

Ergebnisse klinischer Studien zusammenfassen: Metaanalysen

Dr. Waldemar Siemens

Systematische Übersichtsarbeiten erstellen

03.05.2021

Session 4

Zuverlässige Evidenz.
Informierte Entscheidungen.
Bessere Gesundheit.

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages



Interessenskonflikt

- Seit 2021 wissenschaftlicher Mitarbeiter der Cochrane Deutschland Stiftung und Institut für Evidenz in der Medizin
- 2020 bis 2021 Statistiker bei Roche Pharma AG (Grenzach-Wyhlen)



Schritte eines Systematic Review

1. Fragestellung festlegen
2. Auswahlkriterien definieren
3. Methoden definieren
4. Studien suchen
5. Auswahlkriterien anwenden
- 6. Daten extrahieren**
7. Bias-Risiko der Studien bewerten
- 8. Ergebnisse analysieren und darstellen**
- 9. Ergebnisse interpretieren und Schlussfolgerungen ziehen**
10. Review optimieren und aktualisieren

Übersicht

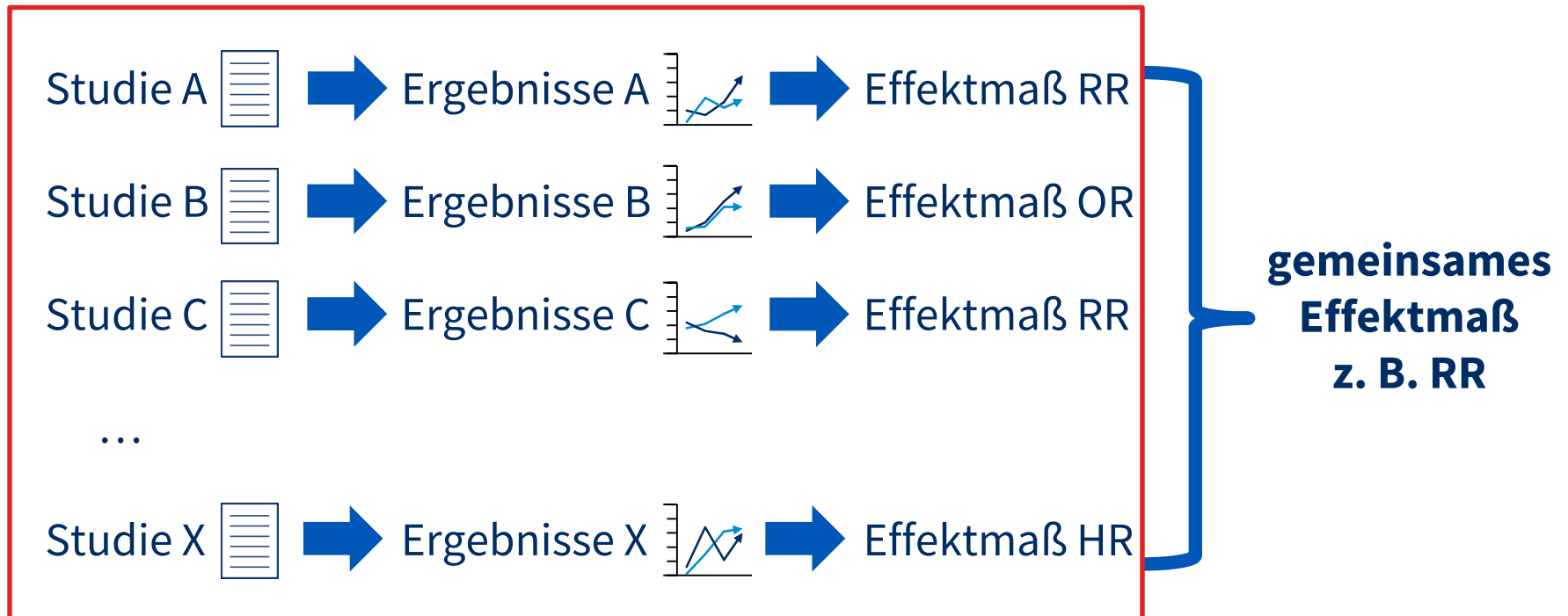
-
- 01** Konzept der Metaanalyse

 - 02** Fixed Effect und Random Effects Modell

 - 03** Heterogenität in Metaanalysen

 - 04** Publikationsbias und Funnel-Plot

1. Konzept der Metaanalyse



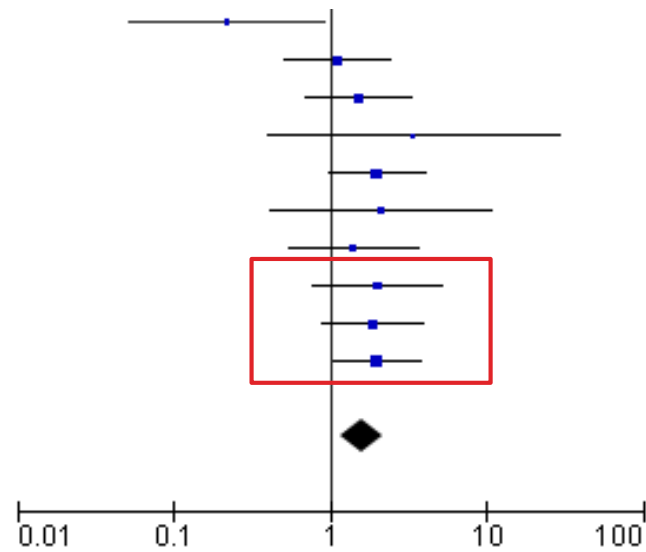
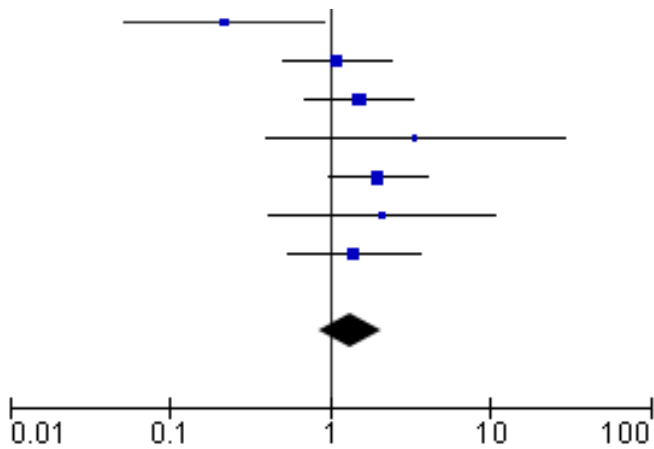
1. Konzept der Metaanalyse

- **Metaanalyse = Zusammenfassung** der Ergebnisse mehrerer Einzelstudien (mit gleicher Fragestellung) zu einem gemeinsamen Schätzer – falls dies sinnvoll ist
- Teil einer **systematischen Übersichtsarbeit** (Systematic Review)
- In jeder Studie wird der Therapieeffekt mittels einer **Maßzahl** geschätzt (bei binärer Zielgröße gewöhnlich RR oder OR)
- **Gemeinsamer Schätzer** für den Therapieeffekt: berechnet aus den Ergebnissen der Einzelstudien, **gewichtet** nach Informationsgehalt
- **Visualisierung** einer Metaanalyse: Forest-Plot

1. Warum wird eine Metaanalyse gemacht?

- Quantifizieren der Wirksamkeit (oder des Schadens) einer Behandlung und der Unsicherheit
- Erhöht Power
- Verbessert Präzision
- Unterschiede zwischen Studien werden anschaulicher
- Widerspruch zwischen Studienergebnissen auflösen
- Neue Hypothesen generieren

1. Erhöhung der Power



1. Wann sollte *keine* Metaanalyse gemacht werden?

Äpfel mit Birnen vergleichen

- Jede Studie muss die gleiche Frage haben
 - ▶ Vergleich und Endpunkte betrachten
 - ▶ Erfordert subjektive Einschätzung (→ klinische Expertise)
- Wenn eine breit gefächerte Auswahl von Studien kombiniert wird, kann auch nur eine weitgefasste Frage beantwortet werden.
- Wenn die Studien zu unterschiedlich sind,
 - ▶ ist die Antwort möglicherweise nicht aussagekräftig oder
 - ▶ tatsächliche Studieneffekte werden nicht sichtbar.

1. Wann sollte *keine* Metaanalyse gemacht werden?

“garbage in – garbage out”

- Metaanalyse ist nur so gut wie die eingeschlossenen Studien
- Ergebnisse der Metaanalyse sind verzerrt, wenn
 - ▶ eingeschlossene Studien verzerrt sind (Risiko für Bias!)
 - ▶ z.B. Publikations-Bias, Reporting-Bias etc. vorliegen

1. Wann kann eine Metaanalyse durchgeführt werden?

- Mehr als eine Studie hat den gleichen Endpunkt gemessen
- Studien sind hinreichend ähnlich, um in der Metaanalyse ein aussagekräftiges und nützliches Ergebnis zu erzielen
- Endpunkte wurden mit vergleichbaren Methoden erhoben
- Daten sind in einem nutzbaren Format berichtet

1. Schritte einer Metaanalyse

- Vergleiche definieren
 - ▶ z.B. Kaffee mit vs. ohne Koffein
- Endpunkte und Effektmaße festlegen
- Daten aus jeder relevanten Studie extrahieren
- Ergebnisdaten kombinieren, um einen zusammengefassten ('gepoolten') Effektschätzer zu erhalten
- Unterschiede zwischen den Studien untersuchen (Heterogenität)
- Ergebnisse interpretieren

1. Vergleiche, Outcome und Effektmaß aussuchen

Hypothetischer Review: Koffein und Kopfschmerzen

koffeinhaltiger Kaffee

vs

Entkoffeinierter Kaffee

- Paarweise Metaanalyse: Fragestellung in paarweise Vergleiche aufteilen.
- (Indirekter) Vergleich >2 Interventionen: Netzwerkmetaanalyse
- Review kann einen oder mehrere Vergleiche enthalten.
- Für jeden Vergleich müssen Outcome(s) und Effektmaß festgelegt werden.
- Klinische Expertise: Welche Behandlungen zusammenfassen / trennen?

1. Beispiel

Hypothetischer Review: Koffein und Kopfschmerzen

koffeinhaltiger Kaffee

vs

Entkoffeiniertes Kaffee

- Vergleich: Koffeinhaltiger vs koffeinfreier Kaffee bei gesunden Erwachsenen
- Outcome: Kopfschmerz, Blutdruck
- Effektmaße:
 - ▶ Kopfschmerz: Risk Ratio (RR), absolute Risikodifferenz (RD)
 - ▶ Blutdruck: Mittelwertdifferenz (MD)

1. Beispiel

Kopfschmerz	Koffein	Koffeinfrei
Amore-Coffea 2000	2/31	10/34
Deliciozza 2004	10/40	9/40
Mama-Kaffa 1999	12/53	9/61
Morrocona 1998	3/15	1/17
Norscafe 1998	19/68	9/64
Oohlahlazza 1998	4/35	2/37
Piazza-Allerta 2003	8/35	6/37

1. Den gemeinsamen Schätzer berechnen

Kopfschmerz	Koffein	Koffeinfrei	Risk Ratio
Amore-Coffea 2000	2/31	10/34	0,22
Deliciozza 2004	10/40	9/40	1,11
Mama-Kaffa 1999	12/53	9/61	1,53
Morrocona 1998	3/15	1/17	3,40
Norscafe 1998	19/68	9/64	1,99
Oohlahlazza 1998	4/35	2/37	2,11
Piazza-Allerta 2003	8/35	6/37	1,41
Gepoolter (gewichteter) Schätzer			1,34

- Für dichotome Maße: Pooling auf Log-Skala!
- Gewichtung der Studienschätzer
 - ▶ Simples Zusammenzählen kann zu Fehlern führen (Simpson Paradoxon)
 - ▶ Unterschiedlich große Studien sollten unterschiedlich gewichtet werden

1. Gemeinsamer Schätzer und Gewichte

Studien, die mehr Informationen beitragen, erhalten mehr Gewicht

- Mehr TeilnehmerInnen → mehr Ereignisse
- Mehr Ereignisse → höhere Präzision (engeres Konfidenzintervall)
- Berechnung des gepoolten Schätzers mittels (Studien)-Schätzer und Varianz
- Gewichtungsmethode: Inverse-Varianz-Methode

$$\text{Gewicht} = \frac{1}{\text{Varianz der Schätzer}} = \frac{1}{SE^2}$$

$$\text{Gepoolter Schätzer} = \frac{\text{Summe von (Schätzer} \times \text{Gewicht)}}{\text{Summe aller Gewichte}}$$

Übersicht

-
- 01** Konzept der Metaanalyse

 - 02** Fixed Effect und Random Effects Modell

 - 03** Heterogenität in Metaanalysen

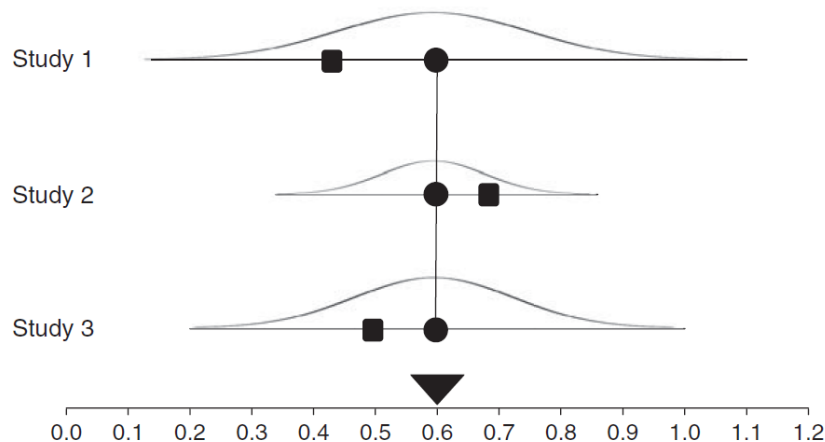
 - 04** Publikationsbias und Funnel-Plot

2. Fixed Effect und Random Effects Modell

Annahmen beim **Fixed (=Common) Effect Modell**:

- ▶ Allen Studien liegt der gleiche „wahre“ Therapieeffekt zugrunde
- ▶ Unterschiede in den geschätzten Therapieeffekten sind ausschließlich auf die Zufallsstreuung zurückzuführen
- ▶ Es gibt mehrere Gewichtungsmethoden, z. B. inverse Varianz Methode

Gewicht z.B. $w_i = \frac{1}{SE(x_i)^2}$ (Intrastudienvarianz)

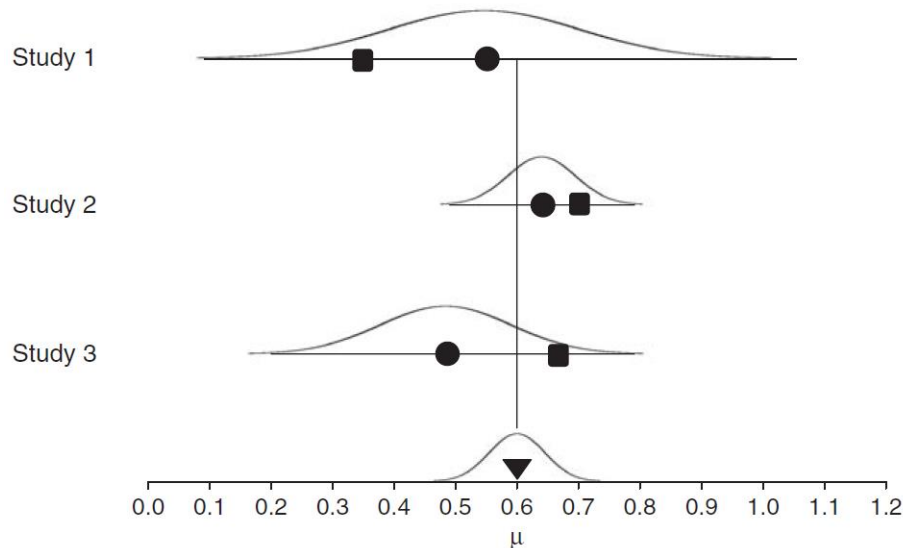


2. Fixed Effect und Random Effects Modell

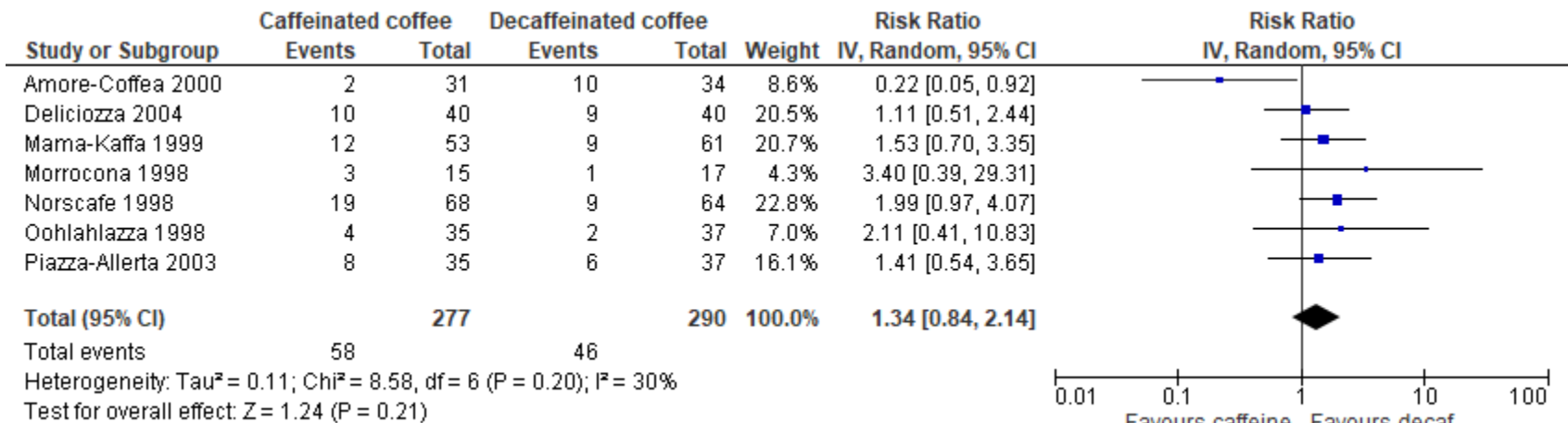
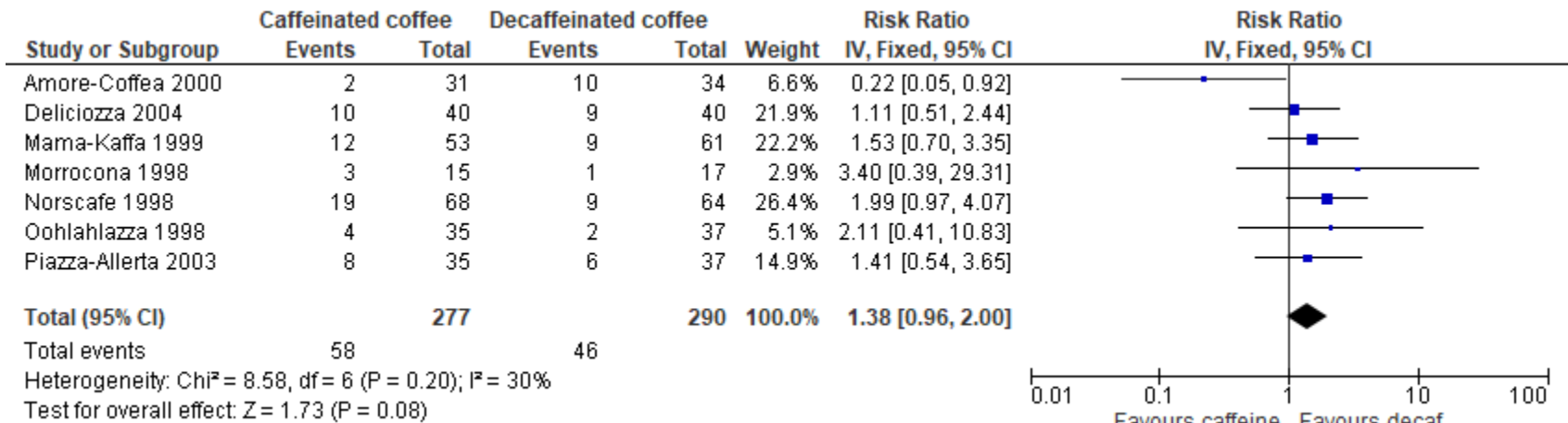
Annahmen beim **Random Effects Modell**:

- ▶ Unterschiedliche Effekte für jede Studie (Normalverteilungsannahme)
- ▶ Kleinere Studien bekommen mehr Gewicht als bei Fixed Effect Modell
- ▶ Es gibt mehrere Gewichtungsmethoden, z. B. inverse Varianz Methode

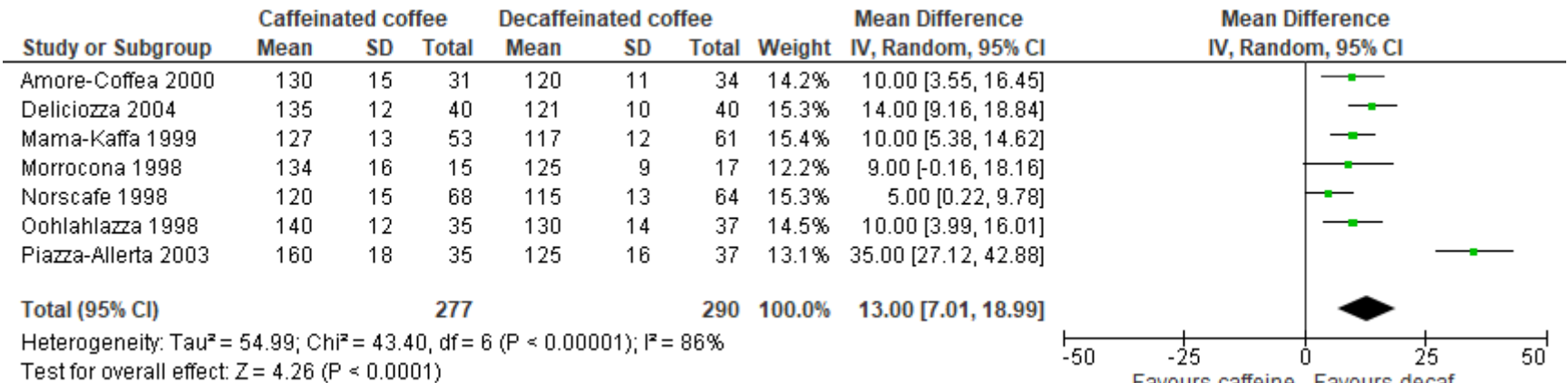
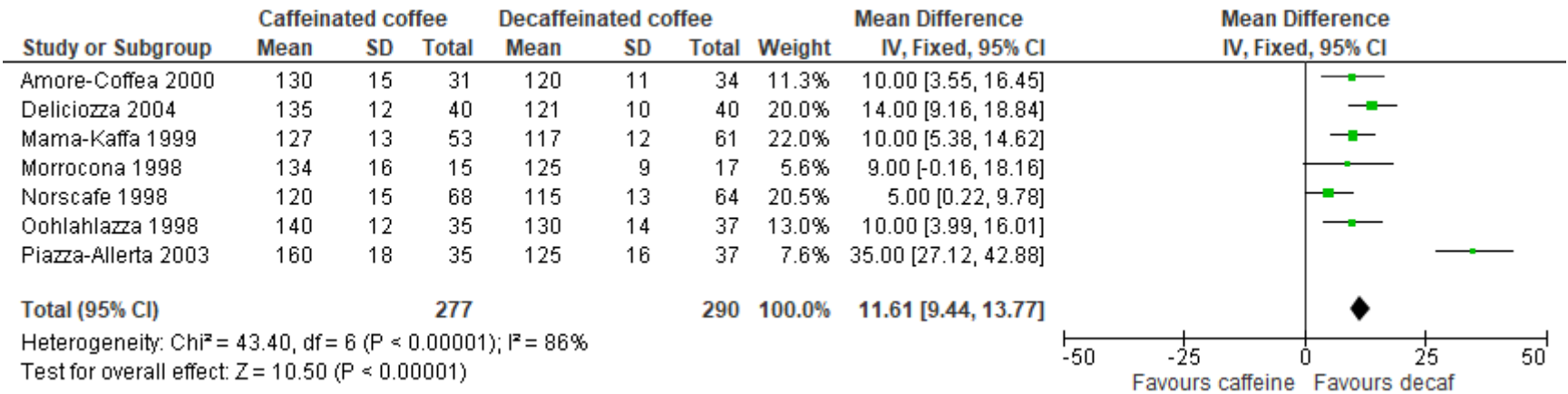
$$\text{Gewicht } w_i = \frac{1}{SE(x_i)^2 + \tau^2} \quad (\text{Intrastudienvarianz} + \text{Interstudienvarianz } \tau^2)$$



2. Fixed Effect und Random Effects Modell



2. Fixed Effect und Random Effects Modell



2. Fixed Effect vs. Random Effects Modell

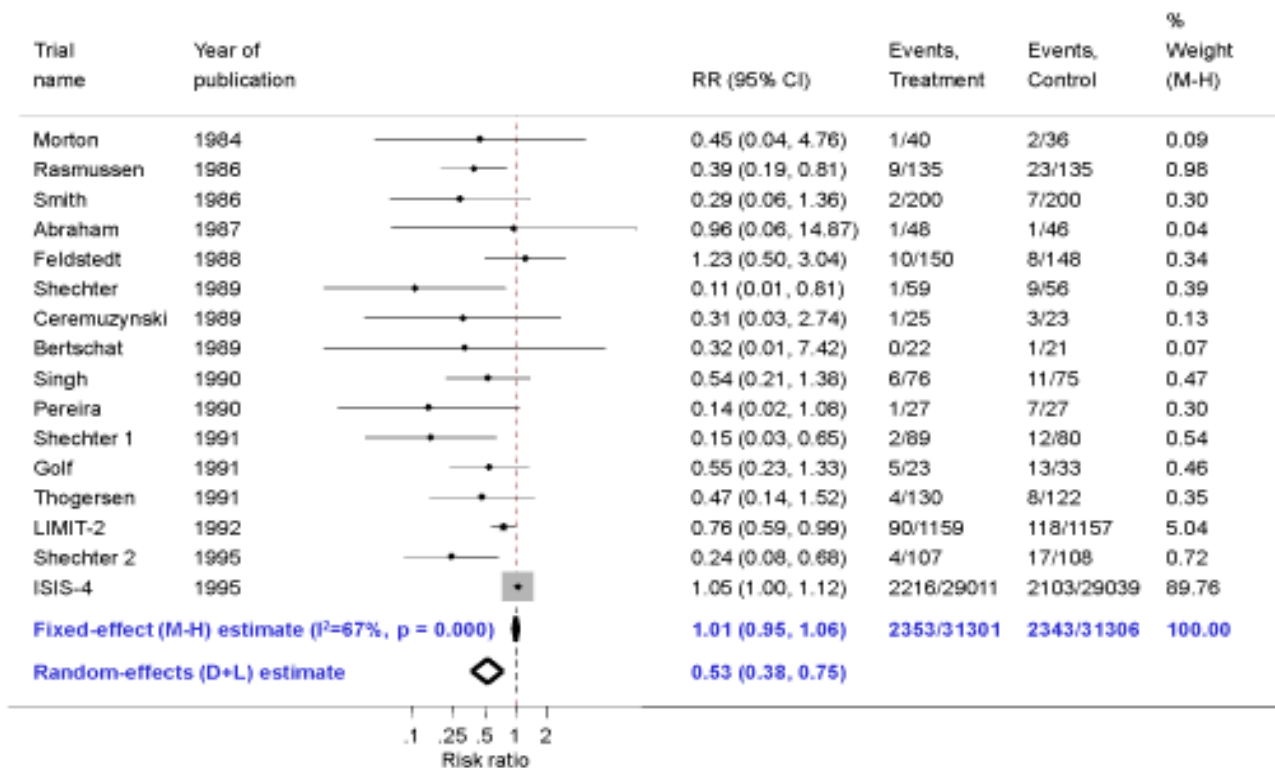
Cochrane Handbook (Higgins & Thomas, 2022):

*„The choice between a fixed-effect and a random-effects meta-analysis should **never** be made on the basis of a statistical test for heterogeneity.“*

→ **vorher festlegen!**

- Modelle oft **vergleichbar / gleich**, wenn wenig / keine Heterogenität
- Wenn Heterogenität vermutet wird, sollte das Random Effects Modell gewählt werden, da das KI im Fixed Effect Modell die Präzision überschätzt
- Kleinere Studien bekommen größeres Gewicht im Random Effects Modell
- Wieso nimmt man dann nicht immer das Random Effects Modell?
- Konfidenzintervalle beim Random Effects Modell i.d.R. **breiter**
 - ▶ Meist **konservativere** Entscheidung
 - ▶ ABER: „small-study effects“

2. Comparison of fixed and random effects meta-analytic estimates of the effect of intravenous magnesium on mortality following myocardial infarction – small-study effects



Übersicht

-
- 01** Konzept der Metaanalyse

 - 02** Fixed Effect und Random Effects Modell

 - 03** Heterogenität in Metaanalysen

 - 04** Publikationsbias und Funnel-Plot



3. Heterogenität in Metaanalysen

In Metaanalysen kann man verschiedene **Arten von Heterogenität** unterscheiden, zum Beispiel

- klinische Heterogenität
- methodische Heterogenität
- zufällige Schwankungen

} Messbar als (statistische) Heterogenität
(Higgins & Thomas, 2021)

Frage:

Ist es **sinnvoll**, die einzelnen Studienergebnisse zusammenfassen?

⇒ dafür betrachten wir die **statistische Heterogenität**, die klinisch relevant und statistisch signifikant sein kann oder nicht

- Cochran's Q (Chi²-Test)
- Higgins' und Thompson's I^2
- Interstudienvarianz τ^2
- 95% Prädiktionsintervalle (basierend auf τ^2 bzw. τ)

3. Heterogenität in Metaanalysen

Higgins' und Thompson's I^2 – Heterogenitätsmaß

- 0% \triangleq Variabilität ausschließlich durch Zufallsschwankungen
 - 100% \triangleq komplett durch Unterschiede zwischen den Studien
 - Vergleich der Werte von I^2 in Metaanalysen mit **unterschiedlicher Anzahl von Studien** möglich
 - misst den **“Überlappungsgrad”** der Konfidenzintervalle
 - bei sehr großen Studien geht I^2 gegen 100% \Rightarrow klinisch relevant?
 - I^2 alleine sollte nicht dafür benutzt werden, zu entscheiden, ob gepoolt werden kann oder nicht!
- **Klinische Einschätzung, τ^2 , 95% Prädiktionsintervalle**

3. Heterogenität in Metaanalysen

Bei **sehr starker Heterogenität** reicht das Random Effects Modell **nicht** aus!

Untersuchungen der Ursache der Heterogenität:

- **Sensitivitätsanalysen**
 - ▶ Untersuchen, ob der Ausschluss einzelner Studien aus der Analyse einen starken Einfluss auf das Ergebnis hat.
- **Subgruppenanalysen**
 - ▶ Studien nach vorab spezifizierten Faktoren gruppieren.
- **Metaregression**
 - ▶ Outcome Variable wird durch erklärende Variable(n) vorhergesagt.

Bei sehr starker Heterogenität sollte in Erwägung gezogen werden, auf eine formale Metaanalyse zu verzichten und die gefundenen Ergebnisse **ohne einen gemeinsamen Schätzer** darzustellen und zu diskutieren.

Übersicht

-
- 01** Konzept der Metaanalyse

 - 02** Fixed Effect und Random Effects Modell

 - 03** Heterogenität in Metaanalysen

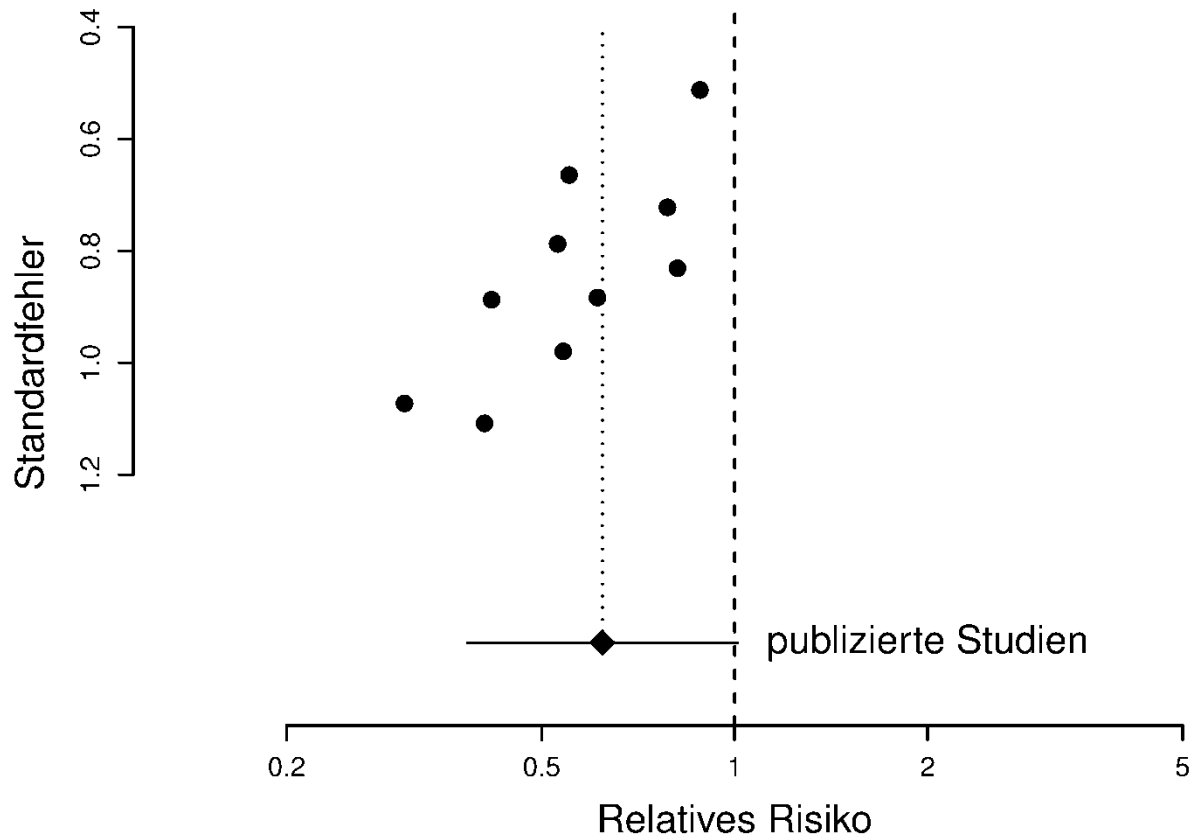
 - 04** Publikationsbias und Funnel-Plot

4. Publikationsbias und Funnel-Plot

- Studien mit **nicht-signifikantem** oder **negativem** Ergebnis werden
 - ▶ mit geringerer Wahrscheinlichkeit / später publiziert
 - ▶ häufiger in nicht englischsprachigen Journalen veröffentlicht
 - ▶ Freiburger Ethikkommission (DFG-Projekt): von den 2000-2002 registrierten Studien wurden bis 2011 nur **52%** publiziert
- Es gibt **graphische Methoden / statistische Tests**, um das Vorhandensein von Publikationsbias zu untersuchen, sowie diesen auszugleichen

4. Funnel-Plot

Asymmetrie im Funnelplot



Fazit

- Die **Metaanalyse** ist ein statistisches Verfahren, um über Studien hinweg einen gepoolten Effektschätzer zu berechnen.
- Dabei werden die einzelnen Studien anhand ihres Informationsgehalts **gewichtet**.
- **Forest-Plots** stellen die Ergebnisse einer Metaanalyse graphisch dar
- Auch in Metaanalysen sind alle Schätzer ungenau und sollten immer mit **Konfidenzintervallen** berichtet werden.
- Je nach erwarteter Heterogenität des wahren Effekts sind verschiedene Modelle (**Fixed Effect** bzw. **Random Effects Modell**) für primäre Analyse vor der Analyse zu definieren.
- Unerwarteter oder **hoher Heterogenität** ist mit verschiedenen Verfahren auf den Grund zu gehen (z.B. Sensitivitätsanalyse, Subgruppenanalyse, Metaregression).

Literaturverzeichnis (1)

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to meta-analysis. Wiley 2009.

Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. Lancet. 1991 Apr 13;337(8746):867-72. doi: 10.1016/0140-6736(91)90201-y.

von Elm E, Röllin A, Blümle A, Huwiler K, Witschi M, Egger M. Publication and non-publication of clinical trials: longitudinal study of applications submitted to a research ethics committee. Swiss Med Wkly. 2008 Apr 5;138(13-14):197-203.

Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet. 1997 Aug 2;350(9074):326-9. doi: 10.1016/S0140-6736(97)02419-7.

Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.

Literaturverzeichnis (2)

IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*. 2016 Jul 12;6(7):e010247.

Meinertz T, Kasper W, Schumacher M, Just H. The German multicenter trial of anisoylated plasminogen streptokinase activator complex versus heparin for acute myocardial infarction. *Am J Cardiol*. 1988 Sep 1;62(7):347-51. doi: 10.1016/0002-9149(88)90956-3.

Olkin I. Statistical and theoretical considerations in meta-analysis. *J Clin Epidemiol*. 1995 Jan;48(1):133-46; discussion 147. doi: 10.1016/0895-4356(94)00136-e.

Rücker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I² in assessing heterogeneity may mislead. *BMC Medical Research Methodology* 2008

Review Manager (RevMan). Version (5.4.0). The Cochrane Collaboration, (2020). Available at <https://training.cochrane.org/online-learning/core-software-cochrane-reviews/revman/revman-5-download>

Sterne JAC, Sutton AJ, Ioannidis JPA et al.. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *British Medical Journal* 2011

**Vielen Dank für Ihre
Aufmerksamkeit!**

Dr. Waldemar Siemens
siemens@cochrane.de

**Zuverlässige Evidenz.
Informierte Entscheidungen.
Bessere Gesundheit.**

Gefördert durch:



aufgrund eines Beschlusses
des Deutschen Bundestages

