

14 Qualitätsanalysen mit Routinedaten

Ansatz und erste Analysen aus dem Gemeinschaftsprojekt „Qualitätssicherung mit Routinedaten“ (QSR)

Günther Heller, Enno Swart und Thomas Mansky

Abstract

„Qualitätssicherung der stationären Versorgung mit Routinedaten (QSR)“¹ ist ein gemeinsames Forschungsprojekt des AOK-Bundesverbandes, der HELIOS Kliniken GmbH, des Forschungs- und Entwicklungsinstituts für das Sozial- und Gesundheitswesen Sachsen-Anhalt (FEISA) und des Wissenschaftlichen Instituts der AOK (WiDO). QSR ist ausgerichtet auf die Unterstützung der Qualitätssicherung in der stationären Versorgung auf Basis von GKV-Routinedaten. Dabei steht die Ermittlung von Ergebnisqualität im Mittelpunkt der Projektbemühungen. Bislang gibt es in Deutschland trotz umfangreicher gesetzlicher Regelungen keine Qualitätssicherungsverfahren, die es Kostenträgern oder Versicherten erlauben, datenbasiert gute Leistungen bzw. Leistungserbringer zu identifizieren. Andererseits bieten die bisher in Deutschland etablierten Qualitätssicherungsverfahren auch den Leistungserbringern keine oder eine nur sehr eingeschränkte Bewertung der Ergebnisqualität. QSR ergänzt sowohl die in Deutschland etablierten Verfahren, die primär auf Struktur- und Prozessqualität abzielen (wie EFQM oder KTQ), als auch die gesetzlich verankerte externe Qualitätssicherung nach § 137 SGB V. In diesem Beitrag wird der Ansatz von QSR kurz dargelegt und mit Rekurs auf die internationale Literatur begründet (Abschnitt 14.1). Nach einer kurzen Darstellung von Datengrundlagen und Methodik (14.2) werden erste Ergebnisse aus dem Projekt am Beispiel des Qualitätsindikators Krankenhaussterblichkeit bei Kolonkarzinom-Operation dargestellt (14.3). Die im Benchmarking-Prozess notwendigerweise auftretenden Probleme werden dabei exemplarisch aufgezeigt, mögliche Lösungswege werden diskutiert (14.4).

“Quality assurance in the hospital sector using administrative data (QSR)” is a joint project of the Federal Association of the AOK, the HELIOS Kliniken

1 Die QSR-Steuergruppe besteht aus: Michael Held, Wulf-Dietrich Leber (AOK-Bundesverband), Stefan Felder, Bernt-Peter Robra, Enno Swart (Forschungs- und Entwicklungsinstitut für das Sozial- u. Gesundheitswesen Sachsen-Anhalt), Wolfgang Krahwinkel, Thomas Mansky, Oda Rink-Brüne, Dieter Waldmann, Josef Zacher (HELIOS Kliniken GmbH), Christian Günster, Günther Heller, Jürgen Klauber, Henner Schellschmidt (Wissenschaftliches Institut der AOK).

Group, the Research Institute for Social Services and Health Care Saxony-Anhalt and the AOK Research Institute (WIdO). This article provides a brief introduction of the QSR approach which focusses on outcome quality research, thereby referring to the relevant international literature as well as existing quality assurance projects in Germany (chapter 14.1). Following a short description of the data basis and methodology (14.2), preliminary results of the QSR project are presented for the tracer diagnosis colon cancer surgery (14.3). The problems necessarily evolving in the benchmarking process are exemplarily described and possible solutions are discussed (14.4).

14.1 Einleitung und Problemstellung

14.1.1 Gründe für Qualitätssicherung mit Routinedaten in Deutschland

Ein Projekt der externen Qualitätssicherung im Krankenhausbereich auf der Basis von Routinedaten lässt sich aus unterschiedlichen Perspektiven begründen. Zwei Argumentationslinien erscheinen dabei von zentraler Bedeutung: der Bedarf an praktikablen und validen Alternativen zur Dokumentation nach § 137 SGB V und ein zunehmender Bedarf an Qualitätstransparenz.

Alternativen zur Dokumentation nach § 137 SGB V

Obgleich eine externe Qualitätssicherung im Krankenhaus über den § 137 SGB V gesetzlich vorgeschrieben und mittlerweile implementiert ist (BQS 2002a/BQS 2002b), wird von verschiedenen Seiten eine deutliche Unzufriedenheit mit diesem Verfahren geäußert. Dabei wird der unverhältnismäßig große Aufwand der Dokumentation bemängelt. Noch wichtiger ist der Vorwurf einer mangelnden Validität verschiedener zentraler Items wie auch die mangelhafte Vollständigkeit und damit einhergehend eine mutmaßlich selektive Dokumentation von Krankheitsfällen.

Ein auf Routinedaten basierendes Verfahren wie QSR kann die beteiligten Kliniken in ihrem Bemühen um sinnvolle interne Qualitätssicherungsprojekte wesentlich unterstützen, indem es Unzulänglichkeiten der existierenden externen Qualitätssicherung überwindet und den Krankenhäusern valide externe Benchmark-Informationen bereitstellt, die wiederum in interne Qualitätssicherungsprojekte einfließen können. Insofern versteht sich QSR als ein Serviceangebot der Kostenträger für die Leistungserbringer.²

Zunehmender Bedarf an Transparenz

Die Zukunft des deutschen Gesundheitswesens ist durch eine Zunahme des Wettbewerbs um die Verbraucherentscheidung zwischen Leistungserbringern, aber auch zwischen Kostenträgern charakterisiert. Vertragswettbewerb als Perspektive heißt

² Für eine weitergehende Bewertung und Einordnung von QSR aus Sicht der Krankenhäuser vgl. den Medizinischen Jahresbericht 2002 der HELIOS-Kliniken GmbH.

dabei nicht nur mehr Konkurrenz, sondern öffnet auch Wege zu neuen Kooperationen zwischen Leistungserbringern und Kassen.

Für die Qualitätsfrage hat ein verstärkter Wettbewerb zwei wesentliche Auswirkungen:

- Zum einen kann und wird Qualität sich zu einem Vertragsinhalt entwickeln. Entsprechend wird bei Leistungserbringern und Kostenträgern gleichermaßen der Bedarf steigen, Qualität transparent und messbar zu machen.
- Zum anderen kann ein zunehmender Wettbewerb im Gesundheitswesen einen Verlust an Qualität medizinischer Leistungen zur Folge haben.

Aus beiden Perspektiven heraus erwächst ein Bedarf an praktikablen, validen, objektiven, vollständigen und zugleich flächendeckenden Verfahren zur Messung der Behandlungsqualität in der stationären Versorgung.

Wer seine eigene Versorgungsqualität verbessern will, braucht Informationen über mittel- und langfristige Versorgungsergebnisse. Wer Qualität zum Vertragsgegenstand machen will, muss diese valide bewerten können. Wer Qualitätsrisiken vermeiden will, muss in der Lage sein, diese zu erkennen und vergleichend zu bewerten.

Transparenz, Sicherung, Bewertung und Vertragsfähigkeit der Behandlungsqualität stellen daher zentrale Herausforderungen im deutschen Gesundheitswesen dar. Verfahren, die diesen Anforderungen genügen, sind von hoher Relevanz.

Die bislang in Deutschland implementierten Qualitätssicherungsverfahren der stationären Versorgung können diesem Anspruch nicht genügen³. Mit QSR wird eine Qualitätsberichterstattung mit Routinedaten („administrative data“) erprobt. Dieser für Deutschland neue Ansatz orientiert sich an internationalen Konzepten. Im Vergleich zu den aktuell implementierten Verfahren zeichnet er sich insbesondere durch folgende Vorteile aus:

- **Geringer Erhebungsaufwand/geringe Kosten:** Der mit dem Verfahren verbundene Erhebungsaufwand ist relativ gering, weil gemäß dem Verfahren nach § 301 SGB V auf ohnehin zu Abrechnungszwecken erhobene und geprüfte Daten zurückgegriffen wird. Die Einführung diagnose-orientierter Fallpauschalen (DRGs) bringt eine umfangreiche und detaillierte Kodierung der behandelten Erkrankungen und Komorbiditäten sowie vorliegender Komplikationen mit sich. Die Angaben werden im Zuge der Abrechnung an die Kostenträger übermittelt und geprüft. Sie können für Qualitätssicherungszwecke verwendet werden.
- **Vollständigkeit der Daten:** Da alle Krankenhäuser alle ihre Fälle abrechnen, liegen sowohl für alle Fälle als auch für alle Institutionen Informationen vor. Eine solch vollständige Erfassung wird aktuell durch keines der bestehenden Verfahren in Deutschland gewährleistet. Aktuelle Qualitätssicherungsverfahren sehen sich mit dem Vorwurf konfrontiert, dass problematische Fälle mit einer geringeren Wahrscheinlichkeit dokumentiert werden als unproblematische Fäl-

³ Neben den im Wesentlichen auf Struktur- und Prozessqualität abhebenden Verfahren von KTQ und EFQM sind vor allen die gesetzlich geregelten Verfahren zur externen Qualitätssicherung zu nennen (BQS 2002a/ BQS 2002b).

le. Bei Verwendung von Abrechnungsdaten ist ein Bias durch unvollständige und möglicherweise selektive Meldung von Daten ausgeschlossen.

- **Individueller Bezug und Langzeitbeobachtung:** Aktuell angewandte Qualitätssicherungsverfahren erlauben meist nur einen Aufenthaltsbezug, d. h. sie können Angaben nur für die Dauer des dokumentierten Krankenhausaufenthalts machen. Mit Routinedaten können dagegen Krankheitsverläufe im Längsschnitt analysiert werden. So sind klinikübergreifende Ereignisse, wie z. B. Wiederaufnahmen in anderen Kliniken, analysierbar. Grundsätzlich möglich ist auch die Messung von Überlebensraten in definierten Zeitintervallen, z. B. 1- oder 5-Jahres-Überlebensraten nach Karzinom-Operation.
- **Ergebnisqualität im Fokus:** Durch die Möglichkeit von Längsschnittanalysen können Indikatoren der Ergebnisqualität medizinischer Behandlung in erheblichem größerem Maße ermittelt werden, als dies bisher in Deutschland der Fall war.
- **Validität der Daten:** Weil komplikationsreiche Fälle mit schweren Verläufen zu höheren Entgelten führen, gleichzeitig aber aufgetretene Komplikationen eine mögliche Abwertung in der Qualitätsbewertung zur Folge haben können, ist eine gezielte Manipulation der Abrechnungsdaten für die Qualitätssicherung unwahrscheinlich. Die Validität erhöht sich dadurch, dass die Abrechnungsdaten routinemäßig einer umfangreichen Plausibilitätsprüfung von Seiten der Kostenträger unterzogen werden. Positiv wirkt sich in diesem Sinne auch aus, dass sich zentrale Indikatoren der Ergebnisqualität (z. B. Wiederaufnahmeraten in anderen Kliniken, Revisionsraten, Überlebensraten, usw.) nicht aus der Dokumentation des Krankenhauses ergeben, welches die zu beurteilende medizinische Leistung erbracht hat. Diese Argumente lassen potenziell im Vergleich zu den bisher verfügbaren Qualitätssicherungsdaten eine bessere Datenqualität mit Routinedaten erwarten.

14.1.2 Stand der Forschung

Ein Blick in die internationale Literatur zeigt, dass eine Qualitätsberichterstattung mit Routinedaten in zahlreichen Ländern gut etabliert ist und intensiv für Zwecke der Qualitätssicherung genutzt wird.

Eine Reihe von Autoren vor allem aus den USA und Kanada untersuchen die Nutzbarkeit von Routinedaten für epidemiologische Untersuchungen und zur Qualitätsbeurteilung der Versorgung. Im Zusammenhang mit der Analyse kleinräumiger Versorgungsheterogenitäten wurde die Nutzbarkeit und Aussagefähigkeit von Routinedaten herausgestellt. Die positive Einschätzung dieser Datenquelle resultiert aus ihrem Bevölkerungsbezug, der routinemäßigen und kostengünstigen Verfügbarkeit, ihrer Vollständigkeit, ihrer Unverzerrtheit und der Eignung für Follow-up-Studien und Monitoring (Wennberg et al. 1987, Wennberg et al. 1988). Zwar enthalten die Routinedaten keine tiefen klinischen Informationen, doch können indikationsspezifische Outcome-Indikatoren bzgl. Mortalität, Rezidiven oder Wiederaufnahmen abgeleitet werden (Mitchell et al. 1994). Regelmäßige Auswertungen von Routinedaten auf kleinräumiger Basis, verbunden mit traditionellen Ansätzen zur Bewertung der Qualität der Leistungserbringung, können Einsichten in bevölkerungsbezogene Effekte der medizinischen Versorgung und Ansatzpunkte erbrin-

gen und die Gesundheit der Bevölkerung durch Umsteuerung von Gesundheitsleistungen verbessern (Black und Roos 1998, Roos et al. 1995).

Kritik an der Nutzung von Routinedaten zur Qualitätsmessung und -bewertung erwächst oftmals aus der Begrenzung auf den stationären Aufenthalt oder ein kurzes Zeitfenster nach der Entlassung. Die Verlängerung des Zeitfensters auf Monate oder Jahre könnte die daraus resultierenden Verzerrungen verhindern (Park et al. 1990). Auch ist die Erfassung von Begleiterkrankungen und Komplikationen schwierig, doch liefern Routinedatenquellen kostengünstig schnell verfügbare Informationen (Geraci 2000). Wray et al. (1995) haben auf der Grundlage von neun Kriterien ein methodisches Konzept für die Nutzung administrativer Daten zur Bewertung der Qualität medizinischer Versorgung erstellt. Wenn diese Kriterien durch spezifische Indikator-Ergebnis-Paare erfüllt sind, lassen sich administrative Daten für eine Qualitätsbewertung heranziehen.

Das am meisten verbreitete Verfahren zur Qualitätsmessung von Leistungserbringern ist die vergleichende Gegenüberstellung von Qualitätsindikatoren in Form von Rankings, den sog. „league tables“ (Jacobson et al. 2003). Diese werden in den USA beispielsweise durch die Health Care Financing Agency (HCFA) eingesetzt. Qualitätsindikatoren können dabei nach Ergebnisindikatoren und Prozessindikatoren unterschieden werden (Pouvoirville/Minvielle 2001, Pouvoirville 2003).

Die Sterblichkeit ist der am weitesten verbreitete Qualitätsindikator. Eine rohe Analyse der Krankenhaussterblichkeit ist allerdings von begrenzter Validität, da der gemessene klinikbezogene Wert wesentlich von der Schweregradverteilung der im Krankenhaus versorgten Patienten und anderen Merkmalen des Patientenmix abhängt. Daher wird vor Krankenhausvergleichen mit Resultatparametern ohne vorherige Adjustierung für wichtige Merkmale der Patienten und des Falles gewarnt (Leyland und Boddy 1998). Eine Risikoadjustierung wird üblicherweise über die Einweisungs- bzw. Nebendiagnosen („comorbidities“) sowie patientenbezogene Faktoren wie Alter, Geschlecht oder Ethnizität vorgenommen (Iezzoni 1997, Pouvoirville/Minvielle 2001, Landon et al. 1996, Mukamel et al. 2001). Je nach Krankheit und Personengruppe können unterschiedliche Faktoren relevant sein (Elixhauser et al. 1998). Jüngere Arbeiten zeigen eindrucksvoll, dass mittels Sterblichkeitsanalysen aus Routinedaten wesentliche Informationen über die Qualität der Krankenhausbehandlung abgeleitet werden können (Birkmeyer 2002).

Die Beschränkung auf Todesfälle während des Krankenhausaufenthalts oder begrenzt auf einen engen poststationären Zeitraum (gebräuchlich sind 30 Tage nach der Entlassung) beinhaltet die Gefahr einer unvollständigen Erfassung dieser Ereignisse (Garnick et al 1995). Strategien der Krankenhäuser zugunsten einer frühzeitigen Entlassung oder Verlegung können die Sterblichkeitsrate im beobachteten Zeitfenster beeinflussen (Pouvoirville und Minvielle 2001). Es ist deshalb sinnvoll, den Betrachtungszeitraum für diesen Indikator auf über ein Jahr auszuweiten. Dies erfordert jedoch die krankenhausesübergreifende Beobachtung und Zuordnung von Patienten, da sie in der Zwischenzeit aus dem Krankenhaus entlassen, in ein anderes verlegt oder außerhalb eines Krankenhauses verstorben sein könnten (Park et al. 1990). Schließlich sollte berücksichtigt werden, dass Überlebensraten u. a. auch von der medizinischen Versorgung nach der Entlassung aus dem Krankenhaus beeinflusst werden können.

Die Langzeitanalyse insbesondere der Sterblichkeit stellt allerdings auch neue Anforderungen an die Dokumentation. Bei der in dieser Arbeit analysierten Sterblichkeit bei Darmoperationen aufgrund eines Kolonkarzinoms ist beispielsweise die hier betrachtete kurzfristige Sterblichkeit vom Risikoprofil zum Zeitpunkt der Operation, welches sich über Nebendiagnosen relativ gut erfassen lässt und von der Art der Operation, die über OPS-Schlüssel hinreichend differenziert dokumentiert wird, abhängig. Dieser Wert ist daher bereits jetzt gut beurteilbar. Die Langzeitsterblichkeit hängt aber hauptsächlich von der Tumorausbreitung bzw. dem Tumorstadium zum Zeitpunkt der Erstbehandlung bzw. Erstoperation ab. Das Tumorstadium ist mit der amtlichen ICD-10 nur begrenzt erfassbar. Dies spricht auch in diesem Bereich nicht gegen die Qualitätsanalyse mit Routinedaten, sondern vielmehr für eine Anpassung vor allem der Tumorverschlüsselung (z. B. Erfassung des vergleichsweise gut standardisierten TNM-Stadiums in der Standarddokumentation). Die Möglichkeiten, die sich bei einer vergleichsweise einfachen Verbesserung der Erfassung ergeben, sind beeindruckend: Langzeitanalysen, die sonst nur in äußerst aufwendigen Studien für eine begrenzte Zahl von Fällen möglich sind, können auf der Basis von Routinedaten umfassend und mit wenig Aufwand aus den vorhandenen Daten erfolgen.

Eine hohe Anzahl von Wiederaufnahmen kann bei bestimmten Krankheitsbildern Hinweise auf Probleme bei der Entlassungspraxis (zu frühe Entlassung) geben, ggf. aber auch auf andere Faktoren zurückführbar sein. Auch dieser Qualitätsindikator bedarf einer Risikoanpassung nach patientenbezogenen Faktoren und dem Case-Mix eines Hauses (Birkmeyer et al. 2001, Iezzoni et al. 1999, Pouvourville und Minvielle 2001). Die valide Bewertung von Wiederaufnahmen als Qualitätsindikator setzt voraus, dass Wiederaufnahmen insgesamt und nicht allein solche in das entlassende Krankenhaus erfasst werden. Als methodisches Problem erweist sich, dass die hier zur Verfügung stehenden Routinedaten nicht zwischen geplanten und ungeplanten Wiederaufnahmen unterscheiden können. Allerdings ist es möglich, die Anlässe der Wiederaufnahmen anhand der kodierten Diagnosen (oder auch Prozeduren) zu analysieren. Somit können typische stationär therapierbare Komplikationen einer vorhergehenden Krankenhausbehandlung nachverfolgt werden.

Weitere Qualitätsindikatoren sind für viele Diagnosen z. B. die Art und Häufigkeit von Komplikationen innerhalb des Krankenhausaufenthalts, wie etwa Infektionen oder erneute ungeplante Eingriffe. Sie treten häufiger als Todesfälle auf, zusätzlich kann die Qualität der Versorgung mitunter spezifischer und anschaulicher dargestellt werden. Allerdings ist eine zuverlässige und vollständige Erfassung dieser Parameter Grundvoraussetzung für eine valide Qualitätsbeurteilung. Auf der Internetseite der Zentralstelle der Deutschen Ärzteschaft zur Qualitätssicherung in der Medizin (ÄZQ) ist eine Übersicht über international verwendete Indikatorensätze mit den entsprechenden Links aufgeführt⁴.

Zusammenfassend scheint es nach internationalen Erfahrungen aussichtsreich zu sein, interne Qualitätssicherungsprozesse mit krankenhausesübergreifenden Prozess- und Resultatvergleichen (Benchmarks) auf der Basis von Routinedaten zu unterstützen. Die vorhandenen administrativen Routinedaten eröffnen bzgl. der

4 www.aezq.de

Darstellung und Bewertung der Versorgungsqualität mehr und weiterreichende Möglichkeiten als klinische Daten, die ggf. mehr Informationen enthalten, die aber auf den Zeitraum bis zum Tag der Entlassung begrenzt sind. Ein Instrument, das einen validen Vergleich der Leistungsergebnisse verschiedener Leistungserbringer erlaubt, gibt diesen Anreize in Qualität zu investieren. „Zusammenfassend legen die veröffentlichten Zahlen nahe, dass Vergleichsdatenbanken über Krankenhausleistungen eine sehr nützliche Wissensbasis zur Leistungsverbesserung darstellen, wenn sie strengen methodischen Prinzipien folgen.“ (Pouvourville 2003)

14.2 Material und Methoden

14.2.1 AOK-Abrechnungsdatenbank

Die Datengrundlage von QSR sind leistungsfall- und versichertenbezogene Informationen der AOK gemäss dem Datenaustauschverfahren nach § 301 SGB V (AOK-Abrechnungsdaten). Zur Validierung der GKV-Daten wurden zudem Krankenhausdaten aus dem Krankenhausinformationssystem der HELIOS Kliniken GmbH eingesetzt. Je Abrechnungsfall liegt ein Datensatz vor. Für das QSR-Projekt sind dabei insbesondere folgende Informationen von Interesse:

- Alter bei Aufnahme;
- Geschlecht;
- Entlassungsgrund;
- Abteilung (inkl. Unterscheidung von Haupt- und Nebenabteilung);
- DRG;
- PCCL (ab 2000);
- Beatmungsstunden;
- Art und Tag des Eingriffs;
- Angaben zu Fallpauschalen und Sonderentgelten;
- Haupt-, Aufnahme-, Entlassungs- und Nebendiagnosen (ICD 9 bzw. 10);
- Sterbefälle im Krankenhaus;
- durchgeführte Eingriffe;
- Verweildauer;
- Zweiteingriffe während des initialen Aufenthalts;
- Wiedereinweisungen;
- Revisionen.

In die hier dargestellten Analysen fließen bundesweite Daten von AOK-Patienten mit Entlassungsdatum in den Jahren 1999–2001 ein. Es wurden nur solche Kliniken berücksichtigt, die sowohl im ersten Quartal 1999 als auch im letzten Quartal 2001 Fälle beisteuerten.

14.2.2 Tracer

Da es unmöglich und z. B. wegen des Vorkommens sehr seltener Diagnosen auch nicht sinnvoll ist, die Gesamtheit der Leistungen eines Krankenhauses abzubilden,

werden üblicherweise bestimmte Krankheitsbilder (Tracerdiagnosen) betrachtet. Wesentliche Grundlage für die Auswahl solcher Tracer sind die Häufigkeit der Erkrankung, deren eindeutige Messbarkeit und die Minimierung von Störeffekten⁵.

In QSR wurden bisher Tracer für Herzinsuffizienz, elektive Hüft-TEP und Operation bei Kolonkarzinom entwickelt und erprobt. Die Entwicklung und Evaluation weiterer Tracer ist Gegenstand aktuell laufender Arbeiten.

14.2.3 Qualitätsindikatoren

In QSR wird ein Qualitätsindikator als ein quantitatives Maß verstanden, welches Informationen über die Versorgungsqualität liefert, dieses komplexe Phänomen aber oft nicht direkt abbilden kann (Blumenstock 1996). Im Projekt QSR werden u. a. folgende Qualitätsindikatoren verwendet:

- Krankenhaussterblichkeit;⁶
- Verweildauer bzw. Anteil Langlieger;⁷
- Wiederaufnahmen;
- Re-Operation, bzw. Revisionsoperationen.

14.2.4 Risikostratifizierung

Da zu erwarten ist, dass die Patienten in einigen Häusern bei Aufnahme durchschnittlich kränker sind als in anderen, muss eine vergleichende Analyse dieses unterschiedliche Ausgangsrisiko angemessen berücksichtigen. Dabei wird durch die Auswahl der Tracer-Diagnosen bereits eine erste Risikoadjustierung durchgeführt, da erwartet werden kann, dass die Beschränkung der Analyse auf Krankenhauspatienten mit einer gut definierten Diagnose/Therapie bereits zu einer Angleichung des Risikoprofils führt. Darüber hinaus kann in QSR eine Risikoadjustierung u. a. nach Alter, Geschlecht, nach relevanten Nebendiagnosen und Krankheitsschweregraden (PCCL) durchgeführt werden.

14.2.5 Validierung der verwendeten Datengrundlagen

Um sicherzustellen, dass die verwendeten AOK-Abrechnungsdaten auch das tatsächliche Geschehen in der Klinik abbilden, wurde für die untersuchten Tracer eine Validierung der AOK-Abrechnungsdaten mit den im Krankenhaus verfügbaren Daten (HELIOS Krankenhausinformationssysteme) durchgeführt. Dabei zeigte sich insgesamt für die behandelten AOK-Patienten eine sehr gute Übereinstimmung der AOK-Abrechnungsdaten mit den HELIOS-Daten. Für einige Kliniken ergaben sich

5 Darüber hinaus sind zahlreiche weitere Kriterien denkbar, wie z. B. ökonomische Relevanz, Durchdringung verschiedener medizinischer Disziplinen, Unabhängigkeit vom Vergütungssystem und die Möglichkeit der Risikoadjustierung (vgl. auch Geraedts et al. 2003).

6 Der Begriff Krankenhaussterblichkeit wird in QSR definiert als die Anzahl der im Krankenhaus verstorbenen Patienten bezogen auf die Zahl der tracerspezifisch betrachteten Krankenhausfälle. Zukünftig werden aber auch Überlebensraten in definierten Zeitintervallen (z. B. 30 Tage, 90 Tage, 1 Jahr) analysiert werden können.

7 Gemeint ist der Anteil von Fällen mit einer Verweildauer oberhalb von 24 Tagen.

aber deutliche Abweichungen in Bezug auf die dokumentierten Nebendiagnosen und PCCL. In den hier ausgewerteten Jahrgängen sind die Nebendiagnosen im AOK-Datensatz gegenüber den in der Klinik dokumentierten noch unterrepräsentiert. Die Gründe können in einer zum damaligen Zeitpunkt noch unvollständigen Erfassung auf Kassenseite, aber auch in einer unvollständigen Meldung seitens der Krankenhäuser bestehen. Es ist damit zu rechnen, dass diese Erfassungslücken mit der Einführung des DRG-Systems und der damit erforderlichen vollständigen Meldung der Nebendiagnosen geschlossen werden, da u. a. das Abrechnungsergebnis der Kliniken von der Vollständigkeit der Meldungen abhängt.

14.3 Ergebnisse

14.3.1 Darstellung Bundesergebnisse am Beispiel Krankenhaussterblichkeit nach Kolonkarzinom-Operation

Im Folgenden wird beispielhaft für den Tracer „Operation bei Kolonkarzinom“ die Verteilung des Qualitätsindikators Krankenhaussterblichkeit dargestellt. Darüber hinaus werden Vorgehensweisen der Identifikation von potenziell auffälligen Kliniken gezeigt und problematisiert.

Als Schwellenwert wurde hier der von den HELIOS-Kliniken verwendete interne Grenzwert von 6 % eingesetzt. Dieser Wert ergibt sich aus den Analysen von Birkmeyer et al. (2002) als beobachtete, ganzzahlig gerundete Sterblichkeit für die beiden besten Mengenquintile (6,1 % und 5,4 %).

Abbildung 14-1 zeigt die Krankenhaussterblichkeit nach Kolonkarzinom-Operation für 708 Kliniken. Für 209 Kliniken (29 %) ergeben sich keinerlei Sterbefälle während des Krankenhausaufenthaltes, danach steigt die Krankenhaussterblichkeit bis zu einem Anteil von 37 % an. Würde ab 6 % Krankenhaussterblichkeit von einer auffälligen Rate ausgegangen, so wären 233 der 708 Kliniken (33 %) auffällig.

14

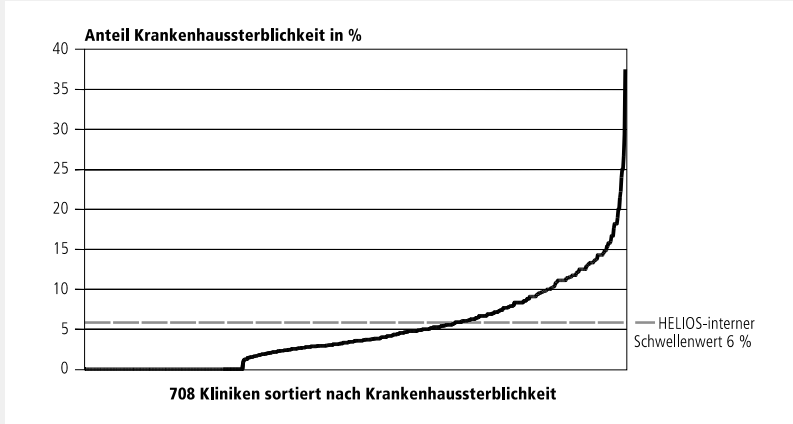
Zufälligkeit der Ergebnisse

Um dem Einwand zu entgegnen, dass rein zufällige Häufungen von Todesfällen in einigen Kliniken bereits als Qualitätsdefizite interpretiert werden, können Verfahren der schließenden Statistik angewendet werden. Meist werden 95 %-Konfidenzintervalle berechnet. Diese sollen anzeigen, ob es sich bei einer beobachteten erhöhten Anzahl von Todesfällen um eine zufällige oder statistisch signifikante Abweichung handelt. Weit verbreitet ist die Darstellung in sog. „league tables“, die graphisch meist als Fehlerbalkendiagramme aufbereitet werden. Wenn allerdings wie in unserem Fall mehr als 500 Kliniken verglichen werden, so ist diese Darstellung wenig übersichtlich (vgl. Abbildung 14-2).

Als ein möglicher Lösungsvorschlag für dieses Problem könnte nur die untere Grenze des 95 %-Konfidenzintervalls für die Krankenhaussterblichkeit über alle Kliniken hinweg aufgezeichnet werden (vgl. Abbildung 14-3). Unter Anwendung der oben bereits verwendeten Auffälligkeitsgrenze von 6 % sind nun nur noch 16 der 708 Kliniken (2,3 %) auffällig.

Abbildung 14-1

Krankenhaussterblichkeit nach Kolonkarzinom-Operation*



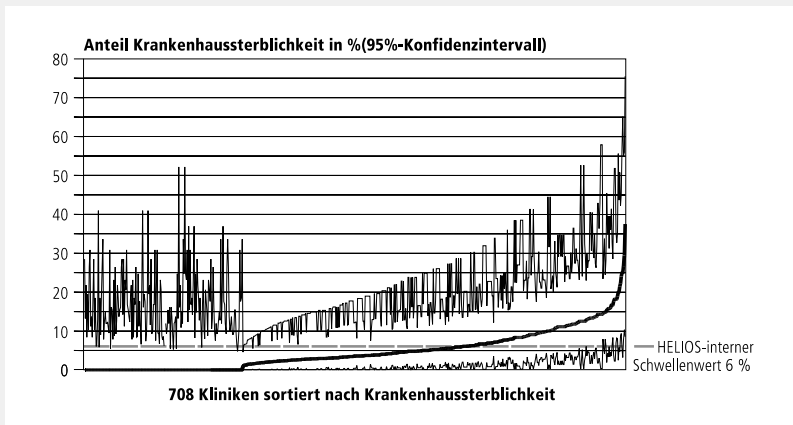
* Bundesauswertung 1999 - 2001, AOK-Patienten

WIdO © 2003

Quelle: WIdO, HELIOS 2003

Abbildung 14-2

Krankenhaussterblichkeit mit 95%-Konfidenzintervall nach Kolonkarzinom-Operation*

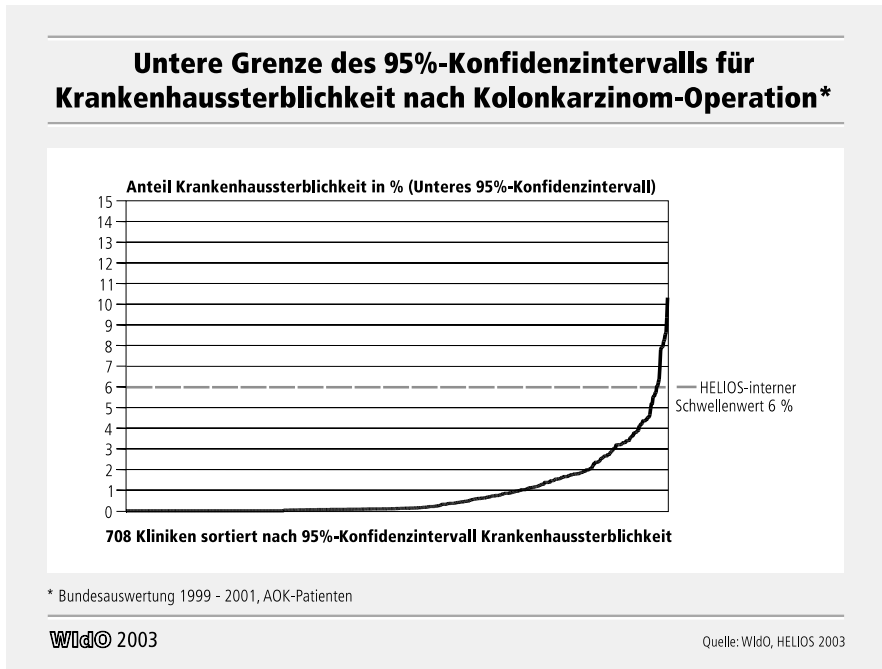


* Bundesauswertung 1999 - 2001, AOK-Patienten

WIdO © 2003

Quelle: WIdO, HELIOS 2003

Abbildung 14-3



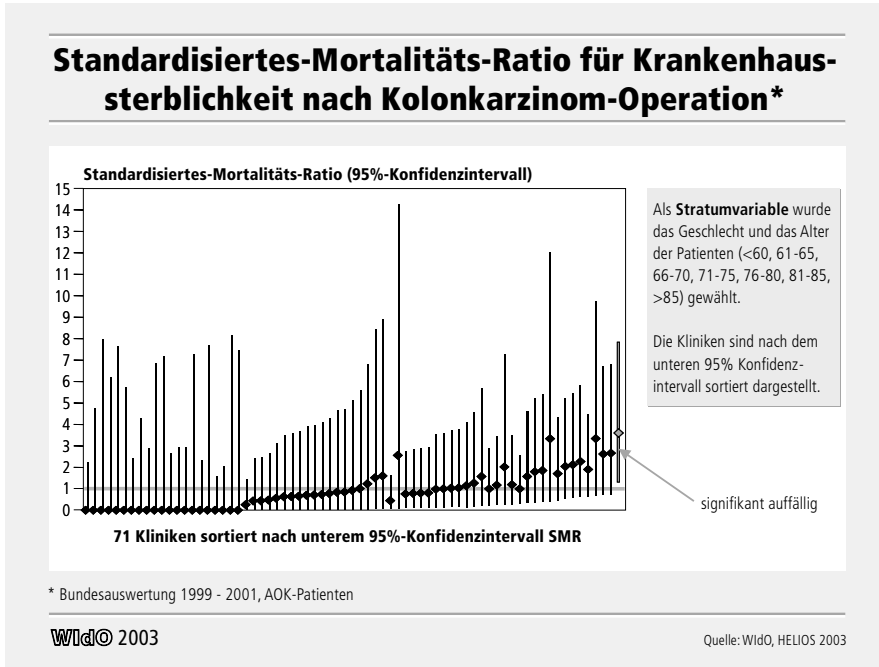
Dieses Ergebnis ist allerdings unbefriedigend, weil eingewandt werden kann, dass diejenigen Kliniken ungerecht behandelt werden, die kränkere Patienten behandeln oder – anders ausgedrückt – einen ungünstigeren Case-Mix aufweisen.

- Um eine Risikoadjustierung zu gewährleisten, werden häufig „standardized mortality ratios“ (SMRs) berechnet. Hier wird die beobachtete Anzahl von Todesfällen in einer Klinik mit der erwarteten Anzahl von Todesfällen in Beziehung gesetzt. Die erwartete Anzahl von Todesfällen wird über eine indirekte Standardisierung berechnet.
- Um keine zufälligen risikoadjustierten Häufungen von Todesfällen auszuweisen, ist es üblich 95%-Konfidenzintervalle für SMRs zu berechnen. Von einer bedeutsamen (signifikanten) Erhöhung einer risikoadjustierten Mortalität ist dann auszugehen, wenn die untere Grenze des 95%-Konfidenzintervall größer als 1 ist.

In der Abbildung 14-4 wurden SMRs (standardisiert für Alter und Geschlecht) für eine 10%-Zufallsstichprobe (71 Kliniken) nebst zugehörigen 95%-Konfidenzintervallen berechnet und in Form einer „league table“ als Fehlerbalkendiagramm dargestellt.

Von den 71 Kliniken in der Stichprobe weisen 26 (37%) ein $SMR > 1$ und 18 (24%) ein $SMR > 1,5$ auf. Bei einer Klinik beinhaltet das 95%-Konfidenzintervall der SMR den Wert 1 nicht mehr. Dies bedeutet, dass diese Klinik bezüglich ihrer Krankenhaussterblichkeit nach Risikoadjustierung signifikant auffällig ist.

Abbildung 14-4



14.3.2 Exkurs: Auffällige Krankenhaussterblichkeit in kleinen Kliniken

Die bisherigen Ergebnisse lassen sich wie folgt zusammenfassen: Wenn eine Krankenhaussterblichkeit von 6% als Auffälligkeitsgrenze definiert wird, liegen etwa ein Drittel der Kliniken über dieser Grenze (Abbildung 14-1). Anhand der analysierten 10%-Stichprobe lässt sich abschätzen, dass auch ein vergleichbar großer Anteil an Kliniken über die Berechnung der risikoadjustierten SMRs auffällig wäre (Abbildung 14-4).

Nach Anwendung von Verfahren der schließenden Statistik zeigt sich allerdings, dass für etwa 2% der Kliniken das 95%-Konfidenzintervall der errechneten Krankenhaussterblichkeit auf bzw. oberhalb der Auffälligkeitsgrenze von 6% liegt. (Abbildung 14-2). Auch die 95%-Konfidenzintervalle der SMRs liegen in nur einem Fall (oder in weniger als 2% der analysierten Stichprobe) oberhalb von 1 (Abbildung 14-4).

Die Ursache dafür, dass so viele Kliniken mit ihrem Zentralwert oberhalb der Auffälligkeitsgrenze liegen, aber nur so wenige auch mit den 95%-Konfidenzintervallen, sind die geringen Fallzahlen der meisten Kliniken in Verbindung mit der großen Seltenheit eines Todesfalls im Krankenhaus. Diese geringe Ereignishäufigkeit führt zu sehr breiten Konfidenzintervallen.

Umgangssprachlich ausgedrückt bedeutet dieser Sachverhalt, dass in kleinen Kliniken, trotz mitunter deutlicher erhöhter Krankenhaussterblichkeit, nicht mit der gleichen Sicherheit von einer Auffälligkeit ausgegangen werden kann wie in grö-

berer Kliniken. Anders formuliert heißt dieses aber, dass es für Kliniken mit geringen Fallzahlen eine wesentlich geringere Wahrscheinlichkeit gibt, über die in der Qualitätssicherung etablierten Verfahren der schließenden Statistik auffällig zu werden.⁸

Vor dem Hintergrund der aktuell geführten Mindestmengendiskussion, in der sich laut internationalen Studien auch für die operative Behandlung des Kolonkarzinoms ein – zwar im Vergleich zu einigen anderen Krankheitsbildern schwächerer, aber doch signifikanter – Mengeneffekt zeigt (Birkmeyer 2002; Schrag 2003), wird hier ein Defizit der angewandten Statistik deutlich. Im Folgenden werden einige Vorschläge für den Umgang mit diesem Problem vorgestellt.

Nur Analysen aufgrund der deskriptiven Statistik

Eine Lösungsmöglichkeit wäre, alle über dem Schwellenwert liegenden Kliniken zu prüfen bzw. zu einer Stellungnahme aufzufordern. Aufgrund der genannten Probleme einer rein statistischen Analyse führen beispielsweise die HELIOS-Kliniken vor allem in den Kliniken, deren Werte über dem HELIOS-intern festgelegten Schwellenwert liegen, Fallanalysen anhand der Aktenlage durch.

Auch aus den USA und beispielsweise aus Schottland sind entsprechende externe Audit-Verfahren bekannt, die bei auffälligen Werten bestimmter Kennziffern eingeleitet werden können oder – wie im schottischen Projekt – alle Fälle betreffen (Stonebridge 1999).

Dieses Vorgehen hat den Vorteil, dass einerseits Risikofaktoren, die in der statistischen Analyse nicht berücksichtigt wurden, in dem Verfahren erkannt und bewertet werden (und somit ggf. dem Einwand, es handle sich um besondere Fallkonstellationen, qualifiziert begegnet oder entsprochen werden kann) und dass andererseits auch Todesfälle in Häusern mit niedriger Fallzahl, die aufgrund statistischer Mechanismen sozusagen nicht auffällig werden „können“, kritisch untersucht werden (HELIOS Kliniken GmbH 2001, 2002).

Dies spiegelt auch die Grundprinzipien statistischer Datenanalyse wider. Denn auch hier ist es *lege artis*, sich zunächst und vor allen Dingen mit den deskriptiven Ergebnissen (hier Sterberaten) zu beschäftigen, diese zu studieren und zu bewerten. Keinesfalls wird empfohlen, sich nur mit signifikanten Ergebnissen zu beschäftigen.

80%-Konfidenzintervalle

Die Konfidenzintervalle könnten enger konstruiert werden. Man könnte sich die Frage stellen, warum eine „Sicherheit von 95%“ gefordert werden muss, bevor eine Klinik als auffällig gilt. Bekanntlich ist es nur eine Konvention, dass dies in der Wissenschaft meist gefordert wird.⁹ Im Sinne der Patienten könnte ebenso gefordert werden, bereits ab einer „80%igen Sicherheit“ aktiv zu werden.

⁸ Dieser Sachverhalt kann jedoch auch als Argument gegen kleine Kliniken benutzt werden.

⁹ Dabei sind auch hier bei speziellen Anwendungen Ausnahmen von der „95%-Regel“ durchaus üblich.

Standardabweichungsbasierte Auffälligkeit

Gelegentlich werden zur Ermittlung von Auffälligkeit auch Verfahren benutzt, die Standardabweichungen der jeweiligen Klinikverteilungen nutzen. Wenn beispielsweise eine Klinik mit ihrer Sterblichkeitsrate mehr als drei Standardabweichungen von dem gesetzten „Benchmark“ entfernt liegt, gilt sie als auffällig. Dieses Verfahren hat den Vorteil, dass kleine und große Kliniken grundsätzlich die gleiche (oder zumindest ähnliche) Wahrscheinlichkeit besitzen auffällig zu werden. Nachteil des Verfahrens ist allerdings, dass kleine Kliniken weiterhin geltend machen können, sie wären rein zufällig auffällig geworden. Damit besitzt dieses Verfahren keinen eigentlichen Vorteil gegenüber der erstgenannten Herangehensweise.

Kleine Kliniken als Gruppe zusammenfassen

Grundsätzlich bietet sich auch die Möglichkeit, anhand der Größe der Klinik¹⁰ verschiedene Gruppen zu bilden, um anschließend Ergebnisindikatoren für die gebildeten Gruppen von Kliniken zu berechnen.¹¹ Allerdings können dann zunächst Rückschlüsse auch nur auf die gesamte Gruppe von Kliniken gezogen werden, nicht auf eine einzelne Klinik. Risikoadjustierte Analysen sind auch mit diesem Ansatz möglich. Sie werden üblicherweise mittels einer logistischen Regression durchgeführt. Zur Illustration ist im Folgenden das Ergebnis einer solchen Analyse dargestellt. Dazu wurden die Kliniken nach der Anzahl der operierten Fälle in vier gleich große Gruppen (Quartile) eingeteilt. Gemeinsam mit den Informationen bezüglich der Nebendiagnose Ileus, des Alters (in sieben Kategorien) und des Geschlechts wurde die kategorisierte Klinikgröße dazu genutzt die Krankenhaussterblichkeit vorherzusagen (vgl. Tabelle 14-1).

Für die Gruppe der Kliniken mit weniger als 30 Kolonkarzinom-Operationen in drei Jahren ergibt sich ein Odds Ratio von 1,25. Die Wahrscheinlichkeit, nach einer Kolonkarzinom-Operation zu versterben, ist also in einer Klinik dieser Größe im Vergleich zu einer Klinik mit mehr als 65 Operationen um den Faktor 1,25 höher. Das Konfidenzintervall zeigt allerdings, dass es sich um ein nicht signifi-

Tabelle 14-1: Logistische Regression. Krankenhaussterblichkeit bei Kolonkarzinomoperation nach kategorisierter Klinikgröße, Bundesauswertung, AOK-Patienten, 1999–2001

Klinikgröße ¹	Odds Ratio	95%-Konfidenzintervall
< 30	1,25	0,90–1,72
30–45	1,04	0,77–1,42
46–65	0,95	0,69–1,32
> 65	1 (Referenzkategorie)	–

© WldO/Helios 2003

¹ Operationalisiert als Anzahl von Kolonkarzinomoperationen im Untersuchungszeitraum

¹⁰ Beispielsweise operationalisiert über die Anzahl der behandelten Fälle in einem gegebenen Zeitintervall.

¹¹ In Deutschland wurden für den Bereich der Geburtshilfe solche Analysen beispielsweise kürzlich von Heller und Koautoren veröffentlicht (Heller et al. 2002, Heller et al. 2003).

kantes Ergebnis handelt. Der Einfluss der übrigen Klinikgrößenklassen ist in dieser Analyse noch erheblich geringer.

Es sei nochmals angemerkt, dass auch ein signifikantes Ergebnis in einer solchen Analyse keinesfalls bedeutet, dass jede Klinik in der betrachteten Gruppe hinsichtlich des Qualitätsindikators schlecht zu beurteilen ist.¹²

Langzeitbeobachtung

Schließlich bietet es sich im Falle von QSR an, die interessierenden Qualitätsindikatoren auch über längere Zeiträume zu betrachten. So würde auch eine „schlechte“ kleine Klinik nach längerer Zeit auffällig werden. Unlängst wurden mehrere statistische Verfahren vorgeschlagen, um eine solche Analyse risikoadjustiert und möglichst effizient durchführen zu können (Tekkis et al. 2003, Spiegelhalter et al. 2003). Allerdings erstrecken sich auch die hier vorgestellten Analysen zur Krankenhaussterblichkeit nach Operation eines Kolonkarzinoms über einen Zeitraum von drei Jahren. So gesehen handelt es sich bei den vorgestellten Ergebnissen bereits um derartige Analysen.

Welche der vorgestellten statistischen Herangehensweisen bei der Beurteilung von Ergebnisqualität angewandt werden sollte, kann und soll hier nicht abschließend beantwortet werden, zumal dies auch stark von der aktuellen Fragestellung abhängt. Allerdings erscheint uns ein mehrstufiges Verfahren der Auffälligkeitsprüfung, welches verschiedene der im Exkurs angesprochenen Herangehensweisen beinhaltet, eine sinnvolle Strategie zu sein. Eine weitergehende Auseinandersetzung zumindest mit den in dieser Analyse unter Maximalkriterien (95%-Konfidenzintervall) auffälligen Kliniken erscheint auf jeden Fall sinnvoll.

14.4 Diskussion und Ausblick

Die vorgestellten Analysen zeigen das große Potenzial des Projekts für eine zeitnahe, kostengünstige, vollständige und valide Qualitätssicherung in Deutschland. Allerdings sind weitere Schritte nötig, um ein valides und aussagekräftiges Instrument der Erhebung von Ergebnisqualität zu erhalten.

Einbeziehung weiterer Daten

Dieser Beitrag dokumentiert den Stand zum Abschluss der ersten Projektphase (QSR I) Anfang Juni 2003. Neben den aktuell bereits entwickelten Tracern (Operation bei Kolonkarzinom, Herzinsuffizienz und elektive Hüft-TEP) ist eine Erweiterung des Projekts auf folgende Tracer vorgesehen: Mammakarzinom, Appendektomie, Hirninfarkt (TIA), Herzinfarkt, diagnostischer und therapeutischer Herzkatheter (jeweils mit und ohne Infarkt), Hüft-TEP nach Fraktur sowie elektive Knie-TEP. Die Verfahren der Risikoadjustierung werden kritisch überprüft.

¹² Darüber hinaus bedarf die Beurteilung von Ergebnissen multivariater Regressionsverfahren im Allgemeinen und logistischer Regressionen im Besonderen weiterer tiefergehender Analysen, bevor endgültige Schlüsse daraus abgeleitet werden können (Hosmer & Lemeshow 2000, Rothman & Greenland 1998).

Ein weiteres Ziel der nächsten Projektphase ist es, getrennt nach Tracern jeweils evidenzbasierte Auffälligkeitsgrenzen für die gewählten Qualitätsindikatoren zu erarbeiten.

Dabei ist absehbar, dass die Aussagefähigkeit von QSR-Daten durch eine Einbeziehung weiterer Daten, wie zum Beispiel 1-Jahres-Überlebensraten, ggf. erheblich erhöht werden kann.

Etablierung der Qualitätssicherung mit Routinedaten?

Wenn die oben genannten Arbeiten abgeschlossen sind, sollen die in QSR gewonnenen Erkenntnisse den beteiligten Kliniken in Form eines Qualitätsberichts verfügbar gemacht werden. So könnten alle an der Regelversorgung in Deutschland teilnehmenden Kliniken zeitnah über ihre Ergebnisqualität informiert werden und diese Informationen für ihr internes Qualitätsmanagement nutzen. Insofern stellt QSR zu wesentlichen Teilen auch eine Serviceleistung der Kostenträger für die Leistungserbringer dar.

Schließlich sollen Möglichkeiten geprüft werden, wie diese Informationen für Verbraucher verständlich aufbereitet und zur Verfügung gestellt werden können.

14.5 Literatur

- Arnold M, Klauber J, Schellschmidt H, Hrsg. Krankenhaus-Report 2002. Stuttgart, New York: Schattauer 2003.
- Birkmeyer JD, Hamby LS, Birkmeyer CM, Decker MV, Karon NM, Dow RW. Is unplanned return to the operation room a useful quality indicator in general surgery? *Arch Surg* 2001; 136: 405–10.
- Birkmeyer JD, Siewers AE, Finlayson EVA, Stukel TAS, Lucas FL, Batista I, Welch HG, Wennberg DE. Hospital volume and surgical mortality in the United States. *N Engl J Med* 2002; 346: 1128–37.
- Black C, Roos NP. Administrative data. Baby or bathwater? *Medical Care* 1998; 36: 3–5.
- Blumenstock G. Qualitätsmanagement im Krankenhaus: Qualitätsindikatoren der stationären Versorgung auf Basis administrativer Daten. St. Augustin: Asgard-Verlag 1996.
- Bundesgeschäftsstelle Qualitätssicherung (BQS). Qualität sichtbar machen. Geschäftsbericht 2001/2002. Düsseldorf 2002a.
- Bundesgeschäftsstelle Qualitätssicherung (BQS). Qualität sichtbar machen. BQS-Qualitätsreport 2001. Düsseldorf 2002b.
- Eichenlaub A, Sangha O, Schellschmidt H, Schneeweiß S. Ergebnisse der stationären medizinischen Versorgung – Das Qualitätsmodell Krankenhaus. In: Arnold M, Klauber J, Schellschmidt H, Hrsg. Krankenhaus-Report 2001. Stuttgart, New York: Schattauer 2002; 209–22.
- Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Medical Care* 1998; 36: 8–27.
- Garnick DW, DeLong ER, Luft HS. Measuring hospital mortality rates: are 30-day data enough? *Health Services Research* 1995; 29: 679–95.
- Geraci JM. In-hospital complication occurrence as a screen for quality-of-care problems. *Medical Care* 2000; 38: 777–80.
- Geraedts M, Selbmann HK, Ollenschlaeger G. Critical appraisal of clinical performance measures in Germany. *International Journal for Quality in Health Care* 2003; 15: 79–85.
- HELIOS Kliniken Gruppen, Hrsg. Kompetenz in Medizin. Medizinischer Jahresbericht 2000. Fulda 2000.

- HELIOS Kliniken Gruppen, Hrsg. Kompetenz in Medizin. Medizinischer Jahresbericht 2001. Fulda 2001.
- HELIOS Kliniken Gruppen, Hrsg. Kompetenz in Medizin. Medizinischer Jahresbericht 2002. Fulda (in Vorbereitung).
- Heller G, Richardson DK, Schnell R, Misselwitz B, Künzel W, Schmidt S. Are we regionalized enough? Early-neonatal deaths in low-risk births by the size of delivery units in Hesse, Germany 1990–1999. *International Journal of Epidemiology* 2002; 31:1061–8.
- Heller G, Schnell R, Richardson DK, Misselwitz B, Schmidt S. Hat die Größe der Geburtsklinik Einfluss auf das neonatale Überleben? Schätzung von „vermeidbaren“ Todesfällen in Hessen 1990–2000. *Dtsch Med Wschr* 2003; 128: 657–62.
- Iezzoni LI. Risk adjustment for medical effectiveness research: an overview of conceptual and methodological considerations. *J Invest Medicine* 1994; 43: 136–50.
- Iezzoni LI, Hrsg. Risk adjustment for measuring healthcare outcomes. Chicago: Health Administration Press 1997.
- Iezzoni LI. Assessing quality using administrative data. *Annals Internal Medicine* 1997; 127: 666–74.
- Iezzoni LI, Mackiernan YD, Calahane MJ, Phillips RS, Davis RB, Miller K. Screening inpatient care using post-discharge events. *Medical Care* 1999; 37: 384–98.
- Jacobson B, Mindell J, McKee M. Hospital mortality league tables. Question what they tell us – and how useful they are. *Brit Med J* 2003; 326: 777–8.
- Landon B, Iezzoni LI, Ash A, Shwartz M, Daley J, Hughes JS, Mackiernan YD. Judging hospitals by severity adjusted mortality rates: the case of CABG surgery. *Inquiry* 1996; 33: 155–66.
- Leyland AH, Boddy FA. League tables and acute myocardial infarction. *Lancet* 1998; 351: 555–8.
- Mitchell JB, Bubolz T, Paul JE, Pashos CL, Escarce JJ, Muhlbaier LH, Wiesman JM, Young WW, Epstein RS, Javitt JC. Using Medicare claims for outcome research. *Medical Care* 1994; 32: JS38–JS51.
- Mukamel DB, Zwanziger J, Tomaszewski KJ. HMO penetration, competition, and risk-adjusted hospital mortality. *Health Services Research* 2001; 36: 1019–35.
- Park RE, Brook RH, Kosecoff J, Kessey J, Rubenstein L, Keeler E, Kahn KL, Rogers WH, Chassin MR. Explaining variations in hospital death rates – randomness, severity of illness, quality of care. *JAMA* 1990; 264: 484–90.
- Pouvoirville G. Kann Qualität ein Wettbewerbsparameter im Gesundheitswesen sein? In: Arnold M, Klauber J, Schellschmidt H, Hrsg. Krankenhaus-Report 2001, Stuttgart, New York 2002; 175–88.
- Pouvoirville G, Mivielle E. Measuring the Quality of Hospital Care: The State of the Art. Konferenz „Measuring up“, Ottawa 5-7.11.2001: 251–75.
- Roos NP, Black CD, Roos LL, Tate RB, Carriere KC. A population-based approach to monitoring adverse outcomes of medical care. *Med Care* 1995; 33: 127–38.
- Roos NP, Roos LL, Mossey J, Havens B. Using administrative data to predict important health outcome: entry to hospital, nursing home, and death. *Med Care* 1988; 26: 221–39.
- Schneeweiss S, Eichenlaub A, Schellschmidt H, Wildner M. Qualitätsmodell Krankenhaus (QMK – Ergebnis-Messung in der stationären Versorgung). Abschlussbericht. WIdO-Materialien. Bonn 2003.
- Schrag D, Panageas KS, Riedel E, Hsieh L, Bach PB, Guillem JG, Begg CB. Surgeon volume compared to hospital volume as predictor of outcome following primary colon cancer resection. *J Surg Oncol* 2003; 83: 68–78.
- Spiegelhalter D, Grigg O, Kinsman R, Treasure T. Risk-adjusted sequential probability ratio tests: applications to Bristol, Shipman and adult cardiac surgery. *Int J Qual Health Care* 2003 ;15: 7–13.
- Statistisches Bundesamt. Diagnosestatistik der Krankenhäuser 1998. Wiesbaden 2000.
- Stonebridge PA, Thompson AM, Nixon SJ. Completion of the journey of care: Scottish audit of surgical mortality (SASM). *J R Coll Surg Edinb* 1999; 44: 185–6.
- Tekkis PP, McCulloch P, Steger AC, Benjamin IS, Poloniecki JD. Mortality control charts for comparing performance of surgical units: validation study using hospital mortality data. *Brit Med J* 2003; 326: 786–8.

- Wennberg JE, Mulley AG, Hanley D, Timothy RP, Fowler FJ, Roos NP, Barry MJ, McPherson K, Greenberg ER, Soule D, Bubolz T, Fisher E, Malenka D. An assessment of prostatectomy for benign urinary tract obstruction: geographic variation and the evaluation of medical care outcomes. *JAMA* 1988; 259: 3027–30.
- Wennberg JE, Roos N, Sola L, Schori A, Jaffe R. Use of claims data systems to evaluate health care outcomes. *JAMA* 1987; 257: 933–6.
- Wray NP, Ashton CM, Kuykendall DH, Hollingsworth JC. Using administrative databases to evaluate the quality of medical care: A conceptual framework. *Soc Sci Med* 1995; 40: 1707–15.