

Unter welchen Voraussetzungen kann die Steuerung durch Qualität funktionieren?

Max Geraedts
Institut für Versorgungsforschung und
Klinische Epidemiologie
Philipps-Universität Marburg



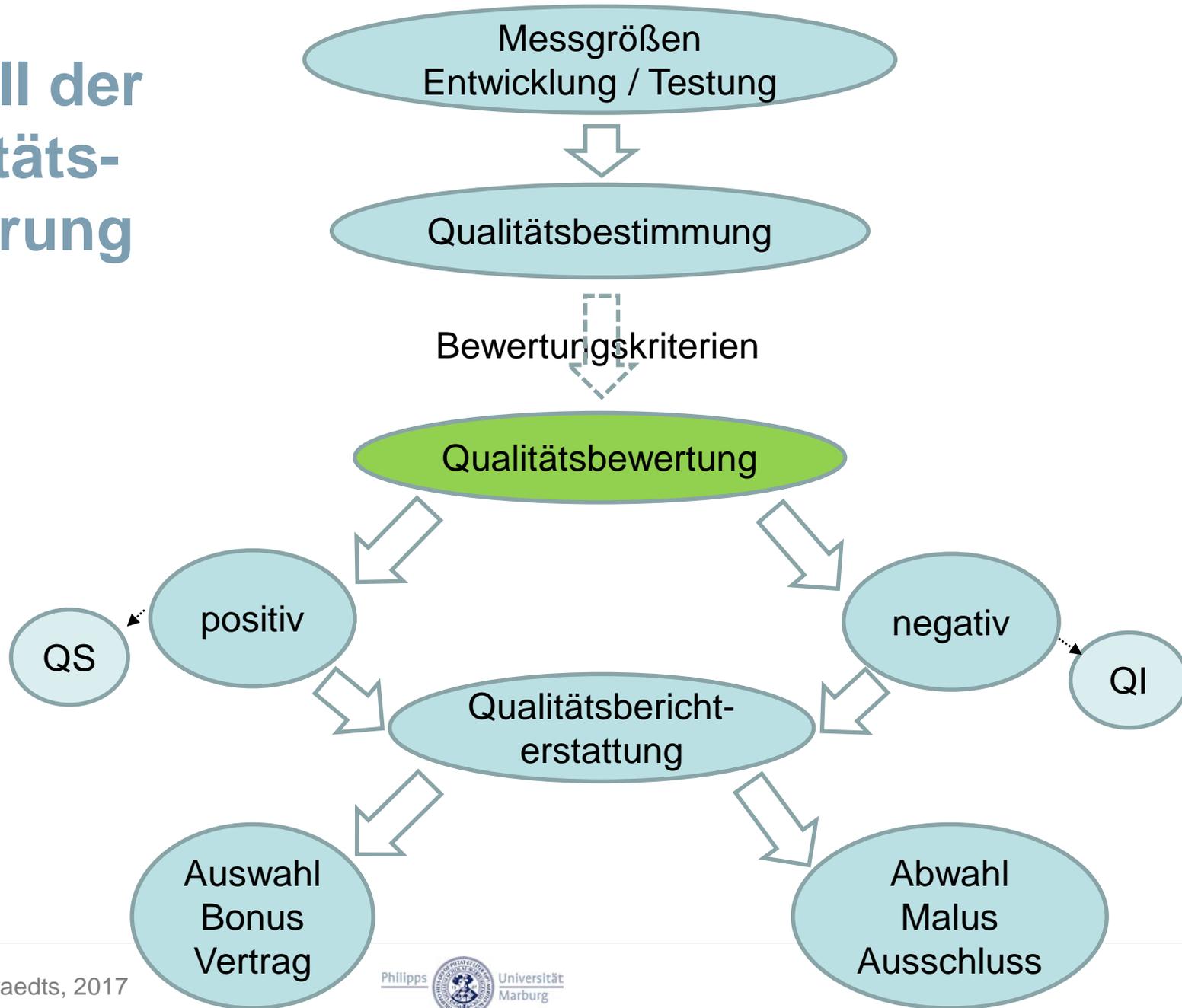
Agenda

- In Deutschland eingeführte und neue Maßnahmen der Steuerung durch Qualität
- Theoretische Anmerkungen
- Probleme und Steuerungswirkungen einiger eingeführter Maßnahmen
- Fazit

In Deutschland genutzte qualitätsorientierte Steuerungsinstrumente

- Externe Qualitätssicherung (eQS)
- Mindestmengen
- Qualitätsberichterstattung
- Krankenhausplanung auf der Basis von Qualitätsparametern u.a. NRW
- Selektivverträge im Rahmen der IV („value based purchasing“)
- Strukturanforderungen (Personal / Ausstattung, Fortbildung/QZ))
- Dokumentationsüberprüfungen
- Ringversuche
- Förderung Onkologischer Zentren
- Peer Review Verfahren
- Zertifizierung
- Planungsrelevante Qualitätsindikatoren
- Zu- und Abschläge auf der Basis der eQS

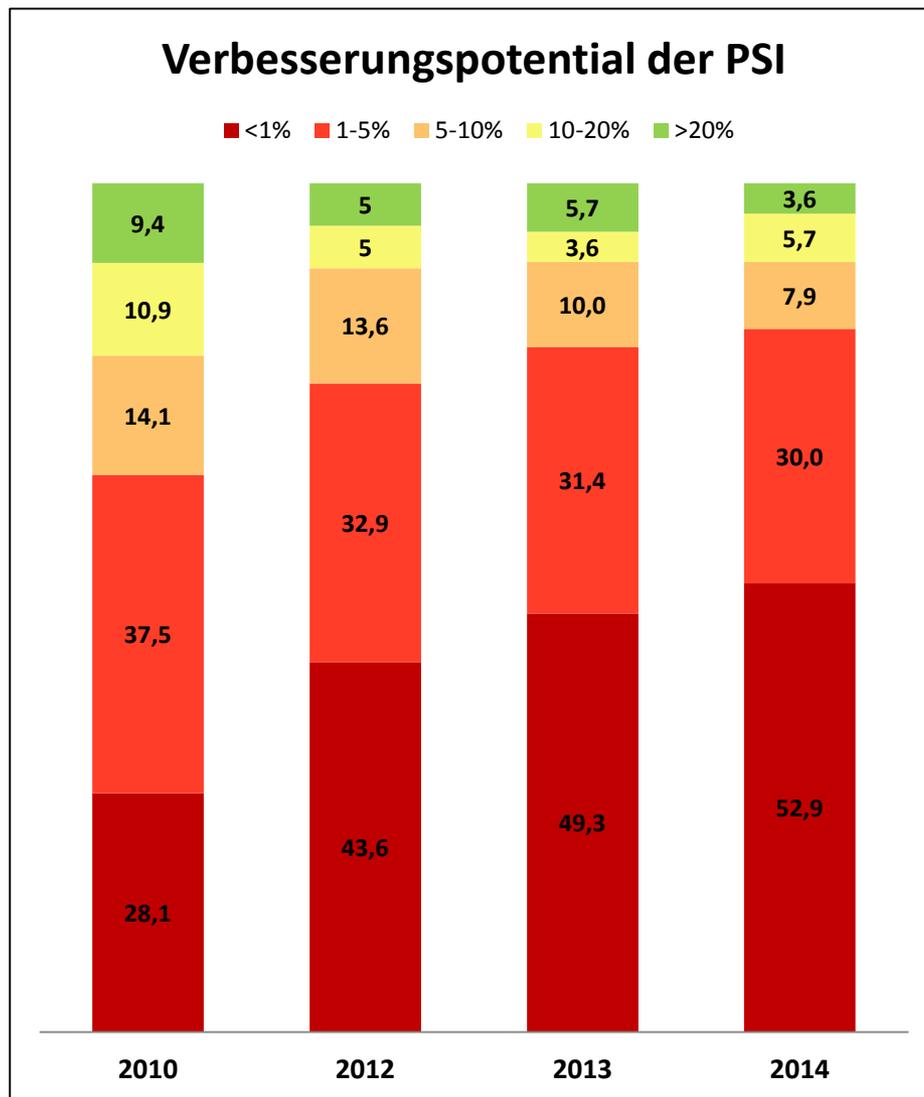
Modell der Qualitäts- steuerung



eQS-Indikatoren

- Historisch zu verstehende Leistungsbereiche (FP/SE) und Indikatoren:
 - Ausgangspunkt: von Fachgesellschaften für freiwillige Qualitätsvergleichs- und -verbesserungsprojekte entwickelt
 - spiegeln nicht Versorgungsrelevanz / Optimierungspotenzial wider
- (wiss. Auswahlkriterien u.a.: hohe Frequenz, hohes Risiko für die Patienten, oft mit Problemen einhergehend [JCAHO, 1990]; Verdacht auf Über-, Unter- oder Fehlversorgung; hohe Versorgungsvariabilität, Versorgung kürzlich stark verändert, hohe finanzielle Bedeutung, Verbesserungspotenzial)
- Testung unzureichend / wie viele hätten heute Bestand?

Verbesserungspotential bei PSI in der eQS



- **Material:** Erfassungsjahr 2014: 416 QI aus 31 Leistungsbereichen, davon 270 QI als PSI; identifiziert (Dank an K. Döbler, KCQ)
- insgesamt 279 QI berichtspflichtig, davon 175 PSI
- **Methode:** längsschnittliche Analyse
- **Ergebnis: das Verbesserungspotential der PSI sinkt stetig mit der Zeit**
- Bei den beobachteten/gemesenen PSI wird es immer unwahrscheinlicher, weitere Verbesserungen zu erreichen
- Verbesserungspotential wird weitgehend ausgeschöpft

Konstanz der in den QB berichteten QI 2006 - 2012

Richtig gemeinsam (gleiche Definition etc.)				
	2006	2008	2010	2012
2006	30			
2008	16	29		
2010	9	21	182	
2012	8	16	126	289

Qualitätsbestimmung

Qualitätsreport 2015 des IQTIG:

- 728 rechnerische Auffälligkeiten bei den indirekten Verfahren; 74,3% der auffälligen Krankenhausstandorte (n=541 = knapp ein Drittel aller Standorte) zur schriftlichen Stellungnahme aufgefordert; 49,4% (n=267) begründen mit fehlerhafter Dokumentation
- auch bei direkten Verfahren 49% fehlerhafte Dokumentation

Table 2: Ergebnisse der Statistischen Basisprüfung: Auffälligkeitskriterien zur Plausibilität und Vollständigkeit

	Anzahl	Anteil
Rechnerische Auffälligkeiten (gesamt)	911	100,0 %
Maßnahmen		
Keine Maßnahmen	1	0,1 %
Hinweise	216	23,7 %
Stellungnahmen	694	76,2 %
Sonstiges	0	0,0 %
Weiterführende Maßnahmen (pro rechnerische Auffälligkeit)		
Anlässe für Besprechungen	7	-
Anlässe für Begehungen	0	-
Zielvereinbarungen	36	-
Einstufungen der rechnerischen Auffälligkeiten, zu denen eine Stellungnahme angefordert wurde		
Korrekte Dokumentation bestätigt	269	38,8 %
Fehlerhafte Dokumentation bestätigt	347	50,0 %
Keine erklärenden Gründe genannt	30	4,3 %
Sonstiges	29	4,2 %

Qualitätsbestimmung

Qualitätsreport 2015 des IQTIG:

- Herzschrittmacher 5% - Nieren-TX 10% - Leber-TX 0% der Datenfelder mit Problemen (waren das die relevanten Datenfelder?)
- Herzschrittmacher: bei 26 von 43 Krankenhausstandorten unterschieden sich die Angaben in der QS-Doku von der Patientenakte; Niere-TX: Abweichung bei 2 von 4 geprüften Standorten

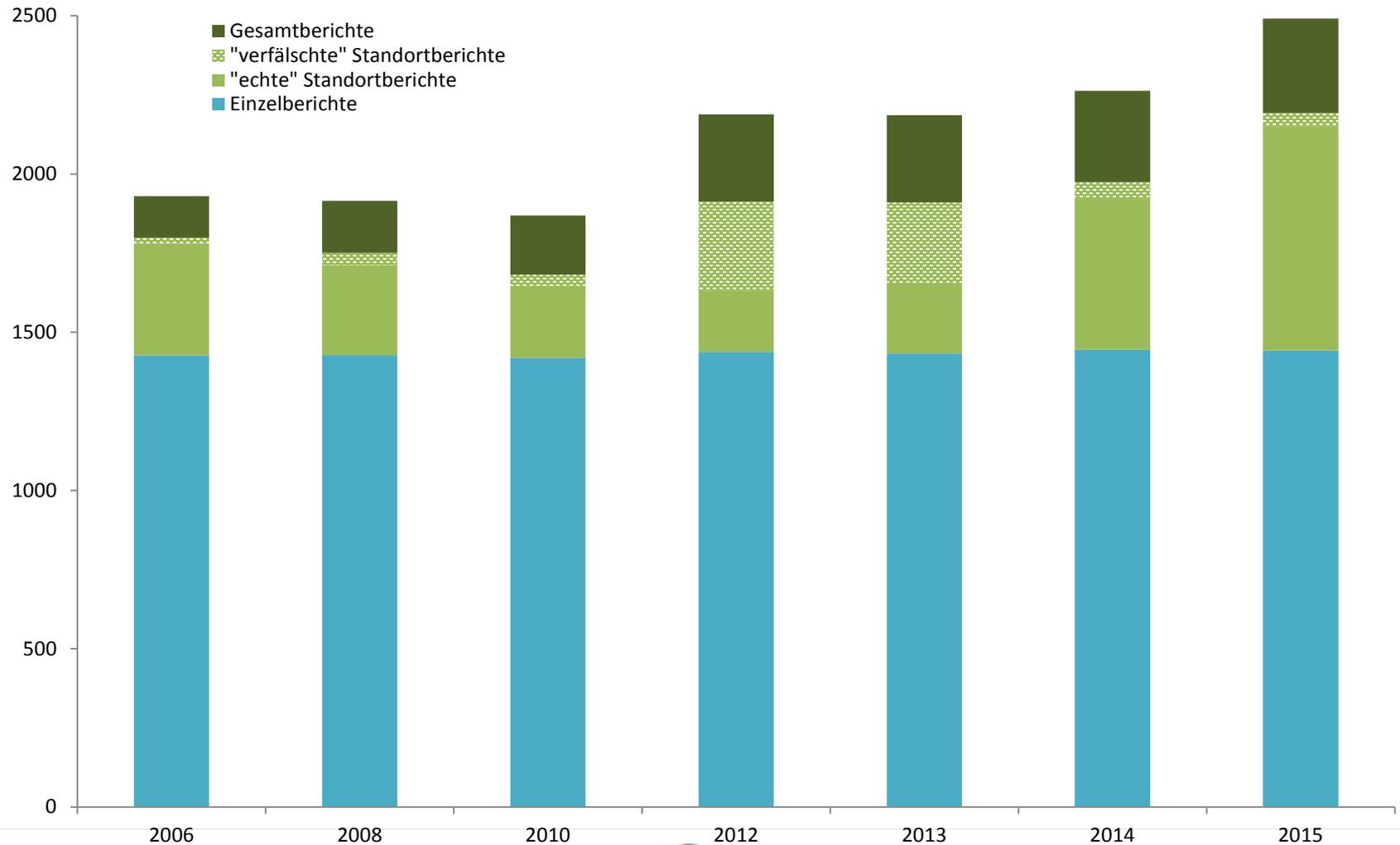
Tabelle 3: Umfassende Datenvalidierung: Ergebnisse der Stichprobenprüfung mit Datenabgleich

QS-Verfahren	Besuchte Standorte	Untersuchte Fallakten	Geprüfte Datenfelder	Datenfelder mit verbesserungsbedürftiger Datenvalidität
Herzschrittmacher-Revision/ -Systemwechsel/-Explantation	43	496	40	2
Nierentransplantation	4	79	19	2
Leberlebendspende	4	41	25	0

Eignung der Qualitätsberichte für die Qualitätsbestimmung?

Problem	Häufigkeit* 2006-2012	Lösungsansatz
1. Fehlende und unvollständige QB		
Fehlende QB	2-15%	Zum Teil Ersatz fehlender Werte
Unvollständige QB (Schätzung)	2-6%	
2. QB von Krankenhausverbänden nicht eindeutig zuordenbar		
Nicht erkennbar, dass es sich um einen Verbund-/ oder Standortbericht handelt	2-16%	Verknüpfung mehrerer Berichtsjahre
Verfälschte Standortberichte (Verwendung von Verbunddaten)	2-12%	Vergleich der Daten aller Standortberichte
3. Datenverknüpfung innerhalb eines Berichtsjahres unmöglich		
Verwendung unterschiedlicher „IK-Standort-Kombinationen“ in den Dateien	2-3%	Dateiname als zusätzliche Schlüsselvariable
4. Fehlende Schlüsselvariable für die Verknüpfung mehrerer Berichtsjahre		
Datenverlust bei Verknüpfung allein über die IK ohne Standortnummer	23%	Kombination aus "IK-Standortnr.", "Adresse" und "Bettenanzahl"
Datenverlust bei Verknüpfung über IK-Standort-Kombination	20%	
Verwendung der gleichen Adresse	18%	

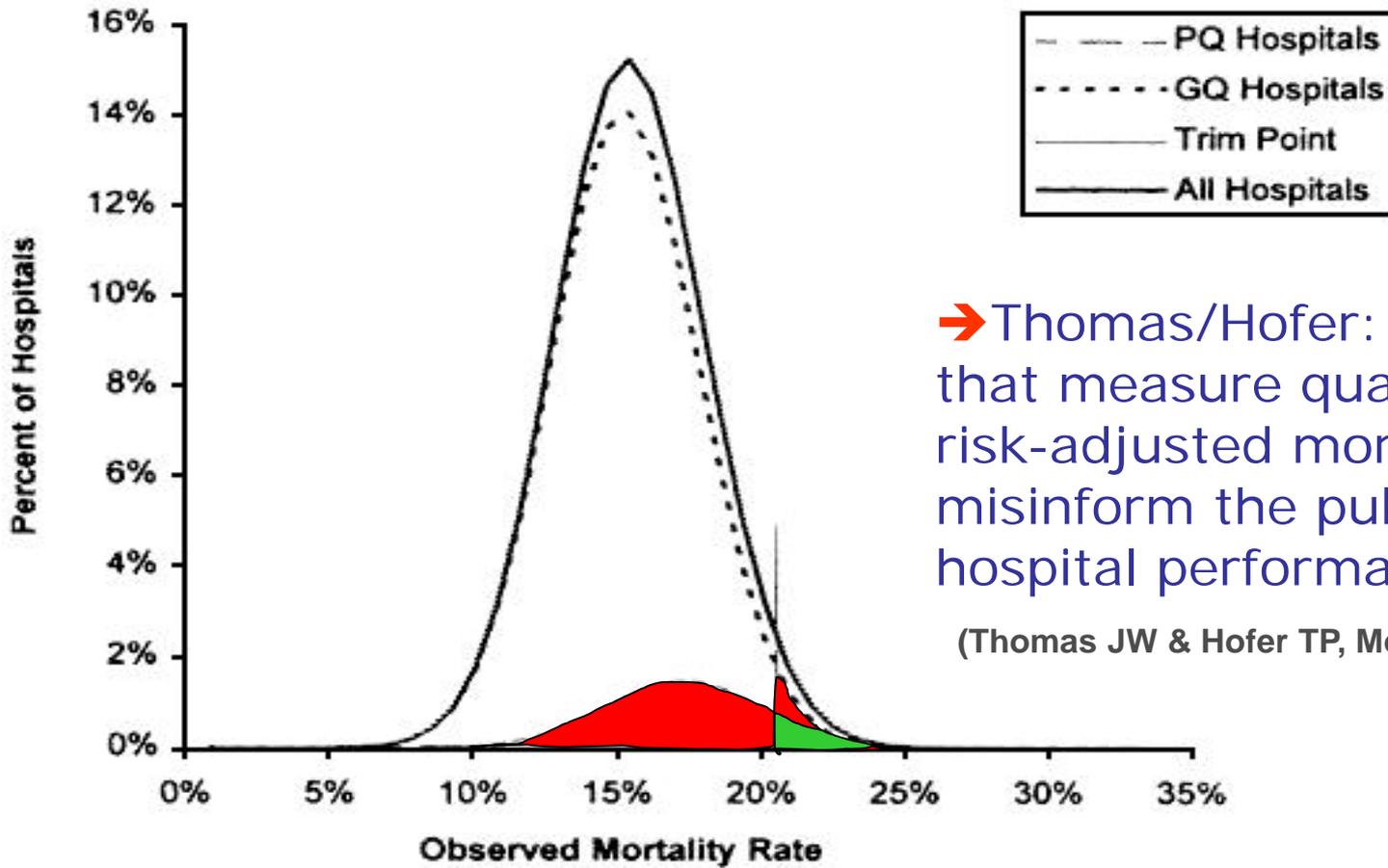
Aktualisierung 2015-Daten: 209 QB erstmalig im Datensatz (88% davon Standortberichte)



Methodische Probleme der vergleichenden Darstellung von Versorgungsergebnissen

- Zumeist seltene Ereignisse
- Langzeitergebnisse oft besonders relevant
- Beeinflussung durch vielfältige externe Faktoren
- Zur Adjustierung notwendige sowie Langzeit-Daten oft nicht vorhanden
- Assoziation zur medizinischen Versorgung unsicher
- Ergebnisdarstellungen in Form von Wahrscheinlichkeiten werden selten verstanden
- Allein aufgrund statistischer Überlegungen kaum sichere Aussagen möglich

Präzision adjustierter Letalitätsraten als Qualitätsmessgröße: Verteilung richtig und falsch positiver sowie falsch negativer Krankenhäuser



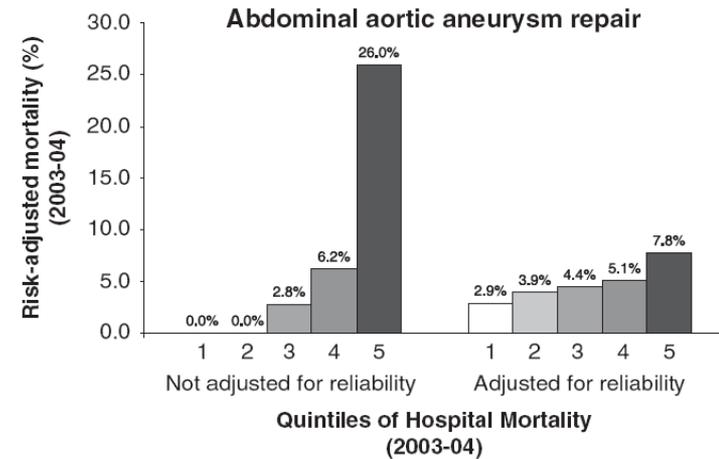
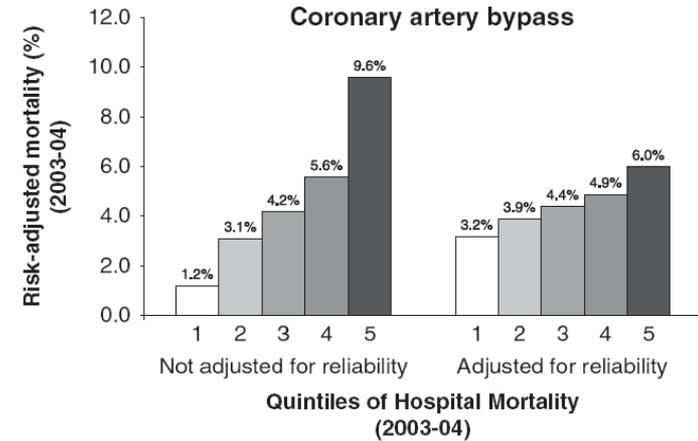
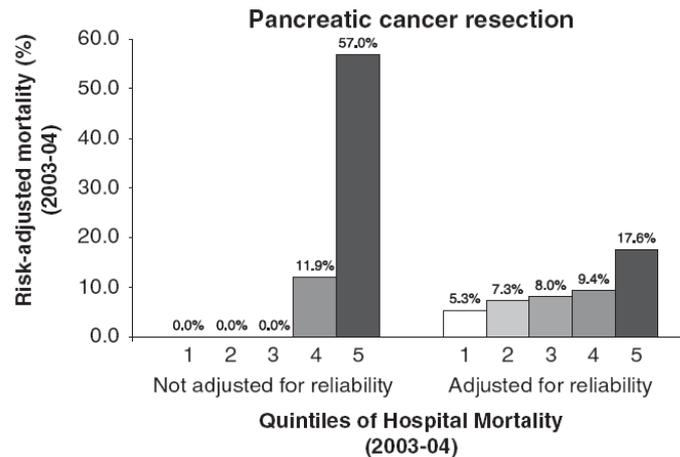
→ Thomas/Hofer: "Reports that measure quality using risk-adjusted mortality rates misinform the public about hospital performance."

(Thomas JW & Hofer TP, Med Care 1999)

Ranking Hospitals on Surgical Mortality: The Importance of Reliability Adjustment

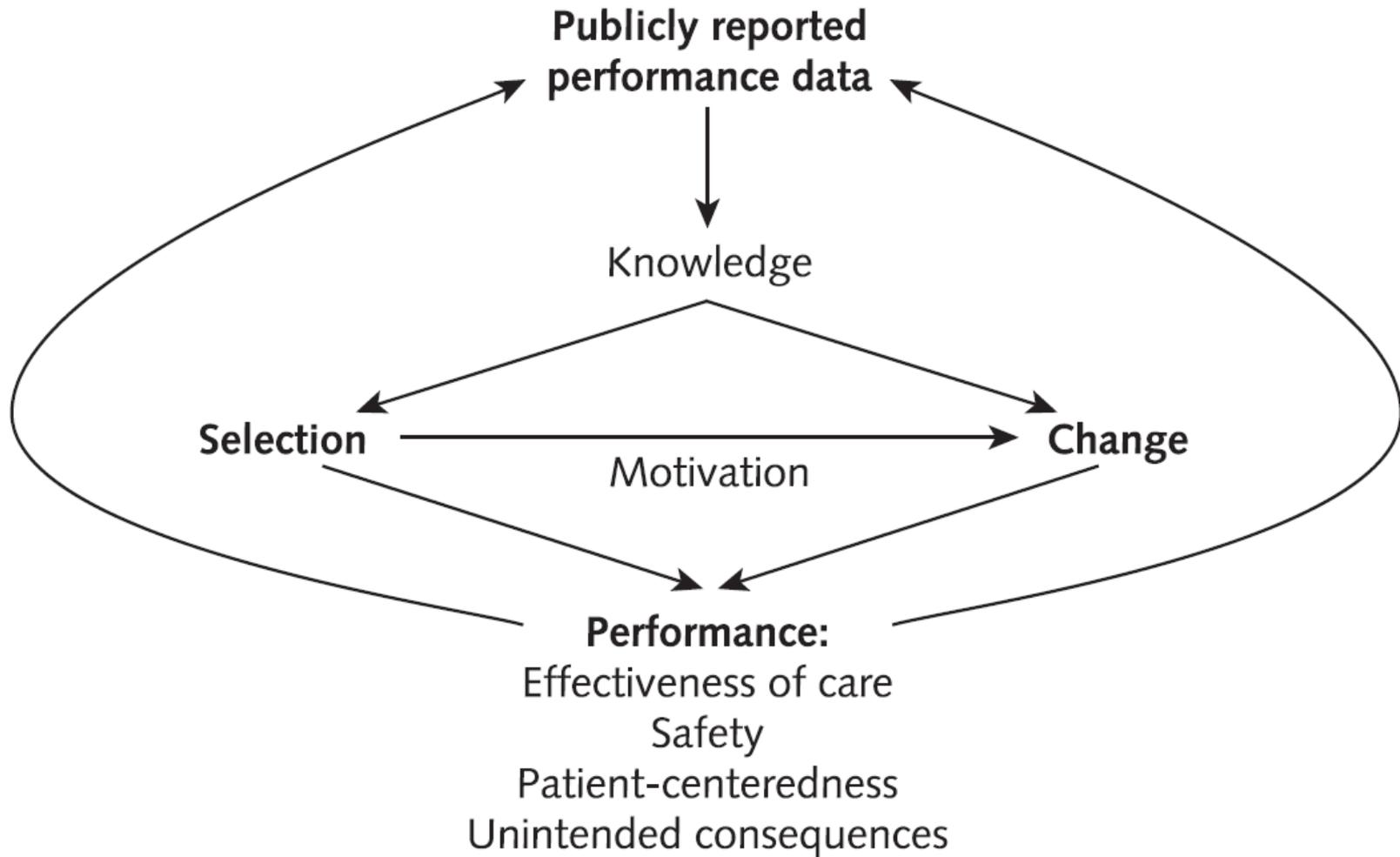
Dimick JB, Staiger DO, Birkmeyer JD.
Health Serv Res 2010

Figure 2: Variation in Mortality Rates across Hospital Quintiles, before and after Adjusting for Reliability



Modell der Qualitätsberichterstattung

(Fung et al. 2008 nach Berwick et al. 2003)



Effekte der Qualitätsberichterstattung: Systematische Literaturbewertung - Totten et al. 2012

- 11.809 Titel/Abstracts (1980-2011) gefunden; 1.632 bewertet; 97 quantitative und 101 qualitative Studien eingeschlossen
- Evidenzstärke gering
- öffentliche Berichterstattung regt Leistungserbringer zu Qualitätsverbesserungen an und führt zu besseren Performanzdaten
- keine oder kaum Evidenz für Auswahleffekte auf Seiten der Patienten/Angehörigen
- Einfluss von Charakteristika der Berichte kaum erforscht
- keine Hinweise auf Schäden durch Berichte

Totten et al. „Public Reporting as a Quality Improvement Strategy“; AHRQ: Evidence Report/Technology Assessment Number 208, 2012

Effekte der öffentlichen Berichterstattung in Deutschland



Hermeling P, Geraedts M.
Kennen und nutzen Ärzte den strukturierten Qualitätsbericht?
Das Gesundheitswesen 2012



de Cruppé W, Geraedts M.
Wie wählen Patienten eine Krankenhaus für elektive operative Eingriffe?
Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2011



Auras S, de Cruppé W, Blum K, Geraedts M.
Mandatory quality reports in Germany from the hospitals' point of view: a cross-sectional observational study.
BMC Health Serv Res 2012



Kraska RA, de Cruppé W, Geraedts M.
Probleme bei der Verwendung von Qualitätsberichtsdaten für die Versorgungsforschung.
Das Gesundheitswesen 2015.



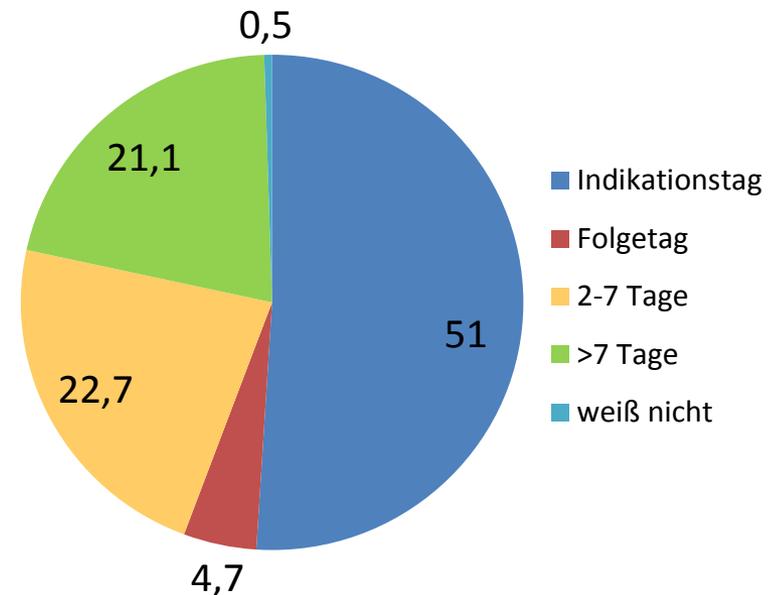
Kraska RA, Krummenauer F, Geraedts M.
Impact of public reporting on the quality of hospital care in Germany: A controlled before-after analysis based on secondary data.
Health Policy 2016.

Nutzung von Qualitätsdaten durch Bürger / Patienten / Krankenhäuser

- Voraussetzung:
 - Bekanntheit- / Verstehen der Informationen / Bewertung als nützlich / Nutzung (Faber 2009 nach Hibbard 2002 und Rogers 1995)
- Realität: Bürger / Patienten / Ärzte kennen die gesetzlichen Qualitätsberichte kaum, finden diese wenig hilfreich (viele präferierte Kriterien fehlen), nutzen sie kaum (Hermeling & Geraedts 2013; de Cruppé & Geraedts 2011)
- Krankenhäuser nutzen die QB, halten aber Aufwand für zu groß (Auras, de Cruppé, Blum & Geraedts, 2012)

Können Patienten überhaupt QB nutzen, um zu wählen?

- BMBF-Studie Krankenhauswahlverhalten [de Cruppé & Geraedts 2014]
- 1925 stationäre Patienten befragt in 17 Krankenhäusern (3 Größenklassen) in NRW, je 50 konsekutiv in 46 Fachabteilungen / disproportional geschichtete Stichprobe aus 11 Fachgebieten mit 92% aller stationären Fälle
- Patienten im Durchschnitt 59 Jahre (22 J. SD); 51% Frauen
- 6% erster Krankenhausaufenthalt, 42% 1-5 Mal, 52% mehr als fünf Voraufenthalte
- Zeit für die Entscheidung (%): siehe Abbildung
- Nur 9% nutzten Internet vor KH-Wahl



Studie „Impact of public reporting on the quality of hospital care in Germany: A controlled before-after analysis based on secondary data.“

Hintergrund:

Internationale Studien belegen, dass die öffentliche Berichterstattung

- Krankenhäuser zu Qualitätsverbesserungen anregt
- wahrscheinlicher zu Verbesserungen führt, wenn die Leistungsanbieter in einem wettbewerbsorientierten Markt agieren

Ziel:

Analyse der Effekte der QB auf die Qualität der Krankenhausversorgung mit Berücksichtigung der Krankenhausträgerschaften (Profitorientierung)

Methode

Design: kontrollierte Prä-Post-Interventionsstudie

- Basis: Qualitätsindikatoren (QI) der externen Qualitätssicherung (eQS)
- Untersuchungszeitraum: 2006 bis 2012
- Intervention: erstmalige Bekanntgabe der veröffentlichungspflichtigen QI
- Interventionsgruppe: veröffentlichungspflichtige QI (Quelle: QB)
- Kontrollgruppe: nicht-veröffentlichungspflichtige QI (Quelle: eQS-Berichte)

Vorgehen

1. Extraktion und Verknüpfung der Daten
2. Statistische Analyse der Interventionsgruppe: zweifaktorielle Varianzanalyse mit Messwiederholungen
3. Vergleich der Interventionsgruppe mit der Kontrollgruppe mittels Kriterium „prozentuale Verbesserung zum Optimum“

(„prozentuale Verbesserung zum Optimum“ = „absolute Verbesserung“ / „Ausschöpfungspotential“)

Ergebnisse

Interventionsgruppe: 6 QI (QI-A bis QI-F) aus 4 Leistungsbereichen

MANOVA mit Messwiederholungen:

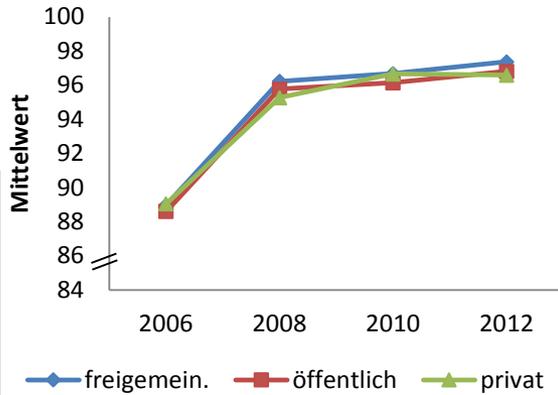
	Zeit						Trägerschaft	
	MANOVA		Differenz zwischen den Jahren				MANOVA	
	p	η^2	2006-2012	2006-2008	2008-2010	2010-2012	p	η^2
QI-A	0,000	0,22	+8,2**	+7,1**	+0,6	+0,6	0,961	0
QI-B	0,000	0,11	+4,9**	+2,4**	+1,7**	+0,8**	0,820	0
QI-C	0,000	0,07	+6,3**	+3,8**	+2,5**	0	0,872	0
QI-D	0,000	0,04	+4,2**	+3,6**	+1,4	-0,8	0,560	0,01
QI-E	0,000	0,13	+13**	+10,8**	+2,1	+0,1	0,719	0,01
QI-F	0,000	0,02	+1,7*	+1,2	+0,4	+0,1	0,163	0,01

**signifikante Verbesserung zum multiplen Niveau 1%; *signifikante Verbesserung zum multiplen Niveau 5%

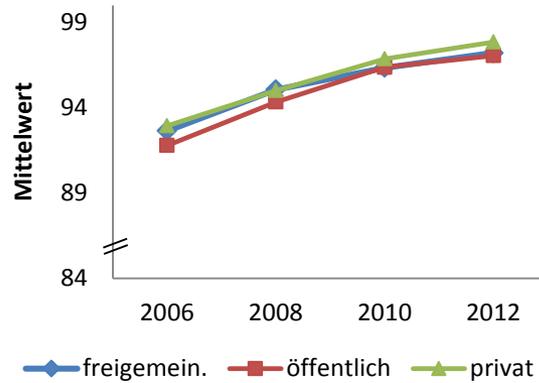
- signifikante Verbesserung bei allen QI
- größter Anteil an Verbesserung direkt nach der Intervention
- keine signifikanten Unterschiede zwischen den Trägergruppen

Ergebnisse

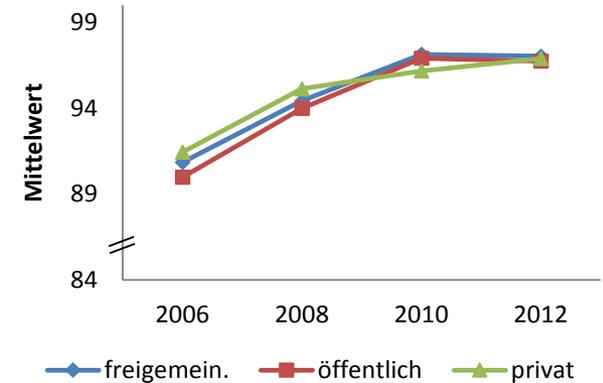
HSM-IMPL2



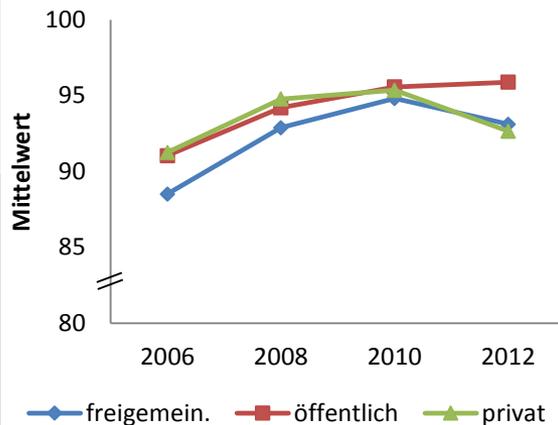
HSM-IMPL4



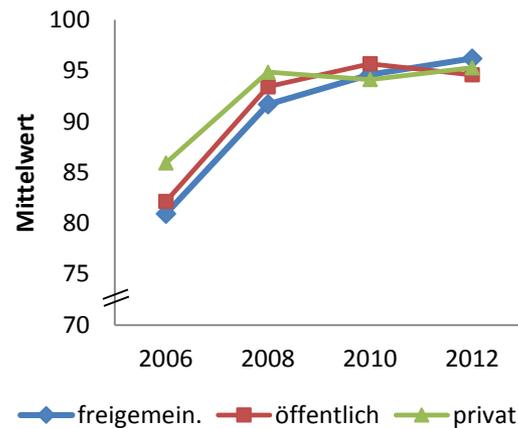
GYN-OP



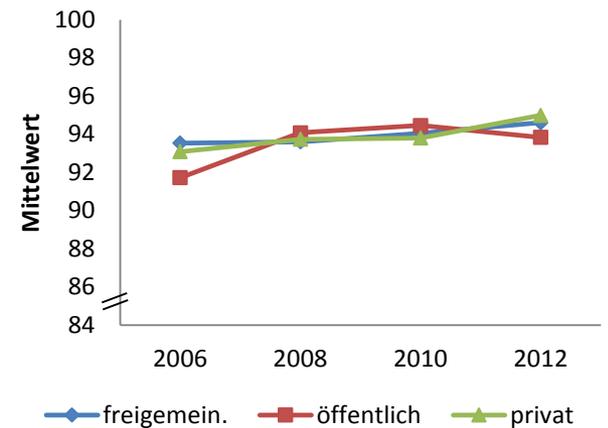
GEBH1



GEBH2



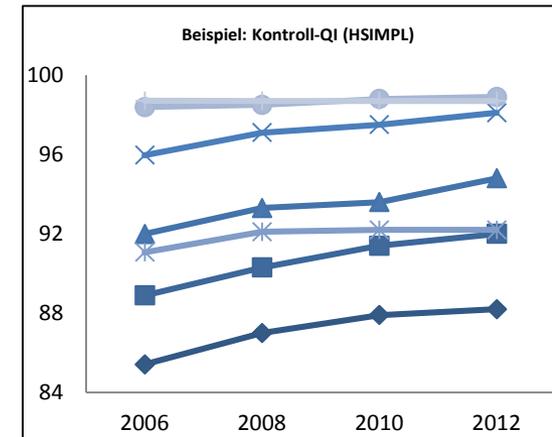
KORO-PCI



Ergebnisse

Kontrollgruppe (31 QI):

- 18 der 31 QI zeigten eine tendenzielle Verbesserung zwischen 2006 und 2012
- 24 der 31 QI wurden zwischen 2008 und 2012 veröffentlichungspflichtig



Vergleich „Interventions- vs. Kontrollgruppe“:

Rang	HSIMPL		GYN-OP		GEBH		KORO-PCI	
	ID	Verbesserung	ID	Verbesserung	ID	Verbesserung	ID	Verbesserung
1	QI(I)-A	73,3% (8,2)	QI(P)-C	66,9% (6,3)	QI(E)-1059	100% (0,02)	QI(E)-F	23,1% (1,7)
2	QI(I)-B	64% (4,9)	QI(P)-666	62,6% (1,2)	QI(P)-E	73,1% (13)	QI(E)-2064	16,8% (1,1)
3	QI(P)-10249	52,9% (2,1)	QI(I)-12874	54,4% (1,9)	QI(P)-D	43,3% (4,2)	QI(I)-2061	3,9% (2,4)
4	QI(P)-10223	35,1% (2,8)	QI(I)-10211	35,5% (7,5)	QI(P)-319	31,8% (0,4)	QI(E)-416	0% (-0,3)
5	QI(E)-583	31,7% (0,5)	QI(P)-665	31,5% (1,6)	QI(E)-322	15,6% (0,2)	QI(E)-414	0% (-0,4)
6	QI(P)-10178	28% (3,1)	QI(P)-612	25,3% (3,2)	QI(E)-324	13,9% (0,4)	QI(E)-417	0% (-0,7)
7	QI(P)-10148	19,1% (2,8)	QI(E)-557	16,7% (0,2)	QI(E)-321	0% (-0,02)	QI(E)-415	0% (-1)
8	QI(E)-582	12,7% (1,1)	QI(I)-672	0% (-0,2)	QI(E)-323	0% (-0,1)	QI(E)-2232	0% (-2,1)
9	QI(E)-1100	0% (-0,01)	QI(E)-51417	0% (-0,1)	QI(E)-331	0% (0)		
10			QI(E)-51418	0% (-0,1)				
11			QI(E)-553	0% (0)				

Fazit zur Wirkung der QB in der eQS

- Generelle, kontinuierliche Verbesserung der QI-Ausprägungen
- die Veröffentlichung von QI-Ausprägungen ist mit einer direkteren und stärkeren Verbesserung der jeweiligen QI assoziiert
- Keine Unterschiede zwischen den Trägergruppen
- Verbliebenes Ausschöpfungspotential vielfach gering – sollte vorhanden sein, damit sich Publikation zum Zweck der Verbesserung lohnt

Gestaltung der Methoden: Auswahl/Abwahl – Bonus/Malus – Vertrag/Ausschluss

- Verschiedenste Optionen:
Zuschläge/Abschläge/Kombination; geteilte Einsparungen; ...
- Finanzierung aus Vorabzug / als Zusatzvergütung
- Höhe der Zu-/Abschläge
- Qualitätsbestimmung aus einzelnen Parametern / Scores, deren Gewichtung
- Zu-/Abschläge wenn einzelne / alle zutreffenden Parameter erfüllt
- Übertragbarkeit der int. Erkenntnisse auf Deutschland?
→ Sozialpsychologie / Verhaltensökonomie einbeziehen

The Effects of Pay-for-Performance Programs on Health, Health Care Use, and Processes of Care

A Systematic Review

Aaron Mendelson, BA; Karli Kondo, PhD; Cheryl Damberg, PhD; Allison Low, BA; Makalapua Motuapuaka, BA; Michele Freeman, MPH; Maya O'Neil, PhD; Rose Relevo, MLIS, MS; and Devan Kansagara, MD, MCR

Background: The benefits of pay-for-performance (P4P) programs are uncertain.

Purpose: To update and expand a prior review examining the effects of P4P programs targeted at the physician, group, managerial, or institutional level on process-of-care and patient outcomes in ambulatory and inpatient settings.

Data Sources: PubMed from June 2007 to October 2016; MEDLINE, PsycINFO, CINAHL, Business Economics and Theory, Business Source Elite, Scopus, Faculty of 1000, and Gartner Research from June 2007 to February 2016.

Study Selection: Trials and observational studies in ambulatory and inpatient settings reporting process-of-care, health, or utilization outcomes.

Data Extraction: Two investigators extracted data, assessed study quality, and graded the strength of the evidence.

Data Synthesis: Among 69 studies, 58 were in ambulatory settings, 52 reported process-of-care outcomes, and 38 reported patient outcomes. Low-strength evidence suggested that P4P programs in ambulatory settings may improve process-of-care outcomes over the short term (2 to 3 years), whereas data on

longer-term effects were limited. Many of the positive studies were conducted in the United Kingdom, where incentives were larger than in the United States. The largest improvements were seen in areas where baseline performance was poor. There was no consistent effect of P4P on intermediate health outcomes (low-strength evidence) and insufficient evidence to characterize any effect on patient health outcomes. In the hospital setting, there was low-strength evidence that P4P had little or no effect on patient health outcomes and a positive effect on reducing hospital readmissions.

Limitation: Few methodologically rigorous studies; heterogeneous population and program characteristics and incentive targets.

Conclusion: Pay-for-performance programs may be associated with improved processes of care in ambulatory settings, but consistently positive associations with improved health outcomes have not been demonstrated in any setting.

Primary Funding Source: U.S. Department of Veterans Affairs.

Ann Intern Med. 2017;166:341-353. doi:10.7326/M16-1881

Annals.org

For author affiliations, see end of text.

This article was published at Annals.org on 10 January 2017.

Fazit: Voraussetzungen für eine qualitätsorientierte Steuerung

- Relevantes Optimierungspotenzial vorhanden
 - Ist-Zustand reliabel und valide messbar (geprüfte und über die Zeit stabile Messgrößen)
 - Diskriminierbarkeit guter / schlechter Qualität (relevante Anteile)
- Optimierung grundsätzlich möglich (behaviorable) und als Ziel allgemein akzeptiert
- Optimierung unter (komplexen) Routinebedingungen in Form einer komplexen Multikomponenten-Intervention belegt
 - Optimierung effektiv und effizient möglich
 - Wirkungen/Nebenwirkungen, Vorteile/Nachteile, Kosten/Nutzen
 - Analyse der Wirkungen bei den verschiedenen Akteuren/Zielgruppen: Patienten / Leistungserbringer / Politik / Gesellschaft
- Keine Einführung ohne vorherige Erprobung

D. Berwick 2015

(Keynote four. Presented at Annu. Natl. Forum on Quality Improv. in Health Care, 27th, Orlando; zitiert nach Doran et al. Annu Rev Public Health 2017;38:449-65)

„ushering in the next era of health care will require reducing excessive measurement, complex incentives, and profit maximization”