

Indirekte Vergleiche von Therapieverfahren - Eine methodische Bestandsaufnahme und Validitätsprüfung

Workshop „indirekte Vergleiche“ – Berlin 12.11.09

Ben Schöttker, Dagmar Lühmann , Heiner Raspe
Universität zu Lübeck

Kontakt: b.schoettker@dkfz.de



Die Ergebnisse, die in diesem Vortrag präsentiert werden, entstammen einem vollständig vom DIMDI finanzierten Projekt. Der entstandene HTA-Bericht „Indirekte Vergleiche von Therapieverfahren“ ist über das DIMDI oder GMS frei verfügbar: http://portal.dimdi.de/de/hta/hta_berichte/hta243_bericht_de.pdf

DIMDI
deutsche agentur für HTA des
Deutschen Instituts für Medizinische
Dokumentation und Information

Agenda

Teil I - Methodische Bestandsaufnahme

- Bedarf an indirekten Vergleichen?
- Internationaler Status quo
- Vorstellung der Methoden
- Häufigkeit ihrer Verwendung in systematischen Übersichtsarbeiten

Teil II - Validitätsprüfung

- Voraussetzungen für die Validität von indirekten Vergleiche
- Validitätsprüfung in drei Teilen



Hintergrund

Vergleichende Nutzenbewertungen von Therapieoptionen

- G-BA SGB V § 135 „Bewertung von Untersuchungs- und Behandlungsmethoden“
- IQWiG SGB V § 35b „Bewertung des Nutzens und der Kosten von Arzneimitteln“

„Das Institut (IQWiG) bestimmt auftragsbezogen über die **Methoden** und Kriterien für die Erarbeitung von Bewertungen nach Satz 1 (Nutzen- und Kosten-Nutzen-Bewertungen) auf der Grundlage **der in den jeweiligen Fachkreisen anerkannten internationalen Standards der evidenzbasierten Medizin und der Gesundheitsökonomie**“.

Goldstandard zum Vergleich der Wirksamkeit (.. und Nutzen?)

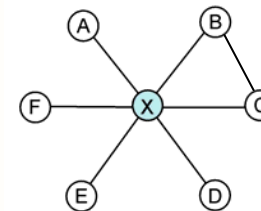
- RCT – bzw. Meta-Analysen
- Head-to-Head-Studien
- Begrenzte Verfügbarkeit
- > Bedarf an indirekten Vergleichen



Direkter Vergleich



Indirekter Vergleich



Mixed Treatment Comparison



Internationaler Status Quo I

AHRQ (Agency for Healthcare Research and Quality; USA):

„Methods Reference Guide for Effectiveness and Comparative Effectiveness Reviews“, Draft, April 2008.

„.....Indirect Comparisons: Placebo-controlled trials **can be helpful** for evaluating absolute rates of benefits and harms associated with an intervention. However, **evidence from head-to-head comparisons is always preferable to adjusted indirect analyses** from placebo or active-controlled trials for evaluating comparative efficacy and harms...“

NICE (National Institute for Health and Clinical Excellence; Großbritannien):

„Guide to the Methods of Technology Appraisal“, Juni 2008.

„There may be situations when data from head-to-head RCTs of the technologies (and/or comparators) are not available. In these circumstances, indirect treatment comparison analyses **should be considered...**“.

IQWiG (Deutschland):

„Vergleichende Nutzenbewertung verschiedener antihypertensiver Wirkstoffgruppen als Therapie der ersten Wahl bei Patienten mit essentieller Hypertonie“, Vorbericht, September 2008.

„.....Für den Vorbericht 2.0 waren – soweit sinnvoll und durchführbar – indirekte Vergleiche **als ergänzende Datenanalyse** zusätzlich zu den direkten Vergleichen geplant. Die Ergebnisse wurden ergänzend in der Diskussion aufgeführt, um mögliche Aspekte, die den direkten Vergleichen nicht zu entnehmen waren, explorativ darzustellen...“



Internationaler Status Quo II

PBAC (Pharmaceutical Benefits Advisory Committee; Australia):

Report of the Indirect Comparisons Working Group to the PBAC: assessing indirect comparisons. Dec. 2008.

„....With regard to common reference based indirect comparisons, the Working Group concluded that **there is not always a strong methodological basis to prefer common reference-based indirect comparisons over other nonrandomised study designs**. The reasons to retain the current preference in the PBAC Guidelines would therefore be pragmatic rather than methodological.

CADTH (Canadian Agency for Drugs and Technologies in Health; Canada):

Wells GA et al. Technology Report: Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis. March 2009.

„.....In the absence of previously performed randomized controlled trials in which two interventions of interest have been compared, **indirect methods may be used**. However, indirect treatment comparisons should be restricted to those situations in which it is not possible to perform a direct head-to-head trial. ...”

IQWiG (Deutschland):

„Allgemeine Methoden zur Bewertung von Verhältnissen zwischen Nutzen und Kosten - Version 1.0“, Oktober 2009.

„..... Um eine Kosten-Nutzen-Bewertung multipler Interventionen zu ermöglichen, kann das Institut unter Inkaufnahme einer – im Vergleich zum Ansatz der reinen Nutzenbewertung – geringeren Ergebnissicherheit **auch indirekte Vergleiche zur Bewertung von Kosten- Nutzen-Verhältnissen heranziehen**. ...”



DIMDI- Projekt: „Indirekte Vergleiche“

Fragestellungen

- Welche Methoden existieren bisher?
- Wie häufig wurden sie bisher verwendet?
- Wie ist die Validität indirekter Vergleiche zu beurteilen?

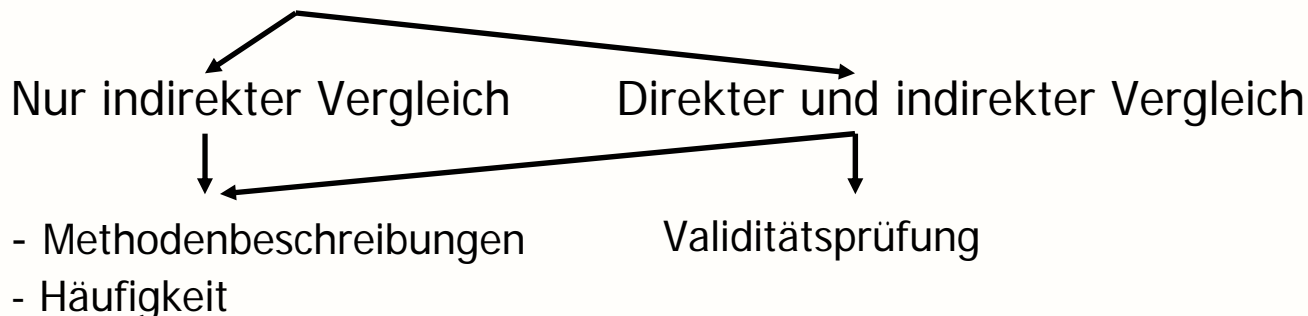
Ausgangslage

- Glenny et al. 2005 mit Literaturrecherche bis 1999

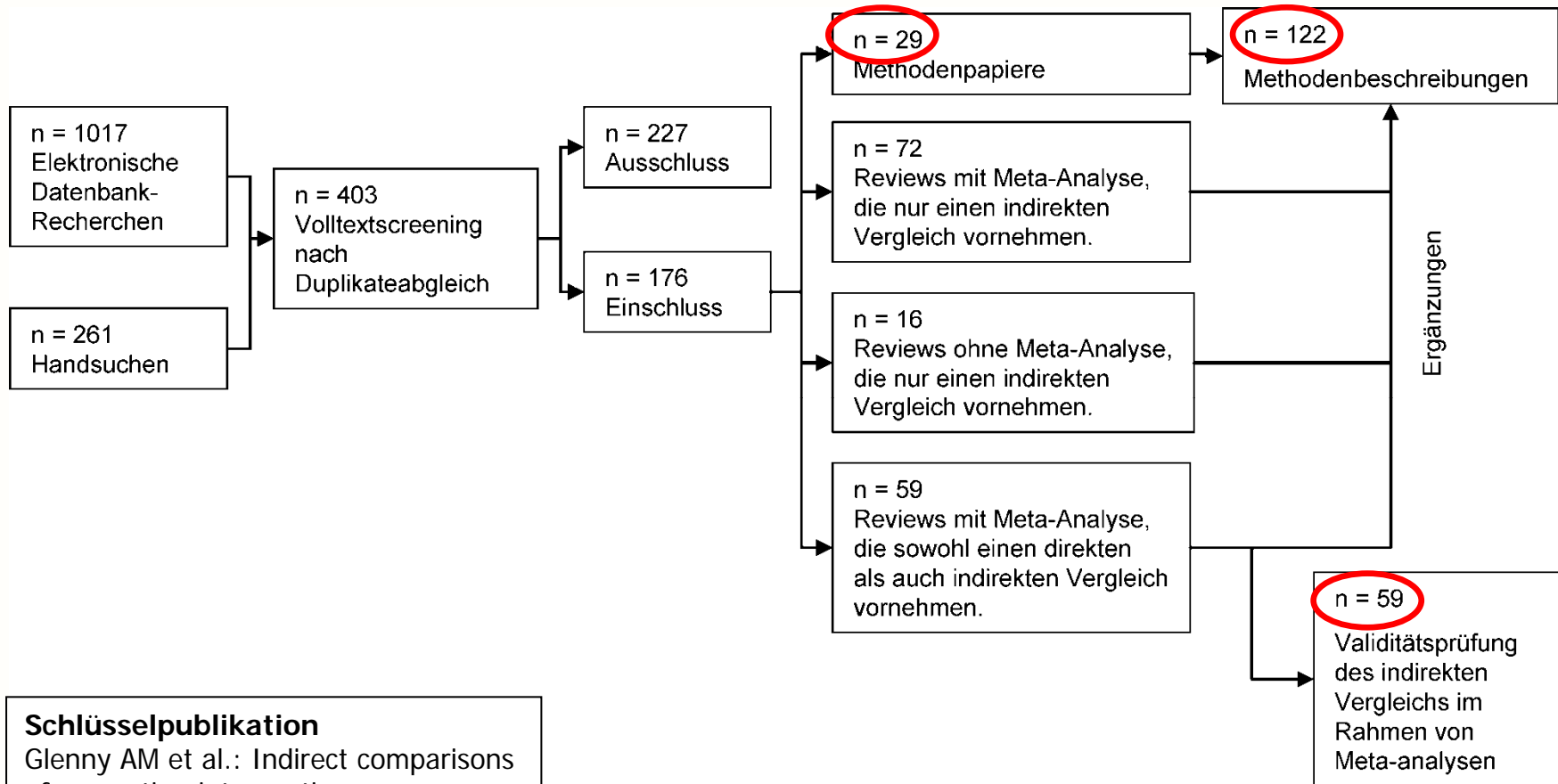
Methodik

Systematische Literaturrecherche:

- Methodenpapiere
- Anwendungsbeispiele (systematische Reviews)



Ergebnisse der Literaturrecherche



Schlüsselpublikation
 Glenny AM et al.: Indirect comparisons
 of competing interventions.
 Health Technol.Assess. 2005
 Recherche bis 1999



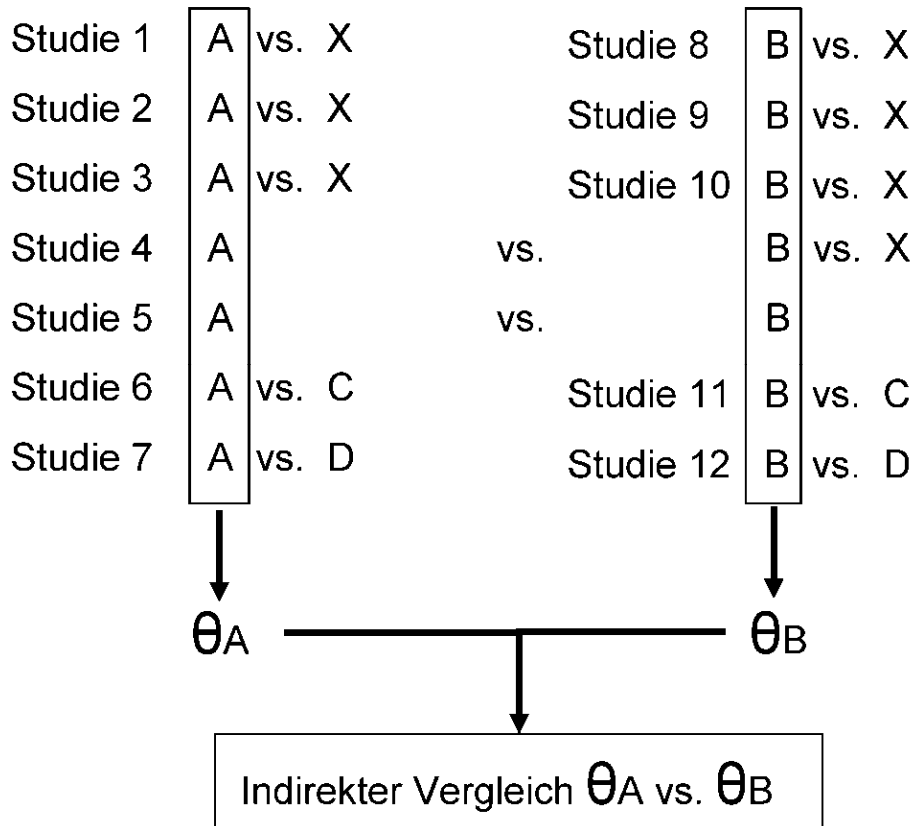
Methoden für indirekte Vergleiche

1. Nicht-adjustierte indirekte Vergleiche
2. Adjustierte indirekte Vergleiche
3. Meta-Regression
4. Mixed Treatment Comparisons

Nicht-adjustierte indirekte Vergleiche I

Studien mit Verumarm A

Studien mit Verumarm B



Legende:

- A, B, C, D : Therapieeffekt von Therapieoption A, B, C oder D
- X: Gemeinsamer Komparator für den indirekten Vergleich
- θ : Gepoolter Effekt
- vs. : Differenz der Therapieeffekte



Nicht-adjustierte indirekte Vergleiche II

Vier Möglichkeiten für die Gegenüberstellung der gepoolten Ergebnisse Θ_A und Θ_B :

1. Differenz bilden: $\Theta_{A \text{ vs. } B} = \Theta_A - \Theta_B$
2. Überlappung oder Nicht-Überlappung der Konfidenzintervalle interpretieren
3. Statistischen Test auf Unterschiedlichkeit der Werte Θ_A und Θ_B (t-Test)
4. Narrativ argumentieren

Nicht-adjustierte indirekte Vergleiche III

Vorteile:

- Einfach und schnell zu berechnen
- Gesamte RCT-Evidenz kann eingeschlossen werden

Nachteile:

- Bruch der Randomisation
- Kein kontrollierter Vergleich
- Hohe Biasanfälligkeit (Alle, die in Beobachtungsstudien auftreten können)

Literatur:

Song et al. 2000, Glenny et al., Gartlehner et al. 2008



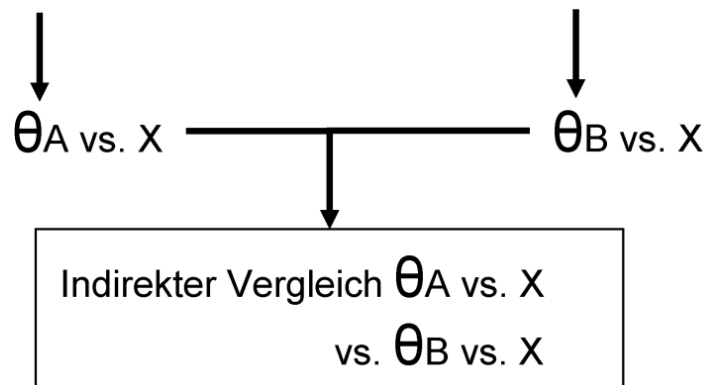
Adjustierte indirekte Vergleiche I

Studien mit Verumarm A

Studie 1 A vs. X
 Studie 2 A vs. X
 Studie 3 A vs. X
 Studie 4 A
 Studie 5 A
 Studie 6 A vs. C
 Studie 7 A vs. D

Studien mit Verumarm B

Studie 8 B vs. X
 Studie 9 B vs. X
 Studie 10 B vs. X
 Studie 11 B vs. C
 Studie 12 B vs. D



Legende:
 A, B, C, D : Therapieeffekt von Therapieoption A, B, C oder D
 X: Gemeinsamer Komparator für den indirekten Vergleich
 θ : Gepoolter Effekt
 vs. : Differenz der Therapieeffekte



Adjustierte indirekte Vergleiche II

Vorteile:

- Eigenschaften der RCT bleiben erhalten.
- Einfach und schnell zu berechnen

Nachteile:

- Heterogenität zwischen den Studien kann zu Ergebnisverzerrung führen
- Nur Vergleiche zwischen zwei Therapieoptionen möglich
- Verfügbare Evidenzbasis wird stark eingeeengt

Literatur:

Bucher et al., 1997; Song et al. 2000, Glenny et al., Gartlehner et al. 2008



Meta-Regression I

Studien mit Verumarm A

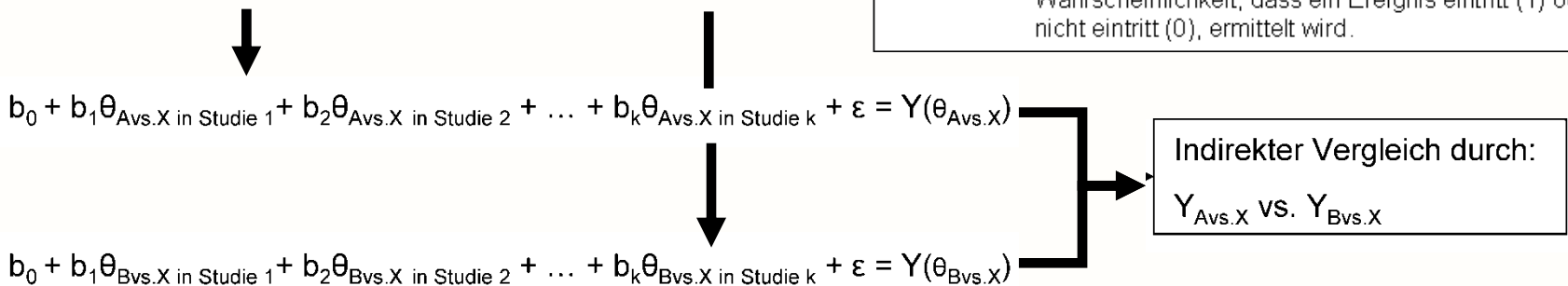
Studie 1	A vs. X
Studie 2	A vs. X
Studie 3	A vs. X
Studie 4	A
Studie 5	A
Studie 6	A vs. C
Studie 7	A vs. D

Studien mit Verumarm B

Studie 8	B vs. X
Studie 9	B vs. X
Studie 10	B vs. X
vs.	B vs. X
vs.	B
Studie 11	B vs. C
Studie 12	B vs. D

Legende:

- A, B : Therapieeffekt von Therapieoption A oder B
- X: Gemeinsamer Komparator für den indirekten Vergleich
- vs. : Differenz der Therapieeffekte
- 1, 2, ..., k: Indices, die Einflüsseffekte aus Studie 1, 2... bis Studie k anzeigen
- Y: Funktion des gepoolten Gesamttherapieeffekts: „Outcomevariable“
- θ : Therapieeffektunterschied einer beliebigen Studie: „Erklärende Variable“ (Bei dichotomen Daten steht θ für den logarithmierten Therapieeffektunterschied)
- ε : Zufallskomponente, in die Messfehler und nicht beobachtete Einflüsse auf Y eingehen.
- $b_0, b_1 \dots b_k$: Abweichungsfaktoren zwischen modellierten und beobachteten Werten
- 1/0: Index, der anzeigt, dass es sich um einen binären Outcome handelt, bei dem das Risiko oder die Wahrscheinlichkeit, dass ein Ereignis eintritt (1) oder nicht eintritt (0), ermittelt wird.



Formeln gelten nur für kontinuierliche Daten. Für dichotome Daten wird logistische (Meta-) Regression eingesetzt



Meta-Regression II

Vorteile:

- Möglichkeit zur Berücksichtigung von Confoundern
-> Ausgleich von Heterogenität
- Eigenschaften der RCT bleiben erhalten

Nachteile:

- Nur einsetzbar, wenn ausreichend Studien zur Verfügung stehen (> 10).
- Unbekannte Confounder können Ergebnisse verzerren.

Literatur:

Thompson et al. 1999 und 2002, Glenny et al. 2005



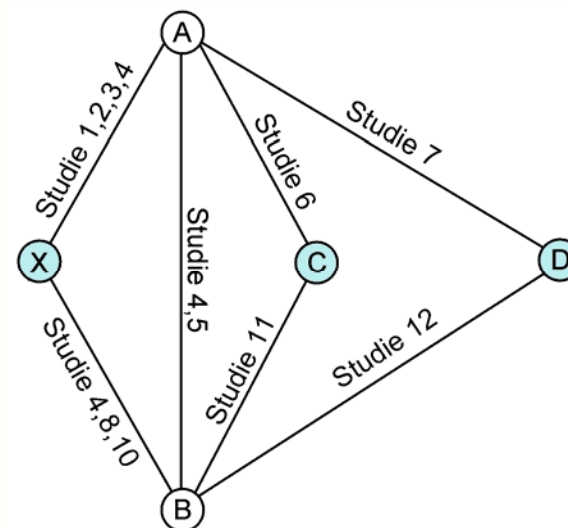
Mixed Treatment Comparison I

Studien mit Verumarm A

Studie 1 A vs. X
 Studie 2 A vs. X
 Studie 3 A vs. X
 Studie 4 A
 Studie 5 A
 Studie 6 A vs. C
 Studie 7 A vs. D

Studien mit Verumarm B

Studie 8 B vs. X
 Studie 9 B vs. X
 Studie 10 B vs. X
 vs. B vs. X
 vs. B
 Studie 11 B vs. C
 Studie 12 B vs. D



Legende:
 A, B, C, D : Therapieeffekt von Therapieoption A, B, C oder D
 X: Gemeinsamer Komparator für den indirekten Vergleich
 θ : Gepoolter Effekt
 vs. : Differenz der Therapieeffekte



Ergebnis einer Mixed Treatment Comparison

Survival benefits with treatment regimens for advanced colorectal cancer¹

Therapie-Regime ²	Hazard Ratio Mortalität ³ in Gegenüberstellung zu Fluorouracil-Monotherapie (95%-Credibility interval)	Wahrscheinlichkeit, das wirksamste Therapie regime zu sein
Fluorouracil+ Irinotecan+ Bevacizumab	0,60 (0,44-0,84)	68%
Fluorouracil+ Irinotecan+ Oxaliplatin	0,72 (0,54-0,97)	14%
Fluorouracil+ Oxaliplatin+ Bevacizumab	0,72 (0,57-0,90)	10%
Fluorouracil+ Bevacizumab	0,78 (0,60-1,03)	5%
Fluorouracil+ Oxaliplatin	0,87 (0,78-0,98)	0%
Irinotecan+ Oxaliplatin	0,89 (0,76-1,19)	0%
Fluorouracil+ Irinotecan	0,92 (0,84-1,01)	0%
Bevacizumab	0,92 (0,66-1,25)	0%
Irinotecan	1,00 (0,86-1,17)	0%
Oxaliplatin	1,00 (0,67-1,54)	0%

(1) Quelle: Golfinopoulos V, Salanti G, Pavlidis N, Ioannidis JP. Survival and disease-progression benefits with treatment regimens for advanced colorectal cancer: a meta-analysis. Lancet Oncol 2007 Oct;8(10):898-911.

(2) Fluorouracil wurde stets mit Leucovorin kombiniert.

(3) Es werden nur die Ergebnisse gezeigt, die alle 47 Therapievergleiche aus den gefundenen Studien einschließen.



Mixed Treatment Comparison II

Vorteile:

- Erfassung der gesamten RCT-Evidenz
- Erlaubt die Bildung von Wirksamkeitsrangfolgen
- Kann bezüglich Confoundern adjustieren
- Eigenschaften der RCT bleiben erhalten

Nachteile:

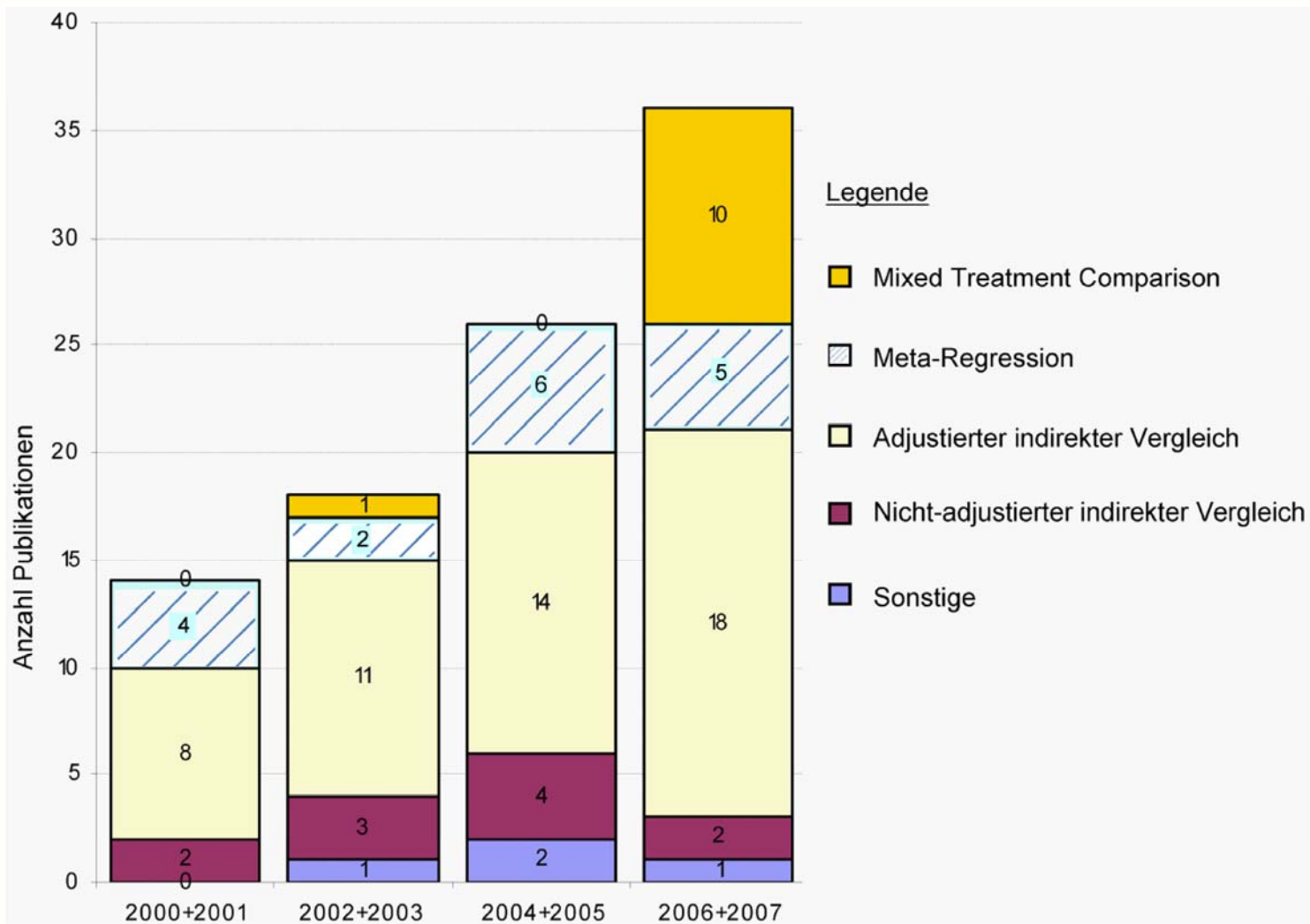
- Aufwändige Statistik
- Für Anwender wenig transparent
- Neue Methode, bisher wenig erprobt

Literatur:

Lumley et al, 2002, Lu und Ades 2004; Lu 2006 und 2007 Eddy 1989, 1990, 1992, Higgins 1996, Gleser und Olkin, 2001, Hirotsu und Yamada, 1999, Dominici, 1999, Hasselblad, 1998, Ades 2003, 2006 und 2007



Häufigkeit der verschiedenen Methoden



Teil II – Validitätsprüfung

- Voraussetzungen für indirekte Vergleiche
- Validitätsprüfung in drei Schritten

Voraussetzungen für indirekte Vergleiche

Die Studien müssen vergleichbar sein.

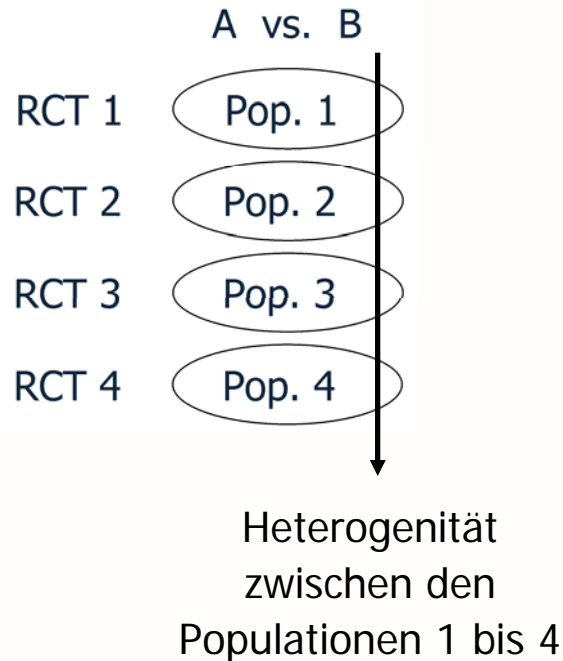
-> Es darf keine Heterogenität zwischen den Studien bestehen.

Was ist Heterogenität?

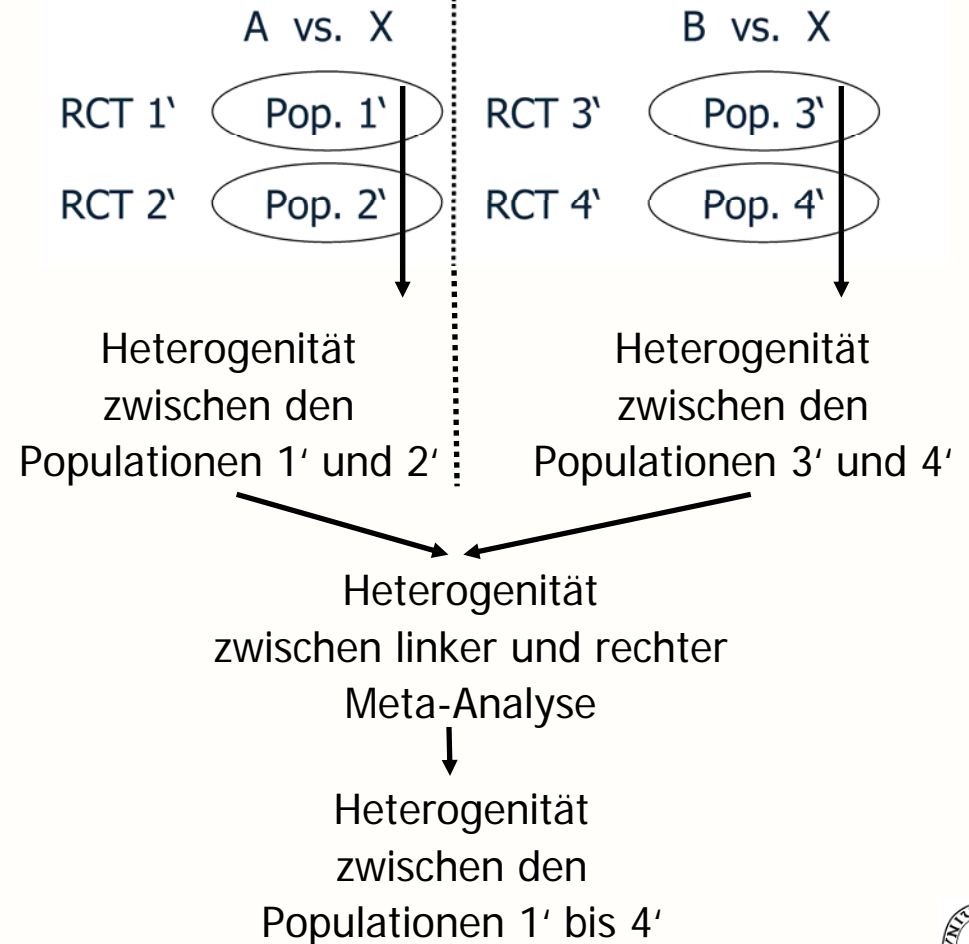
- Methodische Heterogenität
- Klinische Heterogenität

Heterogenität in direkten und indirekten Vergleichen

Direkter Vergleich



Indirekter Vergleich



Validitätsprüfung

Literaturbasis:

- Datensätze aus 59 systematischen Reviews, die sowohl direkte als auch indirekte Vergleiche durchführten (n = 248)

Methodik: z-Statistik

- Differenz zwischen Therapieeffekten von direkten und indirekten Vergleichen

$$Z = \frac{\theta_{\text{Direkt}} - \theta_{\text{Indirekt}}}{SE(\theta_{\text{Direkt}} - \theta_{\text{Indirekt}})}$$

$z = 0$ Übereinstimmung der Ergebnisse

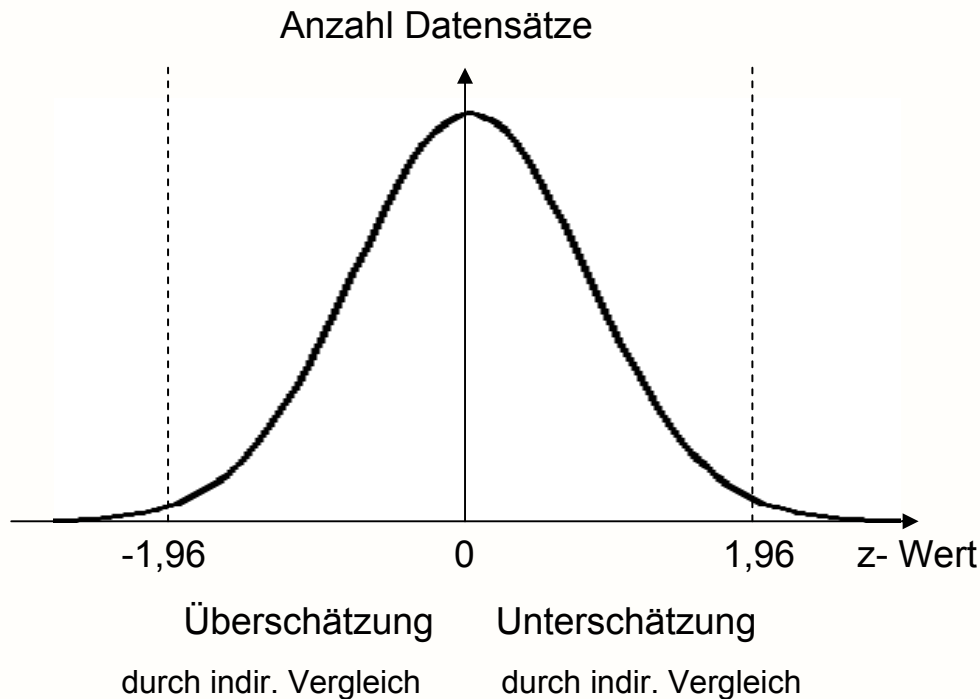
$z > 1,96$ Indirekter Vergleich unterschätzt den Therapieeffektunterschied
statistisch signifikant

$z < -1,96$ Indirekter Vergleich überschätzt den Therapieeffektunterschied
statistisch signifikant



1. Teil der Validitätsprüfung:

Test auf systematische Über – oder Unterschätzung des Therapieeffektunterschiedes durch den indirekten Vergleich



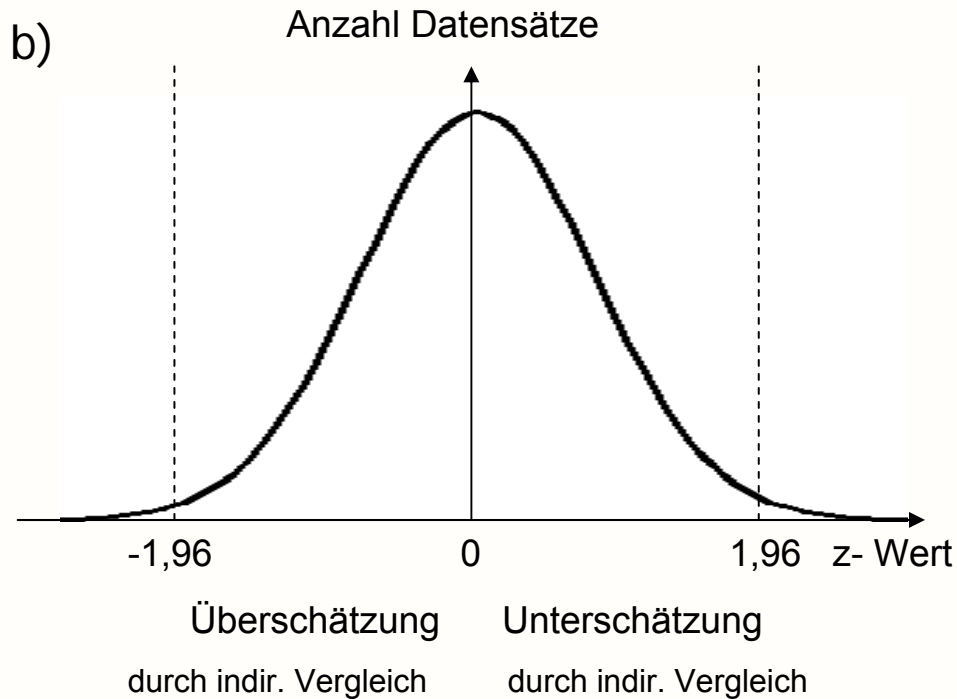
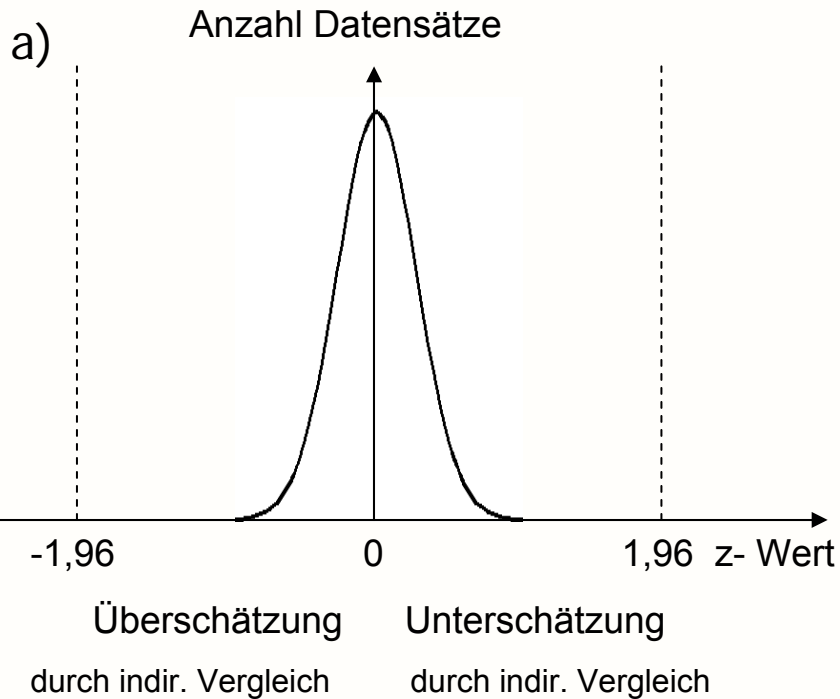
Nullhypothese: Der indirekte Vergleich führt nicht zu systematischer Über- oder Unterschätzung ($\bar{z} = 0$ bei Normalverteilung der z-Werte)

Ergebnis: Nullhypothese wird für keine der Methoden widerlegt.



2. Teil der Validitätsprüfung: Höhe der durchschnittlichen Diskrepanz I

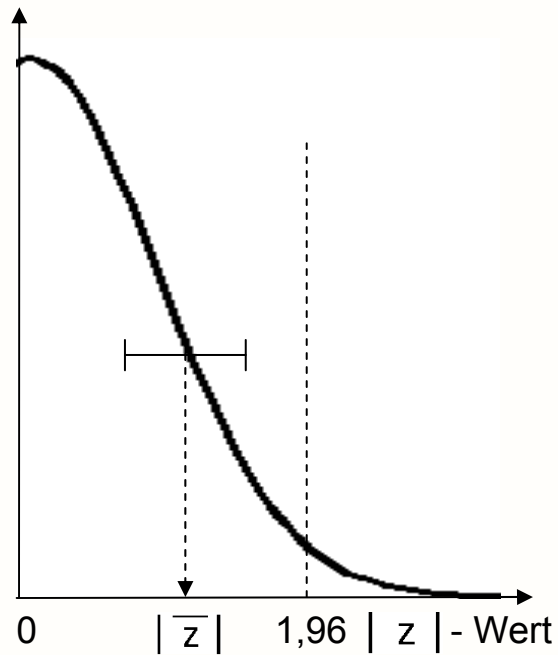
Unterschied, ob z-Werte verteilt sind, wie unter a) oder b) :



2. Teil der Validitätsprüfung: Höhe der durchschnittlichen Diskrepanz II

Berechnung der Beträge der z-Werte und ihrem Durchschnittswert: $|\bar{z}|$

Anzahl Datensätze

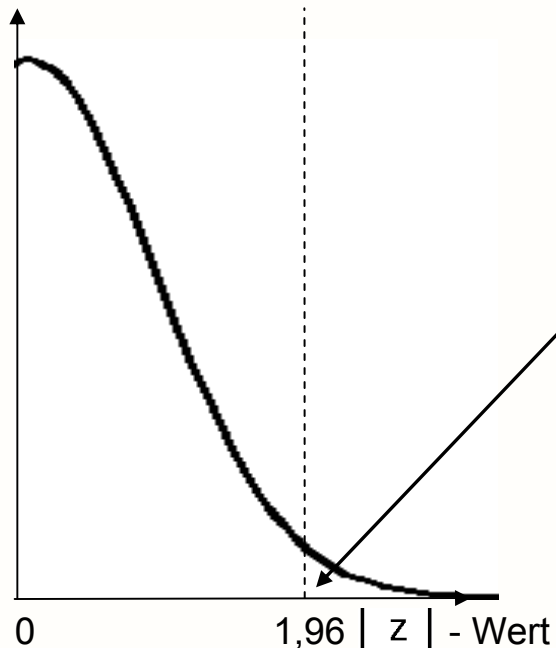


Diskrepanz zwischen indirektem
und direktem Vergleich

Methode	n für Validitätsprüfung	Durchschnittliche Diskrepanz
Nicht-adjustierter indirekter Vergleich	47	$ \bar{z} = 1,63$ (95%CI: 1,20 bis 2,07)
Adjustierter indirekter Vergleich	116	$ \bar{z} = 0,95$ (95%CI: 0,80 bis 1,09)
Mixed Treatment Comparison	57	$ \bar{z} = 0,59$ (95%CI: 0,45 bis 0,73)
Meta-Regression	4	$ \bar{z} = 0,99$ (95%CI: 0,20 bis 1,79)

3. Teil der Validitätsprüfung: Anteil der statistisch signifikant diskrepanten Datensätze

Anzahl Datensätze

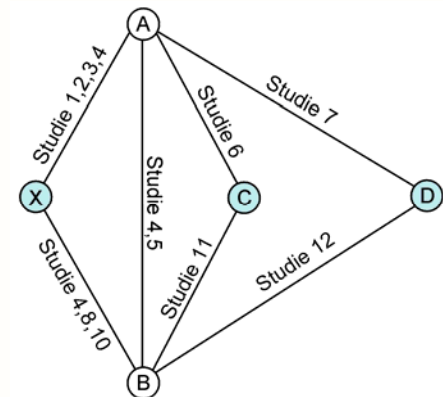
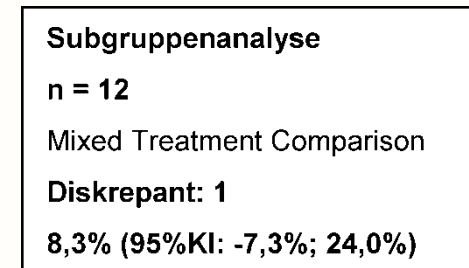
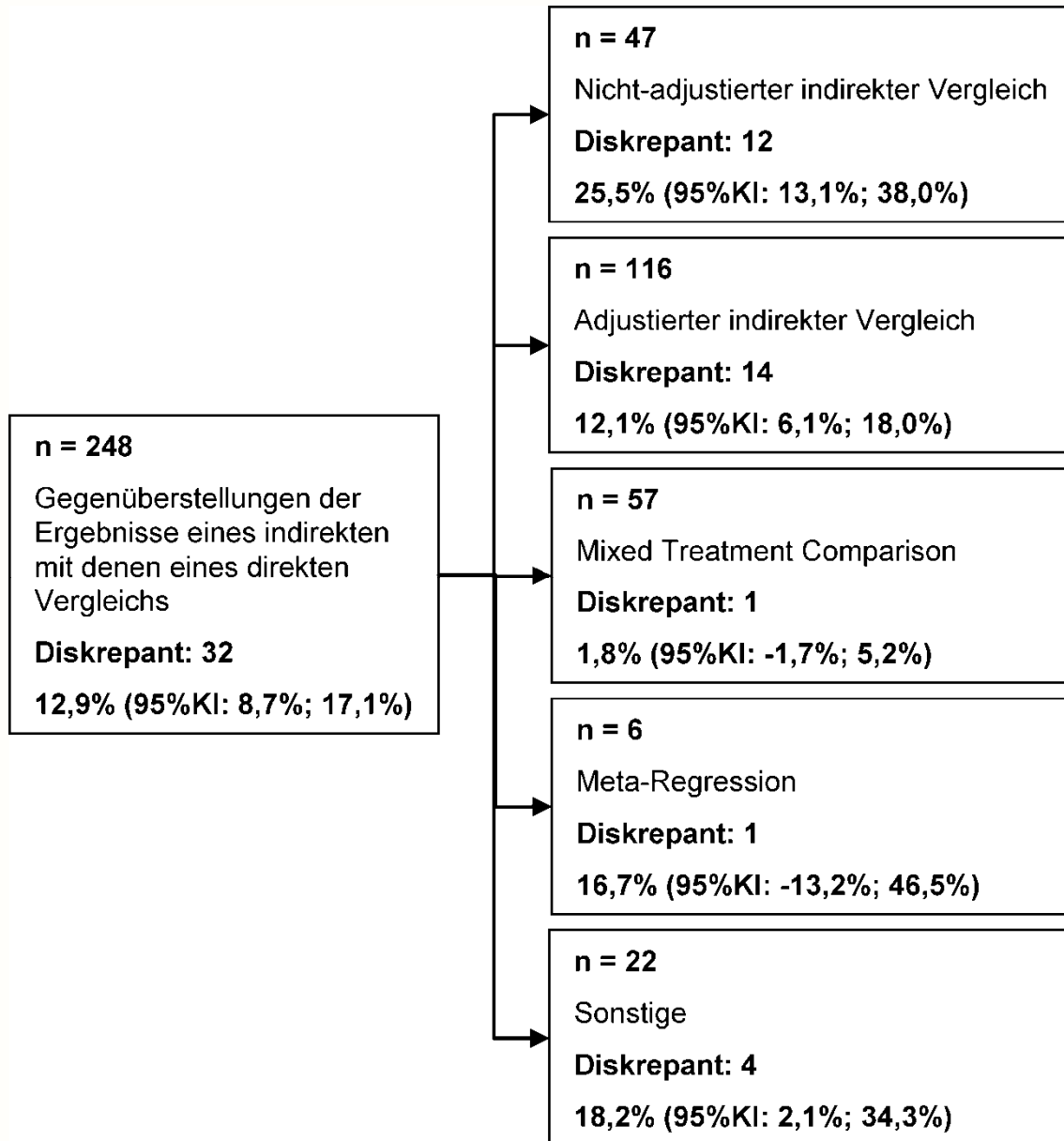


Formel

$$\frac{\text{Anzahl Datensätze mit } | z | > 1,96}{\text{Anzahl Datensätze insgesamt}} \bullet 100\%$$

Diskrepanz zwischen indirektem
und direktem Vergleich

Ergebnisse: 3. Teil der Validitätsprüfung



Limitationen der Validitätsprüfung

- Empirische Evidenz
- Validität des direkten Vergleichs
 - In denjenigen **32 Datensätzen** mit statistisch signifikanter Diskrepanz zwischen direktem und indirektem Vergleich lag bei **8 direkten Vergleichen** statistisch signifikante Heterogenität ($p < 0,10$) vor, und bei **7** wurde die Heterogenität nicht erfasst oder nicht berichtet.

Bewertung der Ergebnisse der Validitätsprüfungen

- Auf Basis der empirischen Daten:

Weist ausreichende Validität auf:	Weist keine ausreichende Validität auf:	Bisher nicht beurteilbar
Adjustierter indirekter Vergleich	Nicht-adjustierter indirekter Vergleich	Indirekter Vergleich mittels Meta-Regression
MTC, die vorhandene Head-to-head-Studien mit einschließt		MTC, die keine Head-to-head-Studien mit einschließt
		Sonstige

- Bestätigung der Ergebnisse des HTA von Glenny et al.



Zusammenfassung

- Head-to-Head- Studien sind der Goldstandard für Wirksamkeitsvergleiche therapeutischer Interventionen.
 - Bei Mangel an Head-to-Head- Studien können indirekte Vergleiche deren Ergebnisse in den meisten Fällen verlässlich abschätzen (empirische Evidenz). Ausnahme: Nicht-adjustierte indirekte Vergleiche.
 - Die Voraussetzungen für die Durchführung eines indirekten Vergleichs unterscheiden sich nicht von denen einer konventionellen Meta-Analyse -> Homogenität des Studienpools
- > Ungeklärte Frage: Ab welchem Ausmaß an Heterogenität sollte man auf indirekte Vergleiche verzichten?



Vielen Dank für Ihr Interesse!

**DIMDI-Bericht frei über das DIMDI oder
German Medical Science (gms) verfügbar:**



http://portal.dimdi.de/de/hta/hta_berichte/hta243_bericht_de.pdf

Kontaktdaten des Vortragenden:

Ben Schöttker
Abt. Klinische Epidemiologie und
Alternsforschung
im Deutschen Krebsforschungszentrum
Bergheimer Straße 20
69115 Heidelberg

Email: b.schoettker@dkfz.de



Zusatzfolie: Ausschlusskriterien des Reviews

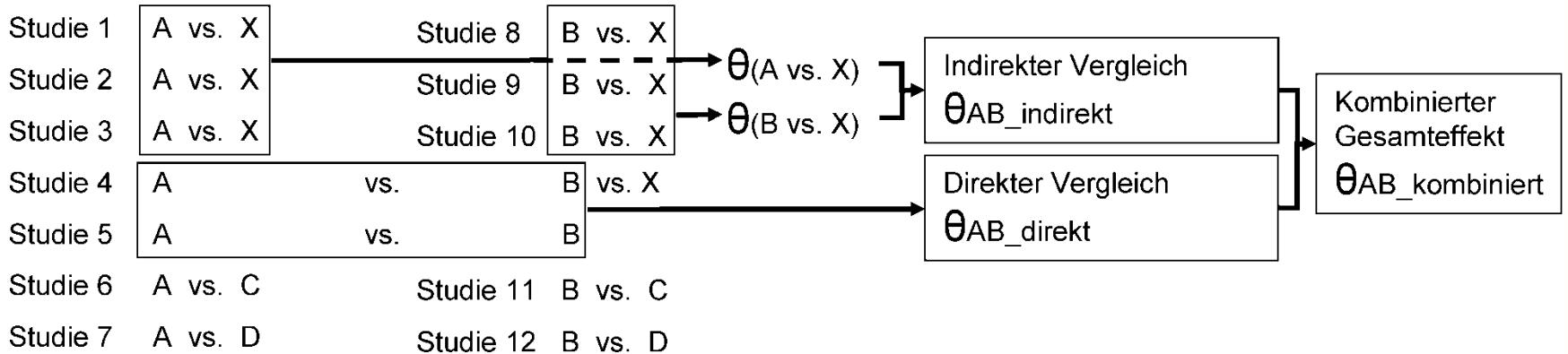
Ausschlusskriterien Volltextselektion

1. Einzelstudien
2. Übersichtsarbeiten, Editorials und Kommentare, bei denen es sich nicht um systematische Übersichtsarbeiten handelt (Definition eines systematischen Reviews im Glossar).
3. Systematische Übersichtsarbeiten zu diagnostischen Testverfahren und komplexen Interventionen (z.B. im Rahmen umfangreicher Präventionskampagnen).
4. Übersichtsarbeiten, in denen nicht über die Durchführung und die Ergebnisse eines indirekten Vergleichs der Wirksamkeit therapeutischer Intervention berichtet wird.
5. Publikationen, ausschließlich methodischen Inhalts, die keine Anleitung geben, wie indirekte Vergleiche therapeutischer Interventionen durchzuführen sind.
6. Mehrfachpublikationen mit identischem Inhalt.
7. Kosten-Nutzen-Betrachtungen, für die keine systematische Literaturrecherche durchgeführt wurde.
8. Der Volltext konnte nicht beschafft werden, da der Artikel nicht über Subito bestellbar war, nur in einem Buch erschienen ist, oder bisher nur als Abstract oder Protokoll publiziert wurde.
9. Systematische Übersichtsarbeiten, die den indirekten Vergleich überwiegend auf Basis von nicht-randomisierten oder nicht-kontrollierten Studien durchführen.

Zusatzfolie - Zusammenfassung von direktem und adjustiert indirektem Vergleich

Studien mit Verumarm A

Studien mit Verumarm B



Legende:

A, B, C, D : Therapieeffekt von Therapieoption
 A, B, C oder D
 X: Gemeinsamer Komparator
 θ : Gepoolter Effekt
 vs. : Differenz der Therapieeffekte