

Interpretation von Ergebnissen geringer Güte zu unerwünschten Ereignissen im Rahmen systematischer Übersichten am Beispiel von Studien zur benignen Prostatahyperplasie

IQWiG im Dialog, 19.06.2009, Köln

Guido Skipka, Katharina Biester, Stefan Lange

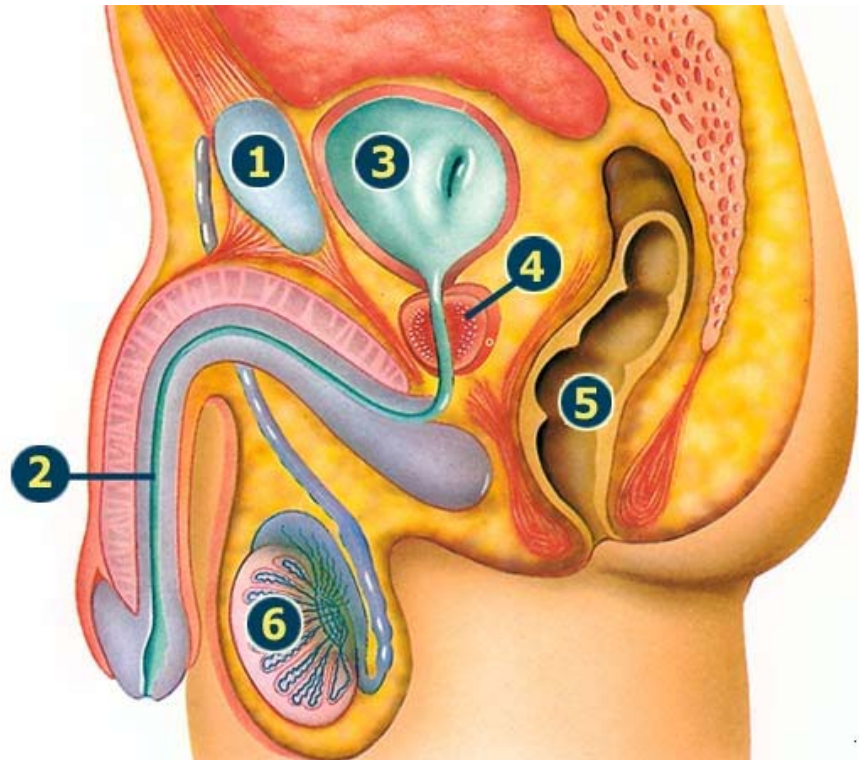
Inhalt

- **Projekt N04-01**
 - **Fragestellung**
 - **patientenrelevante Endpunkte**
 - **Nutzen und Schaden**
 - **Studien- und Ergebnislage**
- **Algorithmus zur Bewertung von Ergebnissen zu unerwünschten Ereignissen**

Begriff

- **Benigne Prostatahyperplasie (BPH) = veralteter Begriff**
- **Heute: Benignes Prostatasyndrom (BPS)**

- ① Schambein
- ② Harnröhre
- ③ Blase
- ④ Prostata
- ⑤ Darm
- ⑥ Hoden



Fragestellung

- **Vergleichende Nutzenbewertung**
 - **nichtmedikamentöser lokaler Verfahren vs.**
 - **Standard**
 - **anderweitige Interventionen (kontrolliertes Zuwarten, Sham, Medikamente)**
 - **nichtmedikamentöser lokaler Verfahren untereinander**

nichtmedikamentöse Verfahren

- **Standard**
 - **TURP (Transurethrale Resektion der Prostata)**
 - **Modifikationen der TURP**
 - **offene Adenomenukleation**
 - **TUIP (Transurethrale Inzision der Prostata)**

- **lokale Verfahren (Prüfverfahren)**
 - **Laser**
 - **Mikrowellen**
 - **hochfrequenter Wechselstrom**
 - **Ultraschall**
 - **heißes Wasser**

Zielgrößen

- **Symptomatik (z. B. gestörte Harnblasenentleerung, Harndrang, Startverzögerung bei der Miktion, häufiges Wasserlassen)**
- **Gesundheitsbezogene Lebensqualität**
- **Unerwünschte Ereignisse (UEs)**
- **Dauer von Krankenhausaufenthalten**
- **Dauer und Notwendigkeit von Katheterisierung**

Studientypen

- **randomisierte Studien (RCTs)**
- **nicht randomisierte Studien (CCTs) mit zeitlich paralleler Kontrolle und erkennbarer Berücksichtigung prognostischer Faktoren**

Studienlage

- **In die Nutzenbewertung eingeschlossen: 56 Studien (RCTs: n=55, CCTs: n=1)**
- **Fallzahlen: 30 – 400 (Median 100)**
- **20 Verfahren: 15 Prüfverfahren, 3 Standardverfahren (+ 2 Modifikationen)**
- **27 Vergleiche**

Zusatznutzen – Herangehensweise A

- **Symptomatik beeinflusst maßgeblich den Leidensdruck der Patienten**
- **i. d. R. Symptomscore primärer Endpunkt in Studien**
- **Zusatznutzen: ein weniger invasives Verfahren führt zu einer stärkeren positiven Beeinflussung der Symptomatik als die Standardtherapie**

Für keines der Prüfverfahren zeigte sich ein Zusatznutzen gegenüber Standard.

Zusatznutzen – Herangehensweise B

- **Rationale zur Entwicklung & Einführung der weniger invasiven Verfahren: bei nur geringfügig schlechteren Ergebnissen hinsichtlich der Symptomatik führen sie zu geringeren Raten von unerwünschten Ereignissen**
- **Zusatznutzen:**
 - **höchstens irrelevante Unterlegenheit** des Prüfverfahrens gegenüber der Standardtherapie bez. der **Symptomatik**
 - **gleichzeitig nachgewiesene Vorteile** hinsichtlich des Auftretens **unerwünschter Ereignisse**.

Für keines der Prüfverfahren zeigte sich eine höchstens irrelevante Unterlegenheit gegenüber Standard.

Betrachtung von UEs beim BPS

- **Senkung der Komplikationen ist (ein) primäres Ziel der lokalen Verfahren**
- **UEs sind Bestandteil der Nutzenbewertung**

unerwünschte Ereignisse

- **Ejakulationsdysfunktion, erektile Dysfunktion**
- **Miktions symptomatik**
- **Inkontinenz**
- **Retention**
- **Rekatheterisierung**
- **Harnwegsinfektion, sonstige Infektion**
- **Striktor**
- **Reintervention**
- **schwere Blutung**
- **Tod**

Datenlage zu UEs → „sehr unübersichtlich“

- **divergierende Definitionen der UEs zwischen den Studien**
- **fehlende Operationalisierung der UEs**
- **fehlende Erklärung der dargestellten Ergebnisse**
- **Intention-to-treat-Prinzip oftmals grob verletzt**
 - *„Patients who required reoperation were excluded from further analysis“ (4/20 vs. 0/20)*
 - *„Patients having repeat laser were included in the follow-up and only those undergoing TURP excluded“ (? vs. ?)*
- **lediglich narrative Angaben zu UEs**
 - *„Most of the patients had suffered from irritative symptoms, which were more prominent in the TURP group“*

- **selektive Berichterstattung (nur für statistisch signifikante Ergebnisse werden Häufigkeiten dargestellt)**
- **Gesamtanzahl an Patienten, die ein UE erlitten hatten, waren oftmals nicht berichtet/bestimmbar**
 - **große Ergebnisunsicherheit**
 - **ungenügende Evidenzlage**
 - **Einschätzung der Evidenzlage sehr schwierig**

unsicher

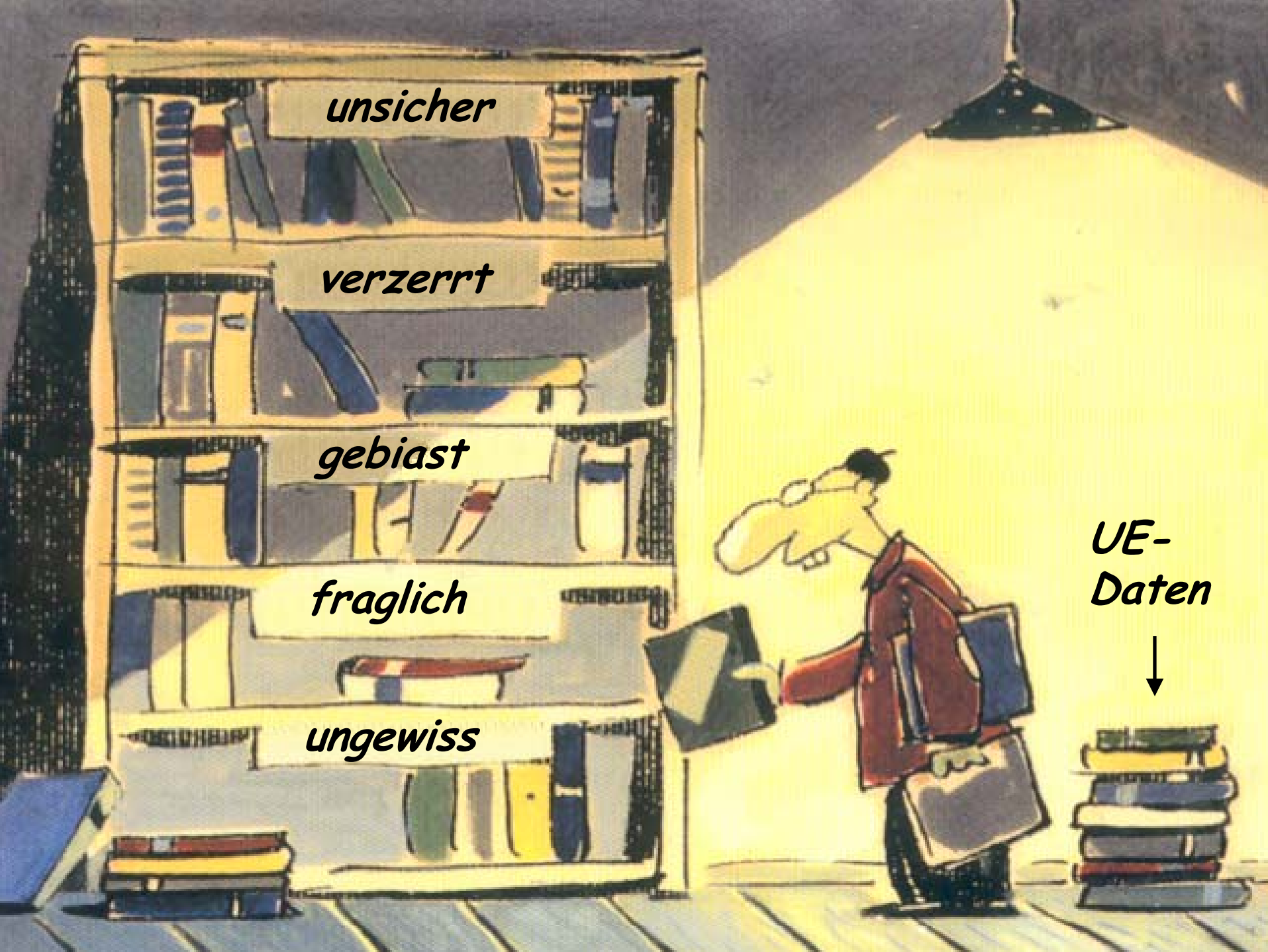
verzerrt

gebiast

fraglich

ungewiss

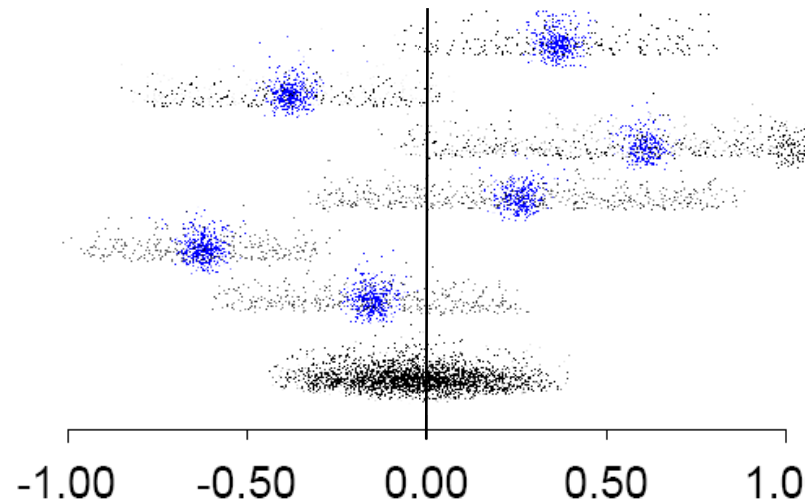
*UE-
Daten*



Wie können – trotz der ungenügenden Datenlage – überhaupt Aussagen zum Auftreten von UEs abgeleitet werden ?

- **Ziel: Algorithmus zur Ableitung von Hinweisen**
- **Berücksichtigung folgender Punkte:**
 - **Reproduzierbarkeit**
 - **Präzision**
 - **Homogenität**
 - **Größe des Effekts**

üblicherweise: **Meta-Analyse** (statistische Berechnung eines Gesamtergebnisses)



verworfen, da

- Verzerrungspotenzial der Ergebnisse zu hoch
- Meta-Analyse suggeriert exakte Ergebnisse

Qualitativerer Ansatz Betrachtung pro Vergleich und UE

- **Betrachtung des Anteils der Patienten pro Gruppe, die das UE erlitten hatten**
- **diskrete Einstufung des Gruppenunterschieds pro Studie (Δ = Differenz der Raten)**
 - **$|\Delta| < 0,05 \rightarrow$ kein Anzeichen für Effekt**
 - **$|\Delta| \geq 0,05 \rightarrow$ Anzeichen für Effekt (inkl. Richtung)**
- **Betrachtung der Anzeichen aller Studien**
 - **Hinweis auf Effekt, falls Anzeichen homogen und gerichtet**
- **Sonderfall: nur eine Studie vorhanden**
 - **Hinweis auf Effekt, falls Fallzahl pro Gruppe mindestens 50 und $|\Delta| \geq 0,20$**

Beispiel: VLAP vs. Standard, Auftreten von schweren Blutungen

Studie	VLAP	Standard	Differenz	Anzeichen
Cowles 1995	0 / 56	2 / 59	3,4%	0
Sengör 1996	0 / 30	2 / 30	6,7%	◀
Suvakovic 1996				—
CLasP Gujral 2000	0 / 36	9 / 44	20,5%	◀
McAllister 2000	0 / 76	12 / 75	16,0%	◀
CLasP Chacko 2001	7 / 74	2 / 74	-6,8%	▶
CLasP Brookes 2003	1 / 117	1 / 117	0,0%	0
Planz 2003	1 / 56	3 / 49	4,3%	0
Razzaghi 2007				—

Bestimmung von Anzeichen für einen Effekt pro Studie.

◀ 0 ▶

Anzahl an Studien

= 1

Reproduzierbarkeit

> 1

Homogenität

Fallzahl pro
Gruppe

Präzision

Größe des Effekts

Sind die Anzeichen
homogen ?

ja

nein

< 50

≥ 50

Größe des
Effekts $|\Delta|$

Sind die
Anzeichen
gerichtet ?

< 0.20

≥ 0.20

ja

nein

kein
Hinweis

Hinweis

kein
Hinweis

Sind die Anzeichen homogen ?

- nein, falls zugunsten beider Gruppen mindestens 2 Anzeichen vorliegen

heterogen	
	0
◀	
◀	
◀	
	▶
	▶
	0
	0
	-

homogen	
	0
◀	
◀	
◀	
	▶
	-
	0
	0
	-

← möglicher
„Ausreißer“
aufgrund hohen
Verzerrungs-
potenzials

Sind die homogenen Anzeichen gerichtet ?

ja, falls für die Studien mit Anzeichen zugunsten einer Gruppe gilt:

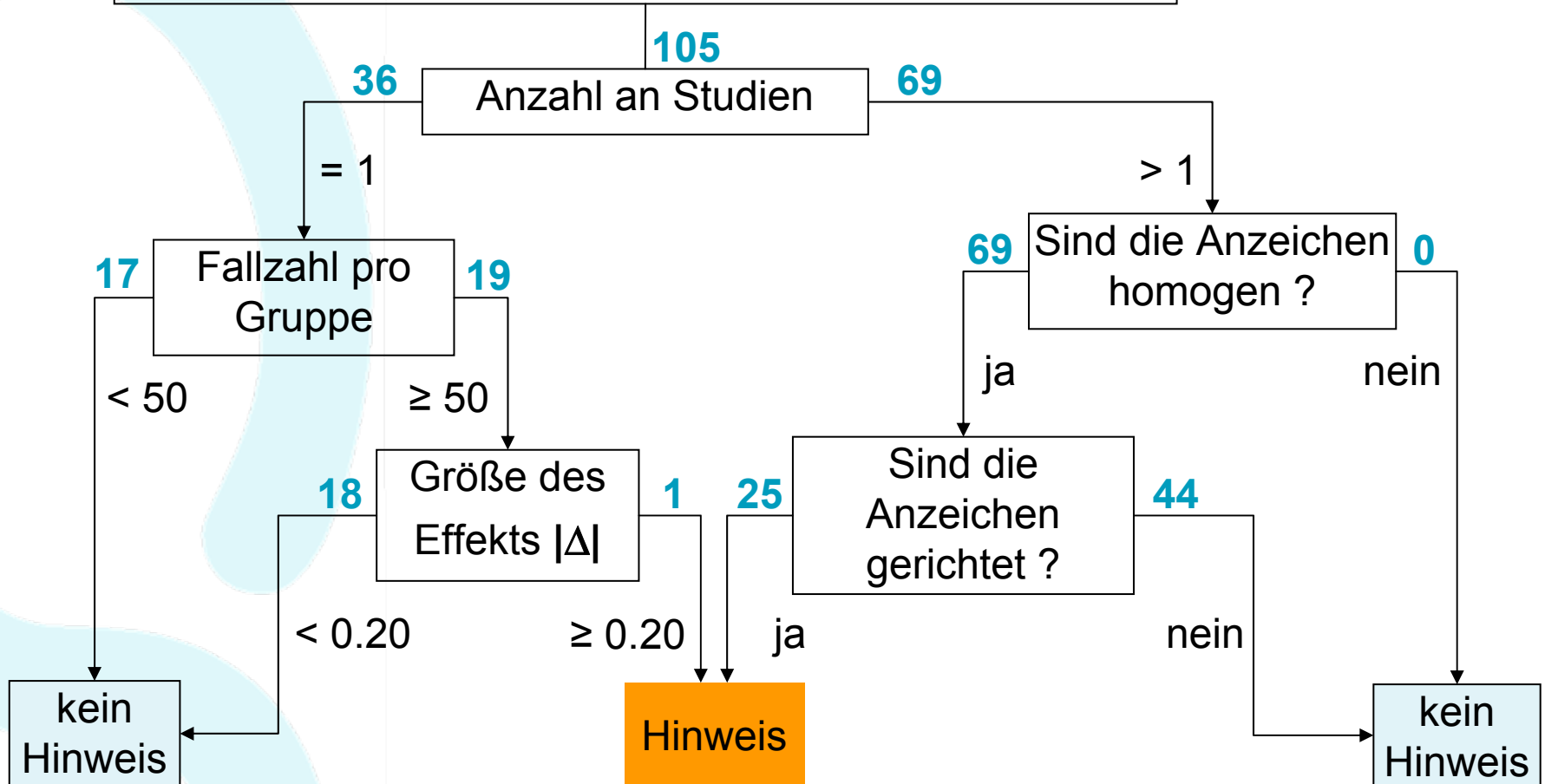
- es sind mindestens 2 Studien
- ihr Anteil an allen Studien mit Ergebnissen ist $\geq 50\%$, wobei ggf. die Studie mit entgegengesetztem Anzeichen ignoriert wird

gerichtet	
	0
▶	
▶	
▶	
	✗
	-
	0
	0
	-

nicht gerichtet	
	0
▶	
▶	
	-
	✗
	-
	0
	0
	-

Bestimmung von Anzeichen für einen Effekt pro Studie.

◀ 0 ▶



Hinweise bezogen auf UEs

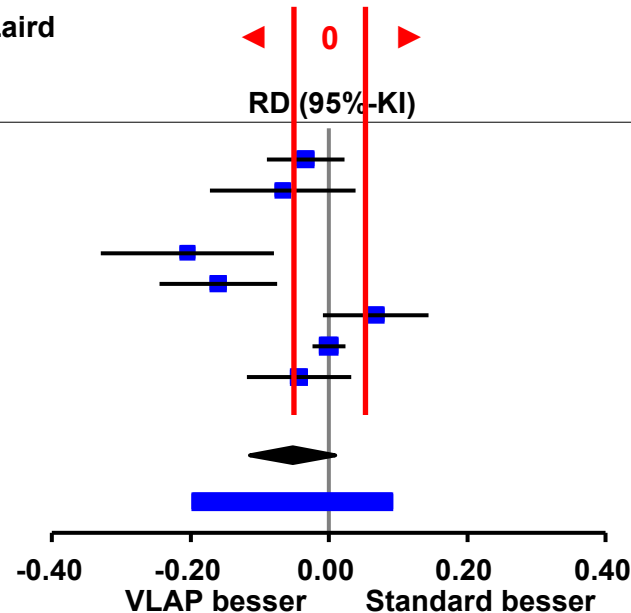
Intervention	Kontrolle	Nicht schwerwiegende UEs					Unklar		Schwerwiegende UEs						
		Ejakulationsdysfunktion	Miktionssymptomatik	Retention	Rekatheterisierung	Harnwegsinfektion	Strikturen	Erektile Dysfunktion	leichte Inkontinenz	schwere Inkontinenz	Infektion anderer Art	Reintervention	schwere Blutung	Tod	
CLAP	Standard				▲										
VLAP		▲									▼	▲			
Hybrid		KTP/Nd:YAG				▼						▼			
		CLAP/VLAP	▲		▼							▼			
ILK		▲				▼		▲			▼	▲			
HoLAP															
HoLRP															
HoLEP			▼										▲		
TUMT		▲	▼	▼	▼	▼							▲	▲	
TUNA		▲					▲	▲				▼	▲		

- ▲ Hinweis auf Vorteil der Intervention
- ▼ Hinweis auf Vorteil des Standards

Vergleich: Meta-Analyse vs. Algorithmus

VLAP vs. Standard
Auftreten schwerer Blutungen
Modell mit zufälligen Effekten - DerSimonian und Laird

Studie	VLAP n/N	Standard n/N	RD (95%-KI)	Gewichtung	RD	95%-KI
Cowles 1995	0/56	2/59		16.2	-0.03	[-0.09, 0.02]
Sengör 1996	0/30	2/30		12.1	-0.07	[-0.17, 0.04]
Suvakovic 1996	—/—	—/—	—	—	—	—
CLasP Gujral 2000	0/36	9/44		10.6	-0.20	[-0.33, -0.08]
McAllister 2000	0/76	12/75		13.8	-0.16	[-0.25, -0.07]
CLasP Chacko 2001	7/74	2/74		14.5	0.07	[-0.01, 0.14]
CLasP Brookes 2003	1/117	1/117		18.2	0.00	[-0.02, 0.02]
Planz 2003	1/56	3/49		14.6	-0.04	[-0.12, 0.03]
Razzaghi 2007	—/35	—/52	—	—	—	—
Gesamt	9/445	31/448		100.0	-0.05	[-0.11, 0.01]
95% Heterogenitätsintervall						[-0.20, 0.09]



Heterogenität: $Q=36.92$, $df=6$, $p<0.001$, $I^2=83.7\%$
Gesamteffekt: $Z\text{ Score}=-1.67$, $p=0.095$, $\tau=0.073$

Zusammenfassung

- **Einbezug von UEs in die Nutzenbewertung bei N04-01 notwendig, da das primäre Ziel der Prüfindervention UEs beinhaltet**
- **Evidenzlage zu UEs im nichtmedikamentösen Bereich oft ungenügend**
- **Algorithmus zur Ableitung von Hinweisen durch „diskretes“ Zusammenfassen anwendbar**