

# Evidenz-informierte Gesundheitspolitik: (wie) kann das funktionieren?

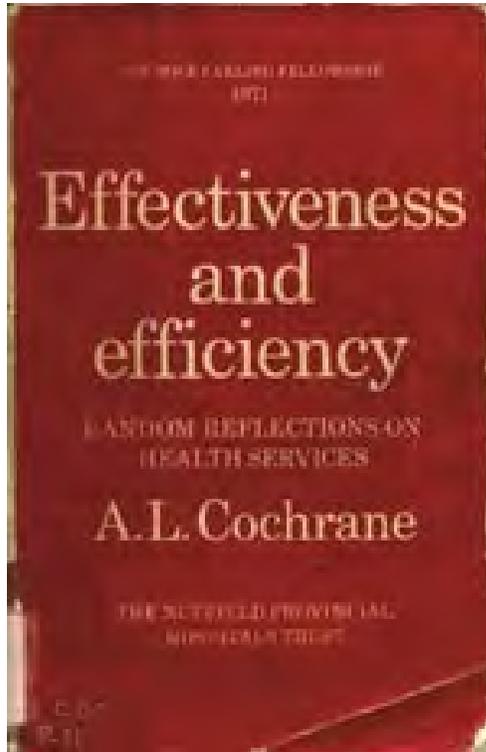
**Reinhard Busse, Prof. Dr. med. MPH**

FG Management im Gesundheitswesen, Technische Universität Berlin  
(WHO Collaborating Centre for Health Systems Research and Management)

&

European Observatory on Health Systems and Policies

# Probleme wie mangelnder Wirksamkeitsnachweis,



**Table.** Estimates of Annual US Health Care Waste, by Category<sup>a</sup>

	\$ in Billions					
	Annual Cost to Medicare and Medicaid in 2011 <sup>b</sup>			Annual Cost to US Health Care System in 2011		
	Low	Midpoint	High	Low	Midpoint	High
Failures of care delivery	26	36	45	102	128	154
Failures of care coordination	21	30	39	25	35	45
Overtreatment	67	77	87	158	192	226
Administrative complexity	16	36	56	107	248	389
Pricing failures	36	56	77	84	131	178
Fraud and abuse	30	64	98	82	177	272
<b>Total<sup>c</sup></b>	<b>197</b>	<b>300</b>	<b>402</b>	<b>558</b>	<b>910</b>	<b>1263</b>
<b>% of Total Spending</b>				<b>21</b>	<b>34</b>	<b>47</b>

<sup>a</sup>Table entries represent the range of estimates of waste in each category from sources cited in the text. The total waste estimates are simply the sums of the category-level estimates. This simple summing is feasible because the categories are defined in such a way that wasteful behaviors could be assigned to at most 1 category and because, like Pacala and Socolow,<sup>9</sup> we did not attempt to estimate interactions between or among the categories.  
<sup>b</sup>Including both state and federal costs.  
<sup>c</sup>Totals may not match the sum of components due to rounding.

Verschwendung,

**Table 3. Variation in number of surgical procedures performed per 10,000 persons for the 13 Vermont hospital service areas and comparison populations, Vermont, 1969. (Rates adjusted to Vermont age composition.)**

Surgical procedure	Lowest two areas	Entire state	Highest two areas
Tonsillectomy	13 32	43	85 151
Appendectomy	10 15	18	27 32
Hemorrhoidectomy	2 4	6	9 10
<b>Males</b>			
Hernioplasty	29 38	41	47 48
Prostatectomy	11 13	20	28 38
<b>Females</b>			
Cholecystectomy	17 19	27	46 57
Hysterectomy	20 22	30	34 60
Mastectomy	12 14	18	28 33
Dilation and curettage	30 42	55	108 141
Varicose veins	6 7	12	24 28

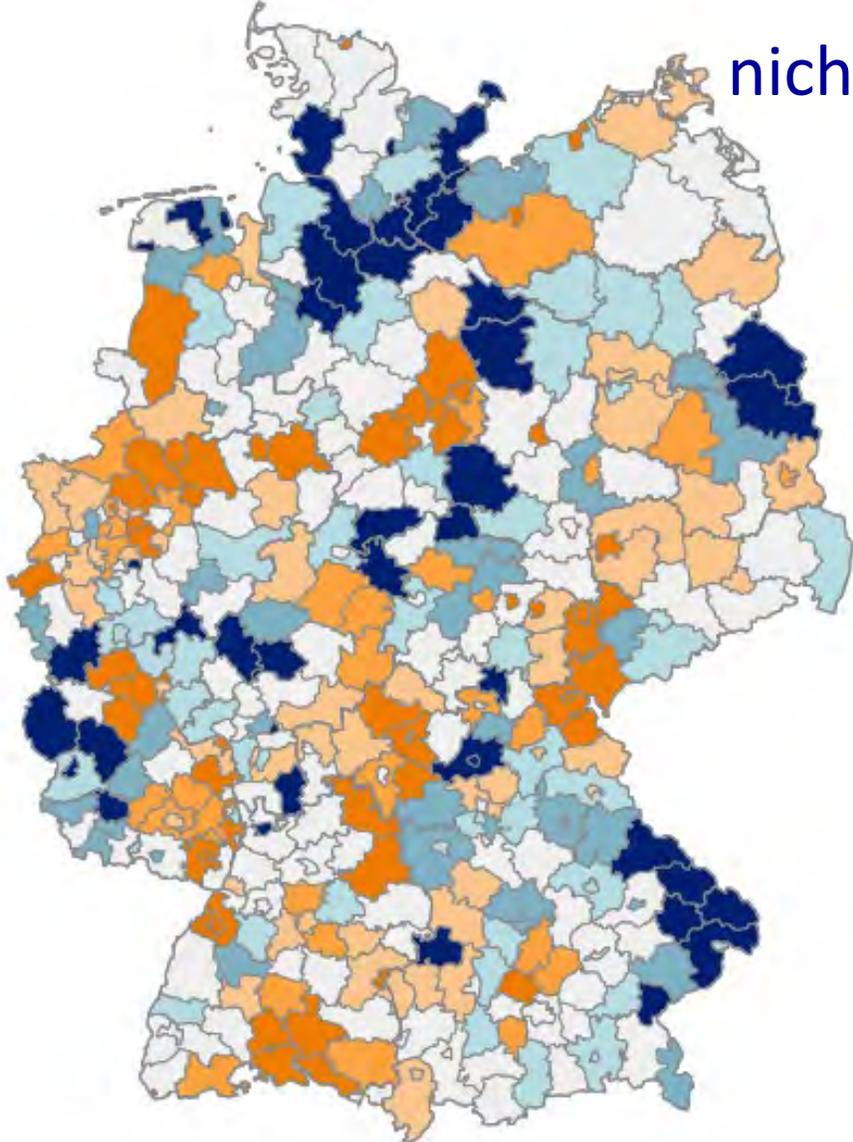


## TO ERR IS HUMAN: Qualitätsmängel, BUILDING A SAFER HEALTH SYSTEM

**H**ealth care in the United States is not as safe as it should be--and can be. At least 44,000 people, and perhaps as many as 98,000 people, die in hospitals each year as a result of medical errors that could have been prevented, according to estimates from two major studies. Even using the lower estimate, preventable medical errors in hospitals exceed attributable deaths to such feared threats as motor-vehicle wrecks, breast cancer, and AIDS.

unerklärbare Variationen ...

# nicht nur woanders ... sondern auch bei uns



Am **häufigsten** werden betroffene Männer operiert aus:

Altenburger Land: 13,0

Kitzingen: 11,8

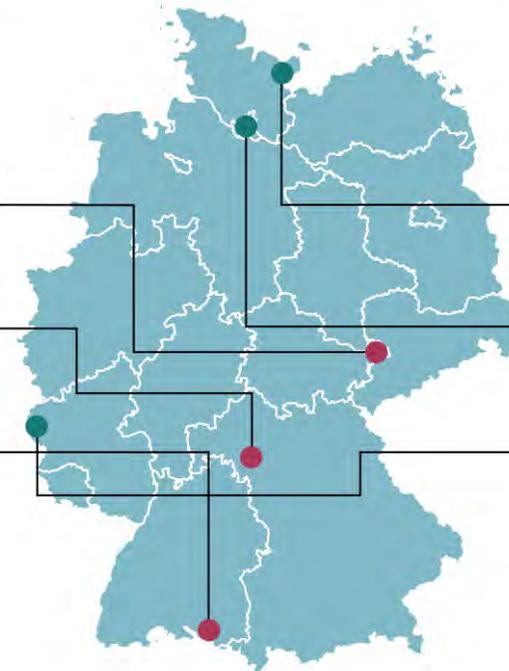
Bodenseekreis: 11,6

Am **seltentsten** werden betroffene Männer operiert aus:

Ostholstein: 2,7

Hamburg: 2,8

Eifelkreis Bitburg-Prüm: 1,7



je 10.000 Männer, 2010 bis 2012  
Berechnungen bezogen auf den Wohnortkreis der Patienten, Daten nach Alter bereinigt.

Warum führt eigentlich niemand solche Erhebungen mehr durch?

LEGENDE: ENTFERNUNG DER PROSTATA JE 10.000 MÄNNER, 2010-2012



## Natürlich hat der Gesetzgeber reagiert ...

- Externe Qualitätssicherung
- Gemeinsamer Bundesausschuss mit Qualitätsmandat
- IQWiG
- Versorgungsforschung
- IQTIG ...

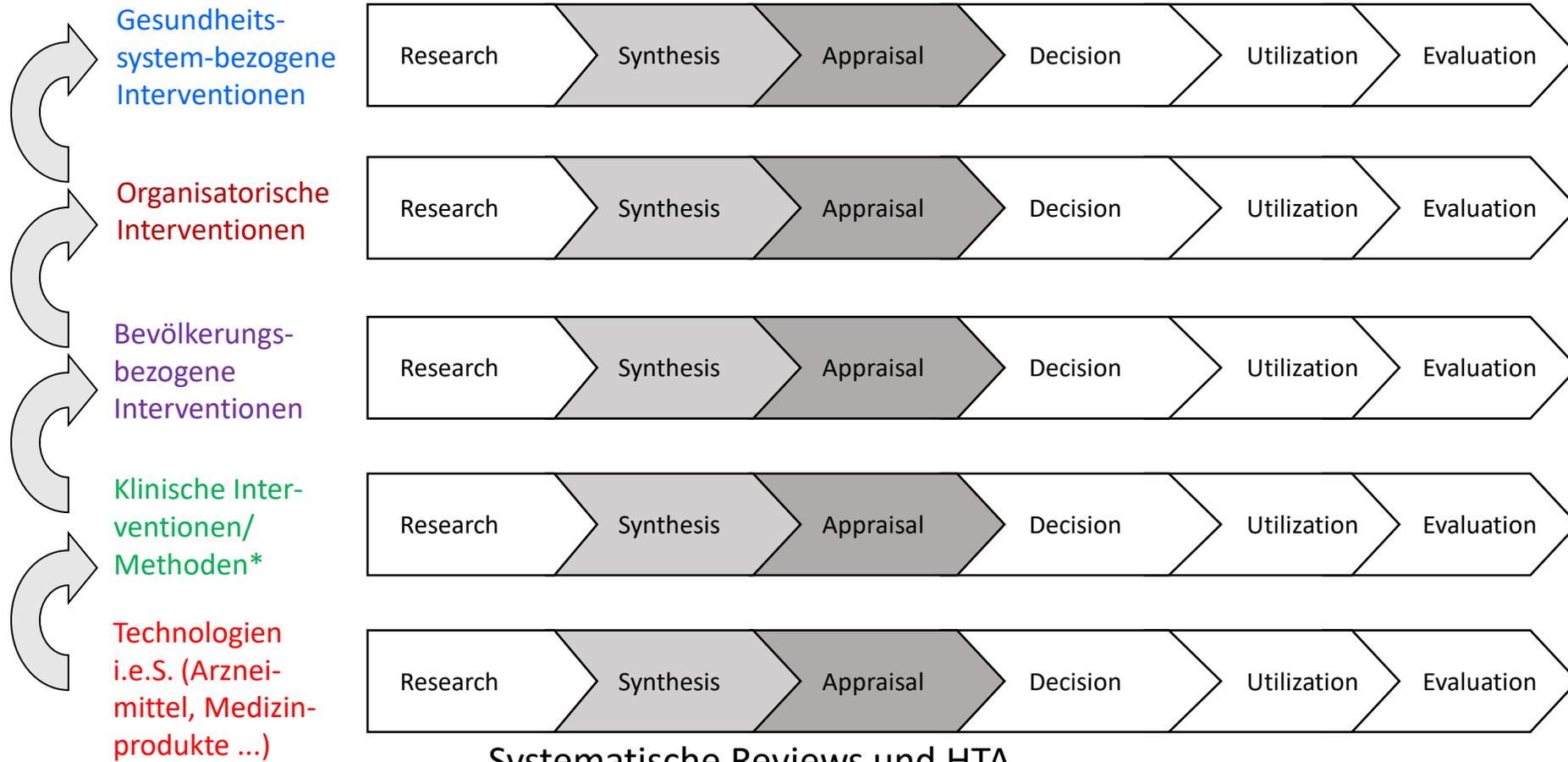
... aber führt das zu evidenzinformierter Gesundheitspolitik?

... trotz begrenzter Einsicht bei Verantwortlichen ...



Zitat Spahn: Wir haben das beste Gesundheitssystem der Welt

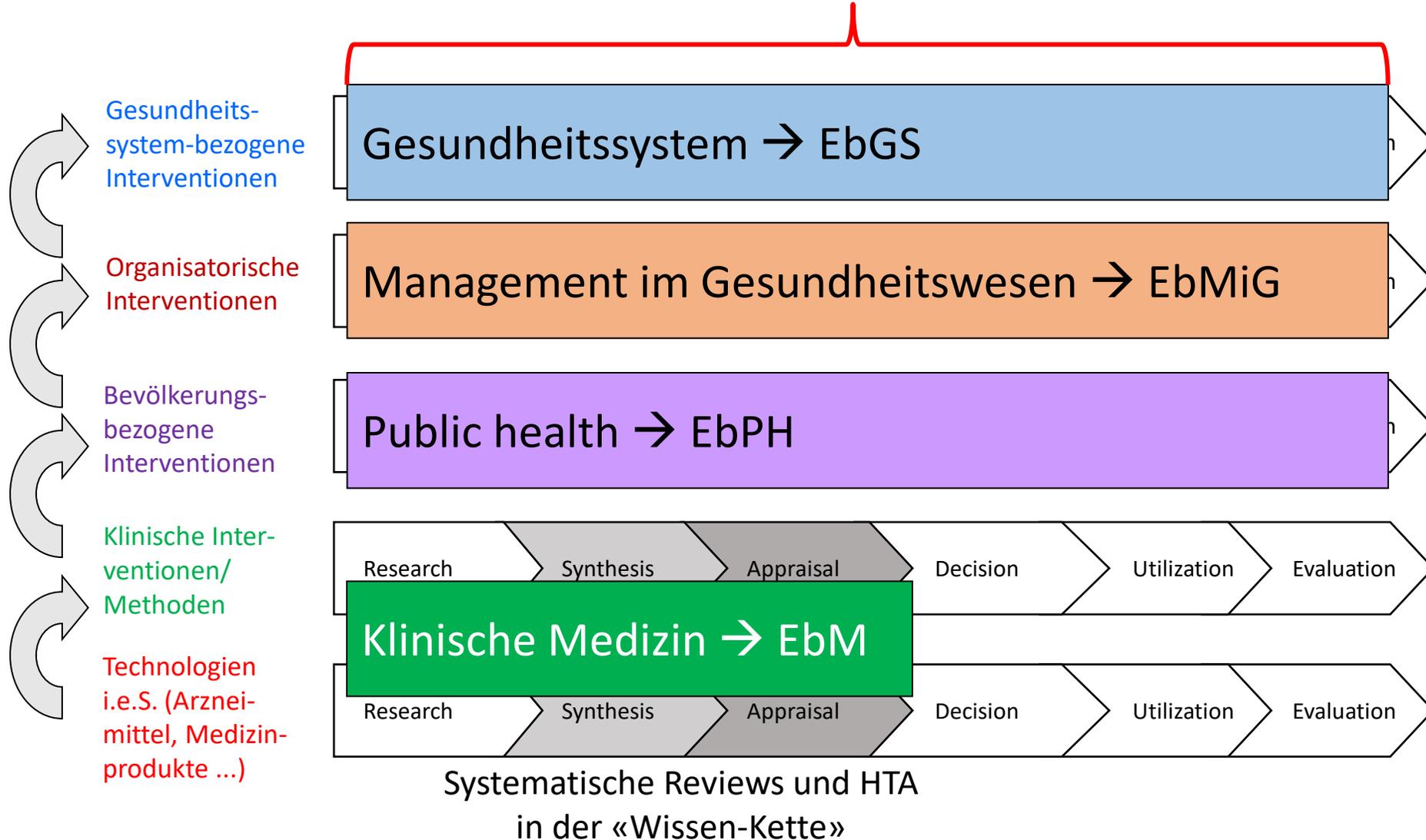
# Problem 1: breites Verständnis von “Technologien” und konsequente Aufarbeitung für deren Evidenz fehlt



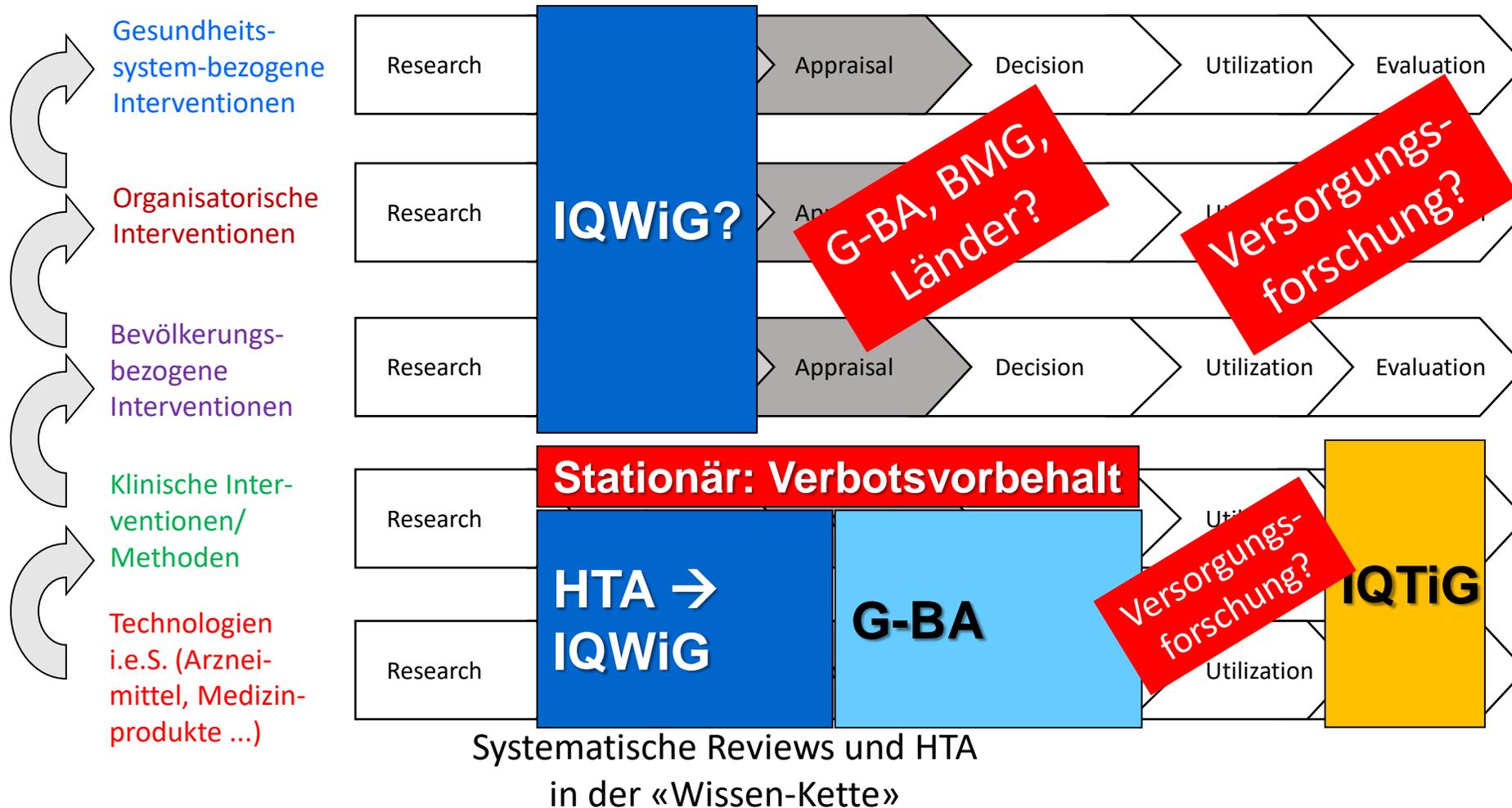
Systematische Reviews und HTA  
in der «Wissen-Kette»

... unabhängig davon, wie wir jede Ebene nennen

## Evidenz-informierte Gesundheitspolitik



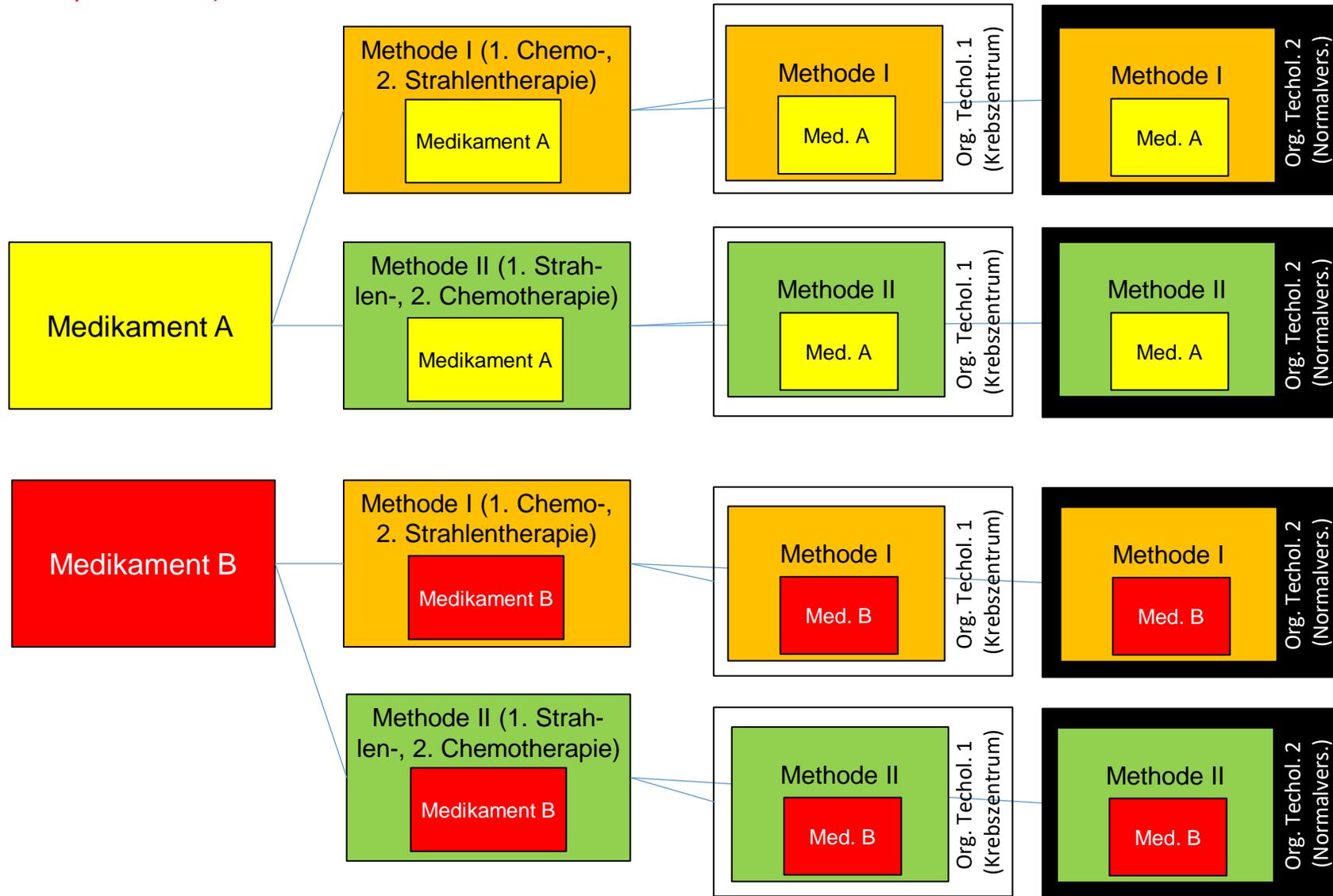
# Punkt 1: Wir müssen unsere Herangehensweise ändern und klären, wer zuständig ist ...



Technologien i.e.S.  
(Arzneimittel, Me-  
dizinprodukte ...)

Klinische Interven-  
tionen/ Methoden

Organisatorische Interventionen



... und dabei  
das  
Zusammen-  
spiel zwischen  
Technologien  
nicht  
vergessen

Gute Evidenz  
müsste für  
alle 8  
Gruppen  
vorhanden  
sein  
→ nicht  
durch 8-  
armige RCTs  
zu lösen!

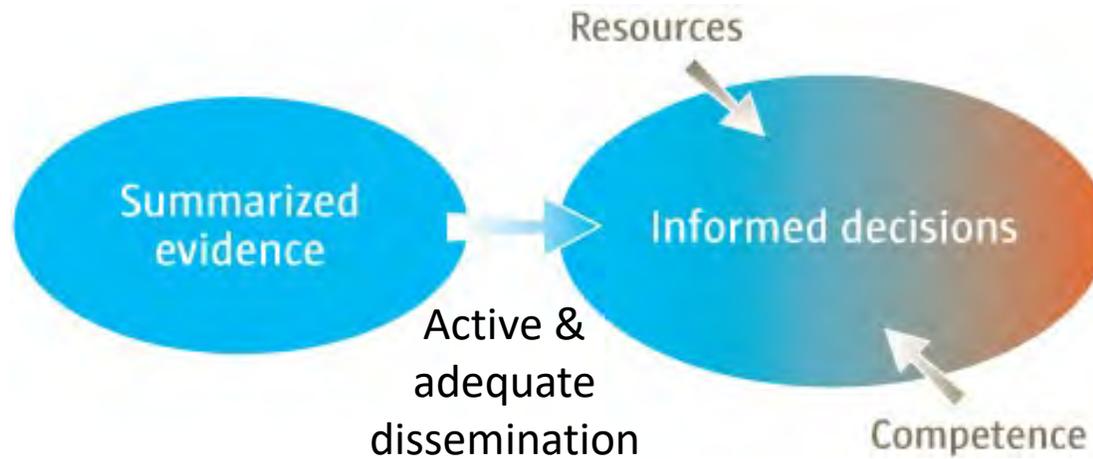
# Erweiterung 2: Erweiterung der akzeptablen Evidenz (und Berücksichtigung der Rückkopplung)

Published literature + “Grey” literature

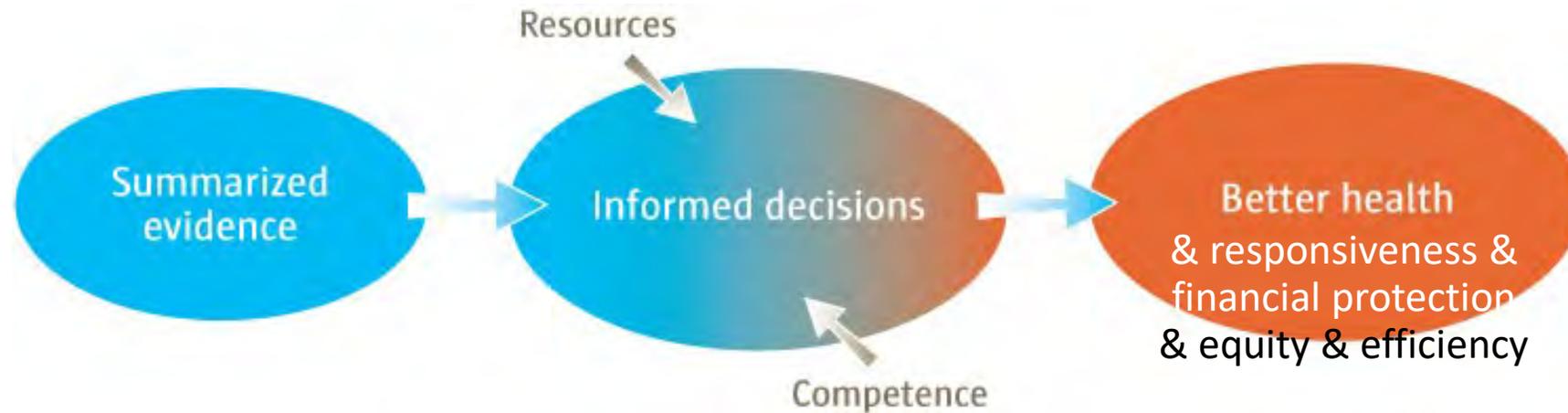


+ study data

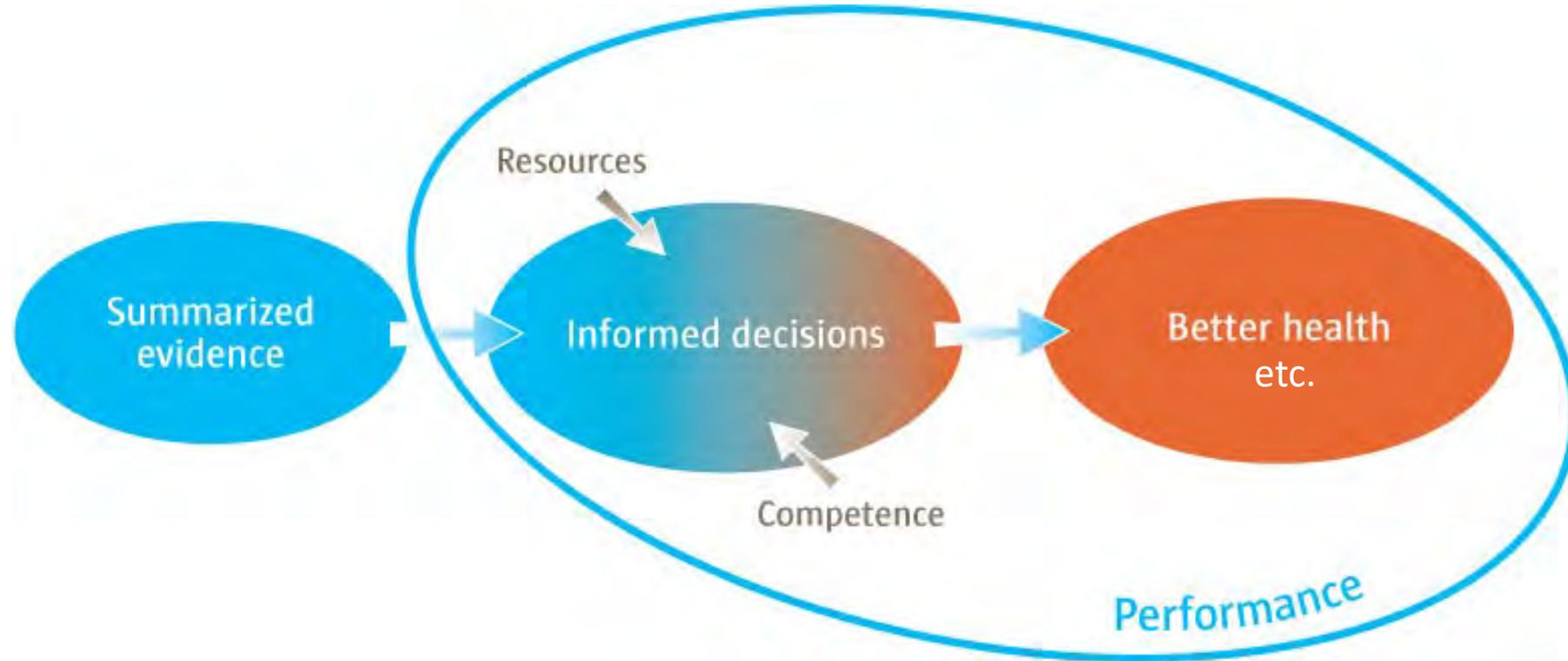
# Erweiterung 2: Erweiterung der akzeptablen Evidenz (und Berücksichtigung der Rückkopplung)



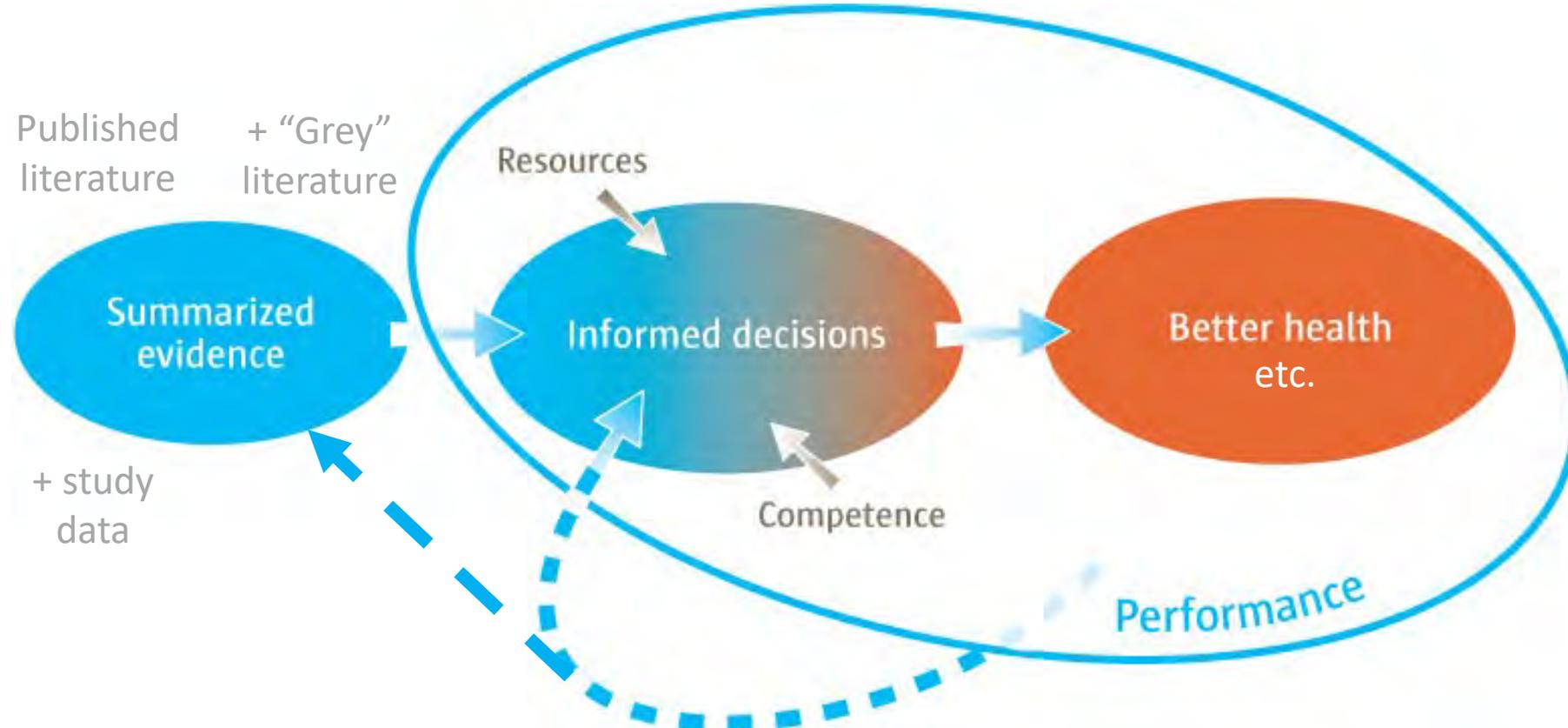
# Erweiterung 2: Erweiterung der akzeptablen Evidenz (und Berücksichtigung der Rückkopplung)



# Erweiterung 2: Erweiterung der akzeptablen Evidenz (und Berücksichtigung der Rückkopplung)



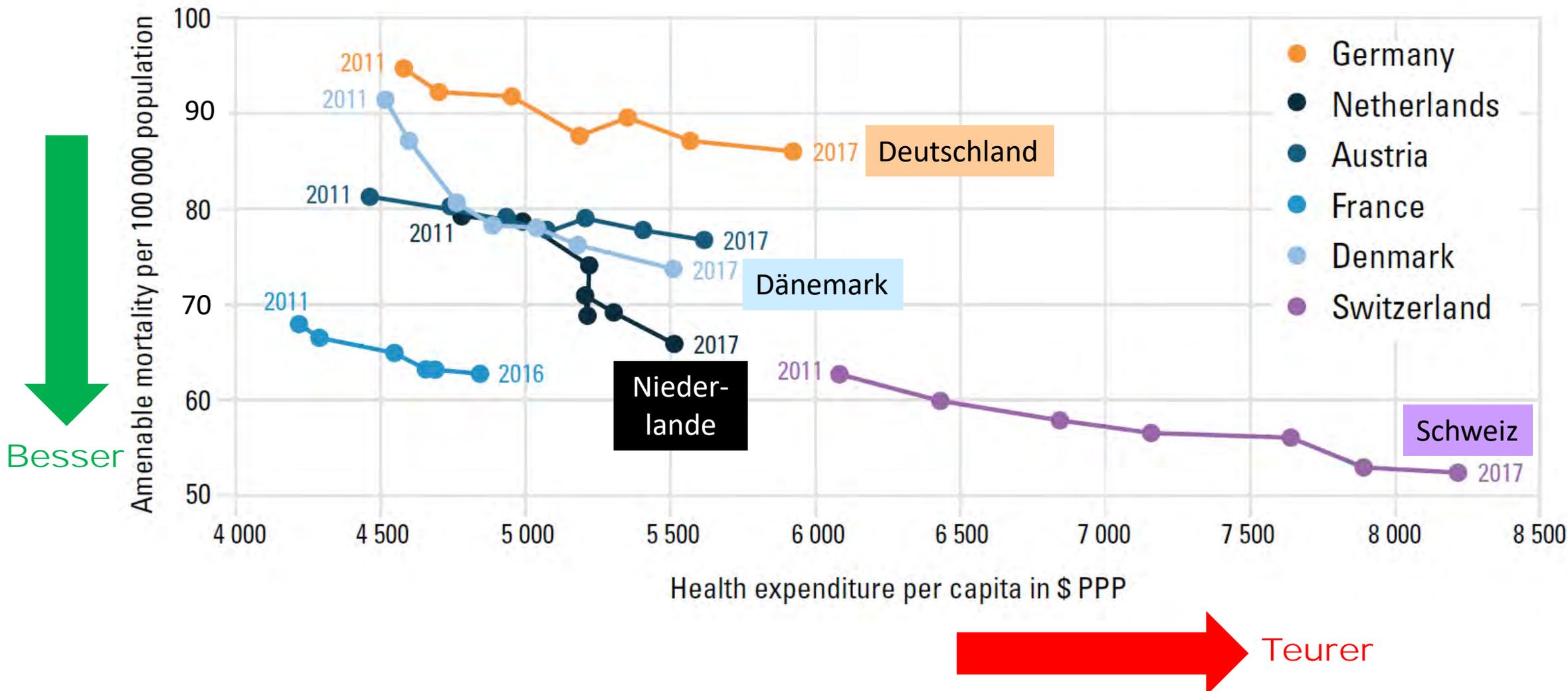
# Erweiterung 2: Erweiterung der akzeptablen Evidenz (und Berücksichtigung der Rückkopplung)



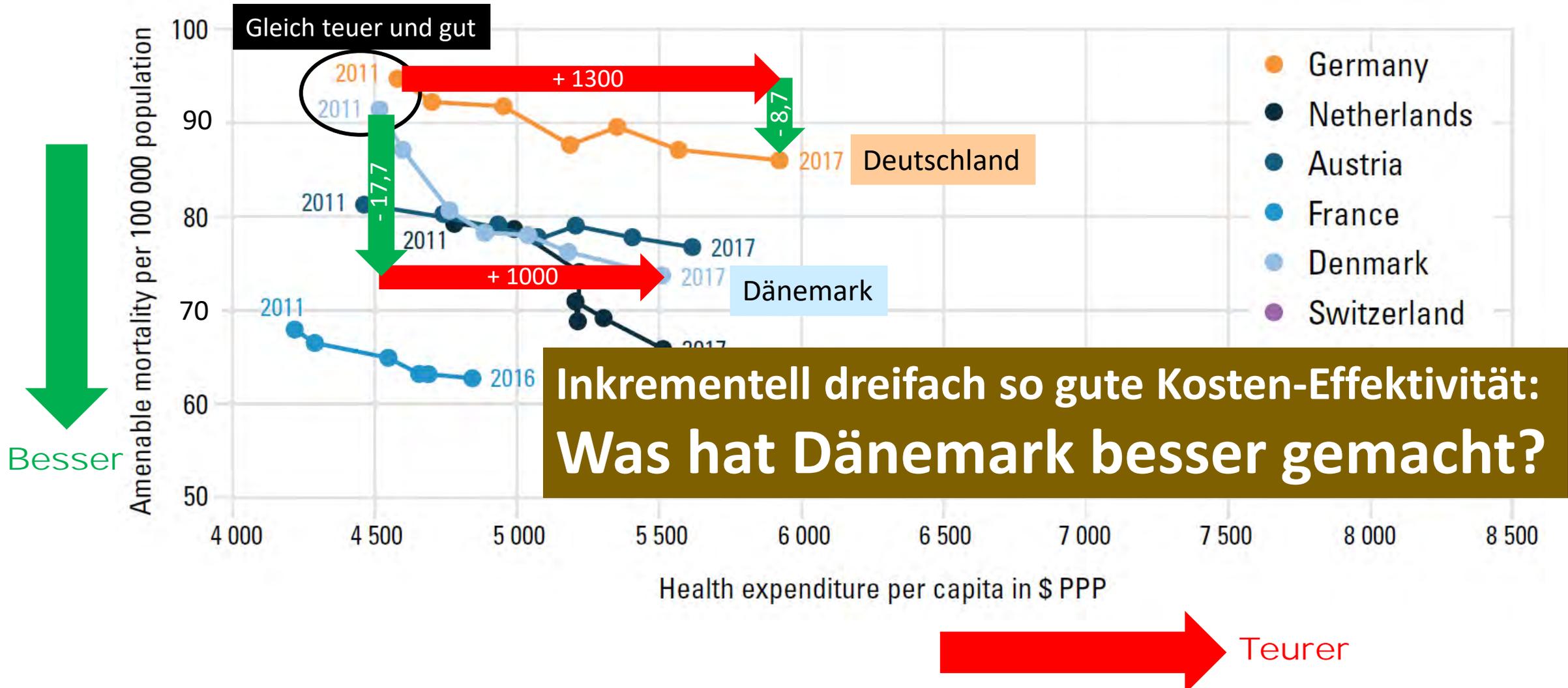
Performance assessment as evidence

Aber in Deutschland werden diese Informationen  
zumeist nicht (systematisch) erhoben bzw. nicht genutzt!

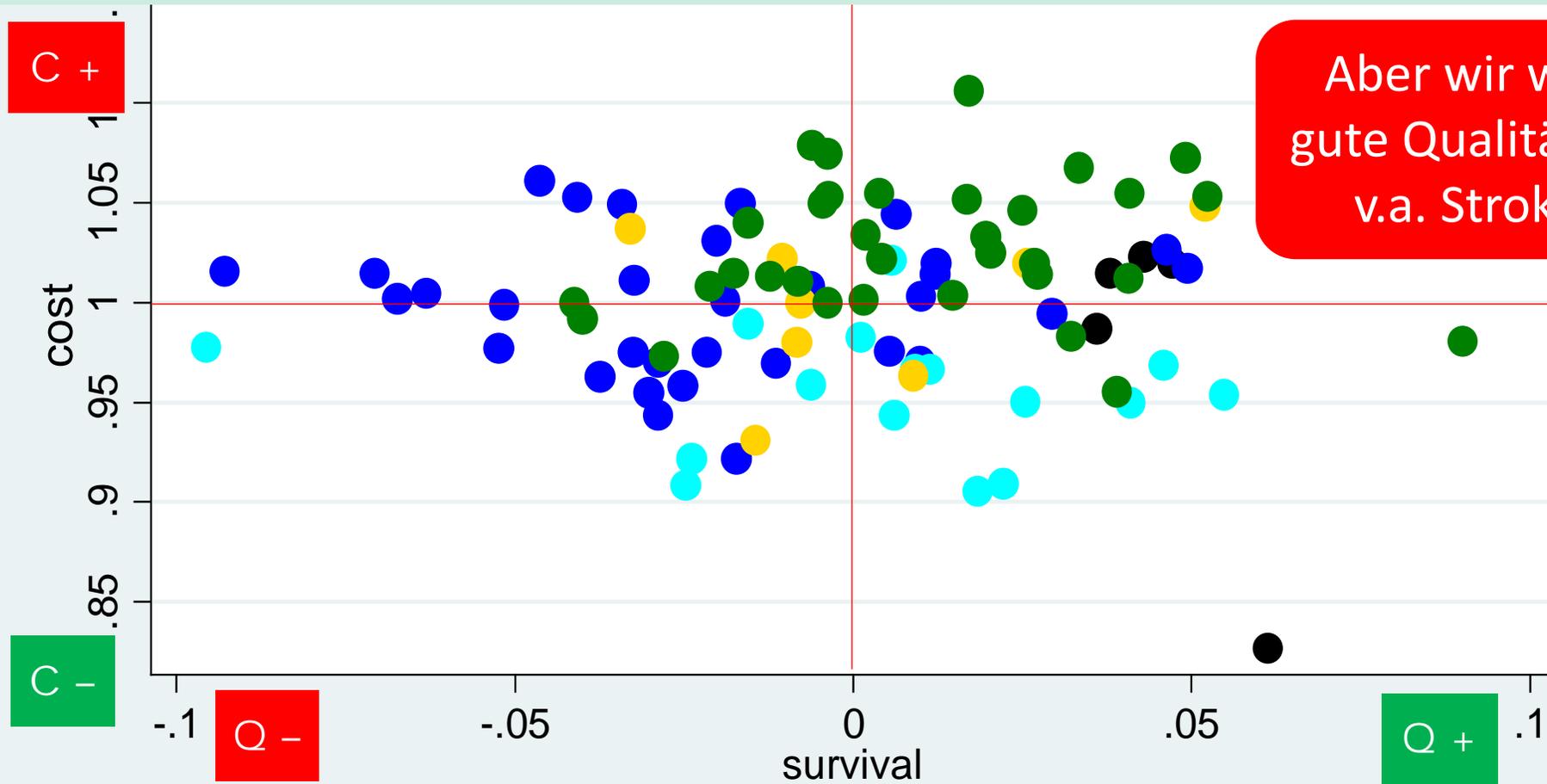
# Performance Assessment des deutschen Gesundheitssystems offenbart einen Rückstand auf andere Länder: Ausgaben pro Kopf vs. “vermeidbare Sterblichkeit”, 2011 bis 2017



# Und Performance des deutschen Gesundheitssystems offenbart einen Rückstand auf andere Länder: Ausgaben pro Kopf vs. "vermeidbare Sterblichkeit", 2011 bis 2017



# Beispiel Schlaganfall-Versorgung: Qualität in deutschen Krankenhäusern stark unterschiedlich



Aber wir wissen, was gute Qualität ausmacht:  
v.a. Stroke Units ...



FINLAND	● FRANCE
GERNANY	● SPAIN
SWEDEN	

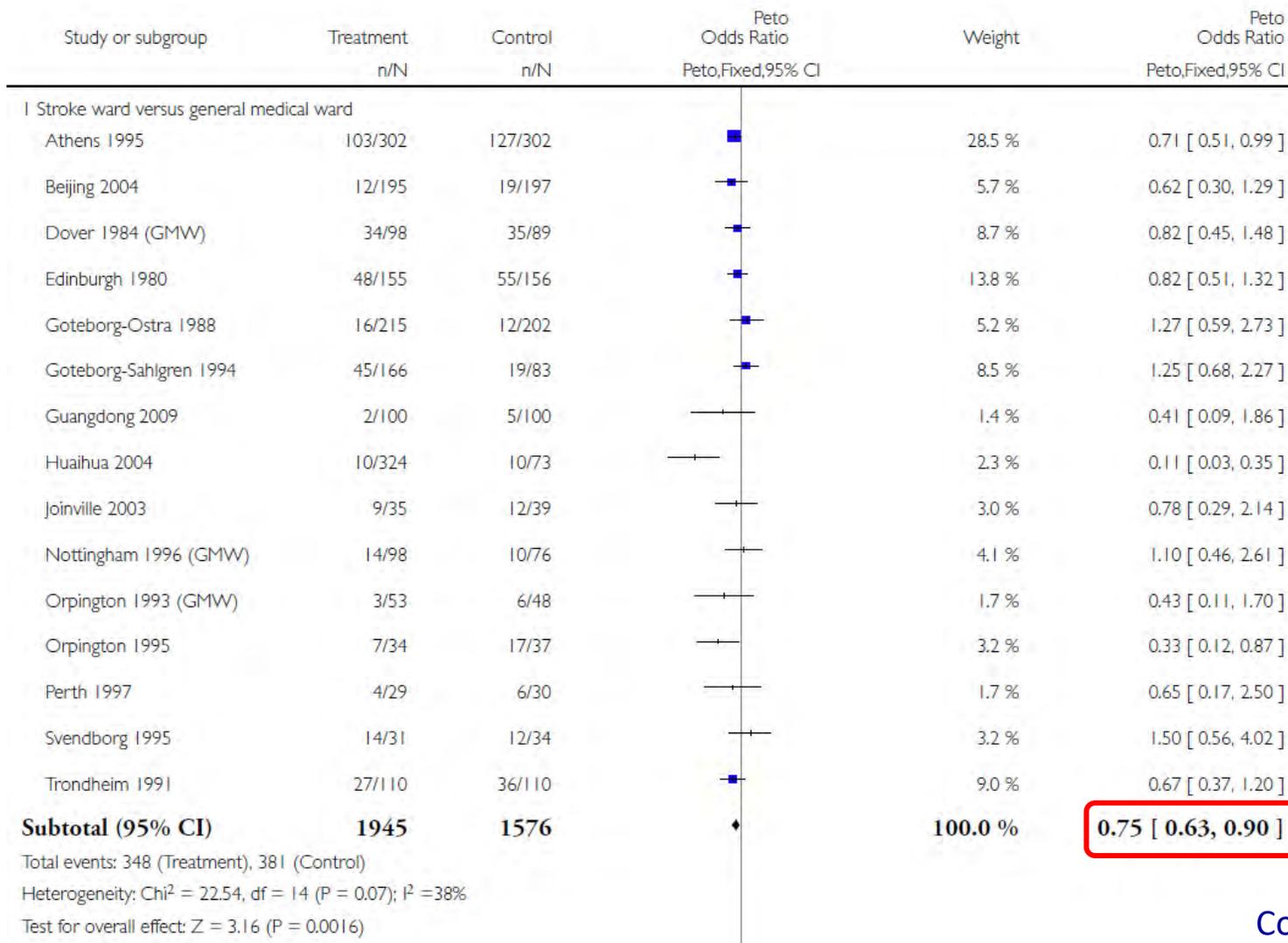
**Analysis 1.1. Comparison 1 Organised stroke unit care versus alternative service, Outcome 1 Death by the end of scheduled follow-up.**

Review: Organised inpatient (stroke unit) care for stroke

Comparison: 1 Organised stroke unit care versus alternative service

Outcome: 1 Death by the end of scheduled follow-up

**... mit gesicherter Evidenz im systematischen Review ...**

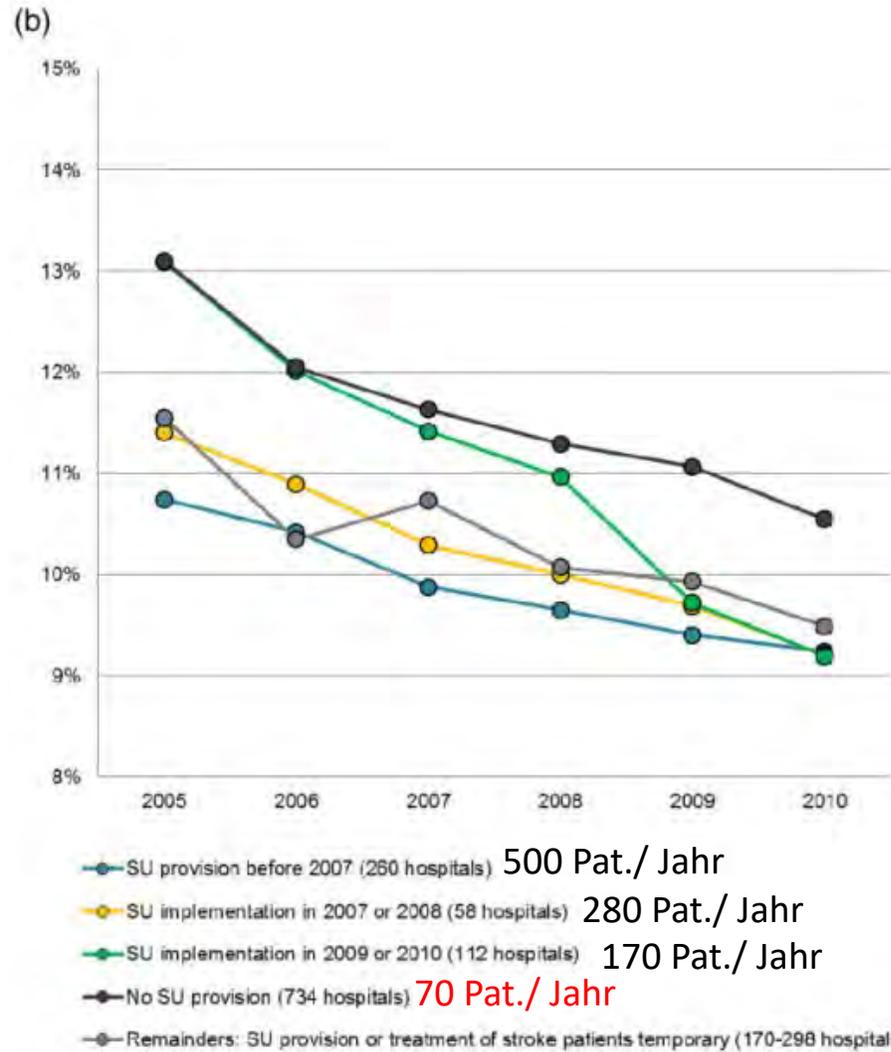


# Schlaganfall-Sterblichkeit in deutschen Krankenhäusern (adjustiert)

... und in Routinedaten

2010: 240.000 Patienten  
(190.000 mit, 50.000 ohne *kodierte*  
Stroke Unit-Behandlung)  
in >1300 Krankenhäusern,  
davon 456 (35%) „mit“ Stroke Unit

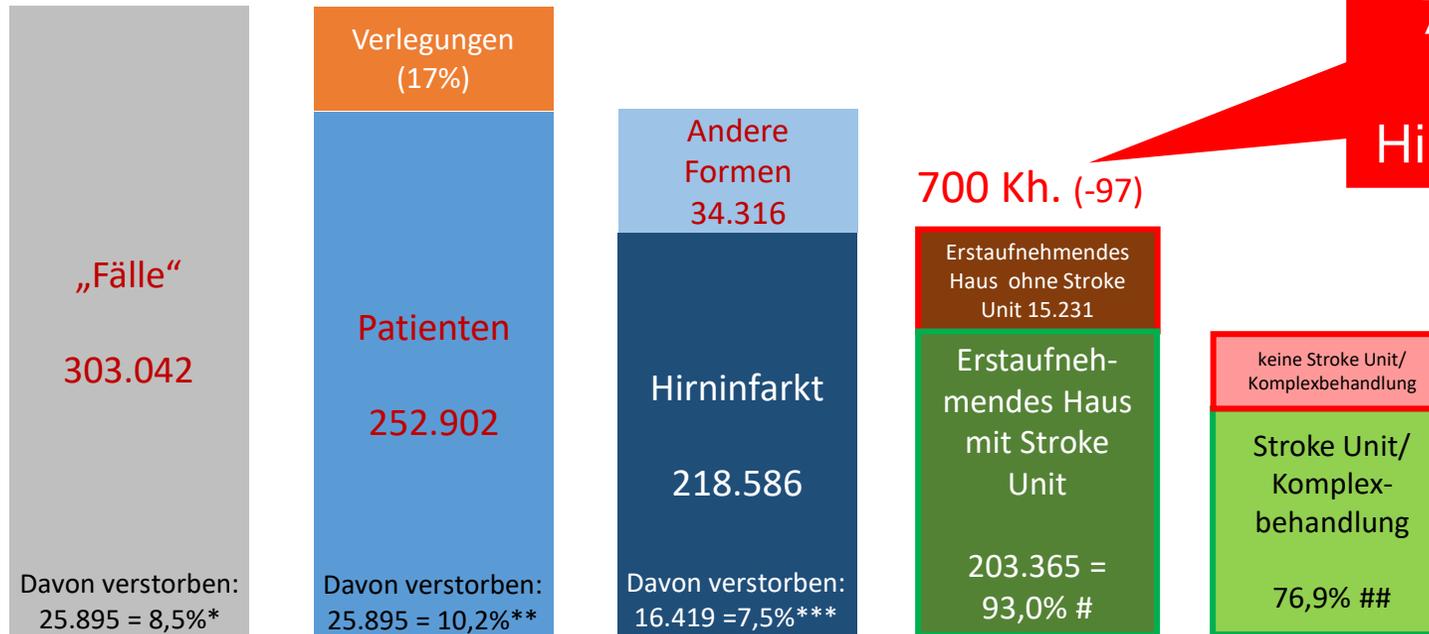
Aber: nur 270  
zertifizierte Stroke  
Units (2015)



Nimptsch & Mansky, Int J Stroke 2014

- SU provision before 2007 (260 hospitals) 500 Pat./ Jahr
- SU implementation in 2007 or 2008 (58 hospitals) 280 Pat./ Jahr
- SU implementation in 2009 or 2010 (112 hospitals) 170 Pat./ Jahr
- No SU provision (734 hospitals) **70 Pat./ Jahr**
- Reminders: SU provision or treatment of stroke patients temporary (170-298 hospitals)

# Schlaganfall (2018) – Sterblichkeit & Anteil mit leitliniengerechter Stroke Unit/ Komplexbehandlung bei Hirninfarkt



Warum dürfen diese 700 Krankenhäuser Patienten mit Hirninfarkt behandeln?

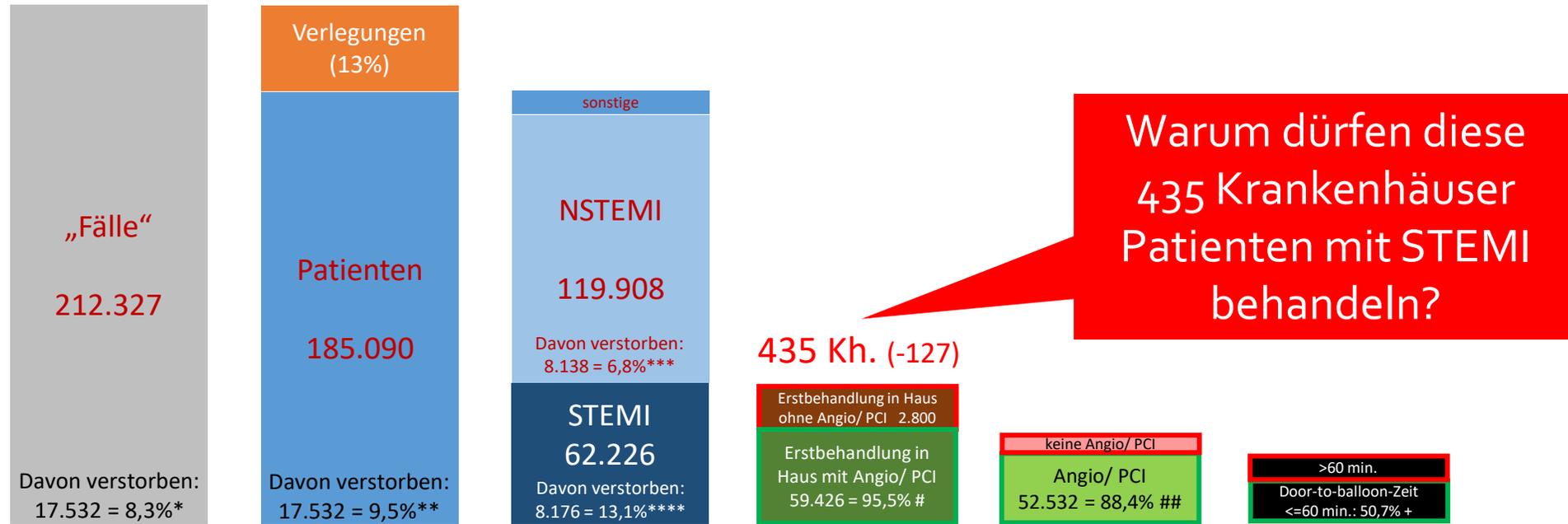
700 Kh. (-97)

473 Kh. (+10)

Stroke Unit-Behandlung bei Hirninfarkt  
 $0,930 \times 0,769 = 72\% (+1\%/ \text{Jahr})$   
 (2014:  $0,884 \times 0,766 = 68\%$ )

2014: \*8,5%, \*\*10,0%, \*\*\*7,3%, # 88,4%, ## 76,6%

# Herzinfarkt (2018) – Sterblichkeit & Anteil STEMI mit leitliniengerechter Koronarangiographie innerhalb von 60 min.



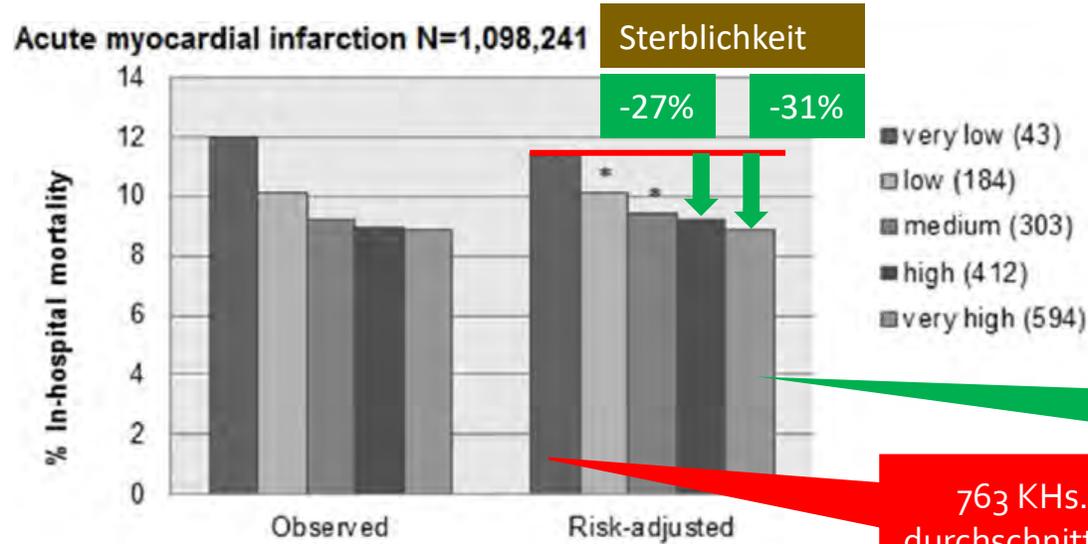
Warum dürfen diese 435 Krankenhäuser Patienten mit STEMI behandeln?

570 Kh. (+38)

Koronarangiographie innerhalb 60 min.  
 $0,955 \times 0,884 \times 0,507 = 43\% (+2\%/Jahr)$   
 (2014:  $0,912 \times 0,860 \times 0,452 = 35\%$ )

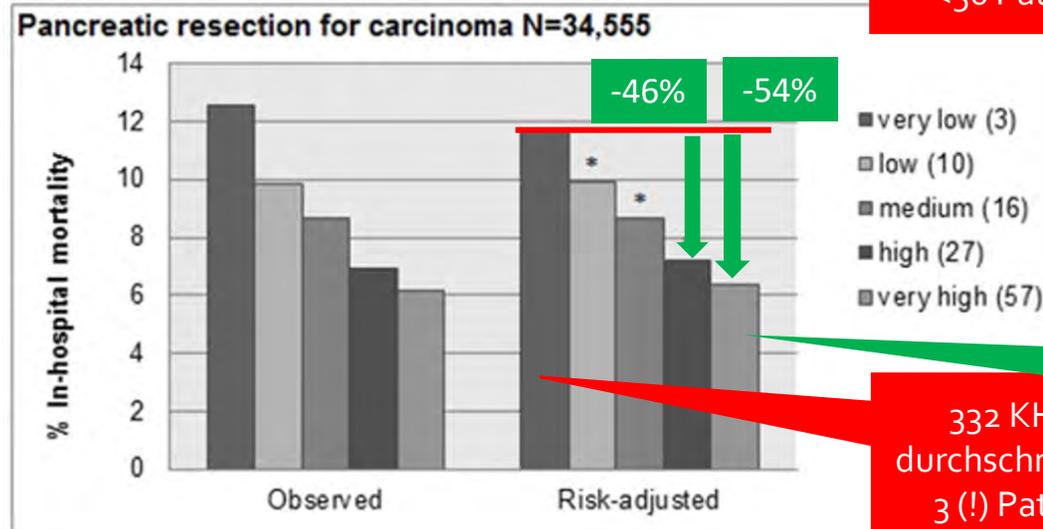
2014: \*8,8%, \*\*10,0%, \*\*\*7,2%, \*\*\*\*13,0%, # 91,2%, ## 86,0%, + 45,2%

# Ganz klar: Menge ist auch eine wichtige „organisatorische Technologie“, die Qualität beeinflusst



763 KHS. mit durchschnittlich <50 Pat./ Jahr

54 KHS. mit durchschnittlich ~ 600 Pat./ Jahr



332 KHS. mit durchschnittlich 3 (!) Pat./ Jahr

→ Basis der Mindestmengen-Regelung

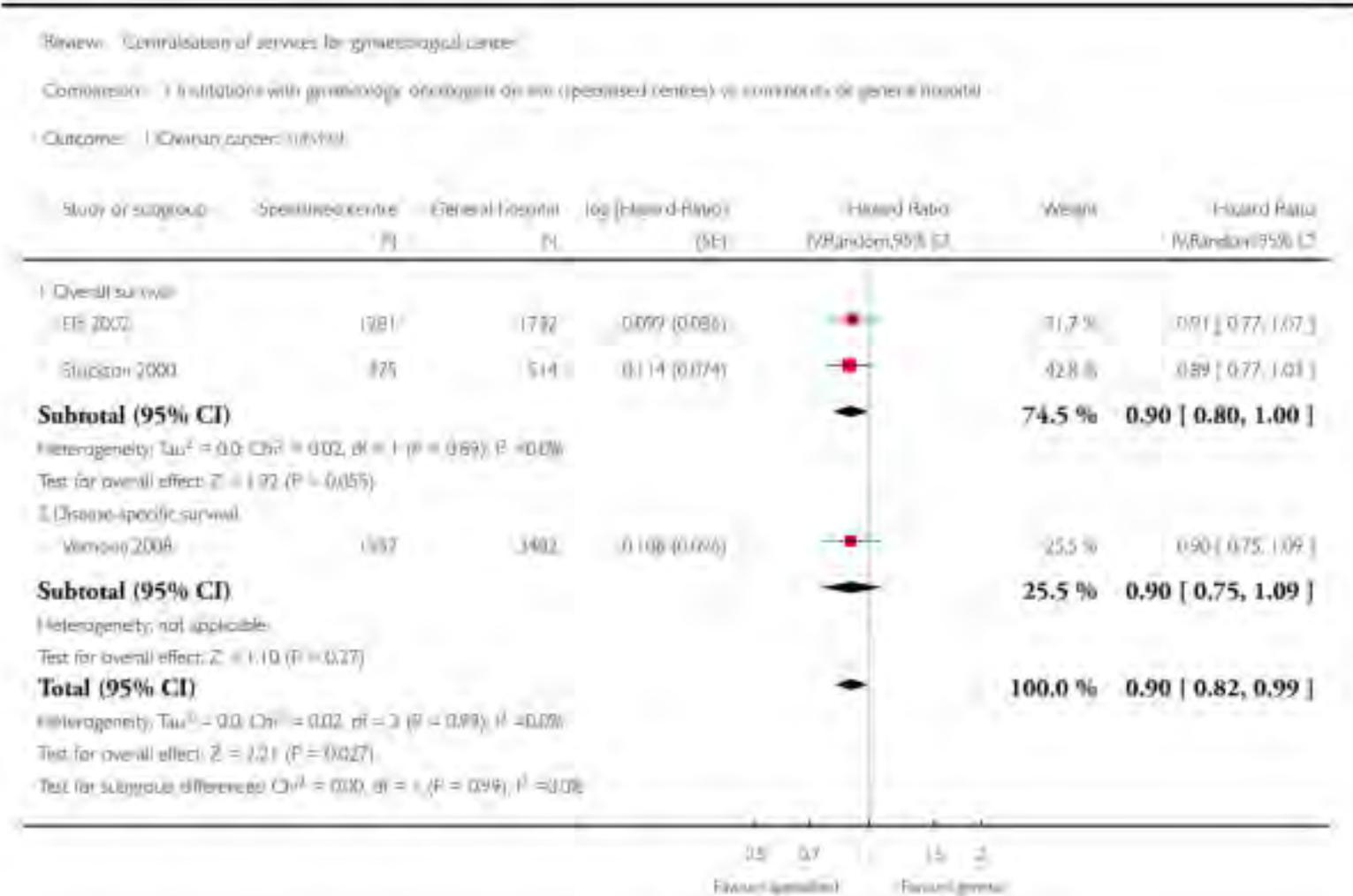
17 KHS. mit durchschnittlich 57 Pat./ Jahr

## Review: Centralisation of services for gynaecological cancer

Comparison: 1 Institutions with gynaecologic oncologists on site (specialised centres) vs. community or general hospital

Outcome: 1 Ovarian cancer: survival

Weiteres Beispiel:  
Onkologische  
Zentrenbildung –  
System. Review  
zeigt Outcome-  
Verbesserung (I)

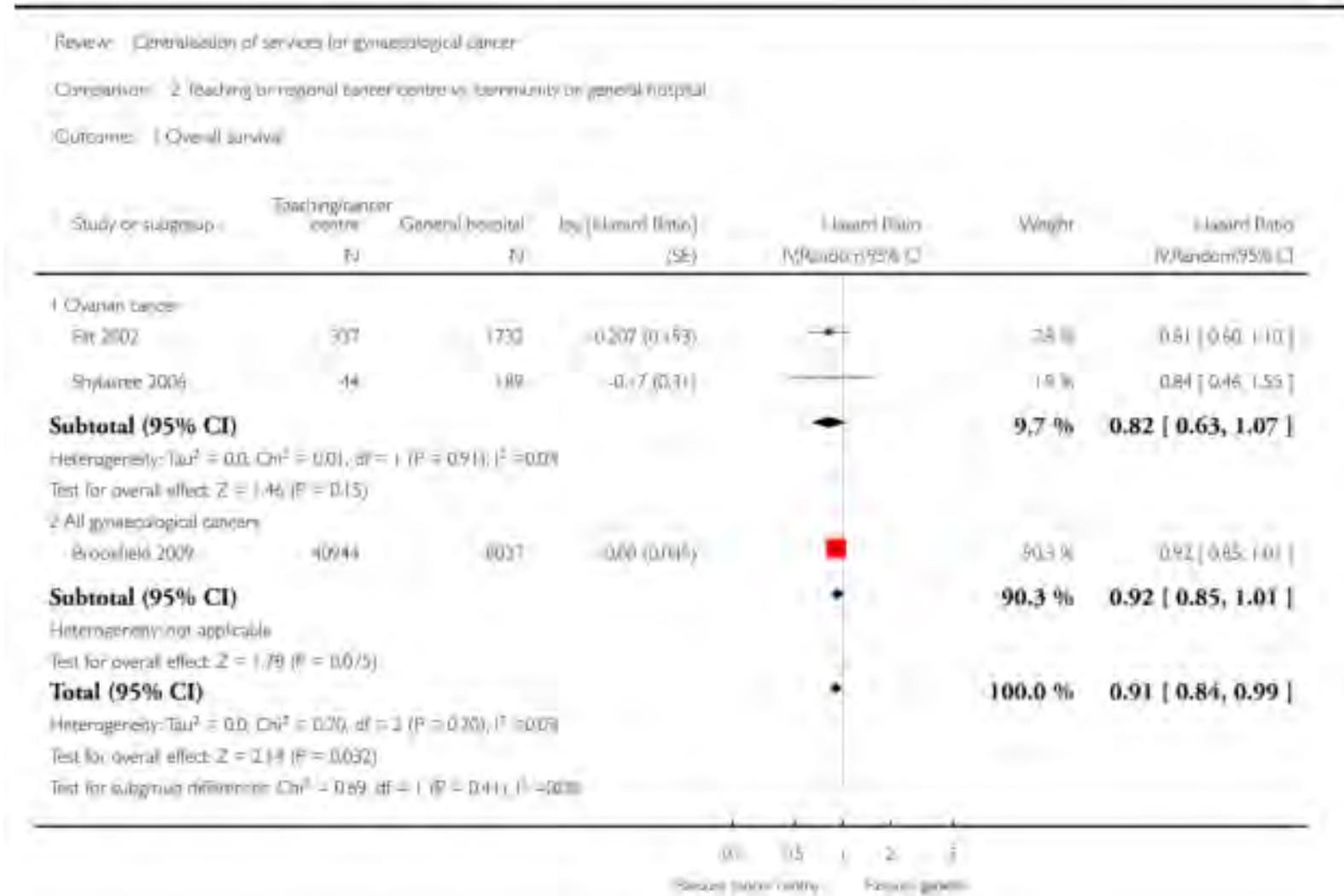


Weiteres Beispiel:  
Onkologische  
Zentrenbildung –  
System. Review  
zeigt Outcome-  
Verbesserung (II)

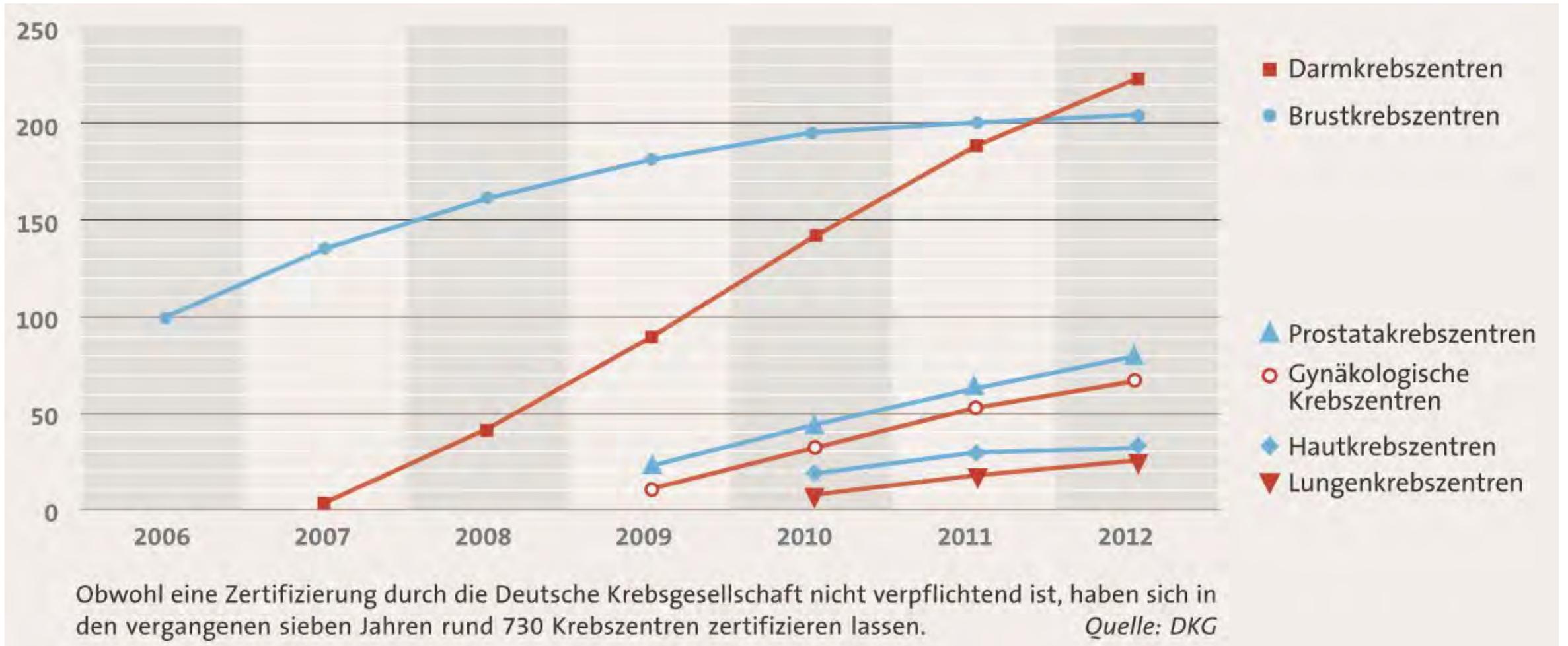
Review: Centralisation of services for gynaecological cancer

Comparison: 2 Teaching or regional cancer centre vs. community or general hospital

Outcome: 1 Overall survival



# Umsetzung sieht auf den ersten Blick vielversprechend aus



→ Inzwischen 285 zertifizierte Zentren für Darmkrebs und 243 für Brustkrebs (d.h. 1 / 300.000 bzw. 350.000 Einw.)

Aber: Wenn Versorgung in Zentren nachweislich besser ist, warum ist sie nicht verpflichtend?

- 56% der Darmkrebs- und 20% der Brustkrebspatientinnen außerhalb dieser Zentren behandelt (2019) - **über alle Krebsarten 55% (in absoluten Zahlen: bei 260.000 von 475.000 Patienten) nicht in Zentren!** Bei Bauchspeicheldrüsenkrebs waren es sogar 69 %!

Und wenn Mindestmengen zur Verbesserung der Outcomes zielführend sind (stellt G-BA ja fest), warum dürfen sie unterlaufen werden?

- 2016 missachteten noch 279 von 618 Krankenhäuser mit komplexe Eingriffe an der Bauchspeicheldrüse und 274 von 402 mit Eingriffen an der Speiseröhre die jeweiligen **Mindestmengen** von jeweils 10 pro Jahr. Regulierung zwar verschärft, wird aber **von den Ländern unterlaufen**, derzeit bei der Speiseröhre z.B. 21-mal und Bauchspeicheldrüse 33-mal (Mindestmengen-Transparenzliste 2022).

*Manchmal haben wir die Evidenz noch nicht ...*

## Unser Monitoring von NUB-Methoden, Studienevidenz und Nutzung in deutschen Krankenhäusern

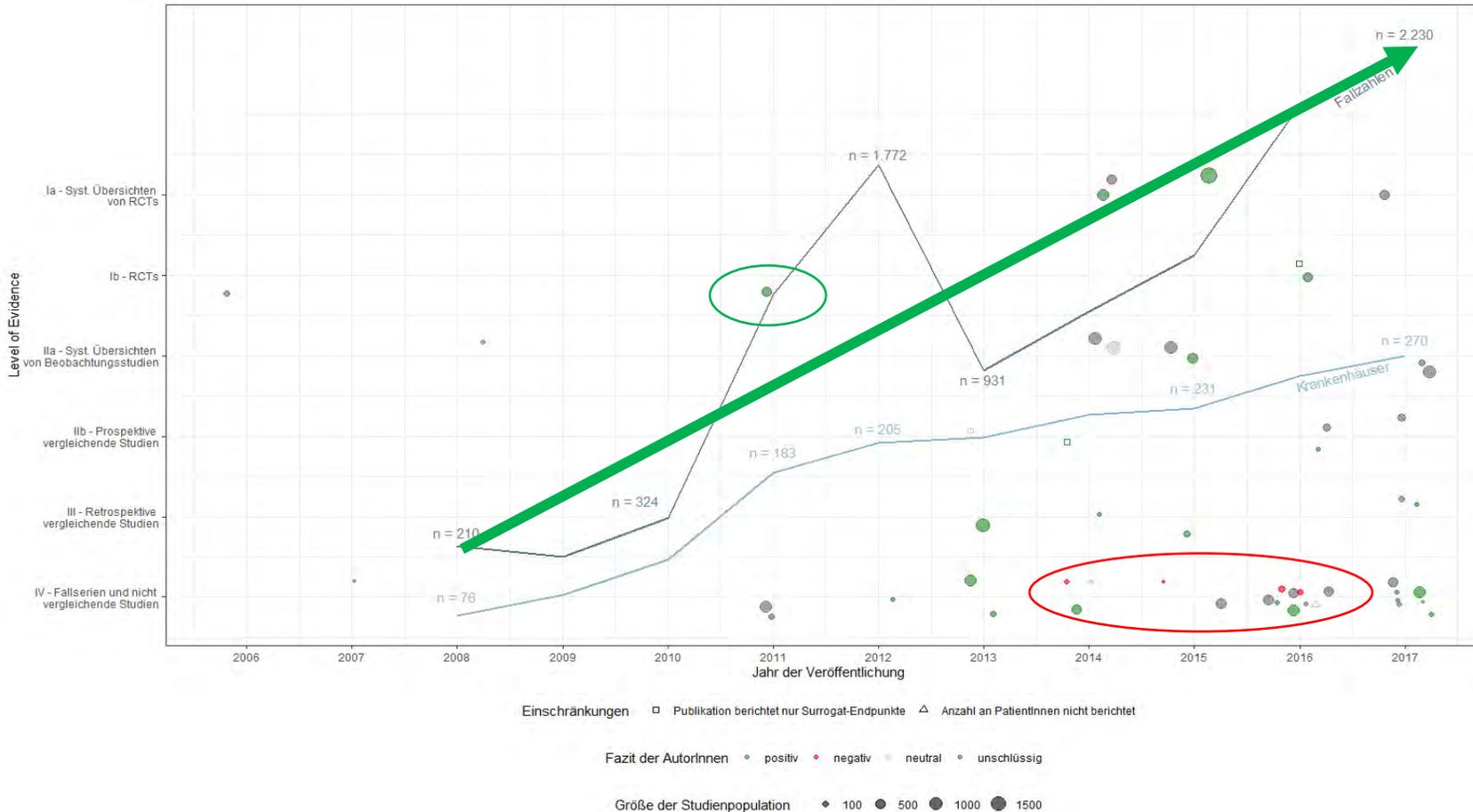
Wie hängen Nutzung (#Krankenhäuser, #Pat. longitudinal) mit der Entwicklung des Evidenzkörpers zusammen?

		EVIDENCE		
		Strong direct evidence (positive)	Strong direct evidence (negative)	Limited Evidence
		Moderate evidence (positive)	Moderate evidence (negative)	No Evidence
UTILIZATION	Widespread	Success	Hazard	Overadoption
	Cautious	Underadoption	Vigilance	Prudence

**Fig. 2** Six-field table for the classification of diffusion patterns based on utilization and evidence. Source: Authors' own, adapted from Denis et al. [12]

# Unser Monitoring von NUB-Methoden, Studienevidenz und Nutzung in deutschen Krankenhäusern

## (1) Medikamentenfreisetzende Stents für Oberschenkelgefäße I



Fallzahlen und Anzahl anwendender KH:

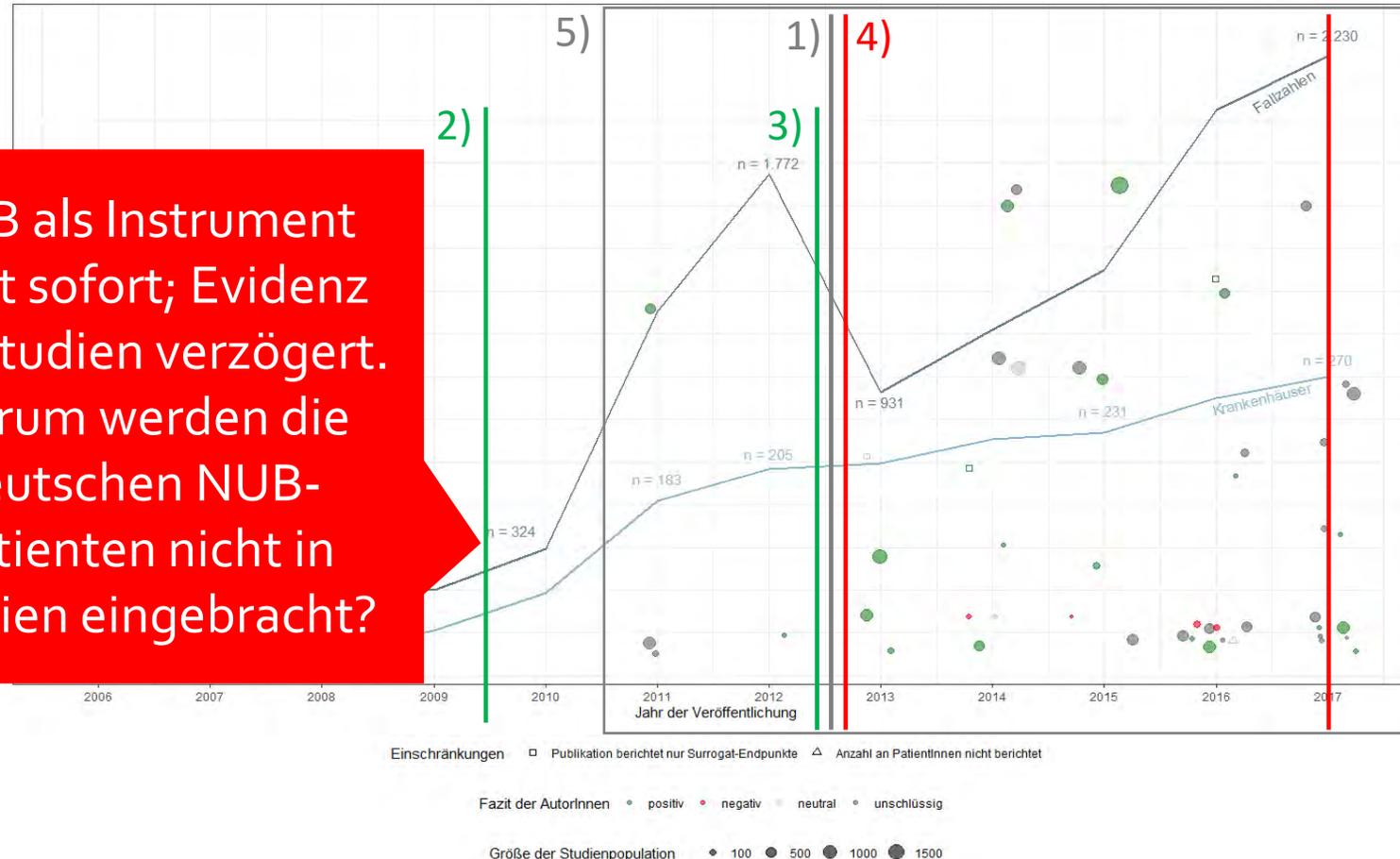
- **Werte steigen konstant** für beide Kennzahlen über Beobachtungszeitraum
- Ausnahme für Fallzahlen: 2011 bis 2012 stärkerer Anstieg, Absinken im Folgejahr (2012 bis 2013)

Entwicklung von Evidenz und Ergebnissen:

- Publikationen zu Studien der LoE Ia–IV, n=52
- meiste: LoE IV (n=27, 52%), wenigste: LoE Ia/Ib (je n=4, 8%), LoE Ia–III nahezu gleich verteilt
- verschiedene Populationsgrößen (n=5 bis n=1.956)
- **Verdichtung der Evidenz** im Zeitverlauf:
  - Quantität der Publikationen in allen LoE ↑
  - Probandenzahlen ↑
- alle Bewertungen vertreten
- v.a. Publikationen mit **positiven Bewertungen**, nur n=4 negativ (2014 bis 2016, LoE IV)
- in 2011: RCT mit positiver Bewertung

# Unser Monitoring von NUB-Methoden, Studienevidenz und Nutzung in deutschen Krankenhäusern: (1) Medikamentenfreisetzende Stents für Oberschenkelgefäße II

NUB als Instrument wirkt sofort; Evidenz aus Studien verzögert. Warum werden die deutschen NUB-Patienten nicht in Studien eingebracht?



## Leitlinien:

- Leitlinie der Deutschen Gesellschaft für Angiologie – Gesellschaft für Gefäßmedizin (2015) [11] → keine ausdrückliche Empfehlung (sondern: andere Ansätze bei PAVK)

## Finanzierung:

- ab 2009 über verschiedene Ansätze (NUB-Entgelte, Fallpauschale/DRG)

## Sicherheitswarnungen und Rückrufe:

- n=12 Meldungen international (davon n=3 aus Deutschland), z.B. Bruch des Kathetersystems bei *Silver PTX<sup>TM</sup>* [12,13], unvollständige Stentfreisetzung bei *Eluvia<sup>TM</sup>* [14]

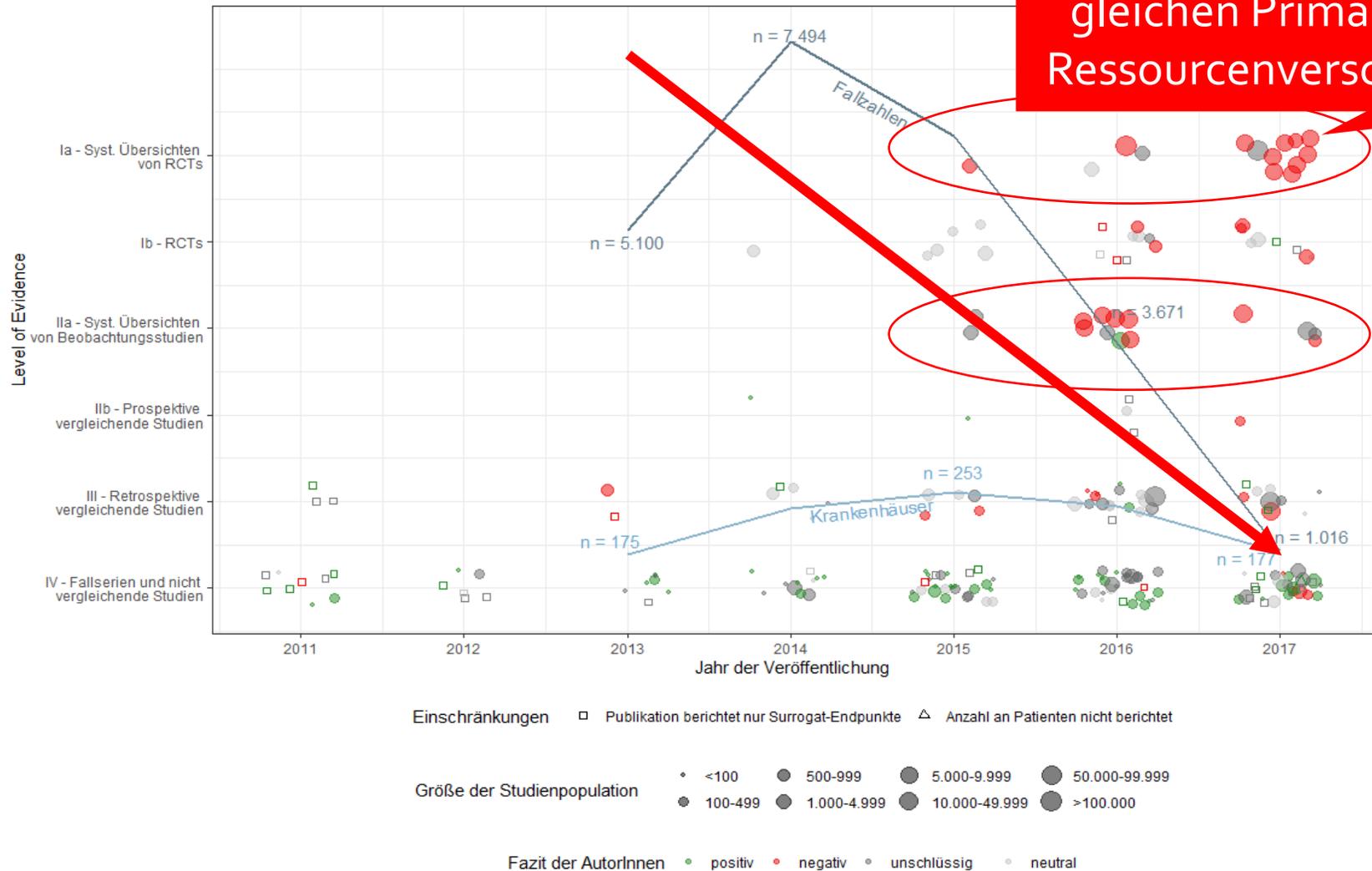
## Zusammenhangsvermutung:

- 1) ab 2013: Entwicklung des Evidenzkörpers und Verdichtung
- die Leitlinie spielt vermutlich keine Rolle
- 2) ab 2009: Finanzierung gegeben
- 3) 2012: Änderung der Finanzierungsart
- 4) 2013 und 2017: Sicherheitswarnungen/Rückruf
- 5) seit 2011: Mehrzahl der Studien zu *Silver PTX<sup>TM</sup>* (kein Off-Label-Use notwendig)

# Unser Monitoring von NUB-Methoden, Studienevidenz und Nutzung in deutschen Krankenhäusern:

## (2) Bioresorbierbare Koronarstents

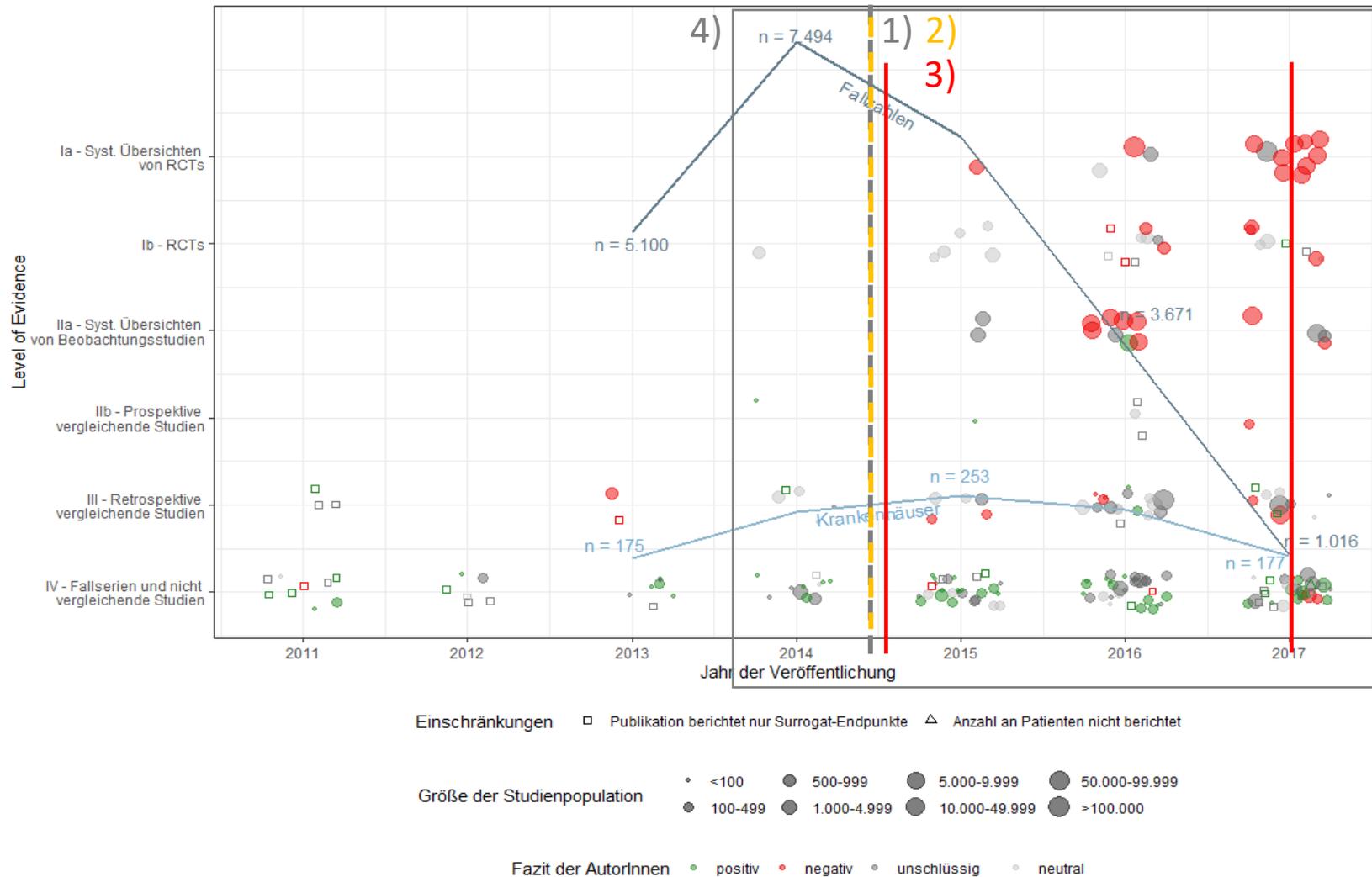
Praktisch immer mit den gleichen Primärstudien = Ressourcenverschwendung



anwendender KH:  
 im ersten Jahr der Beobachtung (2013)  
 und **sinken ab Folgejahr kontinuierlich**  
 steigt zunächst (2013 bis 2015) und  
 verringert sich bis 2017 auf Ursprungsniveau

- Entwicklung von Evidenz und Ergebnissen:
- Publikationen zu Studien der LoE Ia–IV, n=223
  - meiste: LoE IV (n=123, 55%), wenigste: LoE IIb (n=6, 3%), höhere LoE (Ia–IIb) weniger vertreten (n=44, 26%)
  - verschiedene Populationsgrößen (n=59 bis n=126.526)
  - vorrangig Publikationen von Studien niedriger LoE, aber auch n=23 RCTs und n=27 Systematische Reviews (SR)
  - **Verdichtung der Evidenz** im Zeitverlauf:
    - Quantität der Publikationen hoher LoE ↑
  - alle Bewertungen vertreten
  - niedrige LoE: viele Publikationen mit positiver Bewertung (n=50)
  - insgesamt viele Publikationen mit **negativer Bewertung** (n=41, davon n=19 SR)

# Unser Monitoring von NUB-Methoden, Studienevidenz und Nutzung in deutschen Krankenhäusern: (2) Bioresorbierbare Koronarstents II



## Leitlinien:

- Leitlinie der European Society of Cardiology (2014) [15] → keine Empfehlung (Begründung: schlechte Studienlage)

## Finanzierung:

- ab 2012 über verschiedene Ansätze (NUB-Entgelte, unbewertete Zusatzentgelte)

## Sicherheitswarnungen und Rückrufe:

- n=48 Meldungen international (davon n=2 aus Deutschland), z.B. Aktualisierung der Gebrauchsanweisung für *Absorb<sup>TM</sup>* durch Hersteller: Beschränkung auf kleine Gefäße [16], Einsatz nur in ausgewählten Zentren [17]

**Wäre das zur Evidenzgenerierung nicht von vornherein sinnvoll gewesen?**

## Zusammenfassung:

- 4) insgesamt Mehrzahl der Studien zu *Absorb<sup>TM</sup>*, u.a. 2014/15 erste RCT-Ergebnisse nach 12/24-Monats-Follow-up

# Unser Monitoring von NUB-Methoden, Studienevidenz und Nutzung in deutschen Krankenhäusern über alle 27 NUB hinweg:

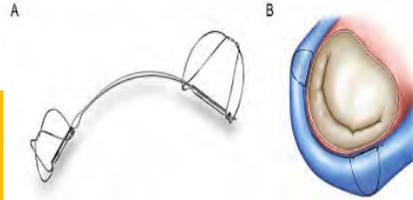
Ein Zusammenhang ...

kann nicht ausgeschlossen werden

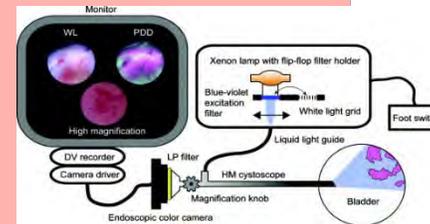
lässt sich nicht herstellen

Fallzahl- und Evidenzentwicklung sind miteinander vereinbar

5/27  
untersuchten  
Technologien



10/27  
untersuchten  
Technologien



12/27  
untersuchten  
Technologien



Quelle: Absorb Abbott®

- Kein eindeutiger Zusammenhang, Richtung der Studienergebnisse steht dem Verlauf nicht entgegen
- Die Richtung der Studienergebnisse und der Fallzahlentwicklung sind unvereinbar (z.B. sinkenden Fallzahlen trotz positiver Studienlage)
- Übereinstimmung der Richtung der Studienergebnisse (positiv, negativ, neutral) mit der Richtung der Fallzahlentwicklung (steigend, fallend)

# Ein notwendigerweise unvollständiges Fazit

- Vermutlich am wichtigsten: die (Selbst-)Erkenntnis, dass wir nicht unbedingt am besten sind und von anderen lernen können. Warum machen wir bei OECD-Studien oder dem European Observatory on Health Systems and Policies nicht mit?
- Wir haben – trotz aller Maßnahmen und Institutionen – keine wahrhaft evidenzinformierte Gesundheitspolitik. Eine solche setzt voraus, dass wir uns von „mikro“ bis „makro“ konsequent an Evidenz orientieren wollen.
- Dazu sollten wir auch bisher wenig genutzte Evidenzquellen – und ihre Zusammenführung mit anderen – nutzen (und daraus neue Fragen generieren); dazu zählen insbesondere Routinedaten.
- Leistungen ohne Evidenz sollten nur dann erbracht werden, wenn sie zur Evidenzgenerierung beitragen.