

14 Qualitätssicherung mit Routinedaten – Aktueller Stand und Weiterentwicklung

Günther Heller

Abstract

Im vorliegenden Beitrag wird die bisherige Entwicklung des Projekts Qualitätssicherung mit Routinedaten (QSR) dargestellt. Darüber hinaus wird das Prävalenz-Fallzahl-Problem erläutert und gezeigt, dass eine sichere Beurteilung der Ergebnisqualität mit den etablierten Verfahren der Ergebnisqualitätsmessung aus einfachen statistischen Gründen bislang nur sehr eingeschränkt möglich war. Anschließend wird eine mögliche Lösung dieses Problems, die Verwendung von Qualitätsindizes nebst deren Konstruktion und Implementierung in QSR vorgestellt. Zusätzlich werden geplante methodisch wissenschaftliche Weiterentwicklungen von QSR skizziert. Sie beziehen sich auf Weiterentwicklungen der o. g. Qualitätsindizes, zusammen mit Patienten, Einweisern, medizinischen Fachexperten und Kliniken, aber auch auf mögliche Weiterentwicklungen methodischer und statistischer Verfahren, inklusive der Entwicklung weiterer sektorenübergreifender Tracer.

This article describes the current development of the project Quality Assurance of Hospital Care with Routine Data (QSR). In addition, it explains the problem with small sample sizes and shows that for simple statistical reasons, a reliable assessment of outcome quality with the established procedures of outcome quality measurement is limited. The article presents a possible solution to this problem, i. e. the use of quality indices, as well as their design and implementation in QSR. It also outlines further methodological developments of QSR. These refer to developments of the above indices in co-operation with patients, hospital referrers, medical specialists and clinics, but also to possible developments of methodological and statistical procedures, including the development of further cross-sectoral tracers.

14

14.1 Einführung – die Entwicklung von QSR

QSR wurde im Jahr 2002 als ein Gemeinschaftsprojekt des AOK-Bundesverbandes, der HELIOS-Kliniken GmbH, des Forschungs- und Entwicklungsinstituts für das Sozial- und Gesundheitswesen in Sachsen Anhalt (FEISA) und des Wissenschaftlichen Instituts der AOK (WiDO) initiiert. Ziel war es, ein Qualitätssicherungsverfahren bzw. genauer ein Qualitätsmessverfahren für den stationären Sektor zu entwickeln, welches auf Routinedaten der Krankenkassen basiert und an Ergebnis-

qualität orientiert ist. Dabei wurde davon ausgegangen, dass valide Qualitätsinformationen (zumindest) für drei Akteure im Gesundheitswesen von großem Interesse sind:¹

- den Kliniken selbst, z. B. zur Verwendung im internen Qualitätsmanagement
- den Kostenträgern, um in selektiven vertraglichen Regelungen Qualitätsaspekte als Vertragsgegenstände einfließen lassen zu können
- Patienten und Krankenhauseinweisern zur Auswahl von Kliniken mit möglichst guter Qualität.

Die erste Entwicklungsphase endete im März 2007 mit der Veröffentlichung des QSR-Abschlussberichts.² In QSR wird allerdings nicht der Anspruch erhoben, die Qualität des gesamten Leistungsspektrums interessierender Krankenhäuser zu beurteilen; vielmehr werden bestimmte Leistungsbereiche (Tracer) definiert und im Detail analysiert. Insgesamt wurden bislang acht operative wie auch nicht operative Tracer entwickelt. Im Einzelnen handelt es sich um: akuten Myokardinfarkt, Herzinsuffizienz, Schlaganfall, Operation bei kolorektalem Karzinom, Appendektomie, Hüftgelenks-Endoprothese bei Hüftfraktur, Hüft-Totalendoprothese bei Coxarthrose und Kniegelenks-Totalendoprothese. Als erstes Produkt dieser Entwicklungsphase wurde ein QSR-Klinikbericht zur Verwendung im internen Qualitätsmanagement entwickelt, welcher über 3M Healthcare vertrieben wird. Aktuell liegt die dritte Jahresauswertung des mittlerweile an zentralen Punkten überarbeiteten QSR-Klinikberichts vor.³

Ein zentraler Vorteil des QSR-Verfahrens gegenüber traditionellen Qualitätssicherungsverfahren besteht darin, dass auch Qualitätseignisse jenseits des aktuell betrachteten Krankenhausaufenthaltes analysiert werden können. Dadurch ist eine Konstruktion von valideren und sinnvollerer Indikatoren der Ergebnisqualität möglich (Heller 2008; Heller et al. 2008; Heller et al. 2004).

14.2 Das Prävalenz-Fallzahl-Problem bei der Messung von Ergebnisqualität

Eine wesentliche Überarbeitung bzw. Erweiterung des QSR-Klinikberichts besteht in der Verwendung von Indizes potenzieller Komplikationen, (sogenannter Qualitätsindizes, vgl. weiter unten).

Vorangegangen war die Beobachtung bzw. die Erkenntnis, dass einige der ursprünglichen A-Indikatoren sehr geringe Prävalenzen in Kombination mit vergleichs-

1 Neben den im Folgenden aufgelisteten Verwendungsmöglichkeiten wurde die aktuelle QSR-Methodik mittlerweile bei mehreren Landesministerien vorgestellt, u. a. unter der Frage, ob Qualitätsinformationen auch zum Zwecke der Krankenhausplanung genutzt werden können.

2 URL: http://wido.de/fileadmin/wido/downloads/pdf_krankenhaus/wido_kra_qsr-abschlussbericht_0407.pdf (Zugriff am 05. August 09).

3 URL: http://solutions.3mdeutschland.de/wps/portal/3M/de_DE/his/drg/product-information/quality-management/qsr/ (Zugriff am 05. August 09).

Tabelle 14–1

Durchschnittliche Prävalenzen der A-Indikatoren in QSR, Berechnung minimal benötigte Fallzahl gemäß Poweranalyse

QSR-Tracer	A-Indikator*	Durchschnittliche Prävalenz AOK-Patienten 2007 (%)**	Minimal benötigte Patientenzahl***	Durchschnittliche Anzahl AOK-Patienten pro Klinik 2007 (Interquartils-grenzen)
Herzinsuffizienz	90-Tage-Sterblichkeit	19,22	31	101,5 (47–78–130)
Herzinfarkt	30-Tage-Sterblichkeit	15,92	39	61,7 (21–40–85)
Hirnfarkt oder intrazerebrale Blutung	30-Tage-Sterblichkeit	14,46	45	82,4 (24–51–120)
Kolon- bzw. Rektum-Operation bei kolorektalem Karzinom	90-Tage-Sterblichkeit	9,92	70	25,9 (14–21–30)
Appendektomie	30-Tage-Sterblichkeit	0,49	1629	38,0 (21–32–49)
Implantation einer Hüftgelenks-Totalendoprothese bei Coxarthrose (elektive Hüft-TEP)	30-Tage-Sterblichkeit	0,34	2353	53,2 (23–37–69)
Implantation einer Hüftgelenks-Endoprothese bei Hüftfraktur	30-Tage-Sterblichkeit	7,75	93	22,7 (14–19–28)
Implantation einer Kniegelenks-Totalendoprothese (Knie-TEP)	30-Tage-Sterblichkeit	0,15	5348	55,0 (24–40–69)

* Quelle: http://wido.de/fileadmin/wido/downloads/pdf_krankenhaus/wido_kra_qsr-abschlussbericht_0407.pdf (Zugriff am 05. August 2009)

** vollstationäre AOK-Patienten mit Entlassungsdatum in 2007, nur Patienten aus Kliniken mit mindestens zehn Tracerfällen

*** Um bei Verdopplung des A-Indikators in einer Klinik ein Signifikanzniveau von 95 % und eine Power von 80 % zu erreichen

Krankenhaus-Report 2010

WIDO

weise geringen Fallzahlen aufwiesen (Tabelle 14–1). Beispielsweise zeigt sich eine durchschnittliche 30-Tage-Sterblichkeit nach Implantation einer Hüftgelenks-Totalendoprothese bei Coxarthrose (elektive Hüft-TEP) von 0,34 %. Betrachtet man die absoluten Häufigkeiten von Todesfällen nach Implantation einer elektiven Hüft-TEP, wird dieses Problem noch deutlicher: Von den insgesamt 906 Kliniken, die im Jahr 2007 mindestens zehn Patienten nach elektiver Hüft-TEP entließen, fanden sich bei 757 (83,6 %) keine Todesfälle innerhalb von 30 Tagen nach der Operation. In 136 Kliniken trat ein Todesfall, in 12 Kliniken traten zwei und in einer Klinik traten drei Todesfälle auf. Dabei ist intuitiv einsichtig, dass eine Anwendung von Verfahren der schließenden Statistik, wie etwa eine Berechnung von Konfidenzintervallen, nahezu immer zu dem Ergebnis führen wird, dass hier kein signifikantes Ergebnis vorliegt. Anders formuliert ist hier – aufgrund der geringen Prävalenz dieses Indikators der

Tabelle 14–2

Anzahl Sterbefälle pro Klinik innerhalb von 30 Tagen nach Implantation einer Hüftgelenks-Totalendoprothese bei Coxarthrose (elektive Hüft-TEP)*

Anzahl Verstorbene innerhalb von 30 Tagen nach Aufnahme	Anzahl Kliniken	Anteil (%)
0	757	83,6
1	136	15,0
2	12	1,3
3	1	0,1
Gesamt	906	100,0

* vollstationäre AOK-Patienten mit Entlassungsdatum in 2007

Krankenhaus-Report 2010

WIdO

Ergebnisqualität – bei den gegebenen Fallzahlen pro Klinik und Jahr von ganz überwiegend zufälligen Ergebnissen auszugehen (Tabelle 14–2).

Noch plastischer wird dies, wenn analysiert wird, wie sehr die Ergebnisse für diese Tracer von Jahr zu Jahr schwanken. Um dies zu illustrieren, wurde die Prognosefähigkeit von klinikspezifischen Sterberaten für klinikspezifische Sterblichkeiten des Folgejahres analysiert. In Tabelle 14–3 sind die Ergebnisse solcher Analysen unter Verwendung von linearen Regressionen wiederum für den Tracer elektive Hüft-TEP und die Jahresübergänge 2004/2005, 2005/2006 und 2006/2007 ausgewiesen. Sowohl für rohe als auch für risikoadjustierte Sterblichkeiten (SMRs)⁴ zeigen sich erklärte Varianzen meist deutlich unter 1%. Dementsprechend stellen die Sterblichkeiten des Vorjahres auch keine signifikanten oder irgendwie sinnvollen Prädiktoren für die Sterblichkeiten des Folgejahres dar. Dieser Sachverhalt ändert sich nur unwesentlich, wenn nur Kliniken mit mindestens 100 Fällen betrachtet werden (Tabelle 14–3).⁵

Hier könnte eingewendet werden, dass aufgrund der Verteilung der analysierten Variablen einiges dafür spricht, dass die Grundannahmen einer linearen Regression verletzt sind. Allerdings finden sich auch nach Rekodierung der verwendeten Variablen ähnliche Ergebnisse. So zeigen auch Kliniken, die im Basisjahr ein SMR von drei und mehr aufweisen – was einer Erhöhung der risikoadjustierten Mortalität auf mindestens das Dreifache entspricht – keine prognostische Kraft für die SMR der Folgejahre. Die Regressionskoeffizienten sind für alle drei analysierten Jahresübergänge negativ, was bedeutet, dass die durchschnittliche risikoadjustierte Mortalitätsrate dieser Kliniken im Folgejahr (im Durchschnitt) niedriger als im Basisjahr

4 In QSR werden risikoadjustierte Sterblichkeiten über standardisierte Mortalitäts-Ratios (SMR) ausgewiesen. Dabei werden über logistische Regressionen erwartete Sterblichkeiten für jede Klinik geschätzt. Der Quotient der tatsächlich beobachteten Todesfälle (O) zu den erwarteten Sterblichkeiten (E) bilden das SMR. $SMR = O/E$.

5 Dies mag als ein Hinweis darauf verstanden werden, dass eine alleinige Verlängerung des Analysezeitraums – etwa eine Berechnung von klinikspezifischen Mortalitätsraten wie bei den HELIOS-Indikatoren (vgl. Tabelle 14–5) – dieses Problem meist nicht lösen kann.

Tabelle 14–3

Prognosefähigkeit von klinikspezifischen rohen Sterblichkeiten, standardisierten Mortalitäts-Ratios (SMRs) und Qualitätsindizes 2004–2007 nach Implantation einer elektiven Hüft-TEP*

Prognose von – bis	2004 nach 2005	2005 nach 2006	2006 nach 2007
Rohe Sterblichkeit			
Mindestfallzahl 10*			
R2 (erklärte Varianz) in %	0,05	0,07	0,03
Mindestfallzahl 100**			
R2 (erklärte Varianz) in %	2,73	0,46	1,46
Standardisiertes Mortalitäts Ratio (SMR)			
Mindestfallzahl 10*			
R2 (erklärte Varianz) in %	0,11	0,04	0,03
Mindestfallzahl 100**			
R2 (erklärte Varianz) in %	4,67	0,33	0,19
Qualitätsindex			
Mindestfallzahl 10*			
R2 (erklärte Varianz) in %	6,26	10,63	12,79
Mindestfallzahl 100**			
R2 (erklärte Varianz) in %	22,42	38,22	55,82

* vollstationäre AOK-Patienten, nur Patienten aus Kliniken mit mindestens zehn Tracerfällen im Basisjahr

** vollstationäre AOK-Patienten, nur Patienten aus Kliniken mit mindestens 100 Tracerfällen im Basisjahr

Krankenhaus-Report 2010

WIdO

14

ist (Ergebnisse nicht dargestellt).⁶ Auch eine Rekodierung sowohl der abhängigen als auch der unabhängigen Variablen mit anschließender kategorialer Datenanalyse ergibt keine signifikanten Prognosen (Ergebnisse nicht dargestellt).

Zwar kann eingewendet werden, dass die Kliniken auch einer gewissen Qualitätsdynamik unterliegen, ein gewisser Anteil nicht aufgeklärter Varianz also bei der Prognose erwartet werden muss. Die hier dargelegten Ergebnisse zeigen aber, dass für das hier analysierte Beispiel faktisch kein Zusammenhang der Mortalitätsraten eines Jahres mit dem folgenden Jahr besteht bzw. sie keine Prognosefähigkeit für das Folgejahr haben. Insgesamt sprechen die vorgelegten Analysen stark dafür, dass keine relevante Qualitätsinformation abgebildet wurde bzw. so stark von zufälligen Effekten überlagert ist, dass sie letztlich nicht erkennbar ist.

6 Dies lässt zunächst an einen „regression to the mean“-Effekt denken. Gemeint ist die Tatsache, dass Analyseeinheiten (hier Kliniken), die zufällig einen Extremwert in einer Analyse annehmen, sich in der Folge wieder in Richtung Gesamtmittelwert bewegen (Bland und Altman 1994a/1994b). Ein Ergebnis, welches in sogenannten empirischen Bayes-Analysen genutzt wird. Hier werden extreme Ergebnisse kleinerer Analyseeinheiten in Richtung des Gesamtmittelwertes korrigiert, wobei gezeigt werden konnte, dass die Prognosefähigkeit dadurch verbessert wurde (Ash et al. 2003).

Tabelle 14–4

BQS-Ergebnisqualitätsindikatoren nach Qualitätsbericht 2006, Berechnung minimal benötigter Fallzahlen gemäß Poweranalyse*

Leistungsbereich	Qualitätsindikatoren	Durchschnittliche Prävalenz Qualitätsindikator (%)	Minimal benötigte Fallzahl je Klinik	Durchschnittliche Fallzahl je Klinik***
Cholezystektomie	Reinterventionensrate	1,11	713	136
Herzschrittmacher-Implantation	chirurgische Komplikationen	1,41	559	64
	Sondendislokation im Vorhof	1,42	555	
	Sondendislokation im Ventrikel	1,42	555	
Hüft-Endoprothesen-Erstimplantation	Reinterventionen wegen Komplikation	2,46	316	124
	Endoprothesenluxation	0,62	1 285	
	Postoperative Wundinfektion	0,82	969	
Karotis-Rekonstruktion**	Perioperative Schlaganfälle oder Tod	2,77	298	49
Kniegelenkersatz	Ungeplante Folgeoperation(en) wegen Komplikation(en)	0,55	1 450	125
	Entzündung des Operationsbereichs nach der Operation	1,89	414	
Herzkatheteruntersuchung und -behandlung	Wiederherstellung der Durchblutung der Herzkranzgefäße	100–93,27	108	857
Operation an den Herzkranzgefäßen	Sterblichkeitsrate während des stationären Aufenthalts	3,41	225	605

* Um bei einer Verdopplung der Rate des Qualitätsindikators in einer Klinik ein Signifikanzniveau von 95 % und eine Power von 80 % zu erreichen

** ohne Berücksichtigung der Risikoadjustierung, vgl. Text

*** BQS-Bundesauswertung 2006, URL: www.bqs-outcome.de

Quelle: BQS-Bundesauswertung 2006, URL: www.bqs-outcome.de

Krankenhaus-Report 2010

WIdO

Allerdings ist zu konstatieren, dass dieses Problem auch andere eingeführte Qualitätsmessungssysteme betrifft, die Ergebnisqualität im Fokus haben: Tabelle 14–4 zeigt durchschnittliche Prävalenzen der für den gesetzlichen Qualitätsbericht uneingeschränkt zur Veröffentlichung empfohlenen Indikatoren für Ergebnisqualität des Jahres 2006 aus dem BQS-Verfahren.⁷ Tabelle 14–5 gibt einen Überblick über die HELIOS-Indikatoren⁸, die sich auf Ergebnisqualität beziehen und Zielwerte ausweisen. In beiden Tabellen sind zahlreiche Indikatoren der Ergebnisqualität im niedrigen Prozentbereich oder auch unter 1 % zu erkennen.

Dabei kann die Frage, ob ein Qualitätsindikator einen tatsächlichen existierenden Unterschied abzubilden vermag, auch mit Hilfe von Poweranalysen bzw.

7 URL: www.bqs-outcome.de (Zugriff am 05. August 09) sowie URL: <http://www.bqs-online.com/download/uebersicht-qi-laien.pdf> (Zugriff am 05. August 09).

8 URL: <http://www.helios-klinikfuehrer.de/kn/klinik/helios-uebersicht/qualitaetsberichte-als-pdf.html> (Zugriff am 05. August 09).

Tabelle 14–5
HELIOS-Ergebnis-Indikatoren (II. Generation)*, Berechnung minimal benötigter Fallzahlen gemäß Poweranalyse

Leistungsbereich	Helios-Zieldefinitionen 2006/2007** für Anteil Todesfälle (%)	Minimal benötigte Fallzahl je Klinik gemäß Poweranalyse***	Anteil HELIOS-Kliniken mit ausreichender Fallzahl auf diesem Indikator (in %) ****
Herzinfarkt	10,7	64	82,6
Herzinsuffizienz	11,4	60	79,3
Behandlungsfälle mit Linksherzkatheter (ohne OP ohne Herzinfarkt)	0,5	1 597	36,8
Schlaganfall	11,4	60	87,0
Pneumonie	11,2	61	84,6
Cholezystektomie	0,4	1 999	0,0
Hemiotomie	0,1	6 687	0,0
Kolonresektion bei kolorektalem Karzinom	6,0	123	9,5
Rektumresektion bei kolorektalem Karzinom	6,0	123	6,7
Kolonresektion bei Divertikel ohne Abszess/Perforation	2,0	391	0,0
Aortenaneurysma nicht rupturiert abdominal, offen operiert	5,0	150	0,0
extrakranielle Gefäßoperationen	2,0	391	0,0
Stentimplantation in extrakranielle Gefäße, Anteil Todesfälle	2,0	391	0,0
Maternale Mortalität	0,005	160 743	0,0
Dammris 3. und 4. Grades	2,4	324	93,3
Hysterektomie bei gutartigen Erkrankungen	0,014	57 401	0,0
Hüft-Endoprothesen-Erstimplantation	0,26	3 081	0,0
Hüft-Totalendoprothesen-Wechsel und Komponentenwechsel (BQS 17/β), Anteil Todesfälle	1,1	720	0,0
Knie-Totalendoprothesen-Erstimplantation (BQS 17/5+17/6), Anteil Todesfälle	0,1	8 027	0,0

Tabelle 14–5

Fortsetzung

Leistungsbereich	Helios-Zieldefinitionen 2006/2007** für Anteil Todesfälle (%)	Minimal benötigte Fallzahl je Klinik gemäß Poweranalyse***	Anteil HELIOS-Kliniken mit ausreichender Fallzahl auf diesem Indikator (in %) ****
Wechsel einer Knie-TEP (nach BQS 17/7), Anteil Todesfälle	0,22	3 643	0,0
Schenkelhalsfraktur, alle Altersgruppen, Anteil Todesfälle	5,3	141	24,0
Nephrektomie, Anteil Todesfälle	3,0	257	0,0
Partielle Nephrektomie, Anteil Todesfälle	0,5	1 597	0,0
Prostata-TUR, Anteil Todesfälle	0,2	4 008	0,0
Radikale Prostatovesikulektomie bei Karzinom, Anteil Todesfälle	0,5	1 597	0,0
Beatmung > 24 Stunden (ohne Neugeborene), Anteil Todesfälle	35,0	12	87,5
Sepsis (DRG T60), Anteil Todesfälle	25,7	20	78,6

* Quelle: <http://www.helios-klinikfuehrer.de/kn/klinik/helios-uebersicht/qualitaetsberichte-als-pdf.html> (eingeschlossen wurden alle Kliniken, für die am 05.08.09 Daten zur Verfügung standen)

** Dies entspricht den Mittelwerten auf diesen Indikatoren, die von HELIOS aus unterschiedlichen Quellen zusammengetragen wurden; http://www.helios-klinikfuehrer.de/fileadmin/user_upload/KF_HELIOS_QI.pdf

*** Um bei einer Verdopplung der Rate des Qualitätsindikators in einer Klinik ein Signifikanzniveau von 95 % und eine Power von 80 % zu erreichen

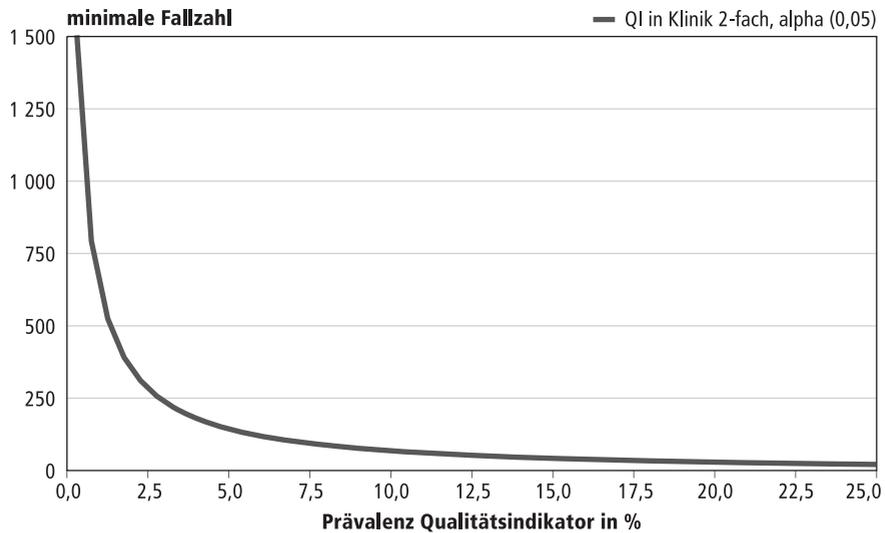
**** bezogen auf diejenigen Kliniken, welche die Leistungen anboten

Krankenhaus-Report 2010

Wido

Abbildung 14–1

Minimale benötigte Fallzahl nach Prävalenzen von Qualitätsindikatoren für Alpha-Niveaus von 0,05 und einer Power von 0,8



Krankenhaus-Report 2010

WlD0

umgekehrt – analog dem Vorgehen bei der Planung von prospektiven Studien – mit Hilfe der Berechnung von minimal benötigten Fallzahlen beantwortet werden.

Abbildung 14–1 zeigt den Zusammenhang zwischen der Prävalenz eines Qualitätsindikators und der minimal benötigten Fallzahl unter den Randbedingungen, dass eine Verdopplung des Qualitätsindikators⁹ gegenüber der Durchschnittsrate des Indikators bei einem Alpha-Fehler von 5 % und einem Beta-Fehler von 20 %¹⁰ sicher erkannt werden sollte. Dabei ist leicht zu erkennen, dass bei Prävalenzen im unteren Prozentbereich mehrere hundert ggf. deutlich über tausend Fälle benötigt werden, um mit ausreichender Sicherheit relevante Qualitätsunterschiede abbilden zu können (vgl. Dimick et al. 2004; Moster et al. 2000). In Tabelle 14–1, Tabelle 14–4 und Tabelle 14–5 wurden die resultierenden minimal benötigten Fallzahlen für die aufgelisteten Qualitätsindikatoren in der jeweils vorletzten Spalte ergänzt. Dabei zeigt sich, dass für die überwiegende Mehrheit der hier aufgelisteten Indikatoren nur wenige Kliniken die benötigten minimalen Fallzahlen pro Jahr aufweisen oder umgekehrt die Power der Indikatoren für die meisten Kliniken unter den gegebenen Fallzahlen nicht ausreichend ist, um die medizinische Ergebnisqualität der Kliniken anhand der genannten Indikatoren mit ausreichender Sicherheit abzubilden.

Dabei sollte beachtet werden, dass die hier durchgeführten Analysen risikoadjustierte Konstellationen nicht berücksichtigen – obgleich für einige Indikatoren

⁹ Da in diesem Fall nur Erhöhungen entdeckt werden sollen, liegt eine einseitige Analyse vor.

¹⁰ Die Wahrscheinlichkeit, dass ein tatsächlich existierender Unterschied dieser Größe als signifikant entdeckt wird (= Power der Studie), soll demnach 1-Beta, also 80 % sein.

risikoadjustierte Ergebnisse vorliegen. Allerdings fehlen für die Berücksichtigung risikoadjustierter Konstellationen im Sinn von multivariaten Poweranalysen – zumindest für die BQS- und HELIOS-Indikatoren – die notwendigen Informationen. Grundsätzlich ist davon auszugehen, dass die Verwendung von multivariaten Verfahren oder auch die Anwendung von Subgruppenanalysen nochmals zu einem Verlust an Power bzw. zu deutlich erhöhten minimalen Fallzahlen führen wird. Insofern sollten die hier durchgeführten Analysen als Mindestanforderungen verstanden werden.

14.3 Qualitätsindizes zur Beurteilung der Ergebnisqualität

Dabei wurde diese Problematik bereits früher erkannt und verschiedene Lösungsmöglichkeiten evaluiert (Heller et al. 2004, Heller 2008). Zuletzt wurden verschiedene potenzielle Komplikationen¹¹ zu Qualitätsindizes zusammengefasst. Da in QSR nicht nur der stationäre Aufenthalt, sondern durch anonymisierte individuelle Verknüpfung auch poststationäre Ereignisse erfasst werden können, stehen für eine Indexbildung sowohl stationäre als auch poststationäre Ereignisse zur Verfügung. Die Entwicklung der Indizes in QSR begann mit der Konstruktion von poststationären Indizes. Als Beispiel sei hier ein Index von potenziellen Komplikationen nach stationärem Aufenthalt wegen Implantation einer Hüftgelenks-Totalendoprothese bei Coxarthrose genannt, der sich aus folgenden Komponenten zusammensetzt:

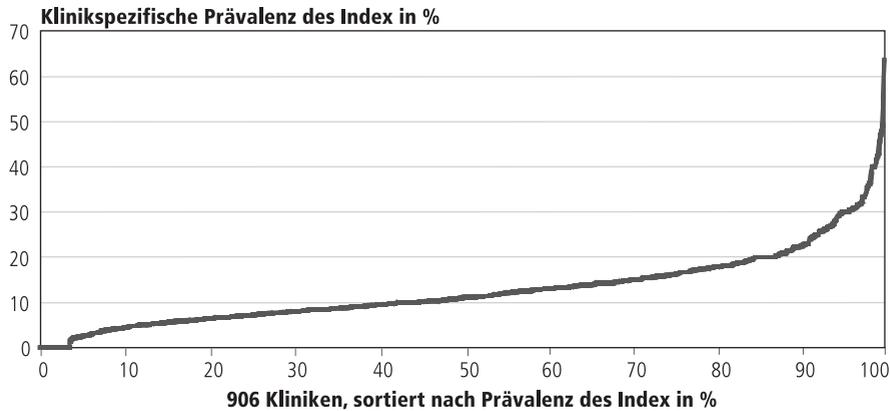
- spezifische Wiederaufnahmen innerhalb von 90 Tagen nach Entlassung. Gemeint sind Wiederaufnahmen mit folgenden Hauptdiagnosen: Komplikationen bei Eingriffen (ICD T81), Komplikationen durch orthopädische Endoprothesen, Implantate oder Transplantate (ICD T84), Thrombose (ICD I80), Lungenembolie (ICD I26), Luxation, Fraktur des Femurs (S72) Verstauchung und Zerrung des Hüftgelenkes und von Bändern der Hüfte (ICD: S73)
- Tod innerhalb von 90 Tagen nach Aufnahme
- Revision der Hüft-TEP innerhalb eines Jahres nach Entlassung.

In einem zweiten Schritt wurden potenzielle Komplikationen während des stationären Aufenthaltes hinzugefügt. Für das Beispiel der elektiven Hüft-TEP sind dies:

- Lungenembolie, Thrombotische Ereignisse, Beatmung über 24 Stunden, Wundinfektionen, sonstige postoperative Komplikationen, Komplikationen durch orthopädische Endoprothesen, Implantate oder Transplantate, Luxation, Verstauchung und Zerrung des Hüftgelenkes und von Bändern der Hüfte, Revision, Wechsel oder Entfernung der Hüft-TEP während stationärem Aufenthalt, Pneumonie.

¹¹ Indikatoren der Ergebnisqualität können auch dadurch beschrieben werden, dass sie entweder auf potenzielle Komplikationen der Therapie hinweisen oder allgemeiner das angestrebte Therapieziel nicht erreicht wurde. Dabei stellt die erste Formulierung (Komplikation) strenggenommen nur eine eingeschränkte und negative Formulierung der letzten Aussage (Nichterreichen des Therapieziels) dar (Heller 2006). Komplikationen werden in gängigen Dokumentationssystemen allerdings oft besser erfasst.

Abbildung 14–2

Index potenzieller Komplikationen (Qualitätsindex) nach Implantation elektiver Hüft-TEPs, AOK-Patienten 2007*

*vollstationäre AOK-Patienten mit Entlassungsdatum in 2007; nur Patienten in Kliniken mit mindestens zehn Tracer-Eingriffen

Krankenhaus-Report 2010

WIdO

Abbildung 14–2 zeigt klinikspezifische Raten des kombinierten (stationären und poststationären) Qualitätsindex für den Tracer Implantation einer Hüftgelenks-Totalendoprothese bei Coxarthrose auf Basis von AOK-Patienten mit Entlassungsdatum im Jahr 2007. Um nur Kliniken zu beurteilen, die eine entsprechende Therapie auch mit einer gewissen Regelmäßigkeit durchführen, wurden nur Fälle aus Kliniken mit mindestens zehn Tracerfällen im Jahr 2007 berücksichtigt.¹² Hier zeigt sich, dass nur 3,4% der Kliniken keinerlei potenzielle Komplikationen aufweisen, etwa 52% weisen mehr als 10% potenzielle Komplikationen auf und knapp 10% weisen mehr als 20% potenzielle Komplikationen auf. Im Durchschnitt treten bei 9,9% der Fälle eine der genannten potenziellen Komplikationen auf. Damit ergibt sich – unter den oben definierten Bedingungen – für diesen Index eine minimal benötigte Fallzahl von 70 Fällen pro Klinik. Demnach kann erwartet werden, dass mit diesem Index ein relevanter Anteil von Kliniken in Deutschland in einem überschaubaren Zeitintervall mit akzeptabler Sicherheit beurteilt werden kann.¹³ Mit Blick auf die oben durchgeführten Analysen sei nochmals darauf hingewiesen, dass

12 In diesen Kliniken werden weniger als 2% der elektiven Hüft-TEPs implantiert. Sie machen allerdings etwa 19% aller Kliniken aus. Die Gesamtprävalenz des Index ist im Übrigen vergleichbar, ob diese Fälle ausgeschlossen werden oder nicht.

13 Wenn also auf Basis der AOK-Zahlen pro Klinik drei Jahre analysiert werden, würden etwa drei Viertel aller Kliniken des analysierten Kollektivs mit mehr als 90% zugehörigen Tracerfällen aufgrund dieses Index mit ausreichender Sicherheit (Power) beurteilt werden können. Es würden etwa gleich große Anteile erreicht, wenn alle Krankenhaus-Tracerfälle aller Kassen eines Jahres ausgewertet werden könnten.

dies mit den bisherigen Indikatoren der Ergebnisqualität oft nicht der Fall ist (vgl. Tabelle 14–1, Tabelle 14–4, Tabelle 14–5).¹⁴

Dies zeigt sich auch, wenn analog zu den oben durchgeführten Analysen die Prognosefähigkeit dieses Qualitätsindex über verschiedene Jahre betrachtet wird. Für den soeben vorgestellten Qualitätsindex der elektiven Hüft-TEP ergeben sich bezüglich einer Prognose für das kommende Jahr erklärte Varianzen von 6,3 % bis 12,8 % (Tabelle 14–3).¹⁵ Dieses Ergebnis verbessert sich nochmals drastisch, wenn nur Kliniken mit mindestens 100 Tracer-Eingriffen betrachtet werden: Dann werden erklärte Varianzen über 50 % erreicht.

Insgesamt zeigen sich also durch Bildung der Qualitätsindizes drastisch bessere Ergebnisse als zuvor für Prognosen, die ausschließlich aufgrund der Mortalitätsraten getroffen wurden. Dies sollte als ein deutlicher Hinweis darauf verstanden werden, dass zurückhaltend formuliert neben zufälligen Ergebnissen zumindest etwas Systematisches, z. B. so etwas wie Ergebnisqualität abgebildet wurde.

Daher wurden in QSR Qualitätsindizes für die Tracer Kolon-/Rektum-Operation bei kolorektalem Karzinom, Appendektomie, Implantation einer Hüftgelenks-Totalendoprothese bei Coxarthrose, Implantation einer Hüftgelenks-Endoprothese bei Hüftfraktur und Implantation einer Kniegelenks-Totalendoprothese entwickelt und stehen interessierten Kliniken als QSR-Klinikbericht, vertrieben über 3M, zur Verfügung.¹⁶

Die Vorteile der hier vorgestellten Indexbildung bestehen darin, dass die Ergebnisqualität breiter erfasst wird. Es wird nicht nur das Überleben, sondern es werden auch weitere relevante medizinische Komplikationen wie Thrombosen, Pneumonien oder nötige Revisionsoperationen berücksichtigt. Dies ist auch deswegen bedeutsam, weil davon ausgegangen werden kann, dass für verschiedene elektive Tracer bei einem entsprechend gesunden und numerisch begrenzten Patientenkollektiv trotz erheblicher Mängel in der medizinischen Versorgung für zahlreiche Kliniken die Mortalität gar nicht erhöht ist. Das Problem einer alleinigen (oder überwiegenden) Orientierung der Qualitätsmessung an Mortalität für zumeist elektive Tracer mit geringen Mortalitätsraten liegt daher mutmaßlich zum größeren Teil darin, dass zahlreiche Kliniken oder genauer Abteilungen mit relevanten medizinischen Qualitätsdefiziten fälschlicherweise die Rückmeldung bekommen, ihre Ergebnisqualität sei gut oder sogar sehr gut.

Andererseits sind auch Nachteile einer Indexbildung möglich: Es werden mitunter Dimensionen mit recht unterschiedlichen Tragweiten abgebildet, z. B. Mortalität und Wundheilungsstörung. Dabei kann mit Recht problematisiert werden, ob das Versterben nach einer elektiven Operation ebenso gewertet werden sollte wie eine ggf. nur marginale oberflächliche Wundinfektion.

14 Insofern ist auch zu erwarten, dass die Korrelationen zwischen unterschiedlichen Qualitätsmesssystemen, etwa zwischen den BQS und den HELIOS-Indikatoren oder zwischen den BQS- und den QSR-Indikatoren, eher marginal sein werden.

15 Die schlechteren Werte für die Prognose der Qualitätsindizes von 2004 auf 2005 liegt in der noch nicht vorhandenen Kodierung der Seitigkeit für TEP-Implantationen/-Revision für 2004 begründet.

16 Siehe Fußnote 3

14.4 Weiterentwicklung der Qualitätssicherung mit Routinedaten

14.4.1 Weiterentwicklung der Qualitätsindizes

Aufgrund der soeben beschriebenen Problematik der Multidimensionalität der Qualitätsindizes und deren unklarer Wertigkeit untereinander wird aktuell untersucht, wie die Wertigkeit dieser Indizes durch unterschiedliche Nutzergruppen beurteilt wird. Patienten werden dabei ebenso einbezogen wie Einweiser oder Fachgruppen aus dem jeweiligen medizinischen spezifischen Fachgebiet. So ist durchaus denkbar, dass in künftigen Versionen einzelne potenzielle Komplikationen (Dimensionen) in unterschiedlichen Gewichtungen in die Indexkonstruktion einfließen, neue Dimensionen hinzukommen oder andere wegfallen.

Darüber hinaus soll eine Risikoadjustierung für die Indizes entwickelt werden. Dies ist deswegen nicht ganz trivial, weil für jede einzelne Dimension ein eigenes Risikoadjustierungsverfahren entwickelt und anschließend eine mathematisch statistisch sinnvolle Verknüpfung durchgeführt werden muss.

14.4.2 Statistische Analyseverfahren

In der Literatur finden sich Hinweise, dass gerade bei der statistischen Schätzung von eher seltenen Ereignissen die Verwendung von Empirical-Bayes-Analysetechniken oder ähnlichen Verfahren einer statistischen Schätzung über konventionelle regressionsanalytische Verfahren überlegen ist (Greenland 2000; Ash et al. 2003; Martens et al. 2008). Mittlerweile liegen auch erste Analysen aus Deutschland zur Thematik vor.¹⁷ Aktuell wird geprüft, ob die Anwendung von Empirical-Bayes- oder vergleichbaren Analysetechniken auch im QSR-Setting zu einer besseren Einschätzung klinikspezifischer Ergebnisqualität führt.

Eine weitere Analysemöglichkeit besteht darin, die longitudinale Information in den Abrechnungsdaten besser zu nutzen, z. B. indem intraindividuelle Vergleiche vor und nach einer Therapie in die Risikoadjustierung einfließen (Tannen et al. 2009). Der Vorteil dieser Herangehensweise besteht darin, dass so jedes Individuum quasi als seine eigene Kontrollgruppe fungiert. Beispielsweise könnten – insbesondere für elektive Leistungen – die indikationsspezifischen Aufwendungen, z. B. für Schmerzmittel, Hilfsmittel, Pflegebedürftigkeit etc., vor Implantation einer elektiven Endoprothese mit den Aufwendungen nach Implantation der Endoprothese verglichen werden. In Kombination mit traditionellen Risikoadjustierungsverfahren bzw. Verfahren zur Ermittlung von Ergebnisqualität wäre dadurch eine deutlich umfassendere Beurteilung der Ergebnisqualität, aber auch eine deutlich verbesserte risikoadjustierte Beurteilung zu erwarten. Durch einen solchen sektorübergreifenden Ansatz wären wesentliche Teile der Indikationsqualität berücksichtigt (Schröder et al. 2008). Durch die Anwendung solcher Verfahren scheinen uns wertvolle

¹⁷ URL: http://www.herzinfarktregister.de/fakten/2008/Symposium/symp08_wegscheider_netz.pdf (Zugriff am 05. August 09).

Analysen im QSR-Kontext z. B. in Pay-for-Performance-Ansätzen umsetzbar (Mahlzahn und Heller 2009).

14.4.3 Weitere Tracer und sektorenübergreifende Qualitätssicherung

Die derzeit entwickelten Tracer bilden nur einen kleinen Teil der stationären Versorgung ab. Daher werden derzeit neue Tracer entwickelt, um einen größeren Anteil der medizinischen Versorgung abbilden zu können. Gleichzeitig wird darauf geachtet, sektorenübergreifende Aspekte der Qualitätssicherung einzubeziehen. Zwar ist dies auch schon heute in QSR angelegt¹⁸, künftig sollen aber auch Leistungsdaten aus anderen Sektoren in die Qualitätsbetrachtung einfließen. So sollen für den aktuell in der Entwicklung befindlichen Tracer „Linksherzkatheter und Linksherzinterventionen“ explizit auch ambulante Linksherzkatheter wie auch Verordnungsdaten analysiert und in die Qualitätsbetrachtung einbezogen werden. Daneben stehen Tracer zur Geburtshilfe oder zur Cholezystektomie in näherer Zukunft zur Entwicklung an.

14.5 Fazit

Im vorliegenden Beitrag wurde die Entwicklung wie auch der aktuelle Stand des Projektes Qualitätssicherung mit Routinedaten (QSR) dargestellt. Darüber hinaus wurde illustriert, warum bisherige Verfahren der Ermittlung von Ergebnisqualität aus einfachen statistischen Gründen oft nicht zielführend sein konnten. Um dieses Problem zu lösen, wurde die Verwendung von Qualitätsindizes vorgeschlagen und deren Implementierung in QSR vorgestellt. Darüber hinaus wurden weitere geplante Entwicklungsschritte in QSR in Bezug auf die statistische Weiterentwicklung, eine Ausweitung der analysierbaren Versorgungsleistungen durch Entwicklung weiterer Tracer mit sektorenübergreifendem Fokus wie auch Aufbereitung von QSR-Inhalten für allgemein zugängliche Qualitätsdarlegungen skizziert.

Ziel der Bemühungen von QSR ist es, einen Beitrag zu einer aufwandsarmen, validen und allgemein akzeptierten sektorenübergreifenden Ermittlung medizinischer Ergebnisqualität zu leisten. Dies geschieht unter der Vorstellung, dass eine transparente Qualitätsdarlegung allen beteiligten Akteuren eine sinnvolle Auseinandersetzung mit der Qualität der medizinischen Versorgung ermöglicht und so dazu beitragen kann, die Medizinische Versorgungsqualität zu verbessern.

¹⁸ So wurden in QSR von Anfang an Ereignisse analysiert bzw. als Qualitätsendpunkte verwendet, die nach dem aktuellen akutstationären Aufenthalt liegen.

14.6 Literatur

- Ash A, Schwartz M, Perköz EA. Comparing Outcomes across Providers. In: Iezzoni L (Hrsg). Risk Adjustment for Measuring Health Care Outcomes. 3rd edition. Chicago: Health Administration Press 2003; 297–333.
- Bland JM, Altman DG. Regression towards the mean. *British Medical Journal* 2004a; 308: 1499.
- Bland JM, Altman DG. Some examples of regression towards the mean. *British Medical Journal* 2004b; 309: 780.
- Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size; in *Journal of the American Medical Association* 2004; 292: 847–51.
- Greenland S. Principles of multilevel modelling. *International Journal of Epidemiology* 2000; 29: 158–67.
- Heller G, Swart E, Mansky T. Qualitätsanalysen mit Routinedaten. Ansatz und erste Analysen aus dem Gemeinschaftsprojekt “Qualitätssicherung mit Routinedaten” (QSR). In: Klauber J, Robra BP, Schellschmidt H (Hrsg). *Krankenhaus-Report 2003*. Stuttgart: Schattauer 2004; 271–88.
- Heller G. Qualitätssicherung der stationären Versorgung – Probleme und Perspektiven. Antrittsvorlesung anlässlich der Habilitation am Fachbereich Medizin der Philipps-Universität Marburg, 26.10.2006.
- Heller G. Zur Messung und Darstellung von medizinischer Ergebnisqualität mit administrativen Routinedaten in Deutschland. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz* 2008; 10: 1173–82.
- Heller G, Günster C, Swart E. Perspektiven der Qualitätssicherung mit Routinedaten (QSR). In: Klauber J, Robra BP, Schellschmidt H (Hrsg). *Krankenhaus-Report 2007*; Stuttgart: Schattauer 2008; 171–84.
- Lütticke J, Schellschmidt H. Qualitätsberichte nach § 137 SGB V – Bewertung und Vorschläge zur Erweiterung. In: Klauber J, Robra BP, Schellschmidt H (Hrsg): *Krankenhaus-Report 2004*; Stuttgart: Schattauer 2005; 197–211.
- Malzahn J, Heller G. Pay for Performance in der stationären Versorgung – Probleme und Lösungen. Diskussionspapier zum Nationalen DRG- Forum. Berlin, April 2009.
- Martens EP, Pestman WR, de Boer A, Belsiter SV, Klungel OH. Systematic differences in treatment effect estimates between propensity score methods and logistic regression. *International Journal of Epidemiology* 2008; 37: 1142–4.
- Moster D, Markestad T, Lie RT. Assessing quality of obstetric care for low-risk deliveries; methodological problems in the use of population based mortality data. *Acta Obstetrica et Gynecologica Scandinavica* 2000; 79: 478–84.
- Schräder P, Boy O, Schleiz W, Dienst R, Reinert C, Sänger V, Schauwecker HH, Siebert W Scharf HP. Indikationsstellung in der primären Hüft- und Kniegelenksendoprothetik. *Der Orthopäde* 2008; 37:1016–26.
- Staiger DO, Dimick JB, Baser O, Fan Z, Birkmeyer JD. Empirically derived composite measures of surgical performance. *Med Care* 2009; 47: 226–33.
- Tannen RL, Weiner MG, Xie D. Use of primary care electronic medical record database in drug efficacy research on cardiovascular outcomes: comparison of database and randomised controlled trial findings. *British Medical Journal* 2009; 338: b81.