the research ecosystem needed for faculty members to embark on scholarly projects. The amount and quality of research outputs rise as a result.

References

Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-based Mental Health, 22*(4), 153–160. https://doi.org/10.1136/ebmental-2019-300117

Cajal, B., Jiménez, R. M. M., Gervilla, E., & Montaño, J. C. M. (2020). Doing a systematic review in health sciences. *Clínica Y Salud, 31*(2), 77–83. https://doi.org/10.5093/clysa2020a15

Chananet, C. (2010). Using R program for meta-analysis. *The Journal of Applied Science, 9*(1), 28-38. https://www.journal.sci.kmutnb.ac.th/journal_files/3_1_2553.pdf

Dechpichai, P., Kongchareon, J., Suwanateep, N., Tuprakay, S.R., Pongsopa, J. (2017). Kham Chuk Nai Lok Phoom Phae Nai Dek Thai: Kan Vi Klo Apiman [The pooled prevalence of allergic disease]. *Advances Science, 17*(2), 78-98. https://sci.bsru.ac.th/sciweb/e-magazine/17-2/chapter-7.pdf (In Thai)

Dettori, J. R., Norvell, D. C., & Chapman, J. R. (2021). Seeing the forest by looking at the trees: how to interpret a meta-analysis forest plot. *Global Spine Journal, 11*(4), 614–616. https://doi.org/10.1177/21925682211003889

Dettori, J. R., Norvell, D. C., & Chapman, J. R. (2022). Fixed-effect vs random-effects models for meta-analysis: 3 points to consider. *Global Spine Journal, 12*(7), 1624–1626. https://doi.org/10.1177/21925682221110527

Feeley, T. H. (2020). Assessing study quality in meta-analysis. *Human Communication Research, 46*(2–3), 334–342. https://doi.org/10.1093/hcr/hqaa001

Gogtay, N. J. & Thatte, U. M. (2017). An introduction to meta-analysis. *Journal of the Association of Physicians in India, 65*, 78-85. https://www.kem.edu/wp-content/uploads/2012/06/14-An-introduction_to_meta_analysis-1.pdf

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R: A hands-on guide.* Chapman & Hall/CRC Press.

Ibrahim, B. (2021). Social media marketing activities and brand loyalty: A meta-analysis examination. *Journal of Promotion Management, 28*(1), 60–90. https://doi.org/10.1080/10496491.2021.1955080

Linden, A., & Hönekopp, J. (2021). Heterogeneity of research results: A new perspective from which to assess and promote progress in psychological science. *Perspectives on Psychological Science, 16*(2), 358–376. https://doi.org/10.1177/1745691620964193

O'Rourke, K. (2007). A historical perspective on meta-analysis: Dealing quantitatively with varying study results. *Journal of the Royal Society of Medicine, 100*(12), 579–582. https://doi.org/10.1177/0141076807100012020

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., &

McGuinness, L. A. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *British Medical Journal, 372*(71). https://www.bmj.com/content/372/bmj.n71

Paul, J., & Barari, M. (2022). Meta-analysis and traditional systematic literature reviews—What, why, when, where, and how? *Psychology & Marketing, 39*(6), 1099–1115. https://doi.org/10.1002/mar.21657

R Core Team (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. https://www.R-project.org/.

S, Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLOS Medicine, 18*(3), Article e1003583. https://doi.org/10.1371/journal.pmed.1003583

Simmonds, M. (2015). Quantifying the risk of error when interpreting funnel plots. *Systematic Reviews, 4*(1). https://doi.org/10.1186/s13643-015-0004-8

Spineli, L. M., & Pandis, N. (2020). Fixed-effect versus random-effects model in meta-regression analysis. *American Journal of Orthodontics and Dentofacial Orthopedics, 158*(5), 770–772. https://doi.org/10.1016/j.ajodo.2020.07.016

Tatsioni, A., & Ioannidis, J. P. A. (2017). Meta-analysis. International Encyclopedia of Public Health, 117–124. https://doi.org/10.1016/b978-0-12-803678-5.00291-5

Wickham, H. &, Bryan, J. (2023). *readxl: Read Excel files.* [R package]. https://readxl.tidyverse.org