

Berliner Herzinfarktregister: **Klinikvergleich 2004/5**

Prof. Dr. Karl Wegscheider

Jan Felix Kersten

Eik Vettorazzi

Klaus Balzer

Universität Hamburg / Berlin

wegsch@econ.uni-hamburg.de

Die nachfolgende Auswertung ist eine Ex-post-Analyse der im Berliner Herzinfarktregister enthaltenen Daten aus den Jahren 2004 und 2005. Sie wurde unternommen, um zu demonstrieren, welche Möglichkeiten in einem mit modernen statistischen Methoden durchgeführten Ranking stecken. Das Ziel war somit nicht der konkrete Vergleich der einzelnen Kliniken. Entsprechend wurden die Kliniken durch Zuweisung einer Codenummer anonymisiert. Auch uns als Auswertern waren und sind bis heute die Codes nicht bekannt.

Als Ex-post-Analyse wurde die Studie natürlich retrospektiv durchgeführt. Wir haben dabei aber eine Ex-ante-Analyse simuliert. Wir haben die Daten prospektiv ausgewertet und vor Beginn der Analyse mit den beteiligten Klinikern die Auswertungskriterien definiert, d.h. nach intensiver Diskussion die Endpunkte (Qualitätskriterien) und die Adjustierungsvariablen (d.h. zu neutralisierenden Einflussgrößen) festgelegt. Dazu haben wir aber bereits über die Kliniken gemittelte Information aus den Daten verwendet. Bei einer „echten“ Ex-ante-Analyse würde man dazu historische Daten, d.h. Daten aus den Vorjahren verwenden.

Gleichwohl gehen wir davon aus, dass diese Vorgehensweise Manipulationen zugunsten der einen oder der anderen Klinik ausschließt.

Empfehlungen für Qualitätsvergleiche

- **Alle Einrichtungen gleich intensiv dokumentieren**
- **Beschränkung auf harte, reproduzierbare Daten**
- **größere Analyse-Intervalle verwenden**
- **hierarchische Modelle verwenden (Korrektur des Varianzexzesses)**
- **sich die statistische Ungenauigkeit vor Augen führen (Konfidenzbereiche) und**
- **man sollte aus Gerechtigkeitsgründen Rankings nur für vergleichbare Patientenkollektive durchführen (Statistische Adjustierung des Patient Mix)**

Diese Folie fasst die allgemeinen methodischen Grundsätze zusammen, die wir für faire Rankings für angemessen halten, vgl. Vortragsfolien Wegscheider vom November 2006.

Die nachfolgende Analyse setzt diese Grundsätze weitgehend, aber nicht vollständig um. Die Intensität der Dokumentation wird zur Zeit noch weitgehend von den Kliniken selbst gesteuert. Es ist denkbar, dass die Kliniken hier unterschiedlich sorgfältig waren. Nicht alle erhobenen Daten waren im methodischen Sinne hart, d.h. unempfindlich gegenüber subjektiven Einschätzungen. Das Analyse-Intervall ist mit zwei Jahren knapp bemessen, drei Jahre sollten für dichotome Endpunkte (ja/nein-Fragen) die Norm sein. Die letzten drei Punkte haben wir (anders als z.B. die BQS) vollständig umgesetzt.

Kollektiv

Infarktpatienten aus 11 Berliner Kliniken 2004-2005

- ohne Intrahospital-Infarkte
- ohne zuverlegte Patienten
- ohne wegverlegte Patienten ohne weitere Angaben

n=1767 Patienten

im Durchschnitt 161 Pat./Klinik, Spanne: 60 bis 344

Die Intrahospital-Infarkte folgen anderen Bedingungen und wurden deshalb ausgeschlossen. Bei zuverlegten Patienten und bei ohne weitere Angaben wegverlegten Patienten gab es Dokumentationsprobleme, die den Ausschluss erforderlich machten.

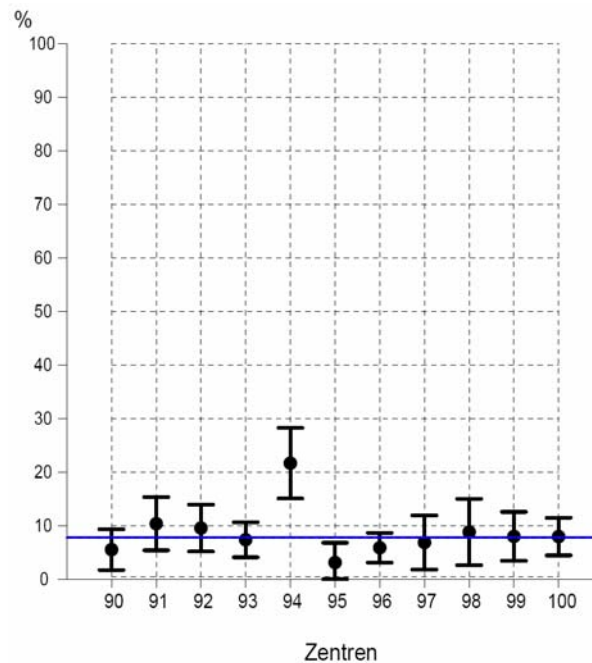
Kollektivbeschreibung

	Gesamt	Spanne	Klinikunterschiede	
	MW		η^2 (Prozent der Gesamtvarianz)	p
Alter [pro Jahr]	67,0 ±13,4	[63,6; 75,3]	6,9%	<0,001
Frau	35,5%	[26,5%; 51,9%]	3,0%	<0,001
BMI > 30	17,2%	[8,8%; 25,0%]	1,7%	0,001
BMI < 20	3,8%	[1,7%; 10,4%]	1,6%	0,002
Raucher	40,2%	[32,4%; 51,0%]	1,5%	0,003
Diabetes Mellitus	28,4%	[19,2%; 44,3%]	2,1%	<0,001
Arterielle Hypertonie	72,8%	[50,0%; 92,5%]	5,5%	<0,001
Hypercholesterinämie	44,5%	[27,5%; 85,8%]	7,5%	<0,001
früherer Infarkt	19,6%	[12,7%; 26,7%]	0,6%	0,433
Bypass-Op	7,2%	[4,7%; 10,0%]	0,2%	0,950
PTCA	14,3%	[7,4%; 18,8%]	0,7%	0,250
Niereninsuffizienz	14,1%	[3,3%; 44,3%]	7,2%	<0,001
ST-Hebungs-Infarkt	54,5%	[23,5%; 70,6%]	6,1%	<0,001
Linksschenkelblock	5,2%	[1,6%; 13,2%]	1,8%	<0,001
Manifeste Herzinsuffizienz	14,2%	[1,9%; 48,1%]	10,7%	<0,001
Ejektionsfraktion < 55%	49,1%	[17,6%; 60,6%]	4,7%	<0,001
Ejektionsfraktion < 35%	11,9%	[4,4%; 20,0%]	1,3%	0,008
Begleittherapie (Rea, Defi, IABP)	6,8%	[0,9%; 12,2%]	2,5%	<0,001

Diese Folie enthält Angaben zu einer Reihe von klinischen Variablen, die als Patienten- bzw. Krankenhauskollektiv-Charakteristika aus klinischer Sicht in Frage kommen. Links ist jeweils das Gesamtmittel, rechts sind die Klinikunterschiede dargestellt. Aus den Zahlen wird die erhebliche Inhomogenität der Patientenpopulationen der einzelnen Kliniken erkennbar. Die Krankenhäuser unterscheiden sich in fast allen Belangen, teilweise um erhebliche Größenordnungen (Herzinsuffizienz, EF). Zu den Unterschieden mögen unterschiedliche Auffassungen über die exakte Definition der Variablen beigetragen haben (z.B. bei der Hypercholesterinämie). Hier wäre in Zukunft eine Vorrunde der beteiligten Ärzte zur Harmonisierung der Variablendefinitionen empfehlenswert. In jedem Fall unterstreicht die Folie, wie wichtig bei solchen Vergleichen eine Adjustierung ist, die einen fairen Ausgleich für den unterschiedlichen Patient Mix der Krankenhäuser schafft. Hier sind wir anderer Auffassung als die BQS, die in der Regel unadjustierte Vergleiche veröffentlicht. In der Diskussionsrunde mit den Klinikern wurde beschlossen, sämtliche hier dargestellten Variablen zur Adjustierung einzusetzen. Aus methodischer Sicht ist es empfehlenswert, für alle Endpunkte stets mit dem gleichen Adjustierungsdatensatz zu arbeiten, auch wenn in einzelnen Jahren oder für einzelne Endpunkte die eine oder andere Variable überflüssig erscheint, weil dadurch Vergleiche zwischen verschiedenen Jahren oder unterschiedlichen Endpunkten möglich werden. Hierzu müssen allerdings genügend Beobachtungen vorhanden sein, damit „die Freiheitsgrade nicht ausgehen“. Im Herzinfarktregister ist das der Fall.

Das Eta-Quadrat misst den Anteil der Gesamtvarianz zwischen den Patienten in dem jeweiligen Kriterium, der auf die Unterschieden zwischen den Klinik-Kollektiven zurückgeführt werden kann, d.h. auf regionale Unterschiede und Unterschiede in der Spezialisierung. Da sich hier auch die Berliner Rettungskette abbildet, dürften diese (epidemiologischen) Zahlen im Einzelnen schwer interpretierbar sein.

Klinikvergleich Stations-Mortalität, beobachtete Raten Konfidenzbereiche aus Modell mit festen Effekten (ANOVA)



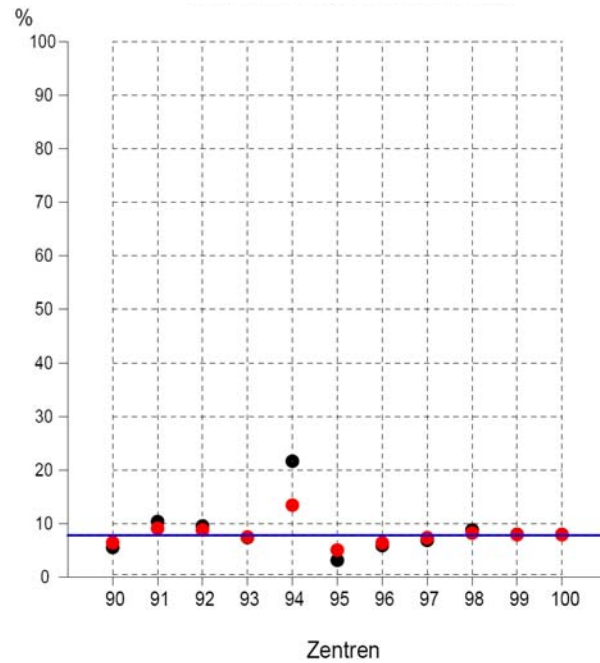
5

Die Abbildung zeigt die Unterschiede zwischen den Kliniken in Bezug auf die Stations-Mortalität, wie sie sich bei einem der verbreiteten „naiven“ Rankings ergeben. Es werden die beobachteten Raten dargestellt, zusammen mit Konfidenzbereichen, wie sie sich aus einer ANOVA ergeben.

Demnach gäbe es eine sehr Klinik mit hoher und eine mit sehr niedriger Stations-Mortalität. Die anderen Klinken scheinen eher zufällig vom Gesamtmittel abzuweichen. Aber Achtung: aus der statistischen Theorie ergibt sich, dass diese Ergebnisse sich bei einem erneuten Ranking unter identischen Bedingungen nicht reproduzieren würden. Die Unterschiede zwischen den Kliniken sind „zu groß“, weil von zufälligen Fehlern überlagert.

Klinikvergleich Stations-Mortalität

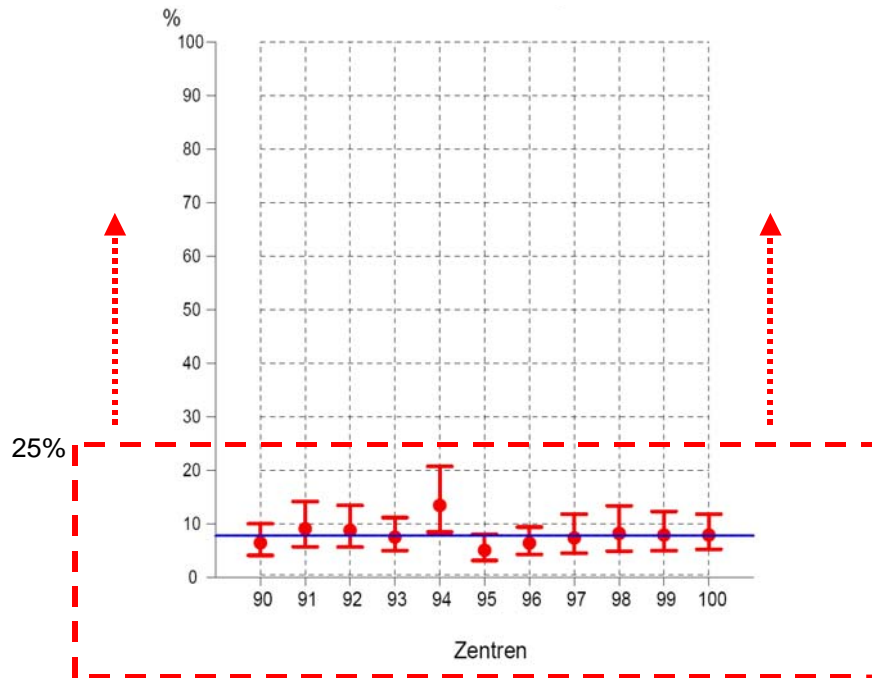
Auswirkungen des Bayesian Shrinkage



6

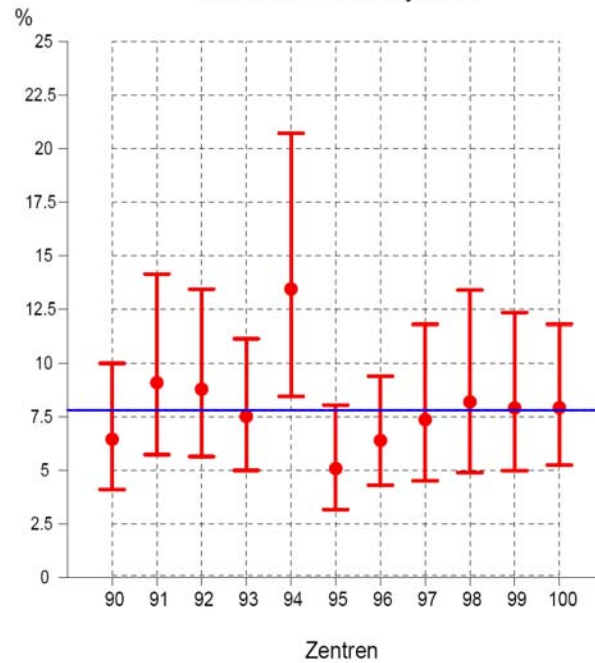
Diese Folie demonstriert, welche Auswirkungen eine bestimmte mathematische Korrektur, das Bayesian Shrinkage (rote Punkte), auf die rohen Mittelwerte (schwarze Punkte) hat. Diese Schätzwerte für die „wahre“, d.h. langfristig reproduzierbare Sterblichkeitsrate minimieren den Vorhersagefehler unter Standard-Annahmen (Normalverteilung der Klinikdifferenzen auf der Log-odds-Skala).

Klinikvergleich Stations-Mortalität, EB-Schätzer und Konfidenzbereiche aus Modell mit zufälligen Effekten



Diese Folie dient lediglich dazu, auf die Skalenanpassung hinzuweisen, die wir zur Verbesserung der optischen Wahrnehmung ab hier einsetzen.

Klinikvergleich Stations-Mortalität, EB-Schätzer und Konfidenzbereiche aus Modell mit zufälligen Effekten



Auch die Bayes-korrigierten Raten weisen bei deutlicher Verringerung der Unterschiede noch auf eine Klinik hin, die eine signifikant höhere Stations-Mortalität aufweist. Aber Vorsicht bei der Interpretation; noch haben wir nicht adjustiert, die Unterschiede in der Mortalität könnten vom unterschiedlichen Patient Mix herrühren.

Adjustierung

Die Unterschiede in den Mortalitätsraten könnten durch einen unterschiedlichen ‚Patient Mix‘ verursacht sein, den die Klinik nicht zu verantworten hat.

Bei einer Adjustierung wird mit Hilfe eines statistischen Verfahrens die (Miss-)Erfolgsquote berechnet, die eine Klinik hätte, wenn sie einen durchschnittlichen ‚Patient Mix‘ zu versorgen hätte.

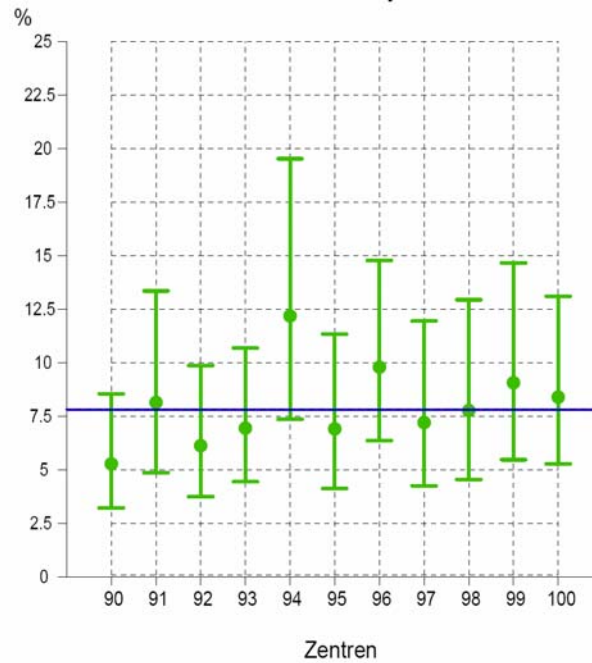
Hier eine „volkstümliche“ Beschreibung des Vorgangs der Adjustierung.

Adjustierungsvariable Stations-Mortalität

	Odds Ratio	p
Alter [pro Jahr]	1,05	<0,001
Frau	1,22	0,394
BMI > 30	0,67	0,253
BMI < 20	1,92	0,132
Raucher	0,67	0,145
Diabetes Mellitus	1,47	0,105
Arterielle Hypertonie	0,57	0,026
Hypercholesterinämie	0,54	0,013
früherer Infarkt	1,16	0,637
Bypass-OP	0,53	0,213
PTCA	1,04	0,917
Niereninsuffizienz	1,75	0,041
ST-Hebungs-Infarkt	1,37	0,189
Linksschenkelblock	1,47	0,303
Manifeste Herzinsuffizienz	1,89	0,024
Ejektionsfraktion < 55	0,72	0,183
Ejektionsfraktion < 35	2,23	0,009
Begleittherapie (Rea, Defi, IABP)	18,82	<0,001

Diese Folie zeigt, in welchem Umfang die ausgewählten Adjustierungsvariablen Assoziationen mit der Mortalität aufweisen. Rot markiert sind mit signifikant erhöhtem Risiko verbundene Kovariate, grün mit signifikant erniedrigtem Risiko verbundene. Letztere sind in der Regel selektionsbedingt. So erleben z.B. Raucher (hier nicht signifikant), Hypertoniker und Patienten mit erhöhten Lipidwerten im Mittel ihren Herzinfarkt - gemessen am alterkorrigierten Allgemeinzustand - früher als andere. Dadurch sind diese Risikofaktoren nicht nur hier, sondern in allen größeren Studien mit einem verminderten Postinfarkt-Überlebensrisiko verbunden.

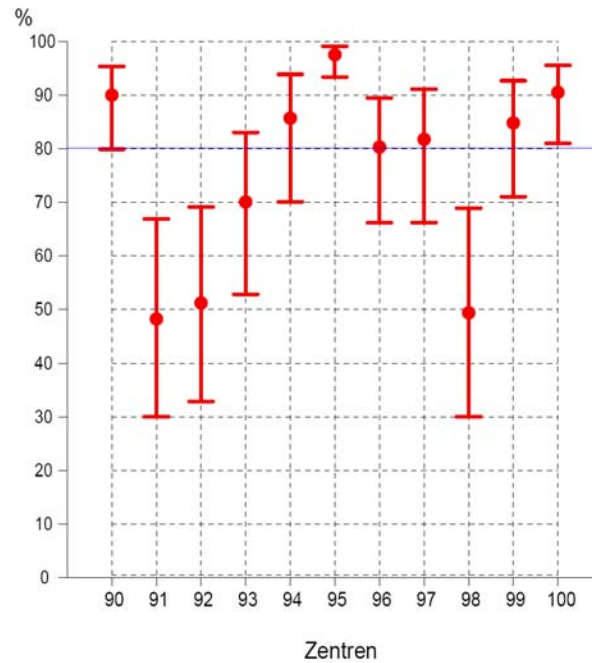
Klinikvergleich Stations-Mortalität, EB-Schätzer adjustiert für 18 Baseline-Variable



p=0,091 (n.s.)

Durch die Adjustierung verändern sich die Positionen der Kliniken untereinander (vgl. mit Folie 8). Die Klinik 94 bleibt die Klinik mit der höchsten Stations-Mortalität, rückt aber näher an die anderen Kliniken heran. Man kann sich nun nicht mehr sicher sein, ob die Unterschiede nicht rein zufällig zustande gekommen sind. Hier fehlt das eigentlich erforderliche dritte Jahr zur statistischen Klärung.

Klinikvergleich Reperfusionrate (n=1726), EB-Schätzer und Konfidenzbereiche aus Modell mit zufälligen Effekten



12

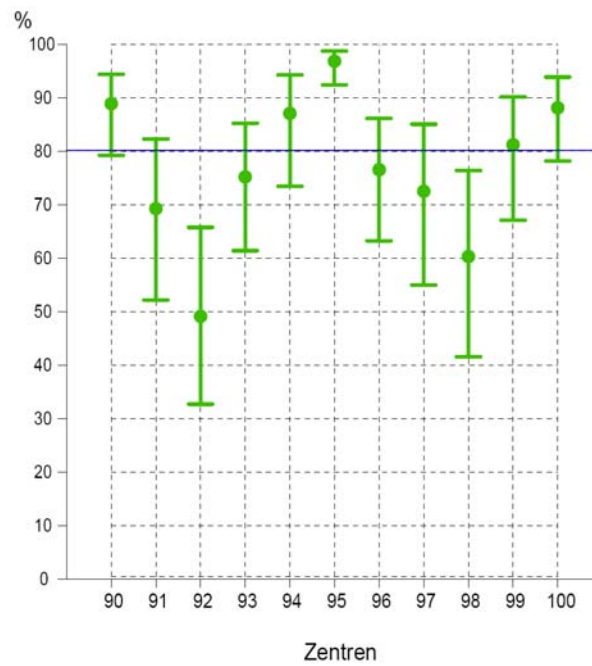
In dieser Darstellung werden die erreichten Reperfusionraten (bezogen auf alle Patienten; nicht nur die für die PTCA erwählten) verglichen, zunächst ohne Adjustierung. Diese Größe ist einerseits Prozess-, andererseits Ergebnisparameter, was die Interpretation erschwert. Die Unterschiede zwischen den Kliniken sind erheblich.

Adjustierungsvariable Reperfusionrate

	Odds Ratio	p
Alter [pro Jahr]	0,95	<0,001
Frau	0,68	0,011
BMI > 30	1,40	0,119
BMI < 20	0,46	0,024
Raucher	0,99	0,949
Diabetes Mellitus	1,27	0,140
Arterielle Hypertonie	0,60	0,009
Hypercholesterinämie	0,96	0,772
früherer Infarkt	0,77	0,217
Bypass-OP	0,86	0,588
PTCA	1,90	0,008
Niereninsuffizienz	0,57	0,004
ST-Hebungs-Infarkt	2,88	<0,001
Linksschenkelblock	0,74	0,288
Manifeste Herzinsuffizienz	0,49	<0,001
Ejektionsfraktion < 55	1,00	0,978
Ejektionsfraktion < 35	0,88	0,577
Begleittherapie (Rea, Defi, IABP)	1,33	0,352

Von den Adjustierungsvariablen sind teils andere Größen mit der Reperfusionrate assoziiert als mit der Mortalität. Hier markiert rot die Assoziation mit niedrigen Reperfusionraten, weil das hier die ungünstigen Ergebnisse sind, grün die mit höheren Reperfusionsraten.

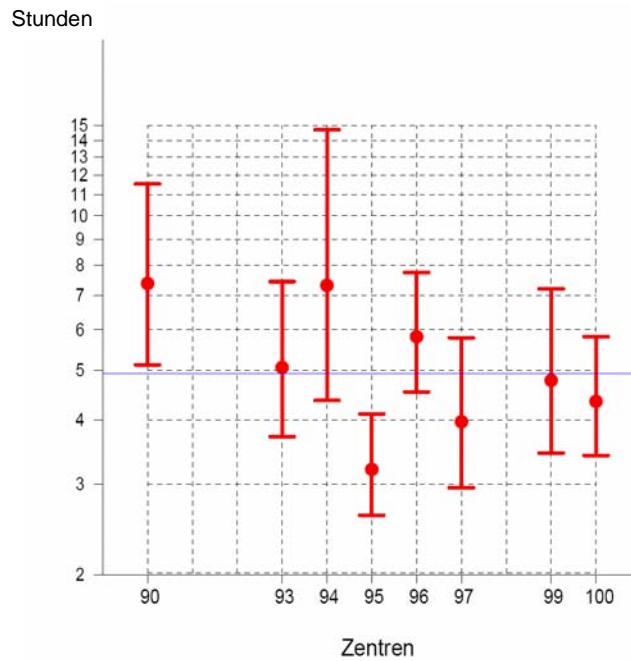
Klinikvergleich Reperfusionrate, EB-Schätzer adjustiert für 18 Baseline-Variable



p<0,001

Nach Adjustierung sind weiterhin beträchtliche signifikante Unterschiede zwischen den Kliniken zu erkennen.

Klinikvergleich Zeitintervall ‚Door-to-reperfusion‘ (nur ST-Hebungsinfarkt, n=722), EB-Schätzer, Konfidenzbereiche aus Modell m. zufälligen Effekten

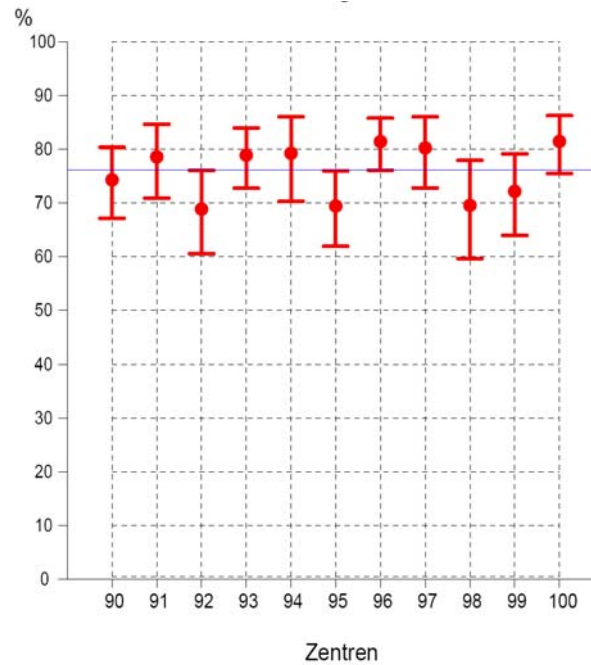


p=0,178 (n.s.)

15

Die Zeiten von Aufnahme bis Reperfusion (nicht: PTCA-Beginn, sondern eröffnetes Gefäß; Selbstangaben) sind nur für ST-Hebungs-Infarkte von Bedeutung. Hier wurde im Einvernehmen mit den Klinikern keine Adjustierung vorgenommen, weil es sich um einen reinen Prozessparameter handelt.

Klinikvergleich Entlassungsmedikation, EB-Schätzer, Konfidenzbereiche aus Modell m. zufälligen Effekten



$p < 0,001$

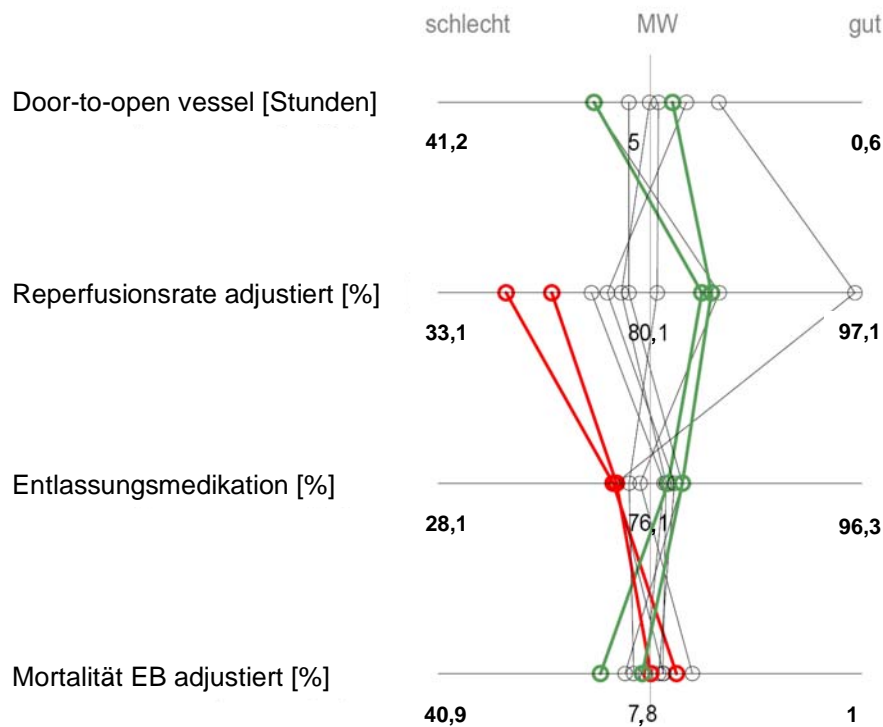
Bei den Entlassungsmedikationsraten (Anteil der Patienten mit leitliniengerechter Vierer-Medikation: Thrombozytenaggregationshemmer, Beta-Blocker, ACE-Hemmer, CSE-Hemmer) scheint es zwei Typen von Kliniken zu geben, die sich um knapp 10% unterscheiden.

Adjustierungsvariable Entlassungsmedikation

	Odds Ratio	p
Alter [pro Jahr]	0,98	<0,001
Frau	0,78	0,068
BMI > 30	1,37	0,078
BMI < 20	0,53	0,031
Raucher	1,18	0,248
Diabetes Mellitus	0,98	0,909
Arterielle Hypertonie	1,32	0,065
Hypercholesterinämie	1,25	0,088
früherer Infarkt	0,83	0,302
Bypass-Op	1,63	0,069
PTCA	1,15	0,498
Niereninsuffizienz	0,75	0,132
ST-Hebungs-Infarkt	1,24	0,099
Linksschenkelblock	1,07	0,825
Manifeste Herzinsuffizienz	0,93	0,713
Ejektionsfraktion < 55%	0,80	0,091
Ejektionsfraktion < 35%	0,91	0,646
Begleittherapie (Rea, Defi, IABP)	0,63	0,115

Die Medikationsgewohnheiten weisen kaum Assoziationen mit den Baseline-Kovariaten auf. Die Gründe für das unterschiedliche Verordnungsverhalten ließen sich nicht aufklären.

Klinikprofile



18

Diese Folie zeigt eine Analyse, die auf der Basis der zur Zeit vorhandenen Daten als eher problematisch anzusehen ist, lediglich zu Demonstrationszwecken. Hier sind die vier Endpunkte parallel dargestellt. Jede Klinik zeigt so ein eigenes Profil in den Endpunkten. Die Schwierigkeit besteht darin, dass hier Prozess- und Ergebnisparameter, adjustierte und nicht-adjustierte Analysen sowie unterschiedliche Skalen gemischt werden. Solche Profile lassen sich aber auf mehrere Arten erstellen, der hier gezeigten Auswahl haftet somit eine gewisse Willkür an. Bitte mit Vorsicht interpretieren.

Immerhin sind die Skalen analog gepolt: rechts sind immer die guten Ergebnisse. Man sieht, dass es z.B. Kliniken mit schlechten Prozessparametern, aber geringer Mortalität gibt (rot, Kliniken ohne eigenes Katheterlabor) und umgekehrt (grün). Profile sind zur Beurteilung der Qualität der Versorgung informativer und augenscheinlich besser geeignet als Rankings.

Schlussfolgerungen

- 1) **Prinzipiell lassen sich mit Hilfe eines sorgfältig geführten vollständigen Registers relevante Klinik-Unterschiede darstellen. Das Register darf nicht zu klein sein oder Lücken aufweisen.**
- 2) **Eine naive Analyse produziert künstliche Unterschiede. Eine saubere statistische Modellierung und Adjustierung ist unerlässlich, wenn man gerechte Vergleiche durchführen will und belastbare Aussagen anstrebt.**
- 3) **Profile sind nützlicher als Rankings, weil Kliniken unterschiedliche Stärken und Schwächen aufweisen.**

- 4) **Unterschiede zwischen den Berliner Kliniken in der Krankenhaussterblichkeit nach Herzinfarkt konnten *nicht* statistisch gesichert werden (Ergebnisqualität).**
- 5) **Hingegen gibt es relevante Unterschiede in der Reperfusionrate und in der Entlassungsmedikation (Prozessqualität)**
- 6) **Die Patientenzufriedenheit ist ein weiteres Kriterium, das bisher im Register fehlt und möglicherweise andere Kliniken vorn gesehen hätte.**

Diese Folie enthält unsere Schlussfolgerungen zu Methodik (oben) und Inhalt (unten).