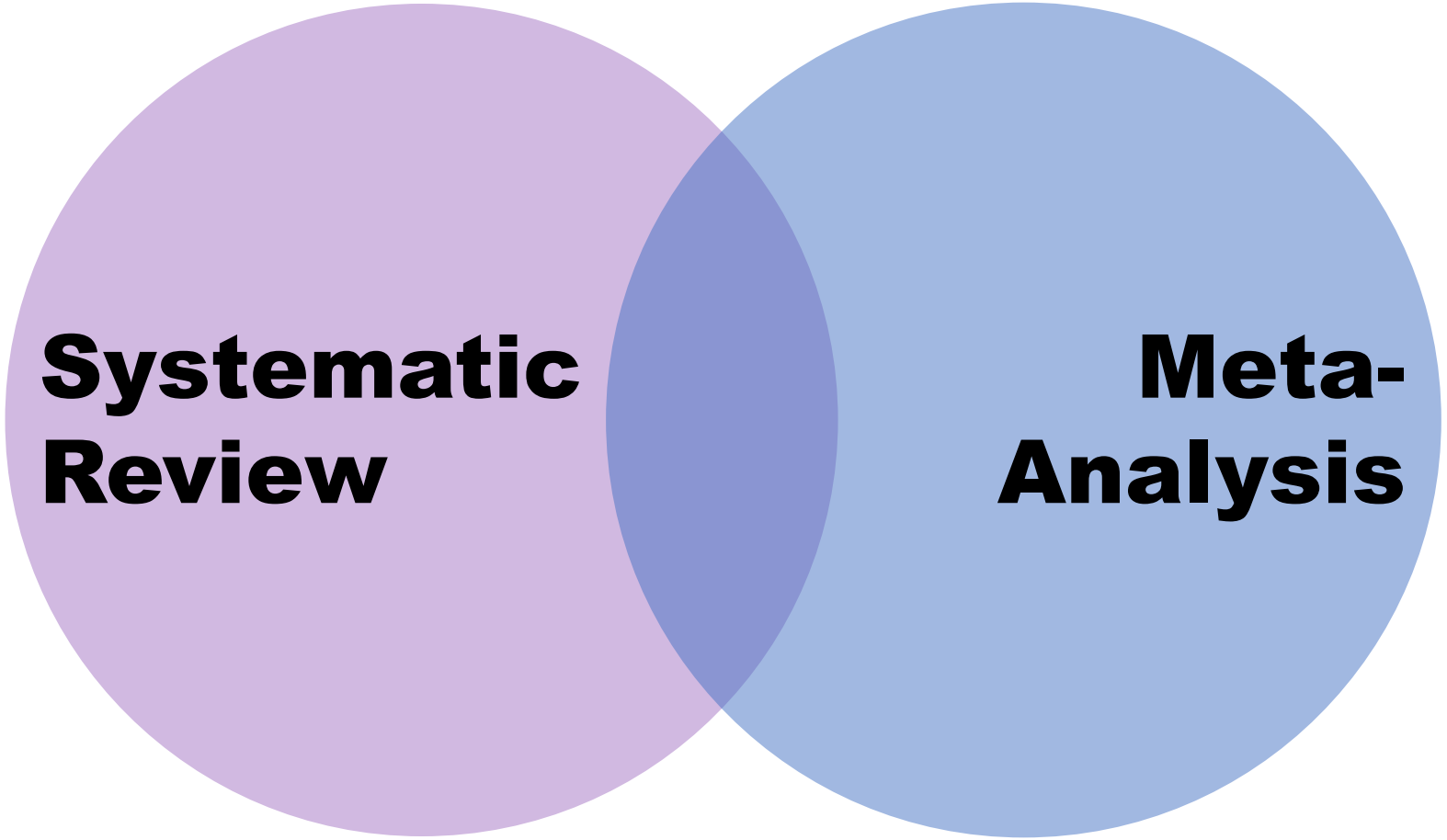
Introduction to Meta-Analysis

Samantha Estrada PhD

ORSSP Data Analysis Lab Consultant

What is Meta-Analysis?

"the ***statistical analysis*** of a large collection of ***analysis results*** from individual studies for the purpose of integrating the findings"
(Glass, 1976)


Systematic Review
Meta-Analysis

Where to start…

- Identify a topic
  - Team
- Keywords
  - Report on the keywords you used, “finney schraw current statistics self-efficacy”, “current statistics self-efficacy CSSE” “CSSE”
- Boolean logic
  - *Statistical self-efficacy OR*
  - *Statistical confidence OR*
  - *Statistical anxiety OR*
  - *Statistical education* OR*
  - *Statistical learning OR*
  - *Statistical self-belief OR*

“Data Collection”

- Identify a popular database within your field to comb through the studies.
  - Google Scholar
  - Web of Science
  - PsycINFO
  - Pubmed/Medline
- Covidence tool: https://libguides.uttyler.edu/c.php?g=1341980&p=9993128

[HTML] **Self-efficacy** beliefs in college **statistics** courses

SJ **Finney**, G Schraw - Contemporary educational psychology, 2003 - Elsevier

… a measure of **statistics self-efficacy** and use it to examine growth in **self-efficacy** over a one-…

First, there is no measure of **statistics self-efficacy** with validity evidence. Instead, previous …

☆ Save Cite Cited by 457 Related articles All 6 versions

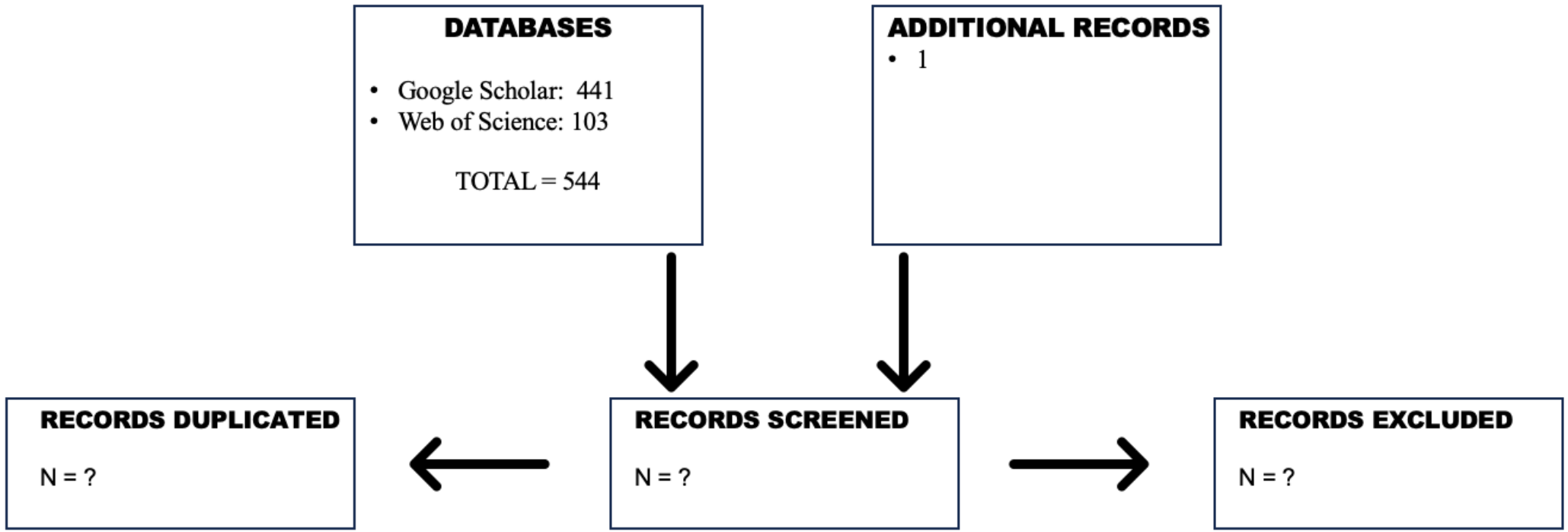
PICOT

- To include the studies there needs to be a similarity in the studies.
- Suggested models for this:
- PICOT (PCR Online, n. d.)
  - Populations
  - Intervention
  - Comparison
  - Outcome
  - Time frame

Guidelines

- PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA)
  - Tricco et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of internal medicine*, *169*(7), 467-473.
  - There are variations of this model
    - PRISMA-IPD for individual, participant data meta-analyses
    - PRISMA-NMA for network meta-analyses.

- REGEMA: REliability GEneralization Meta-Analysis
  - Sánchez-Meca, J., Marín-Martínez, F., López-López, J. A., Núñez-Núñez, R. M., Rubio-Aparicio, M., López-García, J. J., ... & López-Nicolás, R. (2021). Improving the reporting quality of reliability generalization meta-analyses: The REGEMA checklist. *Research Synthesis Methods*, *12*(4), 516-536.


DATABASES
• Google Scholar: 441
• Web of Science: 103
TOTAL = 544
ADDITIONAL RECORDS
• 1
RECORDS DUPLICATED
N = ?
RECORDS SCREENED
N = ?
RECORDS EXCLUDED
N = ?

“Data Collection”

- Including conference proceedings, posters, and others.
  - *Careful of duplicate records*

- Emailing authors to share data.
  - Open Science Framework: https://osf.io/
  - Journal and/or universities databases
- Store digital copies in a reliable place. Things disappear from the internet.
  - Journal Articles, conference proceedings, posters* etc.

Reliability Generalization

- Reliability generalization is a type of meta-analysis
- Focus on the reliability estimates, usually Cronbach's alpha, vary when the test is applied to different samples (Sanchez-Meca et al., 2019)

Coding Process

Method

Participants

The Italian sample consisted of 512 psychology students attending the University of Florence in Italy, who enrolled in an introductory statistics course in 2008 and 2009 ($n = 204$ and $n = 308$, respectively). The course covered the usual introductory topics of descriptive and inferential statistics and their application in psychological research. Participant ages ranged from 19 to 52 ($M = 22.3$, $SD = 5.40$, and median = 20); most of the participants were women (81%). This proportion reflects the gender distribution of the population of psychology students in Italy. The Spanish sample consisted of 336 psychology students attending the University of Huelva and Seville in Spain, who enrolled in an introductory statistics course in 2008 and 2009 ($n = 206$ and $n = 130$, respectively). The course covered the same topics of the Italian one. Participant ages ranged from 18 to 54 ($M = 20.12$, $SD = 3.81$, and median = 19), most of the participants were women (81.5%). This is the gender proportion of the population of psychology students in Spain. All students participated on a voluntary basis after they were given information about the general aim of the investigation (i.e., collecting information to improve students' statistics achievement).

Chiesi et al. 2011

Method

Participants

Participants were 197 undergraduates (79.2% female) in the James Cook University Psychology programs at the Singaporean (70.1%) and Australian (29.9 %) campuses. Their age ranged from 17 to 54 years ($M = 23.80$, $SD = 7.24$). Among these participants, 150 were currently enrolled in a statistics course (66.0% introductory statistics, 30.0% intermediate statistics, and 4.0% advanced statistics) whereas 47 have completed at least one of the aforementioned courses but were not currently enrolled in a statistics course.

Drew & Chillon, 2014

loading. The correlation between the two factors equaled .73. Reliability of the responses to all 20 items remained quite high with Cronbach's coefficient α equaling .95 ($n = 129$). Given the high correlation between the two factors and Spielberger's recommendation, all the items were summed to create a total score.

Coding Sheet Example

A	B	C	D	E	F	G	H	I	J
Study	**Source**	**Sub-groups**	**N**	**# of Items**	**R-Overall**	**R-Exam**	**R-Help**	**R-Interpretation**	**Language**
			ni	**mi**	**ai1**	**ai2**	**ai3**	**ai4**	
1	Cebollero et al (2012)		95	24	0.936	0.898	0.875	0.844	Spanish
2	Cendales et al (2013)		332	10	0.870				Spanish
3a	Chew & Dillon (2015)		204	24		0.890	0.900	0.890	English
3b	Chew & Dillon (2014)		197	24		0.900	0.950	0.880	English
4a	Chiesi et al (2011)	Italian sample	119	24	0.900	0.870	0.920	0.840	Italian
4b		Spanish sample	113	24	0.910	0.910	0.930	0.830	Spanish
5	Guàrdia Olmos et al (2012)		96	24		0.936	0.844	0.898	Spanish
6	Hernandez et al (2015)		397	24	0.920	0.910	0.920	0.810	Portuguese
7	Oliver et al (2014)		256	24		0.870	0.930	0.820	Spanish
8	Sesé et al (2015)		472	24		0.910	0.930	0.840	English
9	Vigil-Colet et al (2008)		159	24	0.910	0.870	0.920	0.820	English
10	Morsanyi, Primi, Handley, Chiesi, & Galli (2012)		105	24	0.880	0.830	0.920	0.830	Spanish
11	Justicia-Galiano et al (2015)		187	24	0.950				English
	Hamid, Shah & Sulaiman (2014)		342	24	0.884	0.82	0.883	0.78	English

A	Study Label	Year	Title
2	Howard & Michael (2019)	2019	Psychometri…
3	Lu et al., (2018)	2018	Psychometri…
4	Bell (2022)	2022	Social Desir…
5	McGrath et al., (2015)	2015	Reducing an…
7	Anonymous Unpublished	TBA	
8	Kaufmann et al., (2022)	2022	Self-efficacy …
9	Brash, M.	2020	Safety in Nu…
10	Hu (2021)	2021	The Impact …
11	Cendales et al., (2013)	2013	Psychologic…

Country	Cronbac...	Number ...	Sample S...
USA	0.920	14	128
USA	0.980	26	186
USA	0.980	26	218
Canada	0.910	15	28
USA	0.980	14	161
Germany	0.900	42	193
	0.907	14	
USA	0.960	15	87
Colombia	0.960	14	332

Type of P...	Database	Coder
Journal Article	Google Scho…	Samy
Journal Article	Google Scho…	Samy
Dissertation	Google Scho…	Samy
Journal Article	Google Scho…	Samy
Unpublished		Samy
		Samy
Dissertation	Google Scho…	Samy
Dissertation	Google Scho…	Samy
Journal Article	Google Scho…	Samy

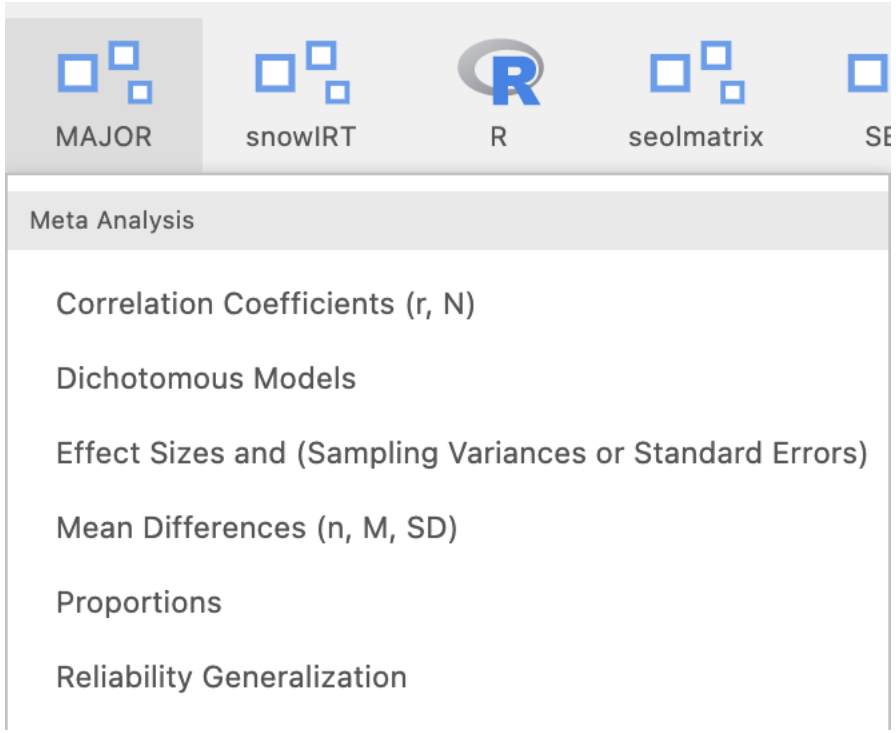
Issues in Meta-Analyses

- Published studies
- The drawer problem
- Fail Safe N

Conducting a Meta-Analysis

Software

- There is specialized software for meta-analysis:
  - Open Meta: https://osf.io/jx2td/wiki/Meta-Analysis%20Tools/
  - R, it's free and open source
    - Meta, metafor, meta-package
    - Too complicated for this presentation
  - Jamovi. Also free and open source.
    - Uses the same package, metafor.


MAJOR
snowIRT
R
seolmatrix
SE
Meta Analysis
Correlation Coefficients (r, N)
Dichotomous Models
Effect Sizes and (Sampling Variances or Standard Errors)
Mean Differences (n, M, SD)
Proportions
Reliability Generalization

Reliability Generalization

- Year
- Title
- G
- Language
- Country
- O
- P
- Age (Mean)
- Age (Standard Deviation)
- Age (Range)
- T

Cronbach's Alpha → Cronbachs Alpha

Number of Items → Number of Items

Sample Size → Sample Size

Study Label → Study Label

Moderator (optional) →

Model Options

Model estimator: Restricted Maximum-Likelihood

Model measures: Raw alpha values

Moderator type: No Moderator

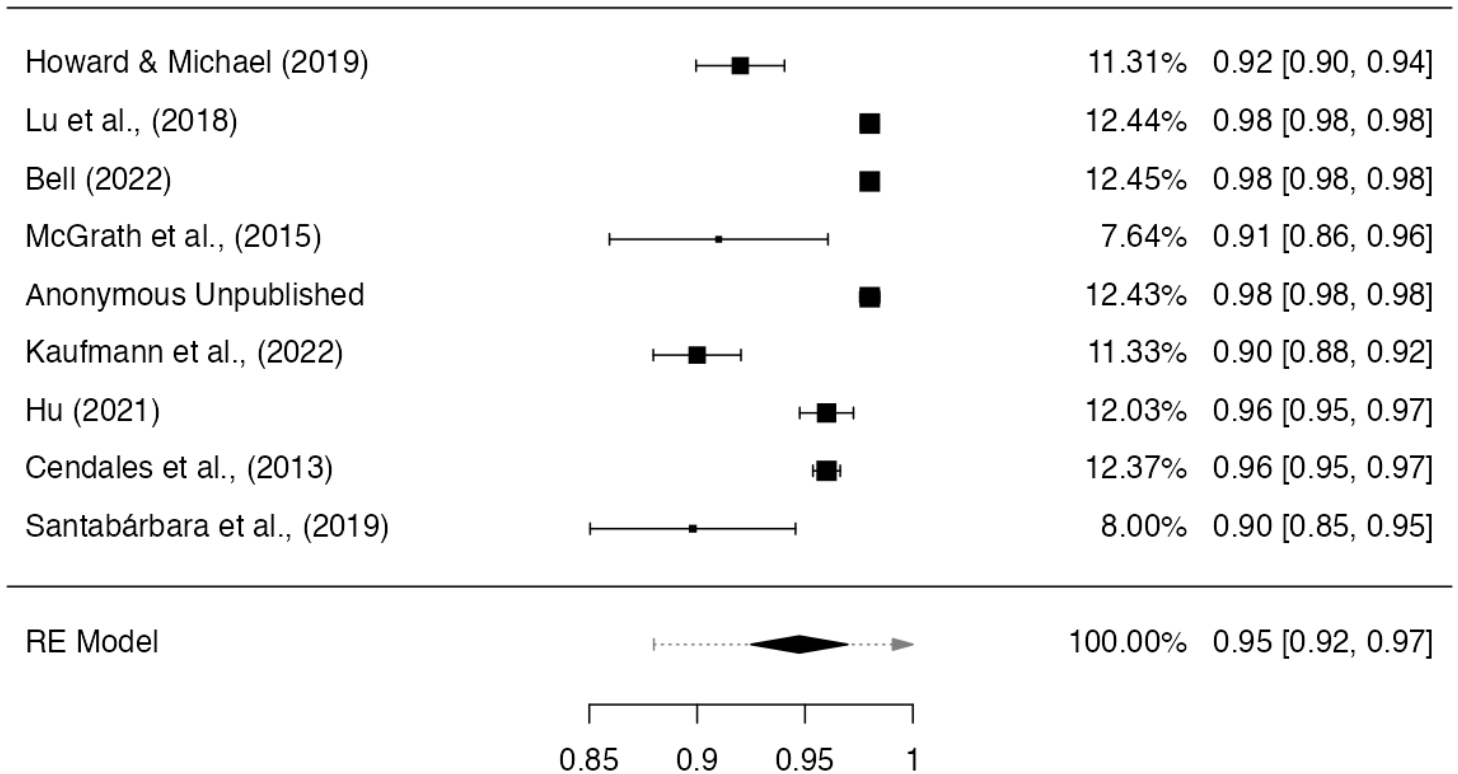
Confidence interval level: 95 %

- [x] Display model fit
- [x] Show Plot of Influence Diagnostics

Forest Plot

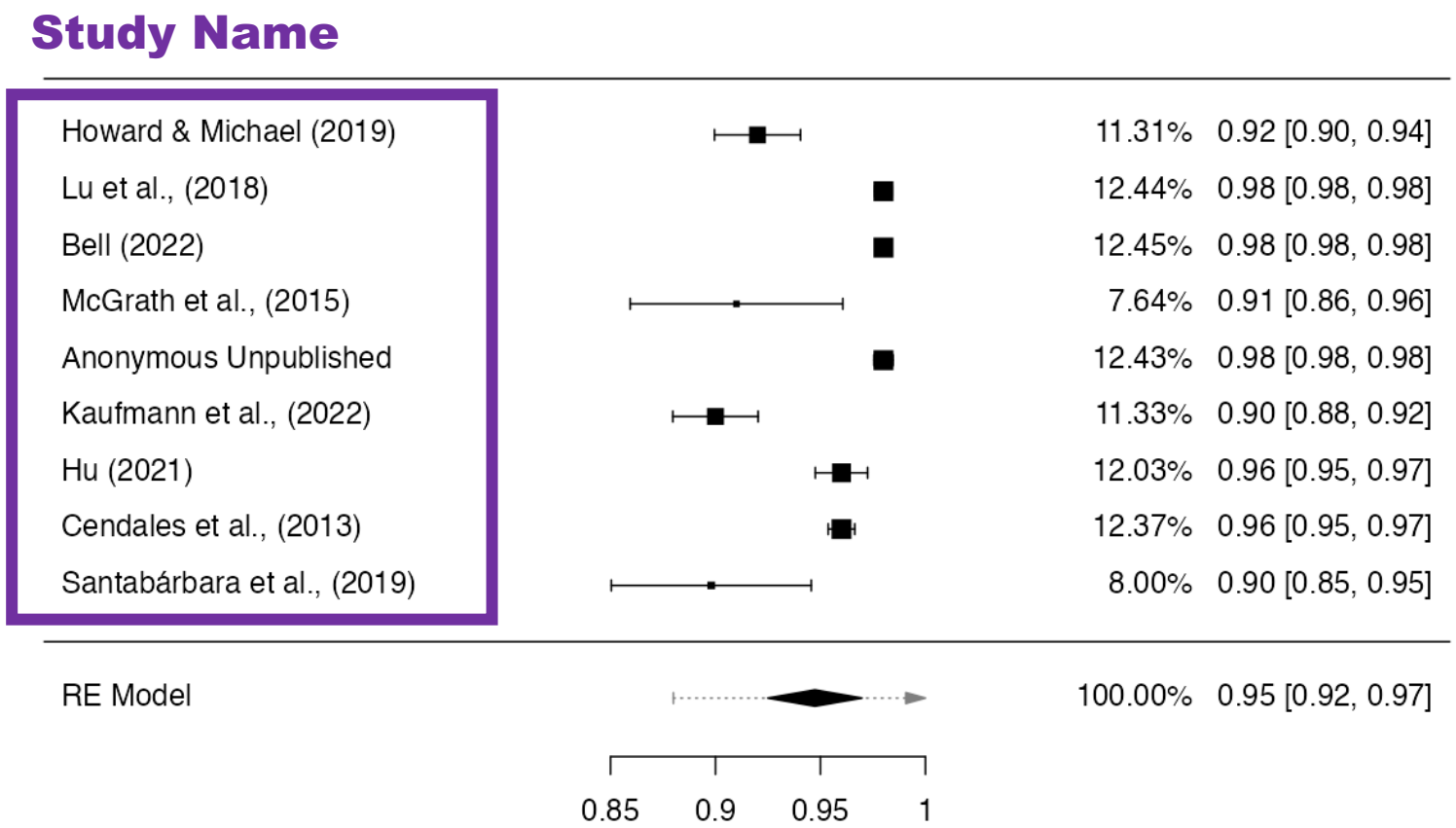
Forest Plot

[3]



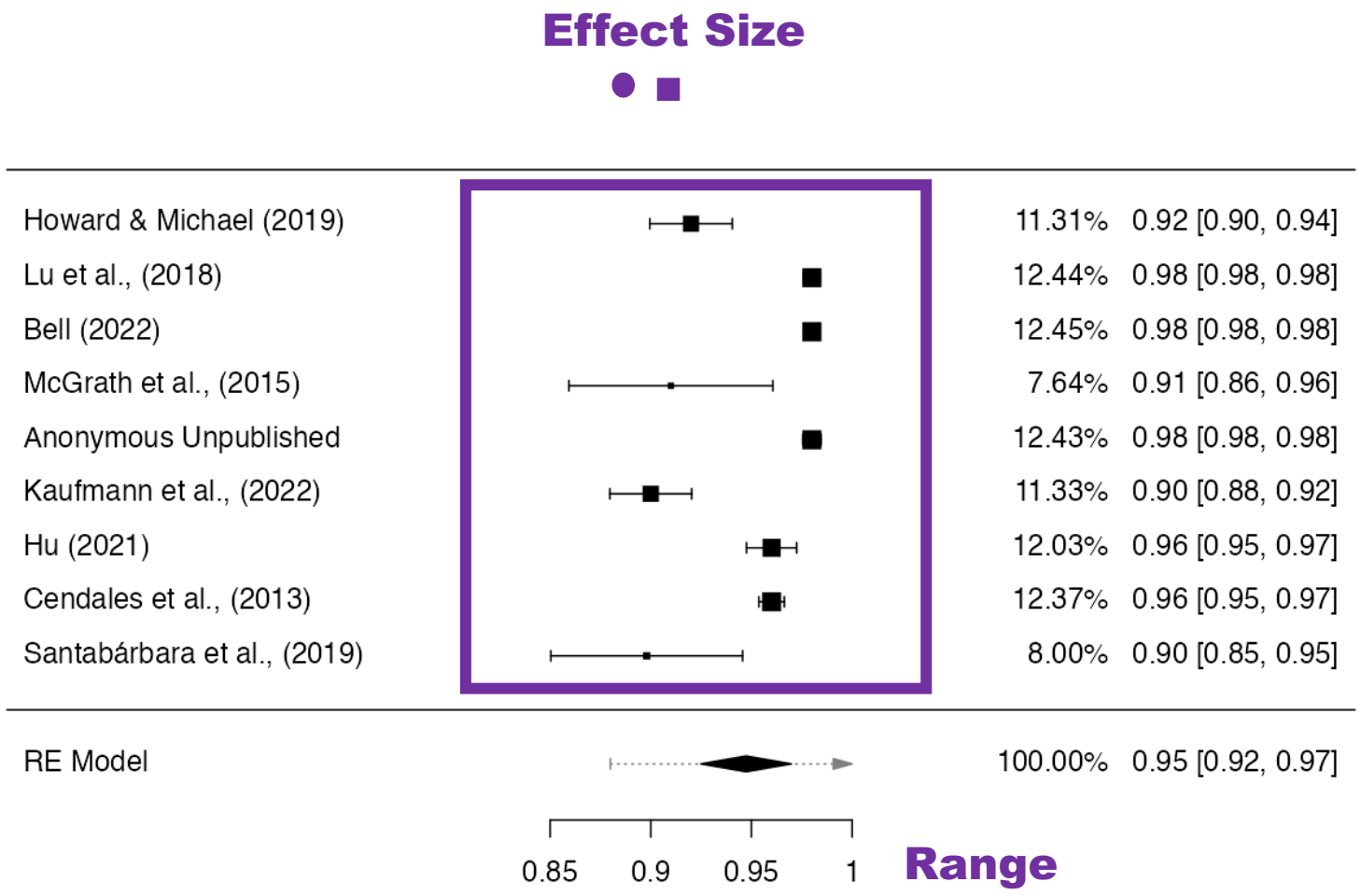
Forest Plot

[3]



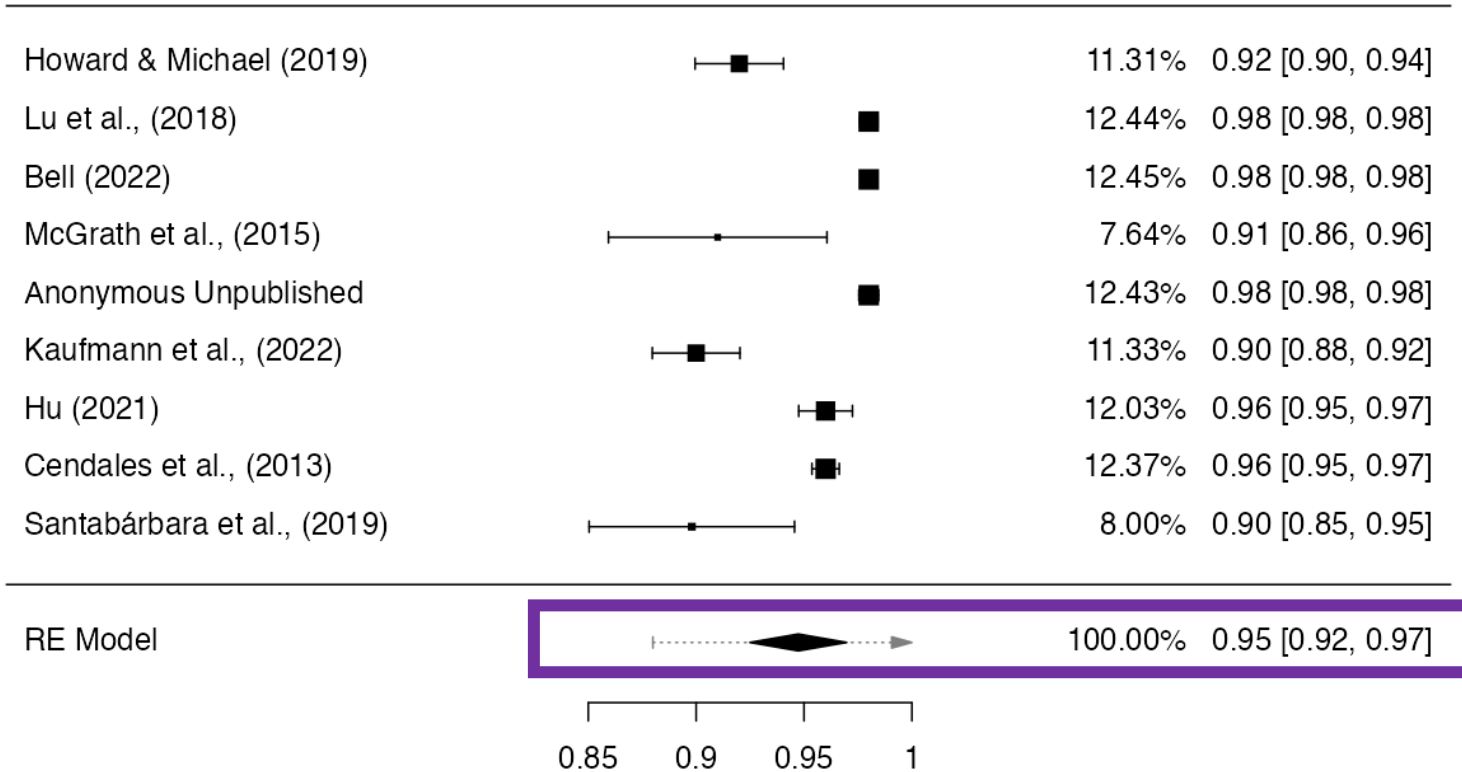
Forest Plot

[3]



Forest Plot

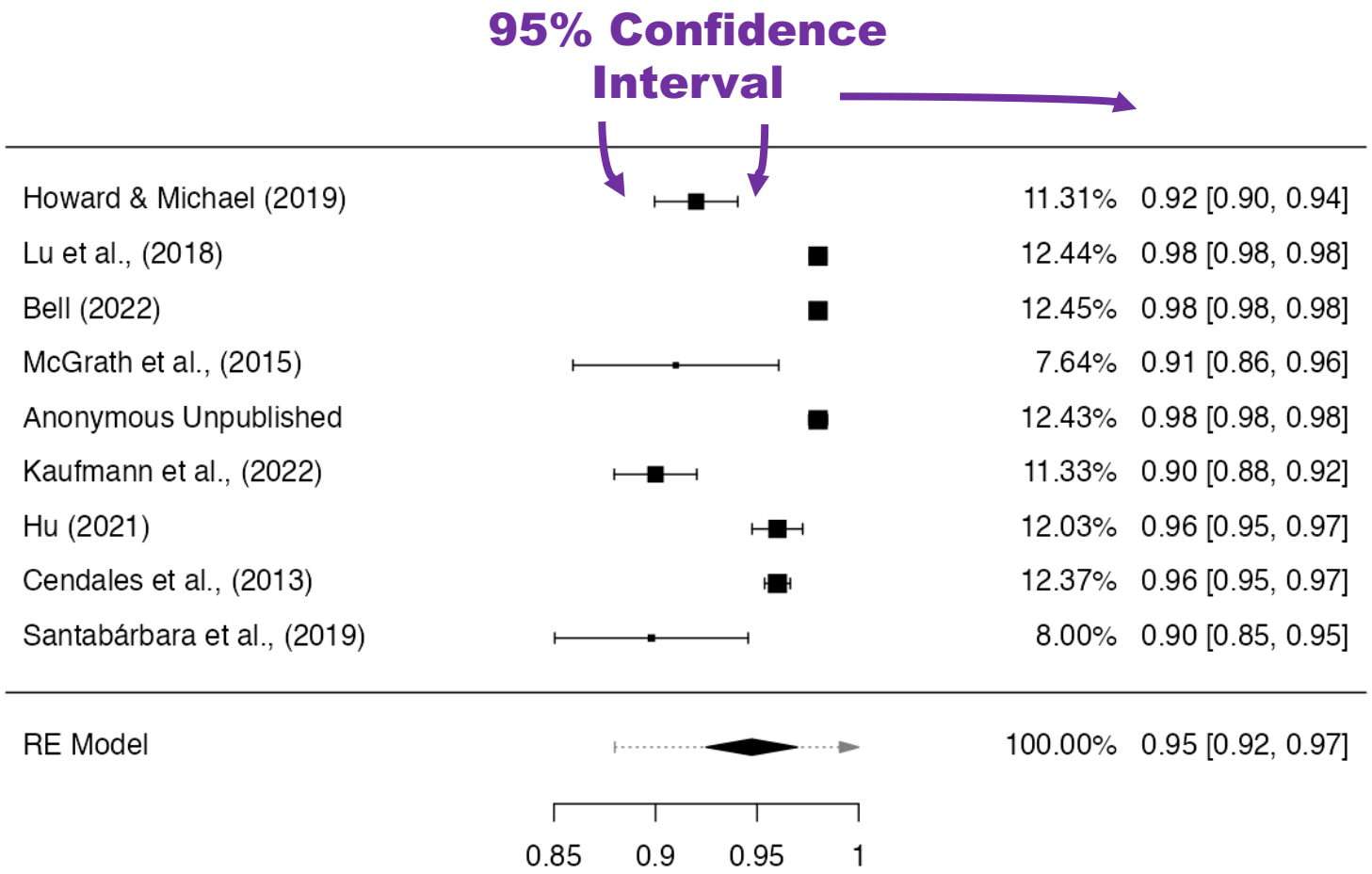
[3]



Pooled Effect

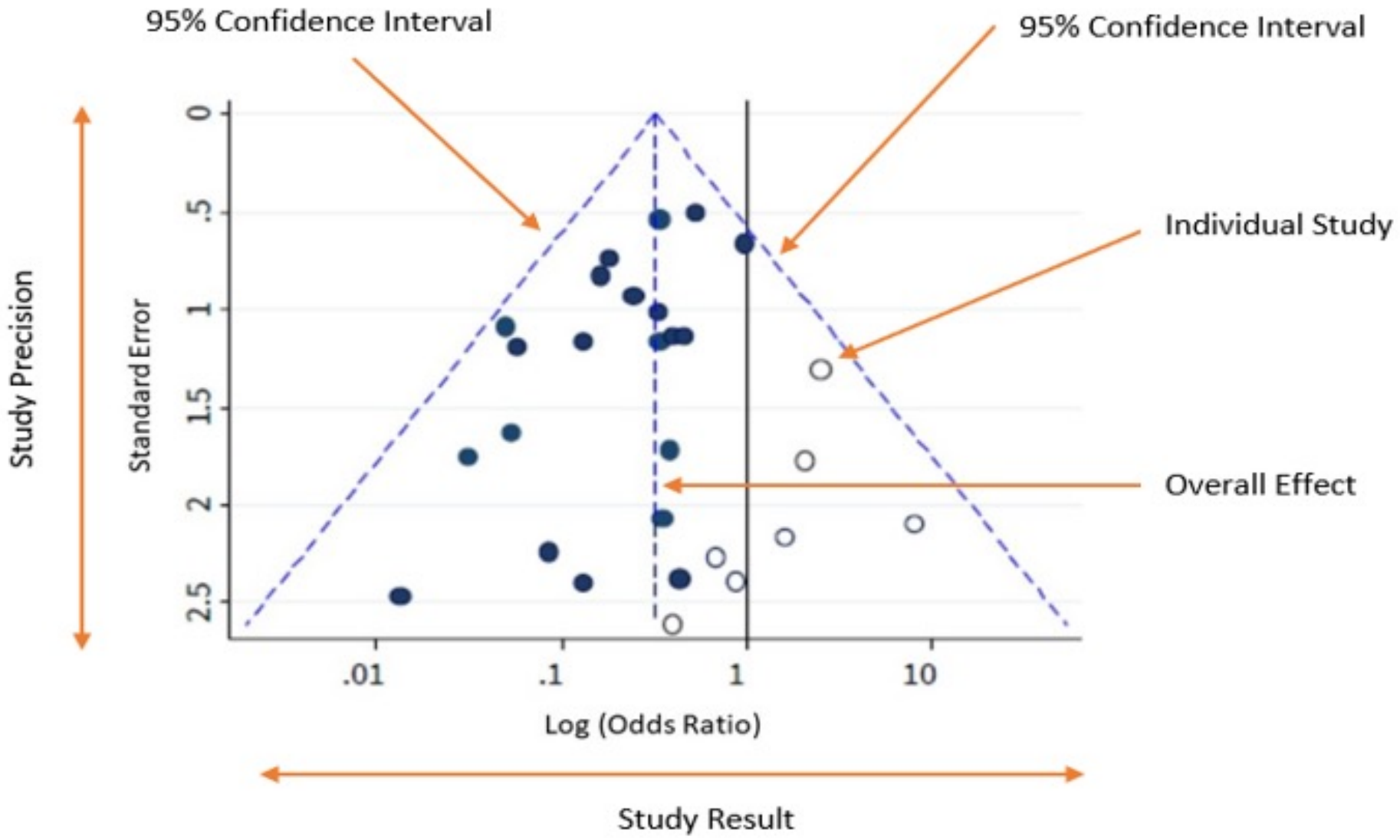
Forest Plot

[3]



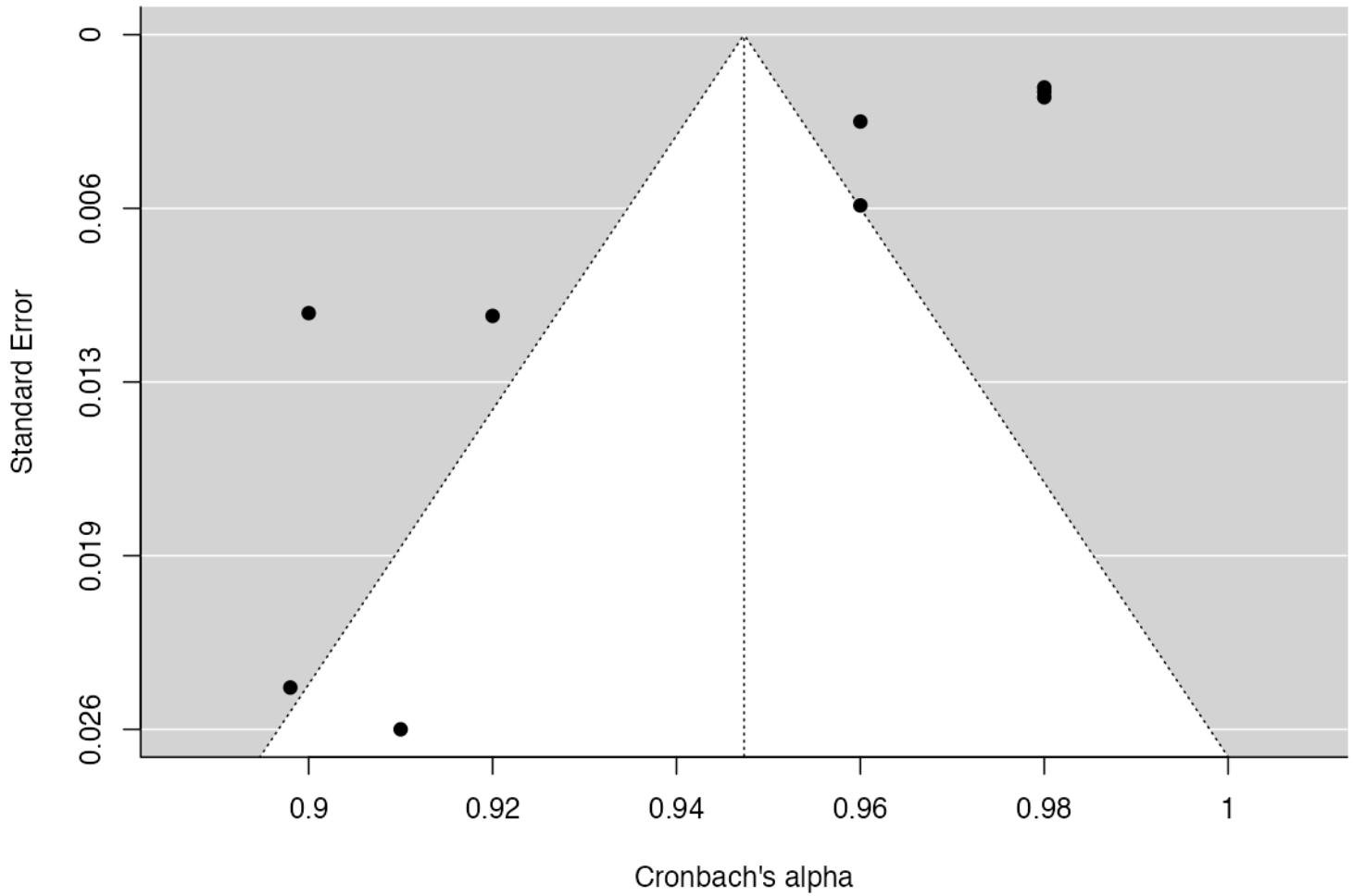
Funnel Plot

- A funnel plot tells you about the variability (standard error) of the individual studies against the mean effect size.
- As the study size increases the SE approaches zero.
- Assumes the plot should be symmetrical (that there are as many studies above / below the mean effect size)
  - Lack of symmetry can suggest publication bias, or ”small study” bias


95% Confidence Interval
95% Confidence Interval
Individual Study
Overall Effect
Study Precision
Standard Error
0
.5
1
1.5
2
2.5
.01
.1
1
10
Log (Odds Ratio)
Study Result

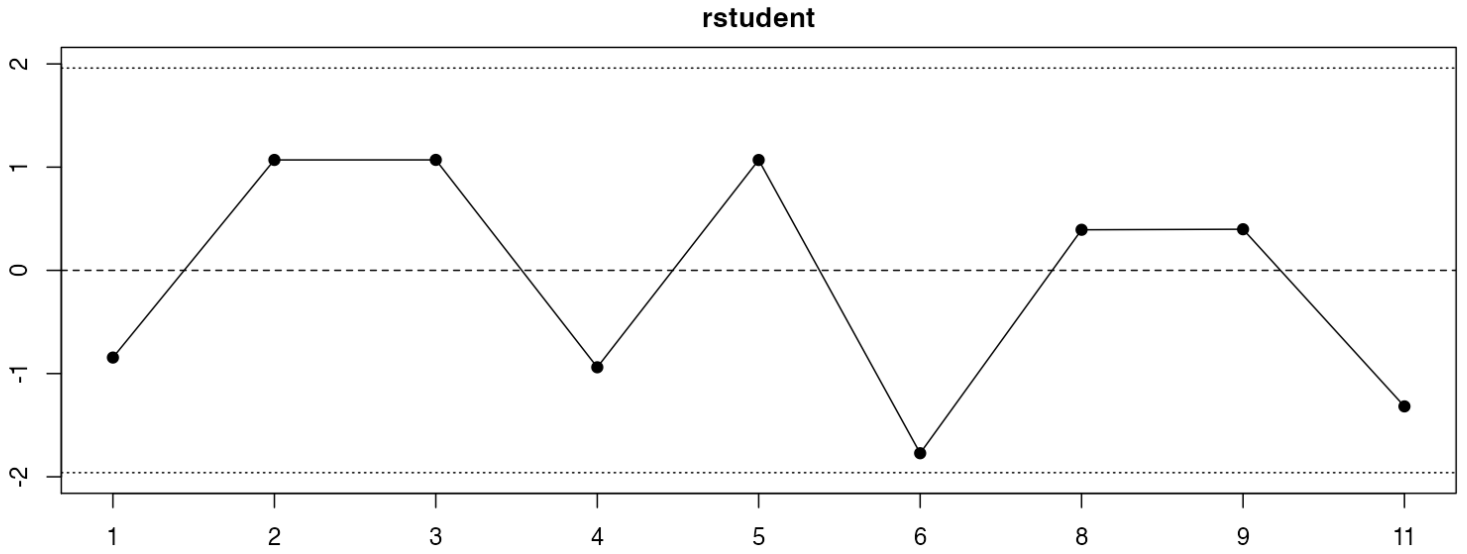
Funnel Plot

[3]

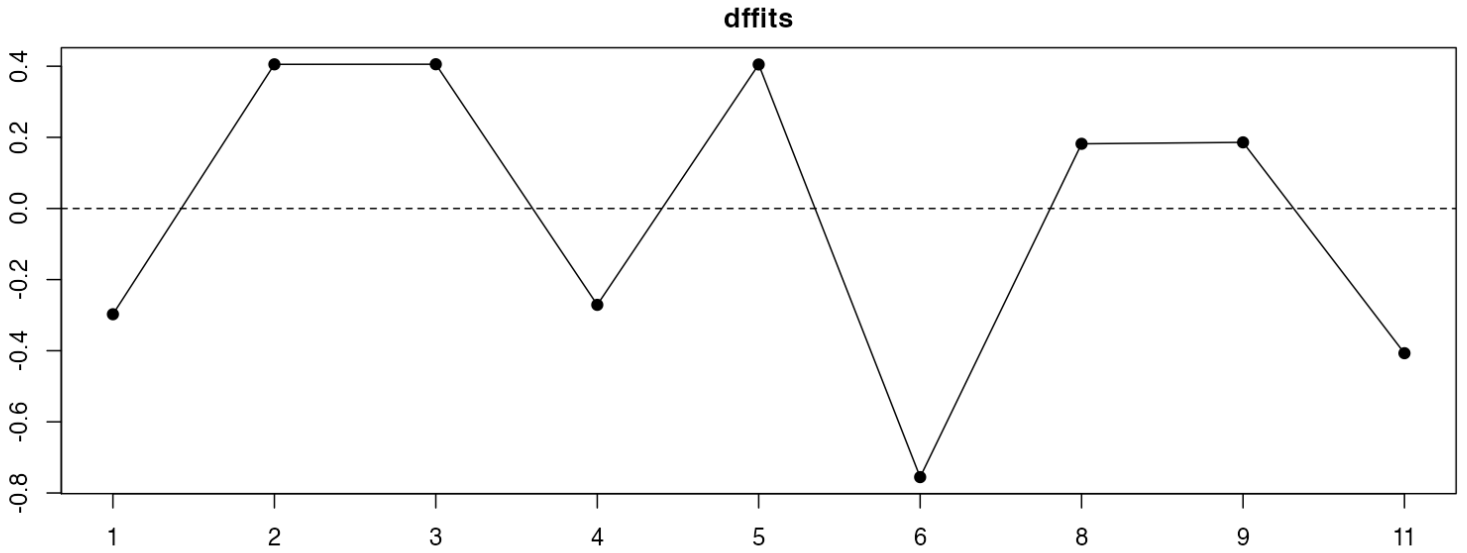


Outlier and Influential Case Diagnostics

Externally Standardized Residual

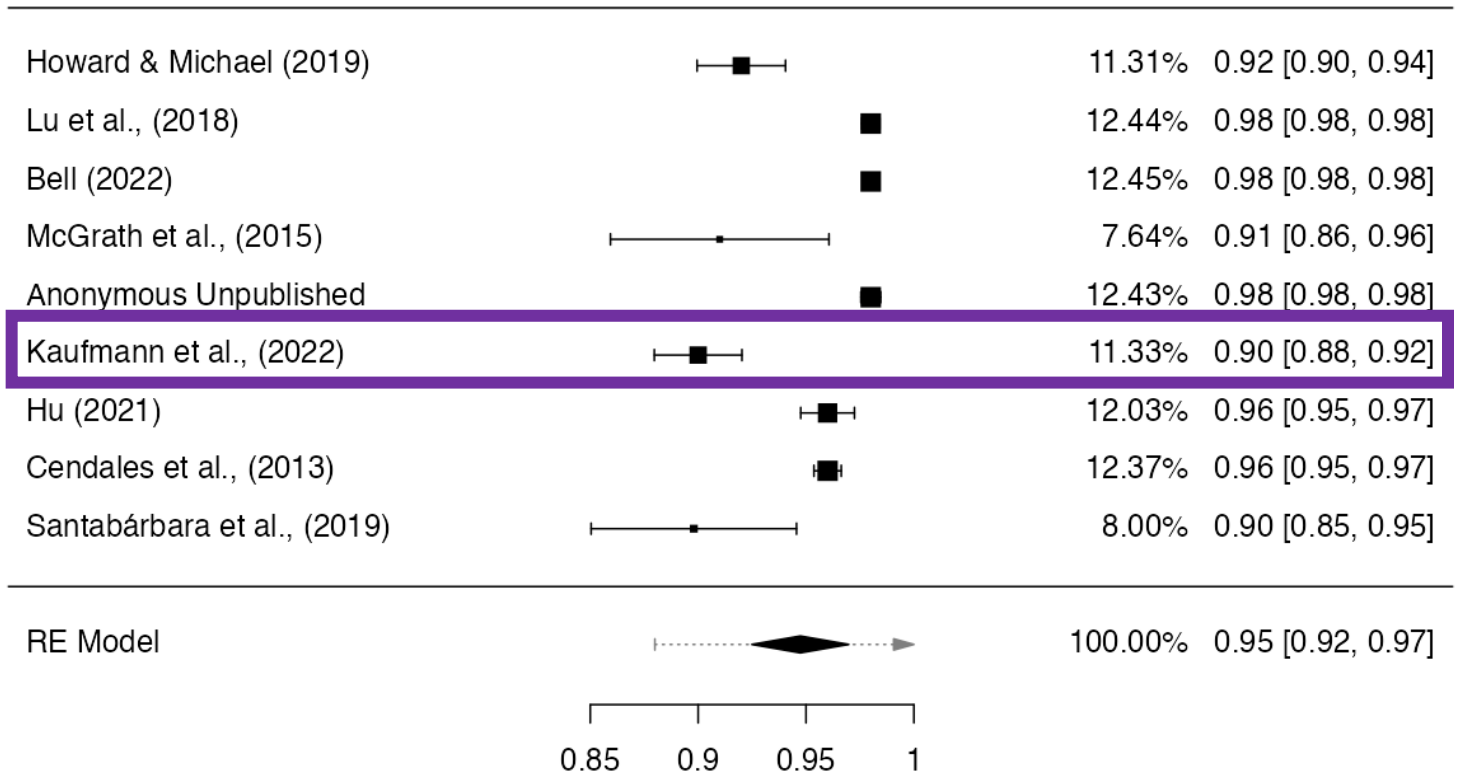


DFFITS Values



Forest Plot

[3]



Heterogeneity Statistics

Tau	Tau²	I²	H²	R²	df	Q	p
0.032	0.0011 (SE= 6e-04)	98.77%	81.046	.	8.000	140.220	<.001

- Cochran's Q
  - Classical measure of heterogeneity
  - The underlying null hypothesis assumes that the true treatment effect is the same across studies and variations are simply caused by chance.
- I^2 statistic
  - The percentage of variation across studies that is due to heterogeneity rather than chance
  - I^2 is an intuitive and simple expression of the inconsistency of studies' results.

Fixed Effects vs Random Effects

- **Fixed Effects**
  - conduct if it is reasonable to assume underlying effect size is SAME for all studies
- Test: test of heterogeneity
  - Pooling
  - If significant, go for random effects model
- If there is very little variation between trials then I^2 will be low and a fixed effects model might be appropriate.

- **Random Effects**
  - Conduct if test of heterogeneity is significant.
  - Q: $p < .05$
- Assumes outcome comes from a normal distribution
- More practical

Random-Effects Model (k = 9)

	Estimate	se	Z	p	CI Lower Bound	CI Upper Bound
Intercept	0.948	0.0107	88.3	<.001	0.927	0.969
	.	.	.	.	.	.

Note. Tau² Estimator: Maximum-Likelihood

Resources

- Covidence Resource: https://libguides.uttyler.edu/c.php?g=1341980&p=9993128
- Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: a practical tutorial. *BMJ Ment Health*, *22*(4), 153-160.
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis*. John Wiley & Sons.
- Del Re, A. C. (2015). A practical tutorial on conducting meta-analysis in R. *The Quantitative Methods for Psychology*, *11*(1), 37-50.

Selected References

- Sánchez-Meca, J., Marín-Martínez, F., López-López, J. A., Núñez-Núñez, R. M., Rubio-Aparicio, M., López-García, J. J., ... & López-Nicolás, R. (2021). Improving the reporting quality of reliability generalization meta-analyses: The REGEMA checklist. *Research Synthesis Methods*, *12*(4), 516-536.
- Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., ... & Higgins, J. P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *Bmj*, *343*.
- West SL, Gartlehner G, Mansfield AJ, et al. Comparative Effectiveness Review Methods: Clinical Heterogeneity [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2010 Sep. Table 7, Summary of common statistical approaches to test for heterogeneity. Available from: https://www.ncbi.nlm.nih.gov/books/NBK53317/table/ch3.t2/#