

Using STATA® for Meta-analysis

Peter Francis Raguindin

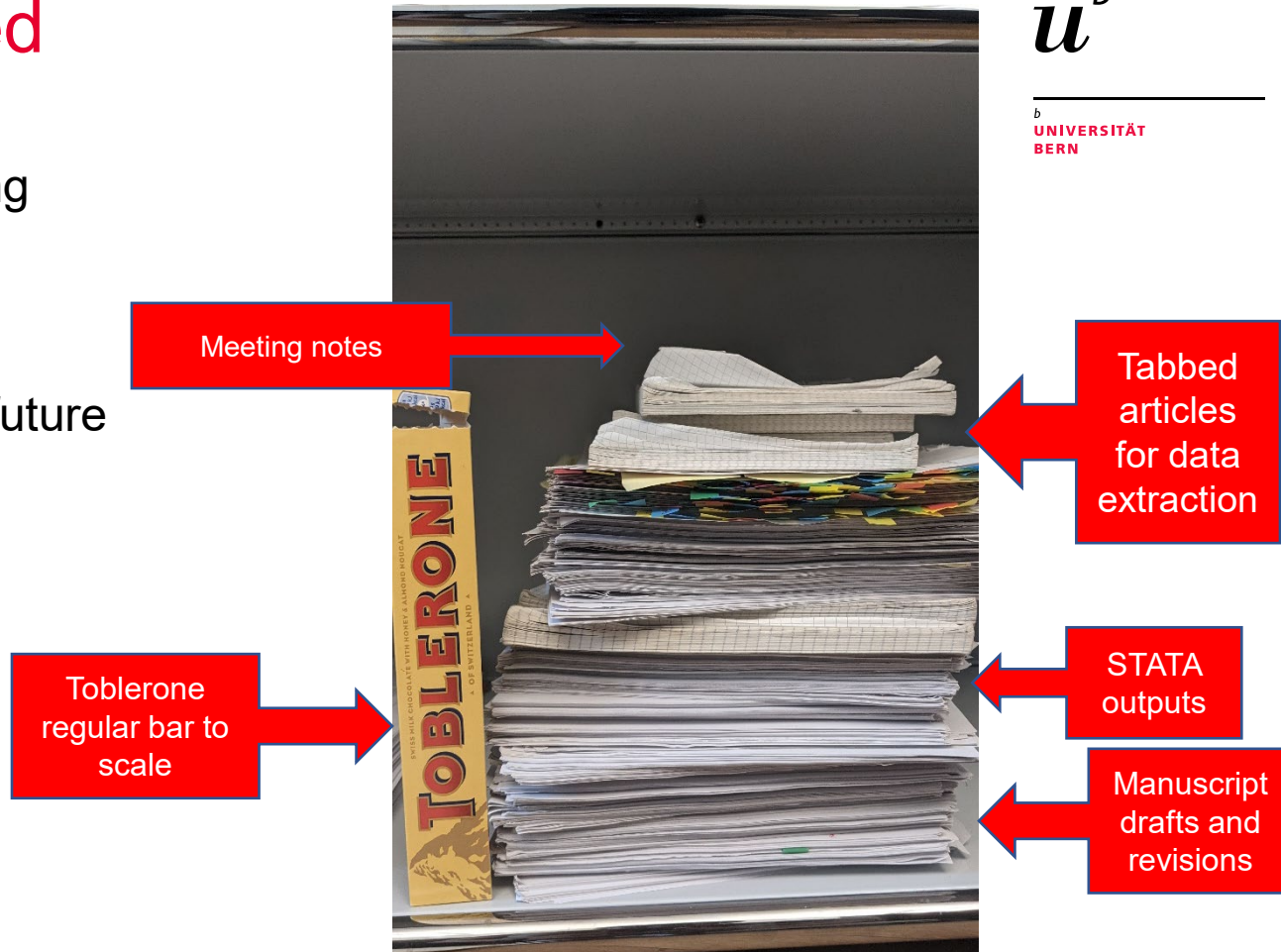
Institute of Social and Preventive Medicine, Bern, Switzerland
Swiss Paraplegic Research, Nottwil, Switzerland

peter.raguindin@ispm.unibe.ch

<https://stata.ispm.ch/statainstallation/>

Lessons learned

- Extremely challenging
- Process oriented
- Organization is key
- Helps develop your future research ideas



A statistical analysis which combines the results of several independent studies considered by the analyst to be “combinable”

Stat Methods Med Res. 2012 Aug;21(4):409-26

- You will learn how to run meta-analysis on STATA (basic)
 - To structure your dataset
 - To install meta-analysis packages
 - To import and describe dataset
 - To compute for pooled estimates
- You will interpret the results and apply basic concepts of fixed-effect and random effect

Features	SPSS	SAS	Stata	JMP (SAS)	R	Python (Pandas)
Learning curve	Gradual	Pretty steep	Gradual	Gradual	Pretty steep	Steep
User interface	Point-and-click	Programming	Programming/ point-and-click	Point-and-click	Programming	Programming
Data manipulation	Strong	Very strong	Strong	Strong	Very strong	Strong
Data analysis	Very strong	Very strong	Very strong	Strong	Very strong	Strong
Graphics	Good	Good	Very good	Very good	Excellent	Good
Cost	Expensive (perpetual, cost only with new version). Student disc.	Expensive (yearly renewal) Free student version, 2014	Affordable (perpetual, cost only with new version). Student disc.	Expensive (yearly renewal) Student disc.	Open source (free)	Open source (free)
Released	1968	1972	1985	1989	1995	2008

Stata 15/16+ screen

Variables in dataset here

History of commands, this window

Output here

Variables in dataset here

Property of each variable here

Write commands here

Files will be saved here

?????

Notes:

1. (/v# option or -set maxvar-) 5000 maximum variables

```
. cd H:
log using mywork.log
import excel "http://dss.princeton.edu/training/mydata.xls", sheet("Sheet1") firstrow clear
summarize
```

Variable	Obs	Mean	Std. Dev.	Min	Max
Year	0				
CountryName	0				
GDPpercapa-200	4542	9402.967	11265.24	161.5976	76319.47
Unemployme-e	4521	.0478866	.0724682	0	.686
Unemployme-b	4521	.0366029	.0544155	0	.546
Unemployme-l	4521	.0425112	.0601523	0	.595
Exportsofg-o	3661	6.49e+10	1.64e+11	4.50e+07	1.78e+12
Importsog-o	3661	6.43e+10	1.74e+11	9.42e+07	2.20e+12
polityorig-l	4542	-.2573756	16.20321	-08	10
polity2adj-d	4498	2.409738	7.03114	-10	10

log on (text)

Command

Properties

Variables

Name	Year
Label	Year
Type	str109
Format	%109s
Value Label	
Notes	

Data

Filename

Label

Notes

Variables

10

Observations

4,546

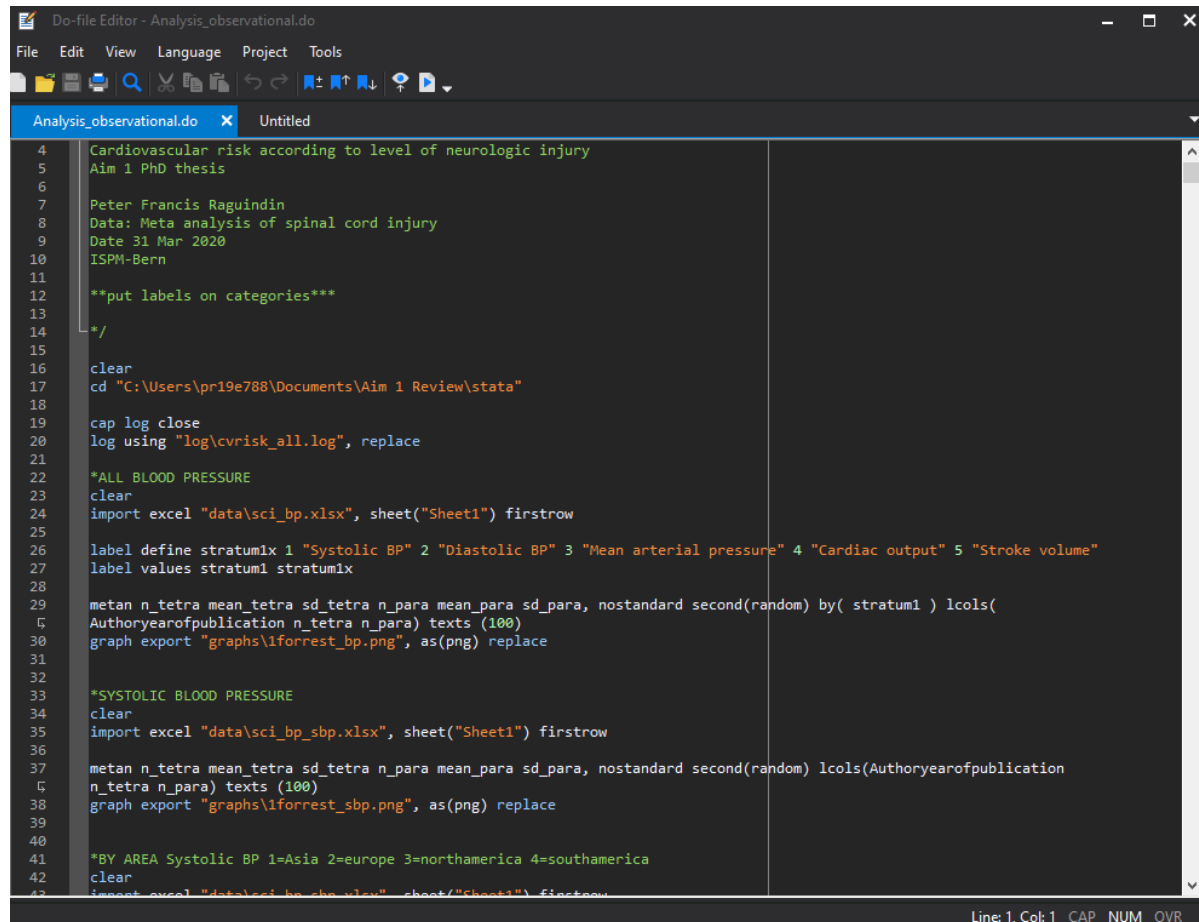
Size

812.42K

Memory

32M

Sorted by

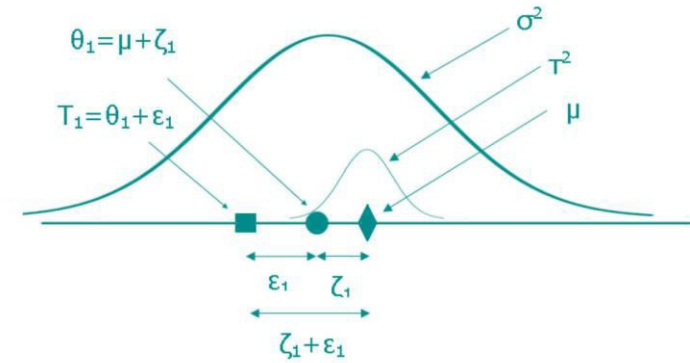
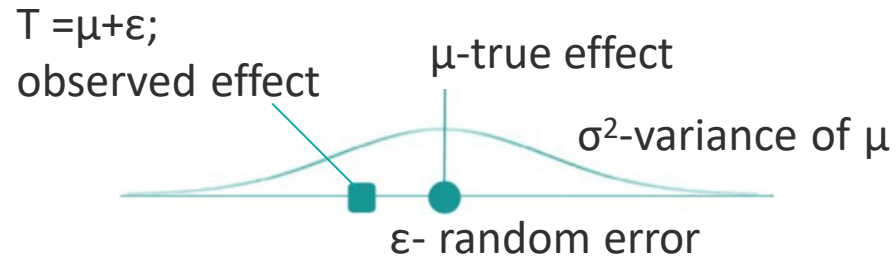


The screenshot shows a Stata Do-file Editor window titled "Do-file Editor - Analysis_observational.do". The window has a menu bar (File, Edit, View, Language, Project, Tools) and a toolbar. The main editor area displays a Stata script with the following content:

```
4 Cardiovascular risk according to level of neurologic injury
5 Aim 1 PhD thesis
6
7 Peter Francis Raguindin
8 Data: Meta analysis of spinal cord injury
9 Date 31 Mar 2020
10 ISPM-Bern
11
12 **put labels on categories**
13
14 */
15
16 clear
17 cd "C:\Users\pr19e788\Documents\Aim 1 Review\stata"
18
19 cap log close
20 log using "log\cvrisk_all.log", replace
21
22 *ALL BLOOD PRESSURE
23 clear
24 import excel "data\sci_bp.xlsx", sheet("Sheet1") firstrow
25
26 label define stratum1x 1 "Systolic BP" 2 "Diastolic BP" 3 "Mean arterial pressure" 4 "Cardiac output" 5 "Stroke volume"
27 label values stratum1 stratum1x
28
29 metan n_tetra mean_tetra sd_tetra n_para mean_para sd_para, nostandard second(random) by( stratum1 ) lcols(
30   Authorityyearofpublication n_tetra n_para) texts (100)
31 graph export "graphs\1forrest_bp.png", as(png) replace
32
33 *SYSTOLIC BLOOD PRESSURE
34 clear
35 import excel "data\sci_bp_sbp.xlsx", sheet("Sheet1") firstrow
36
37 metan n_tetra mean_tetra sd_tetra n_para mean_para sd_para, nostandard second(random) lcols(Authorityyearofpublication
38   n_tetra n_para) texts (100)
39 graph export "graphs\1forrest_sbp.png", as(png) replace
40
41 *BY AREA Systolic BP 1=Asia 2=europe 3=northamerica 4=southamerica
42 clear
43 import excel "data\sci_bp_sbp.xlsx", sheet("Sheet1") firstrow
```

The status bar at the bottom indicates "Line: 1, Col: 1 CAP NUM OVR".

Fixed vs. random effect



Fixed effects vs Random effects

$$W = \frac{1}{V_i}$$

Fixed effect

$$W = \frac{1}{V_i + \tau}$$

Random effects

New Meta-analysis Interface

Setup

Summary

Forest plot

Heterogeneity

Regression

Publication bias

Multivariate

Display meta settings

Modify meta settings

Note: Multivariate meta-analysis does not require any setup. Proceed to the Multivariate pane.

Declare meta-analysis data

☒ Compute and declare effect sizes for two-group comparison of continuous outcomes

☐ Compute and declare effect sizes for two-group comparison of binary outcomes

☐ Declare generic, precomputed effect sizes (in the metric closest to normality)

Main

if/in

Model

Options

Specify group 1 (treatment) variables

Sample size:

n_tetra

Mean:

mean_tetra

Standard deviation:

sd_tetra

Specify group 2 (control) variables

Sample size:

n_para

Mean:

mean_para

Standard deviation:

sd_para

Specify effect size

Effect size:

Mean difference

☒ Assume unequal group variances

Submit

No. of studies: 19
CI level: 95%

Model: Random effects
Method: DerSimonian-Laird

Effect size: _meta_es, Blood pressure dif~e
Std. Error: _meta_se

- Set-up
- Adjust the setting of your data
- Sets the type of analysis
- Cohen and Hedges (standardized)
- Glass delta (standardized with unequal variance)

New Meta-analysis Interface

meta - Meta-Analysis Control Panel

Display meta settings Modify meta settings

Setup

Summary

Forest plot

Heterogeneity

Regression

Publication bias

Multivariate

Note: Multivariate meta-analysis does not require any setup. Proceed to the Multivariate pane.

Declare meta-analysis data

☒ Compute and declare effect sizes for two-group comparison of continuous outcomes

☐ Compute and declare effect sizes for two-group comparison of binary outcomes

☐ Declare generic, precomputed effect sizes (in the metric closest to normality)

Main if/in Model Options

Meta-analysis model

☒ Random effects

☐ Common effect

☐ Fixed effects

Method: DerSimonian-Laird

Submit

No. of studies: 19

CI level: 95%

Model: Random effects

Method: DerSimonian-Laird

Effect size: _meta_es, Blood pressure dif~e

Std. Error: _meta_se

Close

- Choosing your model

New Meta-analysis Interface

meta - Meta-Analysis Control Panel

Display meta settings Modify meta settings

Setup

Summary

Forest plot

Heterogeneity

Regression

Publication bias

Multivariate

Note: Multivariate meta-analysis does not require any setup. Proceed to the Multivariate pane.

Declare meta-analysis data

☒ Compute and declare effect sizes for two-group comparison of continuous outcomes

☐ Compute and declare effect sizes for two-group comparison of binary outcomes

☐ Declare generic, precomputed effect sizes (in the metric closest to normality)

Main if/in Model Options

Meta-analysis model

☒ Random effects

☐ Common effect

☐ Fixed effects

Method: DerSimonian-Laird

Submit

No. of studies: 19

CI level: 95%

Model: Random effects

Method: DerSimonian-Laird

Effect size: _meta_es, Blood pressure dif~e

Std. Error: _meta_se

Close

- Choosing your model

REVIEW ARTICLE

OPEN



The neurological level of spinal cord injury and cardiovascular risk factors: a systematic review and meta-analysis

Peter Francis Raguindin^{1,2,3}, Gion Fränkl^{1,4,6}, Oche Adam Itodo^{1,2,3,6}, Alessandro Bertolo¹, Ramona Maria Zeh¹, Simona Capossela¹, Beatrice Minder⁵, Jivko Stoyanov¹, Gerold Stucki¹, Oscar H. Franco^{1,2}, Taulant Muka² and Marija Glisic^{1,2}✉

© The Author(s) 2021

STUDY DESIGN: Systematic review and meta-analysis.

OBJECTIVE: To determine the difference in cardiovascular risk factors (blood pressure, lipid profile, and markers of glucose metabolism and inflammation) according to the neurological level of spinal cord injury (SCI).

METHODS: We searched 5 electronic databases from inception until July 4, 2020. Data were extracted by two independent reviewers using a pre-defined data collection form. The pooled effect estimate was computed using random-effects models, and heterogeneity was calculated using I^2 statistic and chi-squared test (CRD42020166162).

RESULTS: We screened 4863 abstracts, of which 47 studies with 3878 participants (3280 males, 526 females, 72 sex unknown) were included in the meta-analysis. Compared to paraplegia, individuals with tetraplegia had lower systolic and diastolic blood pressure (unadjusted weighted mean difference, -14.5 mmHg, 95% CI -19.2 , -9.9 ; -7.0 mmHg 95% CI -9.2 , -4.8 , respectively), lower triglycerides (-10.9 mg/dL, 95% CI -19.7 , -2.1), total cholesterol (-9.9 mg/dL, 95% CI -14.5 , -5.4), high-density lipoprotein (-1.7 mg/dL, 95% CI -3.3 , -0.2) and low-density lipoprotein (-5.8 mg/dL, 95% CI -9.0 , -2.5). Comparing individuals with high- vs. low-thoracic SCI, persons with higher injury had lower systolic and diastolic blood pressure (-10.3 mmHg, 95% CI -13.4 , -7.1 ; -5.3 mmHg 95% CI -7.5 , -3.2 , respectively), while no differences were found for low-density lipoprotein, serum glucose, insulin, and inflammation markers. High heterogeneity was partially explained by age, prevalent cardiovascular diseases and medication use, body mass index, sample size, and quality of studies.

CONCLUSION: In SCI individuals, the level of injury may be an additional non-modifiable cardiovascular risk factor. Future well-designed longitudinal studies with sufficient follow-up and providing sex-stratified analyses should confirm our findings and explore the role of SCI level in cardiovascular health and overall prognosis and survival.

Spinal Cord; <https://doi.org/10.1038/s41393-021-00678-6>

Sample Study

$$\mathcal{U}^b$$
b
UNIVERSITÄT[illegible]

Exercise 1: Blood pressure study

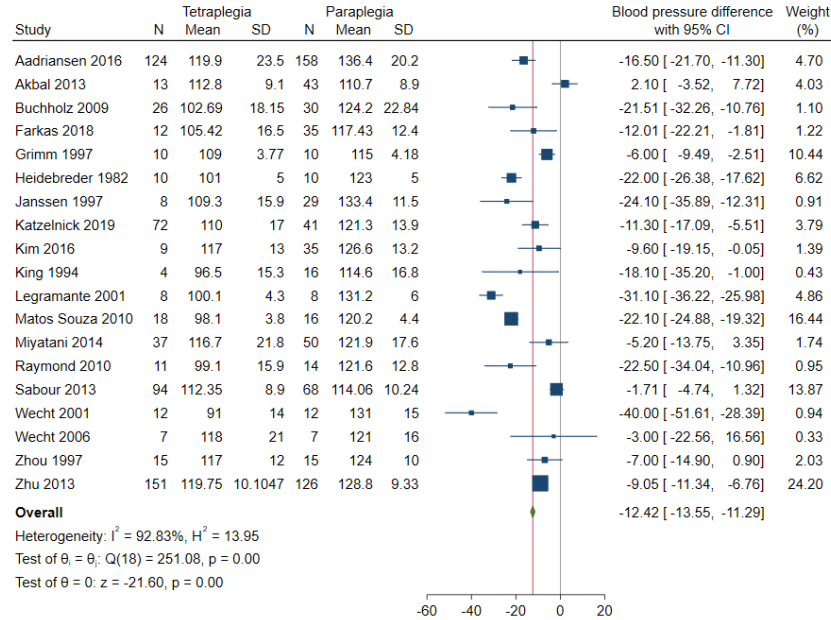
Is there difference in the **blood pressure** of individuals with spinal cord injury according to the **different levels of injury**?

- Meta-analysis of observational studies
 - *Population*: Individuals with Spinal cord injury
 - *Exposure*: Spinal cord injury levels (tetraplegia and paraplegia)
 - *Outcome*: Blood pressure
-
- Data type: CONTINUOUS
 - Data: `sci_bp_sbp.xlsx`

Guide Questions

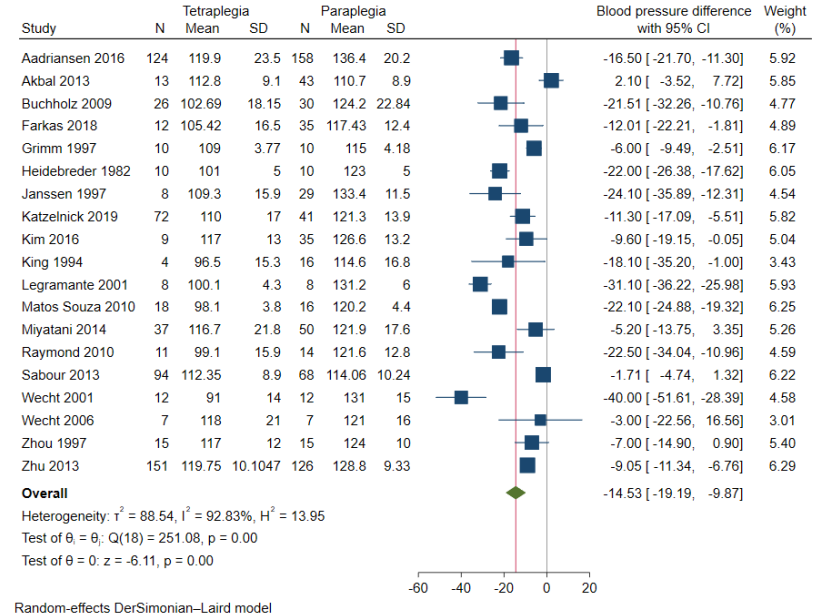
- Do preliminary dataset check
- Is there any difference between the blood pressure of tetraplegia and paraplegia?
 - Compute for mean difference using fixed effect and random effects model
 - Compute for standardized mean difference fixed effect and random effects model
 - Export your forest plot
- Who has higher systolic blood pressure?
- Bonus question: What are the mean difference according to the location of the study
 - Do a subgroup analysis according to the area
- Bonus question: Is there publication bias?
 - Do a funnel plot

Solution



Fixed-effects inverse-variance model

Fixed-effects (WMD)



Random-effects DerSimonian-Laird model

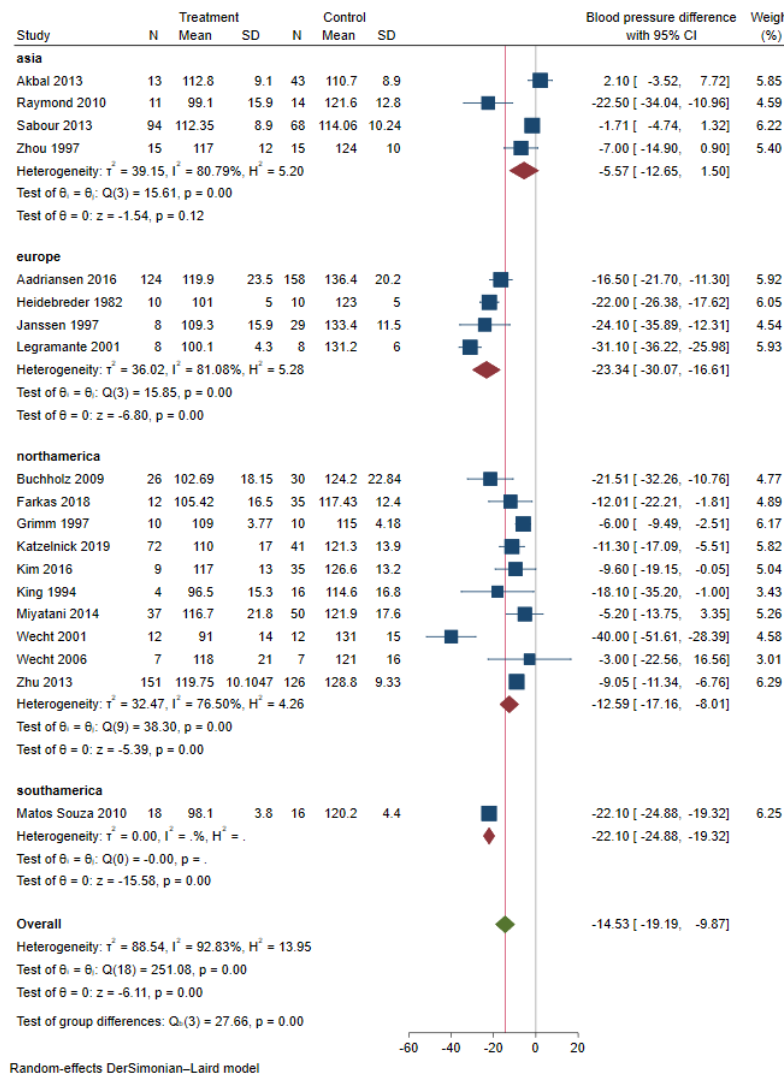
Random effects (WMD)

Test	Results	Interpretation
Test for heterogeneity		
Cochran's Q	$p < 0.05$	There is heterogeneity
I ²	75% to 100%	Considerable heterogeneity
Tau and Tau ²	-	Describes the spread of between study effects
H ²	>1	Ratio (multiplicative factor) There is heterogeneity
Publication bias (small study effects)		
Funnel plot	Assymetrical	There is publication bias (small studies do not mimic large studies)
Egger's test	$p < 0.05$	There is publication bias (small studies do not mimic large studies)

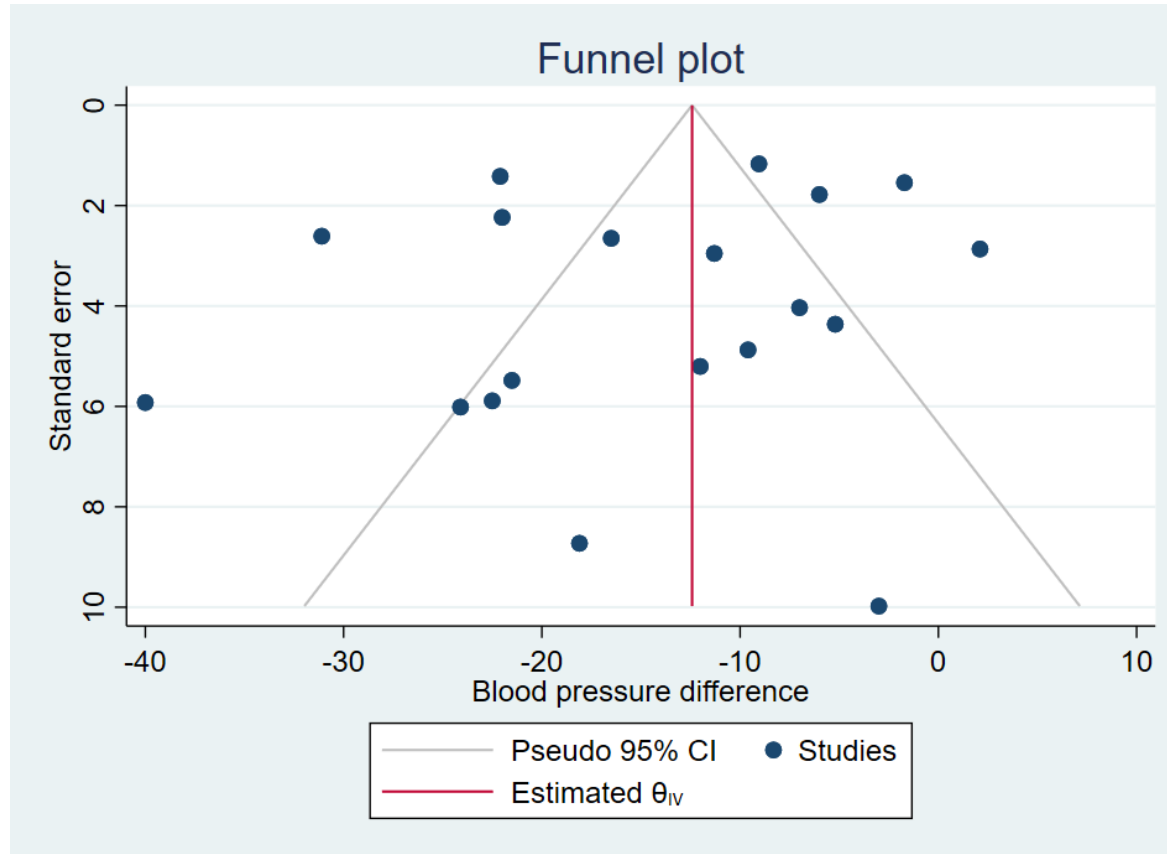
Solution

u^b

^b
UNIVERSITÄT
BERN



Solution



Exercise 2: Streptokinase study

What is the effect of intravenous **streptokinase** on early **mortality** in patients with acute myocardial infarction?

- Meta-analysis of randomized controlled trials
 - *Population*: Adults diagnosed with acute myocardial infarction
 - *Intervention*: intravenous streptokinase
 - *Control*: no streptokinase
 - *Outcome*: mortality
-
- Data type: BINARY or DICHOTOMOUS
 - Data: `strepto.dta`

- Do preliminary dataset check
- Using risk ratio, what is the risk difference between streptokinase and control?
 - WARNING! The command requires to put individuals with and without the outcome.
 - Compute for the risk ratio for each study and pooled effect
 - Compute using fixed effect model using Inverse variance
 - Random effects model using Inverse variance
 - Is streptokinase beneficial?
- Bonus questions
 - Does year of conduct of trials change our results?
 - Does one study influence the whole results?

Solution

meta - Meta-Analysis Control Panel

Display meta settings Modify meta settings

Setup

Summary

Forest plot

Heterogeneity

Regression

Publication bias

Multivariate

Note: Multivariate meta-analysis does not require any setup. Proceed to the Multivariate pane.

Declare meta-analysis data

☐ Compute and declare effect sizes for two-group comparison of continuous outcomes

☒ Compute and declare effect sizes for two-group comparison of binary outcomes

☐ Declare generic, precomputed effect sizes (in the metric closest to normality)

Main if/in Model Options

Specify group 1 (treatment) variables

Number of successes: deaths1

Number of failures: alive1

Specify group 2 (control) variables

Number of successes: deaths0

Number of failures: alive0

Specify effect size

Effect size: Log risk-ratio

Submit

No. of studies: 22
CI level: 95%

Model: Random effects
Method: DerSimonian-Laird

Effect size: _meta_es, Odds ratio
Std. Error: _meta_se

Close

```
generate alive1=pop1-deaths1  
generate alive0=pop0-deaths0
```

Display meta settings Modify meta settings

Setup

Summary

Forest plot

Heterogeneity

Regression

Publication bias

Note: Multivariate meta-analysis does not require any setup. Proceed to the Multivariate pane.

Declare meta-analysis data

☐ Compute and declare effect sizes for two-group comparison of continuous outcomes

☒ Compute and declare effect sizes for two-group comparison of binary outcomes

☐ Declare generic, precomputed effect sizes (in the metric closest to normality)

Main if/in Model Options

Meta-analysis model

☒ Random effects

☐ Common effect

☐ Fixed effects

Method: DerSimonian-Laird

Solution

meta - Meta-Analysis Control Panel

Display meta settings Modify meta settings

Setup

Summary

Forest plot

Heterogeneity

Regression

Publication bias

Multivariate

Summarize meta-analysis data

Main if/in Options Maximization

Options

95 Confidence level

Report effect sizes

☒ Report transformed effect sizes

Transformation: Exponential (exp) Label as:

☐ Sort studies

By: Study ID (_meta_id) Sort order: Ascending

☐ Report t test instead of z test for overall effect size

☐ Suppress output for individual studies

☐ Suppress output header

☐ Suppress meta settings information

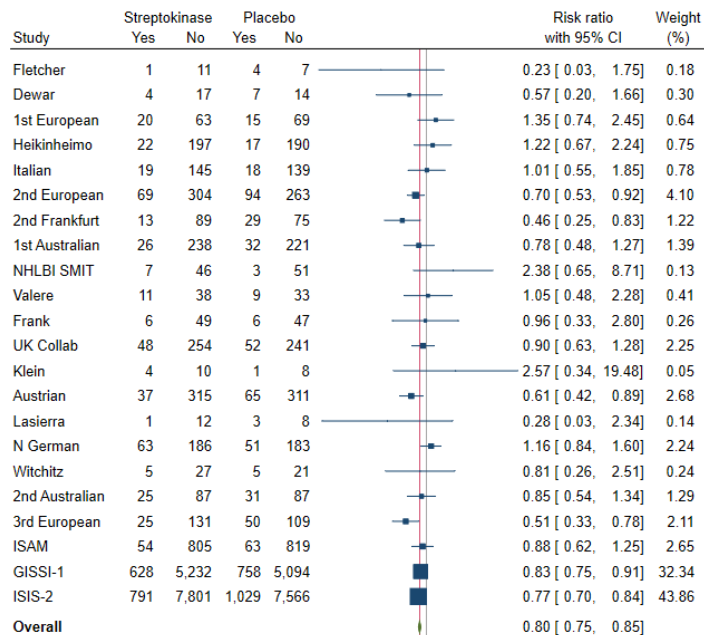
Display options

Submit

No. of studies: 22 Model: Random effects Effect size: _meta_es, Odds ratio
CI level: 95% Method: DerSimonian-Laird Std. Error: _meta_se

Solution

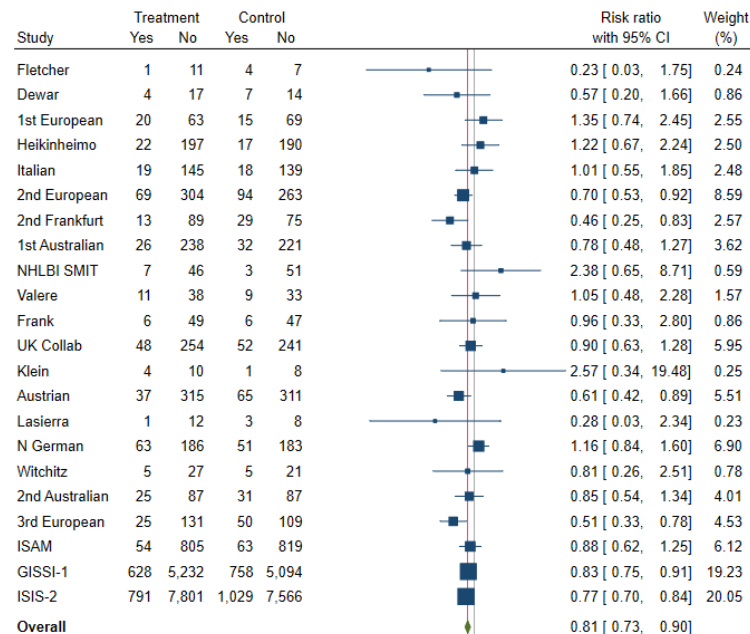
u^b



Heterogeneity: $I^2 = 30.95\%$, $H^2 = 1.45$
 Test of $\theta = \theta_0$: $Q(21) = 30.41$, $p = 0.08$
 Test of $\theta = 0$: $z = -7.75$, $p = 0.00$

Fixed-effects Mantel-Haenszel model

Fixed-effect (RR)

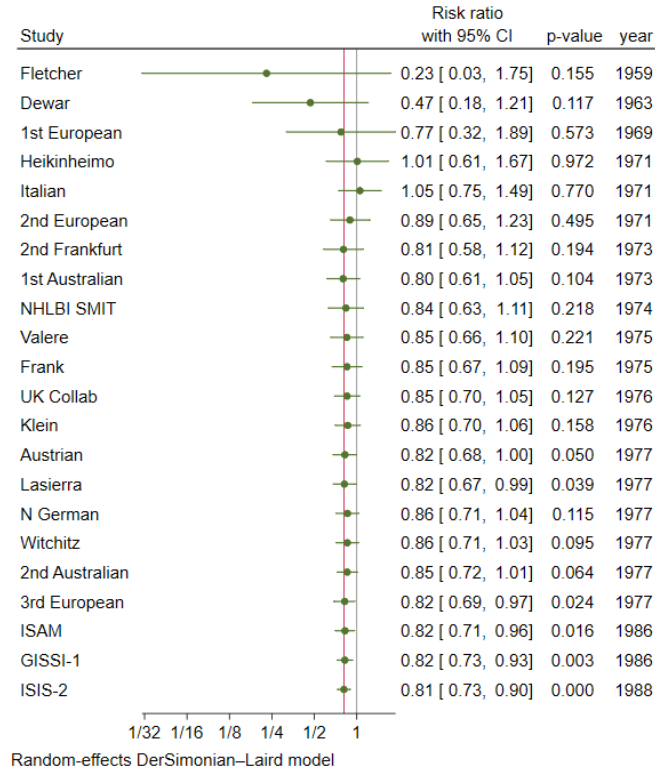


Heterogeneity: $\tau^2 = 0.01$, $I^2 = 30.94\%$, $H^2 = 1.45$
 Test of $\theta = \theta_0$: $Q(21) = 30.41$, $p = 0.08$
 Test of $\theta = 0$: $z = -4.06$, $p = 0.00$

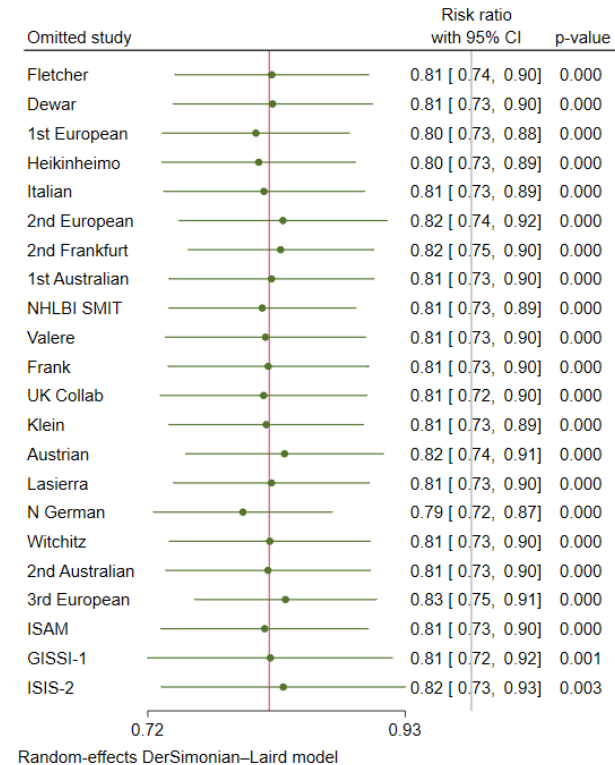
Random-effects DerSimonian-Laird model

Random-effects (RR)

Solution



Cumulative meta-analysis
according to trial year



Leave-one-out analysis (Influence
analysis)