Qualitätssicherung mit Routinedaten

Wissenschaftlicher Streit oder doch eher Poker um die Datenhoheit?

Erneut gibt es Streit um die Berechtigung der Qualitätssicherung mit Routinedaten (QSR), wie sie die AOK seit Jahren vorlegt. Eine neue Publikation stellt die Berechnungsgrundlage hinter den dabei verwendeten Indikatoren in Frage [1]. Wackelt das Verfahren tatsächlich oder munitionieren sich gegnerische Lager im Gesundheitswesen gerade für die nächste Runde im Tauziehen um die Gestaltung der Krankenhauslandschaft?

Für die einen die Zukunft der Qualitätsmessung in der Medizin, für die anderen eher ein rotes Tuch: Die QSR – seit 2002 entwickelt von AOK-Bundesverband, Wissenschaftlichem Institut der AOK (WIdO), Helios-Kliniken und dem Forschungsinstitut FEISA aus Sachsen-Anhalt. Seit 2008 entwickelt das WIdO QSR federführend weiter, flankiert von einem Wissenschaftlichen Beirat und Expertenpanels zu einzelnen Indikatoren. Und es ist, wie sein Auftraggeber der klaren Ansicht: Das kann man so machen – die Abrechnungsdaten der Kliniken taugen für die Bewertung medizinischer Qualität.

Beim Einsatz von Routinedaten für die Qualitätsbewertung in der Medizin ist die AOK nicht alleine. So lassen sich rund 500 Kliniken das eigene Abschneiden bei QSR mittlerweile jährlich für die eigene Arbeit an Qualitätsverbesserungen in – allerdings internen – Klinikberichten vorlegen. Auch die rund 300 Kliniken der Initiative Qualitätsmedizin (IQM) nützen über „300 Ergebnis-, Mengen- und Verfahrenskennzahlen“, alle aus Routinedaten entwickelt, für die Qualitätssicherung. Gepflegt wird das hier zugrundeliegende Set an Deutschen oder Schweizer „Inpatient Quality Indicators“ von der Gruppe um Professor Thomas Mansky an der TU Berlin. Und auch Barmer GEK, TK und viele weitere Kassen werten die ihnen vorliegenden Abrechnungsdaten für Qualitätsfragen der Versorgungsforschung aus.

Wäre da nicht eine Besonderheit des AOK-Konzepts. Nur die größte Kasse stellt Auswertungen seit 2008 öffentlich – Haus für Haus. Sie will Patienten damit ein Ranking-Instrument bieten, um im Vorfeld einer anstehenden Operation ein qualitativ gutes Haus zu finden, etwa für eine Implantation eines künstlichen Knie- oder Hüftgelenks.

Kein Fachkongress der Chirurgie oder von Statistikern des Gesundheitswesens, auf denen seither nicht auch eine Kontroverse zu QSR gegeben wird. Auf der einen Seite die AOK, gelegentlich mit Rückendeckung durch weitere Kassen, auf der anderen vor allem die Leistungserbringer – Krankenhäuser, alias einzelne Landesgesellschaften oder die Deutsche Krankenhausgesellschaft.

Es ist ein Streit auf 2 Ebenen. Im Vordergrund steht die fachliche Kontroverse: Misst QSR wirklich Qualität von Kliniken oder bietet es verzerrte Werte, die Patienten, zumeist Laien zu ganz falschen Entscheidungen verleiten könnten? Grundiert wird die fachliche Auseinandersetzung von einer politischen Kontroverse: Die Tatsache, dass Kassen auf einem Datenfundus sitzen, den sie aus eigenen Stücken für die Qualitätsbewertung der Leistungen von Klinik zu Klinik einsetzen können, schafft den Leistungserbringern ein Problem, wenn sie hier nicht über die Veröffentlichung mitentscheiden können.

Einigkeit über Qualitätswettbewerb

Dabei herrscht über das Ziel Einigkeit. „Wir wollen den Wettbewerb zwischen Kliniken um Qualität“ stellt auch Burkhard Fischer von der Krankenhausgesellschaft Nordrhein-Westfalen klar (siehe das Interview ab Seite 220). Allerdings wünschen die Krankenhausgesellschaften, dass er auf einer „methodisch sauberen Grundlage“ stattfindet – und die sehen sie aktuell dann gewährleistet, wenn er vom Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTiG) neutral begleitet wird. Mithin, wenn er im großen System der externen stationären Qualitätssicherung verortet wird, wo auch die Krankenhäuser über den G-BA (Gemeinsamer Bundesausschuss) mitsteuern können, was gemessen, was öffentlich wird und was nicht.

Die vom Gesetzgeber vorgeschriebene Externe Stationäre Qualitätssicherung – Einzelheiten legt vor allem der §137 des SGBV fest – hat als Datenbasis ein großes Erfassungssystem aus aktuell 279 Indikatoren. Weisse Liste oder vdek-Kliniklotse – das Fundament all dieser „Navigatoren“ bilden die Daten aus den in jährlichem Turnus von rund 2000 zugelassenen Krankenhäusern abzugebenden strukturierten Qualitätsberichten, für deren Archivierung und Weitergabe der G-BA die Fäden in der Hand hält [2]. Auch der AOK-Krankenhausnavigator bietet als Basis diese Daten. Wie andere Kassen auch hat die AOK für deren Aufbereitung die Weisse Liste beauftragt [3].

Die Darstellung des Abschneidens einzelner Kliniken bei medizinischen Parametern, etwa Zahl der Wundinfekte nach Operationen oder auch die Lockerung einer Prothese unmittelbar nach der Implantation, setzt überwiegend auf den Bezug der Einzelwerte eines Hauses zu Durchschnittswerten aller Häuser und zu so genannten Referenzbereichen, bei deren Überschreiten die Qualitätsprüfer mit einem Haus intern Rücksprache nehmen. Der Internetnutzer kann sich ein Ranking aller Kliniken etwa nach Fallzahlen, dem Abschneiden bei einer Patientenumfrage von Barmer und AOK und auch nach dem Abschneiden bei der Gesetzlichen Qualitätssicherung geben lassen. Letzteres bleibt allerdings abstrakt. Messlatte ist, welcher Prozentsatz einer ganzen Reihe an wiederum qualitativ sehr unterschiedlichen Parametern erfüllt wird.

Die Sicht dieses Systems endet an der Kliniktüre, sobald der Patient das Krankenhaus verlässt, ist er für die Klinik vom Radar verschwunden. Denn in diesem Erfas-

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

sungssystem findet keine Verknüpfung der Datensätze eines Patienten über längere Zeiträume statt. Eine Neuaufnahme wird nicht in Bezug zu einer möglichen Erstoperation gesetzt.

Anders bei Routinedaten. Dieser Datenfundus der Kassen speist sich aus dem so genannten 301er-Datensatz (nach dem §301 SGB V), den die Krankenhäuser seit 2003 zu Abrechnungszwecken via DRG-System an die Kassen geben. Bei der Abrechnung muss es nicht bleiben, Kassen, aber auch Kassenärztliche Vereinigungen und unabhängige Forschungsgruppen dürfen die Daten nach Anonymisierung wissenschaftlich weiter verwerten [4]. Und hier erlaubt die Versichertennummer nach Anonymisierung die Verknüpfung der Daten eines Patienten über Jahre und Jahrzehnte hinweg, ein Langzeitmonitoring über die Grenzen der Versorgungssektoren von Krankenhaus und Arztpraxis ist möglich – zumindest so lange der Versicherte brav bei seiner Kasse bleibt.

Streit um Ranking der Kliniken

QSR ist dabei bislang das einzige Verfahren am Markt, das dezidiert eine Einteilung der Kliniken nach der jeweils ermittelten Qualität für konkrete medizinische Parameter bietet und es dem Nutzer mit 2 Mausklicks erlaubt, Häuser zum Beispiel nach der Rate an bestimmten Komplikationen nach Operationen zu sortieren. Jeder Internetnutzer findet QSR als Zusatzservice im AOK-Krankenhausnavigator. Mit ihren rund 24 Millionen Versicherten, rund ein Drittel der über 70 Millionen Kassenpatienten in Deutschland, verfügen die AOKen über reichlich statistische Power für den Anspruch einer vergleichenden Krankenhausbewertung. Veröffentlicht werden aktuell Daten zu 8 Behandlungen. Darunter sind die Implantation künstlicher Knie- und Hüftgelenke, Operationen an Gallenblase, Prostata- und Blinddarmoperationen.

Bei einer Indikation wie der Implantation eines künstlichen Hüftgelenks nach Arthrose oder nach einem gelenknahem Bruch findet der Leser bei QSR nicht nur Ziffern etwa zur Rate an Patienten, die binnen 90 oder gar 365 Tagen nach der Operation in einem bestimmten Haus Komplikationen hatten [5]. Bei jedem Qualitätsparameter alias -Indikator sortiert der Krankenhausnavigator der AOK eine Klinik auch gleich noch in eine von drei Gruppen:

- Ein Plus-Symbol steht für „Überdurchschnittliche Qualität“, die Klinik gehört zu den 20 % der Kliniken mit einer geringen Wahrscheinlichkeit für unerwünschte Ereignisse.
- Ein Kreis steht für: „Durchschnittliche Qualität“, die Klinik gehört zu den 60 % aller Kliniken mit einer mittleren Wahrscheinlichkeit für unerwünschte Ereignisse.
- Ein Minus gibt es für „Unterdurchschnittliche Qualität“, die Klinik gehört zu den 20 % aller Kliniken mit einer höheren Wahrscheinlichkeit für unerwünschte Ereignisse.

Besonders übersichtlich wirkt QSR, indem das Abschneiden einer Klinik bei den Einzelparametern einer Indikation auch noch zusammen in einen Summenwert gegossen wird – eine Note für alles. QSR vergibt dafür bis zu 3 Lebensbaumsymbole, die Anzahl folgt wieder der gleichen Plus-Minus-Systematik: Eins für unterdurchschnittlich, 2 für durchschnittlich und 3 Bäume für überdurchschnittlich.

Und damit Vorhang auf für einen Methodenstreit, der jeden Nicht-Mathematiker schnell in unergründliche Feinheiten der Statistik, von Koeffizienten, Tabellen und Wichtungsfaktoren führt. In der Zeitschrift Gesundheitsökonomie & Qualitätsmanagement [1] kritisieren Professor Harald Englisch vom Verein Gesundheit Mitteldeutschland e. V., Leipzig und Burkhard Fischer von der Krankenhausgesellschaft Nordrhein-Westfalen e. V., Düsseldorf die QSR-Verfahren des WIdO. Die AOK, so das Fazit der Autoren, solle ihren „Alleingang bei der Definition von Krankenhausqualität“ beenden.

Sie stützen ihre These auf die Analyse der QSR-Daten zu 940 Krankenhäusern bei einem orthopädischen „Klassiker“. Im Fachjargon ist es das: „Einsetzen eines künstlichen Hüftgelenks nach hüftgelenknahen Oberschenkelbruch.“

QSR bietet hier zunächst für jede Klinik 3 Einzelindikatoren. Es geht zunächst um das Abschneiden einer Klinik bei:

- Ungeplanten Folge-Operation innerhalb von 365 Tagen nach dem Eingriff
- Chirurgischen Komplikationen innerhalb von 90 Tagen nach dem Eingriff
- Sterblichkeit innerhalb von 90 Tagen nach dem Eingriff

Wie WIdO-Mathematiker Günster im Interview erläutert, berücksichtigt QSR die Tatsache, dass in den Kliniken verschiedene Patienten sind – gerade Häuser der Maximalversorgung haben oft mehr alte und gebrechliche Patienten, bei denen die Risiken bei Erkrankungen und Behandlungen per se höher sind. Um statistisch gesehen hier für Gleichheit zu sorgen, setzt das WIdO für diese Risikoadjustierung auf einen Algorithmus, der von der US-Amerikanerin Anne Elixhauser an der US-Agency for Healthcare Research and Quality, AHRC) entwickelt worden ist.

Danach wird die Zahl der tatsächlichen in einer Klinik erfassten Komplikationen auf die Zahl der statistisch zu erwartenden bezogen. Das Ergebnis ist eine Standard Morbidity (Mortality) Ratio (SMR). Schlichte Regel: Ein Wert höher 1 bedeutet: Es kam häufiger als zu erwarten zu dem Problem. Ein Wert gleich 1 bedeutet – alles im Durchschnitt. Ein Wert unter 1 heißt: Die Komplikation trat sogar seltener auf als sie zu erwarten war. Außerdem kalkuliert das WIdO für jeden SMR-Wert auch noch das so genannte 95 %- Konfidenzintervall (siehe das Interview ab Seite 223).

Ranking nach 95 %-Konfindenzintervall oder nach SMR-Wert?

Das WIdO sortiert die Kliniken in einem Ranking allerdings nicht nach SMR-Werten. Dafür dienen vielmehr die Grenzwerte des jeweiligen 95 %-Konfidenzintervalls um einen SMR-Wert. Und jetzt landen

- die 20 % der Häuser mit den niedrigsten oberen Grenzwerten des 95 %-Konfidenzintervalls in der Gruppe Plus, also bei der Gruppe, die gut abschneidet
- die 20 % der Häuser mit den obersten unteren Grenzwerten des Konfidenzintervalls in der Gruppe Minus
- alle anderen in der Gruppe der Häuser „mittlerer“ Qualität.

Englisch und Fischer lehnen das Verfahren ab. Ihnen seien aus der Literatur keine „weiteren Beispiele für diese ungewöhnliche Klassifikation bekannt“. Ihre These stützen sie durch einen Vergleich des Abschneidens der Kliniken beim Einzelindikator „Sterblichkeit binnen 90 Tagen nach der Operation“ mit dem Gesamtabschneiden, dies außerdem im Vergleich der Fallzahlen in den Kliniken.

Ihr Ergebnis: QSR führe vor allem bei Kliniken mit hohen oder niedrigen Fallzah-

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

len zu verzerrten Ergebnissen – es benachteilige Häuser mit hohen Fallzahlen, die „überproportional oft“ entweder in der Kategorie mit nur 1 oder gleich mit 3 Lebensbäumen landeten. Andererseits landeten von den Häusern mit geringer Fallzahl mehr in der mittleren Qualitätskategorie, also 2 Lebensbäumen, als zu erwarten wäre. Mithin: Die Fallzahl entscheide mit darüber, wo ein Haus bei QSR in der Qualitätseinstufung landet. Die AOK solle wenn überhaupt, dann besser die SMR-Werte darstellen nebst den Konfidenzintervallen – das sei Standard in der Statistik meint Burkhard Fischer, Mathematiker und Bereichsleiter bei der KGNW (Krankenhausgesellschaft Nordrhein-Westfalen). Und der Ort für die Klärung solcher Fragen sei das Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTiG), das seit Anfang des Jahres für den G-BA die Externe Stationäre Qualitätssicherung betreut (siehe das Interview ab Seite 220).

Christian Günster vom WIdO weist die Arbeit seiner Kritiker im ZfOU-Gespräch hingegen als ihrerseits methodisch fehlerhaft zurück. Nebenbei: Prozesse gegen QSR hat die AOK bislang gewonnen. Schon 2014 verloren 2 Kliniken vor Gericht mit dem Antrag auf Einstweilige Verfügung gegen den AOK-Bundesverband. Sie wollten, dass die Veröffentlichung ihrer Daten bei QSR unterbleibt. Klagen im Hauptverfahren gingen bis heute nicht bei Gerichten ein. Der G-BA solle zumindest Elemente aus dem QSR-Verfahren endlich auch als Teil der Gesetzlichen Qualitätssicherung verankern, fordert Günster.

Man könne dem WIdO schlecht zum Vorwurf machen, dass jede Qualitätsbewertung immanente statistische Probleme hat, sekundiert auch Thomas Mansky. Die Konfidenzintervalle derart in die Betrachtung einzubeziehen, sei durchaus Standard. Und dass die SMR-Werte erst ab höheren Fallzahlen auch statistisch signifikant würden, sei Grundgesetz der Mathematik. Mansky: „Das WIdO-Verfahren ist eine Möglichkeit für ein Ranking der Kliniken – es gibt weitere und jedes hat Vor- und Nachteile.“

Mansky war 2002 als Leiter des Qualitätsmanagements bei der Helios-Klinikgruppe Mitentwickler der Qualitätsindikatoren bei QSR, hat allerdings mit den aktuellen Auswertungen des WIdO für den AOK-Krankenhausnavigator nichts zu tun. Ihn stört auch grundsätzlich ein teilweise immer noch vorhandener paternalistischer Ansatz in der Darstellung von medizinischer Qualität. Die Patienten, so Mansky, hätten ein Recht auf derartige Ergebnisse und im Zweifel solle man Ihnen und auch Journalisten, die Ergebnisse weiter aufbereiten, ruhig zutrauen, damit „vernünftig umzugehen und sich aus verschiedenen Informationsangeboten eine eigene Meinung zu bilden.

Ein guter Qualitätsindikator müsse unabhängig von der Fallzahl sein, meint wiederum ein anderer Spezialist auf dem Gebiet der Qualitätsmessung in der Medizin, Professor Hans-Konrad Selbmann vom Vorstand der AWMF (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften). Die QSR-Ergebnisse, so Selbmann, müssten endlich validiert werden (siehe das Interview ab Seite 226).

So ganz undurchlässig sind die Grenzen zwischen beiden Ansätzen auf jeden Fall nicht mehr. Routinedaten haben zaghaft Eingang gefunden – auch in Konzepte für die Externe Stationäre Qualitätssicherung, für die der G-BA seit vielen Jahren neue Konzepte für eine sektorenübergreifende Sicht erarbeiten möchte [6].

Das bis Ende letzten Jahres für die Externe Stationäre Qualitätssicherung beauftragte AQUA-Institut (AQUA: Angewandte Qualitätsförderung und Forschung im Gesundheitswesen) hat schon seit 2012 in mehreren Gutachten dargelegt, dass manche Parameter und Verfahren nach QSR eine zumindest taugliche Erweiterung des Datensets sind [7].

Es ist kein akademischer Streit, um den es hier geht. Neuerdings wünscht sich sogar das Bundesgesundheitsministerium vom G-BA belastbare Qualitätsindikatoren für die Krankenhausplanung. Will sagen: Qualitätsindikatoren, die einem Haus, das deren Messlatte reißt, finanzielle Abschläge, einem anderen, das exzellente Leistung bietet, einen Zuschlag bescheren sollen. Pay for Performance – das seit Januar 2016 gültige Krankenhausstrukturgesetz macht es möglich.

Allein, wonach bewerten? Die Zeit drängt. Seit Anfang 2016 betreut das IQTiG für den G-BA die Externe Stationäre Qualitätssicherung. Bis August 2016 soll das Institut dem G-BA Parameter vorlegen, mit denen die Politik finanzielle Daumenschrauben je nach medizinischer Qualität in die Krankenhausplanung einbringen

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

will. Das Institut bereitet schon den Zeitplan für eine Anhörung durch die Fachgesellschaften der AWMF vor (siehe das Interview ab Seite 226).

Können Routinedaten à la QSR helfen?

Beim IQTiG ist ein guter Teil der ehemaligen Mitarbeiter von AQUA unter Vertrag, darunter sind auch Mitentwickler von QSR. Wo also wäre die Frage um den Stellenwert und den sinnvollen Einsatz von Routinedaten und Methoden nach QSR besser aufgehoben als dort? Um eine Bewertung zur aktuellen Publikation von Englisch und Fischer und der Methodik von QSR gebeten – fällt die Antwort beim IQTiG allerdings ausgesprochen schmallippig aus. Um da eine „belastbare und seriöse Antwort" geben zu können, müsse man die Angelegenheit ausführlich prüfen, mithin tiefer in den Kontext einsteigen, lautet die Auskunft. Für die nötige ausführliche Recherche fehlten momentan die personellen und zeitlichen Ressourcen.

Literatur

1 Englisch H, Fischer B. Worüber sagen AOK-Lebensbäume mehr aus: über Qualität oder über Fallzahlen? Gesundh ökon Qual manag 2016; DOI: 10. 1055/s-0035–1554015

2 https://www.g-ba.de/institution/themenschwerpunkte/qualitaetssicherung/qualitaetsbericht/suche/ Stand: 09.05.2016

3 https://weisse-liste.krankenhaus.aok.de/ Stand: 09.05.2016

4 Epping B. Routinedaten – Zusammen führen, was zusammen gehört! Z Orthop Unfall 2011; 149: 121–124

5 http://www.qualitaetssicherung-mit-routinedaten.de/patienten/ Stand: 09.05.2016

6 https://www.g-ba.de/institution/themenschwerpunkte/qualitaetssicherung/sektorenuebergreifend/ Stand: 09.05.2016

7 https://www.aqua-institut.de/de/aktuelles/komplikation_follow-up_knie-huefte.html Stand: 09.05.2016

Bernhard Epping

Qualitätssicherung

„Mit Lebensbäumen à la QSR kann man Krankenhausqualität nicht bewerten."

Im ZfOU-Interview erläutert der Mitautor einer aktuellen Studie, warum die Ergebnisse bei QSR nicht zuverlässig sind. Vor allem der Einfluss der Fallzahlen auf die Ergebnisse einzelner Häuser lege nahe, diese Daten zumindest nicht in Form eines Krankenhausnavigators zu veröffentlichen.

Burkhard Fischer ist Leiter des Referats Qualitätsmanagement, IT und Datenanalyse bei der Krankenhausgesellschaft Nordrhein-Westfalen (KGNW) in Düsseldorf. Der Medizinische Informatiker ist Mitautor einer aktuellen Publikation, die Methoden des QSR-Verfahrens hinterfragt und kritisiert.

? Sie haben sich in Ihrer neuen Publikation mit QSR befasst [1]. Dabei formulieren Sie eine Kritik an dem Verfahren auf mehreren Ebenen. So monieren Sie einmal die Zuweisung von Daten nach längerem Follow-up zu einem Krankenhaus, in dem die Erstoperation stattfand, warum?

Wir haben uns die QSR-Daten zu 940 Krankenhäusern und –verbünden angeschaut. Und dies zunächst für die Indikation Implantation eines künstlichen Hüftgelenks nach einem hüftgelenksnahem Bruch. QSR nennt da unter anderem die Häufigkeit von Komplikationen bis zu 365 Tage nach der Implantation des künstlichen Hüftgelenks. Komplikationen, die nach so langer Zeit bei einem Patienten, lange nach Verlassen des Krankenhauses entstehen, gehen aber womöglich gar nicht auf das Verschulden des Krankenhauses zurück. Das kann man nicht eins zu eins dem Klinikum zuordnen.

? Sollte man solche Daten nicht erheben, da sie keine Aussagekraft haben?

Nein, darum geht es nicht. Wir brauchen solche Daten natürlich. Was ist 30 Tage später, 90 Tage, was ist 365 Tage nach einer Operation mit dem Patienten? Das sind sehr wohl wichtige Informationen für die Krankenhäuser, denn letztendlich machen wir Medizin, damit es den Menschen gut geht und nicht nur dafür, dass es ihnen gut geht zu dem Zeitpunkt, zu dem sie aus dem Krankenhaus entlassen werden.

? Also ist der Ansatz von QSR da doch richtig?

Nein – wir teilen die Art der Auswertung bei QSR nicht. Das langfristige Ergebnis einer Operation hängt wesentlich von Faktoren ab, die ein Krankenhaus, in dem die Operation stattfand, nicht mehr steuert. Etwa, wie gut jemand danach betreut wird, wie gut sein Gesundheitszustand insgesamt ist und wie engagiert er selbst ist. Und mal angenommen, ein Patient verstirbt binnen eines Jahres an einem Autounfall – dann wäre das nach dieser Art von Darstellung über Routinedaten auch ein Tod nach Hüftgelenksfraktur – hätte aber mit der Erstoperation wohl kaum etwas zu tun.

Wir nennen das in der Statistik die Attribuierbarkeit eines Ereignisses in Bezug auf die mögliche Ursache, die Zuschreibbarkeit. Und da sind nach wie vor viele

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

Fragen offen. Man kann außerdem ja kaum davon ausgehen, dass jedes Krankenhaus die gleiche Patientenklientel hat, die immer ähnlich gesundheitsbewusst lebt, sobald sie aus dem Krankenhaus entlassen ist. Da gibt es Unterschiede im sozialen Umfeld und vieles mehr.

? Das WIdO erklärt aber, dass es Dinge wie Begleiterkrankungen, Alter des Patienten bei seiner Kalkulation der SMR adjustiert – eine so genannte Risikoadjustierung.

Das sozioökonomische Umfeld der Patienten geht in den SMR aber nicht ein. Es werden ausschließlich zu anderen Zwecken dokumentierte medizinische und demographische Daten genutzt. Das Verfahren halten wir aus diesem und weiteren Gründen aber nicht für ausreichend. Wir sprechen hier ja von den Abrechnungsdaten der Krankenhäuser. Und in denen fehlen oft ganz entscheidende Angaben zu Begleiterkrankungen. Wenn ein Krankenhaus eine DRG-Ziffer abrechnet, muss es bestimmte Diagnosen, die Haupt- und die Nebendiagnose liefern. Was dabei konkret eingetragen wird, richtet sich in der Praxis oft danach, welche DRG zur Abrechnung möglich ist. Fehlt in den Abrechnungsdaten die Angabe eines essentiellen Bluthochdrucks oder eines Diabetes mellitus, ist das keine Garantie, dass dieser nicht vorgelegen hat, sondern lediglich, dass er für die akute Behandlung und deren Abrechnung keine Relevanz hat. Die Datengrundlage ist daher für eine saubere Risikoadjustierung nur mit großen Einschränkungen geeignet. Ich bin absolut dafür, dass man diese Abrechnungsdaten nutzt, um Orientierungswerte zu generieren, damit man weiß, wo man steht. Aber man muss dabei immer wissen, wo die Unsicherheiten stecken.

? Ein Beispiel?

Bei den Daten zum Herzkatheter, der PCI, nutzt QSR Angaben zur Risikoadjustierung, ob ein Patient einen Bluthochdruck hat oder nicht. Und das Ergebnis bei QSR ist, dass Bluthochdruck in Bezug auf die Überlebenswahrscheinlichkeit bei einer Herzkatheteruntersuchung oder mit Einsetzen eines Stents nach Herzinfarkt protektiv ist.

? Patienten mit Hochdruck überleben solche Eingriffe eher?

Ja. Da steckt allerdings vermutlich ein krasser Dokumentationsbias dahinter. Denn für jene Patienten, die tragischerweise während solcher Untersuchungen versterben, für die wird ein womöglich bestehender Bluthochdruck erst gar nicht mehr in den Abrechnungsdaten dokumentiert. Abrechnungsdaten haben gar nicht das Ziel, solche Dinge noch eigens zu dokumentieren.

? Dann müssten solche Werte aus der Statistik fallen?

Ja, sie bleiben aber bei QSR drinnen. Ein anderes Beispiel betrifft Diabetes bei Komplikationen innerhalb von 30 Tagen. Da ist dann plötzlich ein Diabetes protektiv. Was aus medizinischer Sicht einfach Quatsch ist. Auch da gibt es wieder den Punkt, dass diese Begleiterkrankungen gar nicht mehr in den Abrechnungsdaten der Krankenhäuser erfasst werden.

? Kurzum, Sie halten die Risikoadjustierung bei QSR für nicht valide?

Ja. Wenn dann noch das Abschneiden aus den Einzelindikatoren in eine Gesamtnote für ein Krankenhaus fließt – die 1–3 Lebensbäume, an dieser Stelle weiß man spätestens dann nicht mehr, welche Unsicherheiten noch alles darin stecken.

? Das läuft auf die Forderung hinaus, eine Auswertung wie bei QSR kann man vielleicht für interne Diskussionen machen, die Werte sind für uns Krankenhäuser interessant – wir wollen aber nicht, dass sie veröffentlicht werden, weil viele der Ergebnisse nicht eindeutig sind?

So ist es. Komplikationsraten bis 365 Tage nach der Erstimplantation eines Kunstgelenks zu zählen, das ist viel zu lang, als dass man dafür allein das operierende Haus verantwortlich machen könnte. Die objektive Qualitätsmessung zu veröffentlichen und nicht dazu zu schreiben, dass die Daten mit Vorsicht zu interpretieren sind, ist verantwortungslos und irreführend.

AOK und WIdO generieren ja Klinikberichte, die den Häusern - meist erst auf Anfrage - zur Verfügung gestellt werden. Das ist durchaus wertvoll, wenngleich auch da wieder nicht mit dieser Absolutheit, die im Anspruch von WIdO und AOK zum Ausdruck gebracht wird. Jeder Qualitätsindikator gibt Hinweise bei einer Analyse auf Verbesserungsmöglichkeiten. Es misst aber nicht Qualität. Schon gar nicht die Qualität in der Medizin.

Und aufsetzend auf diesen schon für sich unreifen Datensätzen, gibt es dann im AOK-Krankenhausnavigator eben noch das Lebensbaumkonzept, das in sich problematisch ist. Unser Artikel setzt sich daher im Wesentlichen damit auseinander.

? Dabei stören Sie sich vor allem am Verfahren des WIdOs, die Qualität der Häuser nach den oberen und unteren Grenzwerten des so genannten 95 %-Konfidenzintervalls zu berechnen. Können Sie das etwas erläutern?

Wir kritisieren nicht die grundsätzliche Kalkulation eines Konfidenzintervalls um die SMR-Werte. Da gibt es in der Wissenschaft zwar unterschiedliche Methoden, keine davon ist wirklich genau, aber da steckt nicht unsere Kritik.

Wir kritisieren vielmehr den Einsatz des Konfidenzintervalls zum Vergleich der Krankenhäuser.

? Was meinen Sie?

Nehmen wir ein Beispiel. Angenommen ein Haus landet bei einem beliebigem Qualitätsindikator bei einem SMR von 1,2 mit einem 95 %-Konfidenzintervall von 0,8–1,6, dann weicht das Haus, wenn man die Werte an sich betrachtet, statistisch nicht signifikant vom Durchschnitt aller Häuser ab. Der Durchschnitt liegt bei 1,0 – und der wird vom Konfidenzintervall ja hier mit umfasst. Wenn Sie einen p-Wert von 0,05 ansetzen, wäre der Unterschied zu einem anderen Haus mit SMR von 1,0 nicht signifikant.

? Die beiden Häuser unterscheiden sich also statistisch gesehen bei diesem Qualitätsparameter gar nicht?

So ist es.

? Dabei spielt auch das 95 %-Konfidenzintervall des zweiten Hauses keine Rolle mehr?

Ja. Aber das WIdO nimmt jetzt die Grenzen der jeweiligen Konfidenzintervalle und nutzt sie, um doch ein Ranking aufzumachen. Und das geht gar nicht.

? Das WIdO reiht ein Haus zu der Gruppe jenes Fünftels an Krankenhäusern mit vergleichsweise guter Qualität, wenn es einen relativ niedrigen oberen Grenzwerte für das 95 %-Konfidenzintervall hat.

Ja, und wenn das Intervall bei hoher Fallzahl schmal ist, und der SMR in der Nähe von 1,0 liegt, kann das dazu führen, dass ein Haus positiv bewertet wird, zugleich aber einen hohen unteren Grenzwert auf-

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

weist, was wiederum zur Abwertung führen wird. Mehr noch – es ist bei diesem seltsamen Verfahren sogar denkbar, dass ein Haus sowohl in der Gruppe Plus als auch bei Minus landet. Also gleich zweimal in den beiden Qualitätsgruppen auftaucht.

? Haben Sie solch einen Fall gefunden?

Nein. Auf jeden Fall liefert das Verfahren aber Werte, die von der Fallzahl des Krankenhauses abhängen.

? Was meinen Sie?

Es benachteiligt eher Kliniken mit hohen Fallzahlen. Bei ihnen wird das Konfidenzintervall vergleichsweise schmal ausfallen. Je größer die Klinik, je mehr Fälle sie hat, desto kleiner ist dieses Intervall.

? Was ja im Prinzip ein gutes Zeichen ist, die Werte erlauben eine zuverlässigere Prognose.

Ja, aber aus statistischen Gründen führt das WIdO-Verfahren mit den Konfidenzintervallen dazu, dass die Häuser mit hoher Fallzahl unverhältnismäßig oft extrem bewertet werden – also 1 oder gleich 3 Lebensbäume erhalten.

? Wie haben Sie das gezeigt?

Wir haben dafür einmal durchgerechnet, wie die Gesamtnote eines Hauses mit dem mit Abstand wichtigsten Parameter für Qualität steht, nämlich der 90-Tages-Mortalität nach der Implantation einer Hüftprothese nach hüftnahem Oberschenkelbruch. Das ist eine Indikation mit einer hohen Sterblichkeit, sie liegt bei 14,4%. Wir haben die 3 Gruppen bei den Lebensbäumen genommen und in so genannten Kontingenztafeln geschaut, wo diese Häuser jeweils beim Abschneiden für die Sterblichkeit bis 90 Tage nach der Operation stehen. Dafür haben wir die SMR-Werte genommen. Einmal SMR-Werte unter 0,8, dann zwischen 0,8 und 1,2 und dann über 1,2.

? Sie argumentieren hierbei, dass die Sterblichkeit der wichtigste Qualitätsindikator überhaupt ist und sich ein gutes Abschneiden dabei auch in der Gesamtnote widerspiegeln müsste?

Ja.

? Wobei in die Kalkulation der Lebensbäume auch noch 2 weitere Indikatoren eingehen. Das Abschneiden beim Indikator 365-Tage-Revisionsrate und die Komplikationsrate binnen 90 Tagen.

Schon richtig. Aber wir finden auf jeden Fall heraus, dass längst nicht alle Häuser, die bei der Sterblichkeit sehr gut abschneiden, dann auch in der Gesamtbilanz 3 Lebensbäume erhalten. Und auch hier deutet sich eben an, dass die Kategorisierung nach Lebensbäume, die eine Klinik erhält, mehr mit der Fallzahl zu tun haben könnte, als mit der medizinischen Qualität. Häuser mit großen Fallzahlen landen überproportional oft in der Gruppe mit nur 1 Lebensbaum oder mit gleich 3 Bäumen. Wir lehnen diese Art der Auswertung zumindest bei den Lebensbaumbewertungen ab.

? Was sollte man stattdessen rechnen?

In Ordnung wäre es, wenn Sie den SMR-Wert für eine Qualitätseinschätzung nehmen. Und das 95%-Konfidenzintervall als das verwenden, was es ist: Als Maß der Unsicherheit. Es gibt den Bereich an, in dem der wahre Mittelwert des SMR noch liegen könnte. Das kann man, so wie es ist, darstellen.

? Am Ende sollte bei einer Darstellung auch nach QSR der SMR-Wert im Vordergrund stehen, wenn man Kliniken vergleichen möchte.

Richtig. Natürlich kann der Patient seine Entscheidung aus einer Risikoperspektive treffen. Wer ein Krankenhaus sucht, kann sich durchaus überlegen – bin ich bereit, ein relativ großes Risiko einzugehen, um ein relativ wahrscheinliches gutes Ergebnis zu erzielen? Anders gesagt – nehme ich ein Krankenhaus, das einen niedrigen SMR-Wert hat, auch dann, wenn das Konfidenzintervall dazu groß ist? So macht die Betrachtung eines Konfidenzintervalls Sinn. Als Maß der Unsicherheit, die hinter dem SMR-Wert steckt.

? Was schlagen Sie vor, um Qualität von Krankenhäusern zu messen?

Unser besseres Modell ist das, was wir gemeinsam mit den Krankenkassen im G-BA beschlossen haben, nämlich das, was BQS und AQUA entwickelt haben und das IQTiG jetzt weiterentwickeln soll. Dazu brauchen wir sicher eine breite fachliche Debatte. Zumal jetzt mit dem neuen Krankenhausstrukturgesetz, Qualitätsindikatoren auch für die Krankenhausplanung herangezogen werden sollen. Wichtig ist es dafür, Erhebungsdaten aus der gesetzlichen Qualitätssicherung zu nutzen.

? Also keine Routinedaten? Viele Ärzte monieren, dass der Dokumentationsaufwand für diese gesetzliche Qualitätssicherung schon heute hoch bis zu hoch ist, auf keinen Fall noch steigen darf.

Das ist durchaus eine legitime Forderung. Vielleicht können wir da auch Routinedaten zur Entlastung mitnützen. Nur müssen wir uns dann überlegen, ob wir bei ICD- und OPS-Dokumentationen nicht eben bestimmte weitere Angaben verpflichtend sein müssen. Wir haben ja eben über die Beispiele von Diabetes und Bluthochdruck gesprochen, wo die Daten einfach fehlen. Ich habe schon vor 5 Jahren gesagt, wir sollten die AOKen davon überzeugen, dass sie ihre Indikatoren an eine neutrale Stelle, heute etwa das IQTiG, geben, damit die dort auf Herz und Nieren geprüft und weiterentwickelt werden.

? Gegen einen Wettbewerb der Häuser um Qualität sind Sie nicht?

Wir sind unbedingt dafür. Aber je stringentere Entscheidungen auf der Basis von Qualitätsindikatoren getroffen werden, demnächst sogar Entscheidungen zur Krankenhausplanung, desto wichtiger ist, erst die zugrunde liegenden Indikatoren sauber zu prüfen.

Das Interview führte Bernhard Epping

Weitere Informationen

Positionen der DKG zur Qualitätsentwicklung:

http://www.dkgev.de/dkg.php/cat/38/aid/11836/title/DKG_stellt_Positionen_zur_Qualitaetssicherung_im_Krankenhaus_vor

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

Qualitätssicherung

„QSR zeigt Wege für die gesetzliche Qualitätssicherung“

Im ZfOU-Interview bewertet der Leiter der zuständigen Gruppe im WIdO Christian Günster eine aktuelle Kritik an der Methodik hinter QSR (Qualitätssicherung mit Routinedaten). Das Verfahren sei vielmehr so ausgereift, dass es Blaupausen für die Nutzung von Routinedaten in der gesetzlichen Qualitätssicherung liefert.

Christian Günster ist Leiter des Bereichs Qualitäts- und Versorgungsforschung beim Wissenschaftlichen Institut der AOK (WIdO) in Berlin. Die Gruppe des Diplom-Mathematikers betreut das Projekt Qualitätssicherung mit Routinedaten (QSR), das seit 2008 Qualitätsdaten für Kliniken aufbereitet.

? Was sagen Sie zur neuen Kritik an der Methodik von QSR.

Die weisen wir als methodisch fehlerhaft zurück. Lassen Sie mich das im Einzelnen erläutern

? Könnten Sie uns zunächst erklären, wie Sie bei QSR vorgehen?

Grundlage sind die Abrechnungsdaten, die die einzelnen AOKen von den Kliniken erhalten und danach in anonymisierter Form an das WIdO übermitteln. Mit diesen Längsschnittdaten können wir damit über viele Jahre hinweg schauen, welche Versorgung unsere Versicherten wo und wie erhalten haben.

? Wie lange heben Sie Routinedaten auf? 10 Jahre?

Sogar länger. Bei den von uns veröffentlichten Daten aus QSR werden Sie aber immer nur die Auswertung der anonymisierten Daten zu einem Haus über einen Zeitraum von 3 Jahren finden. Dabei werden die Ergebnisse jährlich aktualisiert. Im Moment sind in QSR die Daten von 2011 bis 2013 mit Follow-up bis 2014 enthalten. Wir haben ja Indikatoren dabei, die den Zeitraum bis zu 1 Jahr nach einem Eingriff betrachten.

? Die Autoren besagter Publikation kritisieren Ihre Verfahren auf mehreren Ebenen. So sei die Zuordnung von Daten aus einem Follow-up noch 1 Jahr nach der Erstoperation, etwa nach Implantation eines Kunstgelenks, nicht statthaft. Ein Krankenhaus könne längst nicht immer verantworten, was mit dem Patienten passiert, sobald er das Haus verlassen hat. Was sagen Sie dazu?

Bei QSR geht es nicht um individuelle Einzelfälle sondern um die statistische Häufung von Folgeereignissen. Uns beraten in solchen Fragen medizinische Fachgruppen, hier speziell das Expertenpanel Orthopädie und Unfallchirurgie [1]. Nach Einschätzung der Gruppe stehen Revisionen binnen eines Jahres sehr wohl oft im Zusammenhang mit der Erstoperation. Eine Infektion des Hüftgelenks zum Beispiel führt oft erst nach Monaten zur Gelenkrevision und zum Austausch der Prothese. Daher nehmen wir für diesen Indikator einen Zeitraum von 1 Jahr.

? Welchen Einfluss hat das Gremium der Experten generell?

Die Definition und Auswahl der Qualitätsindikatoren, die Risikoadjustierung und die Eignung der Indikatoren für die Berichterstattung sind Dinge, die ausdrücklich von den externen Mitgliedern in diesen Panels entschieden werden. Die Anwendung der Definitionen auf die Daten und deren statistische Aufbereitung inklusive der Klassifizierung in 3 Wertungskategorien (Lebensbaum-Systematik) übernimmt das WIdO.

? Und wie gehen Sie dazu vor?

Wir rechnen für jedes Haus und jeden Parameter zunächst einmal die so genannte Standard Morbiditäts-Rate oder Standard Mortalitäts-Rate (SMR) aus. Damit wollen wir erreichen, dass man die Werte in den Kliniken überhaupt vergleichen kann – denn die Patienten sind in jedem Haus unterschiedlich. Wir machen also eine so genannte Risikoadjustierung.

? Und das geht wie?

Wir kennen aus den Routinedaten die Zahl der tatsächlichen Fälle an Komplikationen in einem Haus und setzen die ins Verhältnis zu einer Zahl an Komplikationen, die angesichts des Risikoprofils der Patienten im Mittel zu erwarten sind. Der SMR-Wert kalkuliert sich dann durch den Bezug: Im Zähler stehen die beobachteten Werte und im Nenner die erwarteten. Ein SMR von 1 wäre damit genau der Durchschnitt, so viele Komplikationsfälle wie erwartet. Ein SMR von 2,0 heißt doppelt so viele Komplikationen oder Sterbefälle wie im Durchschnitt der Häuser. Und ein SMR von 0,5 heißt halb so viel wie erwartet.

? Wie kommen Sie auf die Zahlen für die zu erwarteten Fälle?

Dafür wird der komplette Datensatz aller deutschen Kliniken ausgewertet und dann in einer Regressionsmodellierung berechnet, welche Faktoren in welcher Stärke komplikationserhöhend wirken.

? Vermutlich Dinge wie Übergewicht, Vorerkrankungen?

Ja, ein Diabetes, Adipositas und andere Faktoren. Anhand der Daten zu solchen Risikofaktoren bei den Patienten können wir dann für jedes Krankenhaus hochrechnen, mit wie vielen Komplikationen anhand der Begleiterkrankungen der Patienten zu rechnen ist. Dieses Verfahren der Regressionsmodellierung ist international etabliert und verbreitet sich zunehmend auch in der gesetzlichen externen Qualitätssicherung. Es ist ein wesentlicher Schritt hin zu überhaupt vergleichbaren Daten.

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

? Sie nehmen ein Verfahren zur Risikoadjustierung nach Elixhauser – was heißt das?

Anne Elixhauser hat 1998 erstmals 31 Erkrankungen definiert, die sie für die Risikoadjustierung von Routinedaten verwendet. Diese Einteilung nehmen auch wir, haben dabei aber einige Dinge in Abstimmung mit unseren Expertenpanels aktualisiert und erweitert. Sie können alle Details auf unseren Seiten nachlesen (siehe Link unter Weitere Infos).

? Die Kritiker werfen Ihnen jetzt aber vor, dass Sie für den Vergleich der Kliniken gerade nicht mit den Werten für den SMR arbeiten, sondern mit denen des so genannten 95 %-Konfidenzintervalls. Was ist das?

Bei dem SMR handelt es sich um einen Schätzwert mit einer Streubreite. Daher gibt man ein Intervall an, in dem der tatsächliche Wert mit einer entsprechenden Wahrscheinlichkeit liegt. Das 95 %-Konfidenzintervall gibt an, dass der wahre Wert für einen SMR mit 95-prozentiger Wahrscheinlichkeit in diesem Intervall liegt. Eine ausschließliche Orientierung am SMR-Wert ohne Konfidenzintervall kann zu irreführenden Ergebnissen führen.

? Und wie entwickeln Sie die Einstufung von Kliniken nach ihrer Qualität?

Wir wollen, dass der Leser die Information in einer einfachen Form erhält. Denn nur so kann der Patient sie bei der Suche nach einem guten Krankenhaus nützen. Für jeden einzelnen Qualitätsindikator sortieren wir daher ein Haus in eine von drei Gruppen Plus, Minus oder Null. Die 20 % der Häuser, mit der geringsten Wahrscheinlichkeit für unerwünschte Ereignisse kriegen das Plus, 20 % für die die Wahrscheinlichkeit einer Komplikation hoch ist, kriegen das Minus, die 60 % in der Mitte kriegen die Null.

Wir nehmen für diese Kategorisierung die Lage des 95 %-Vertrauensbereichs. Und zwar mit folgender Grundüberlegung: Ein Patient, der bei uns die Daten von 1000 Kliniken vergleichen will, möchte wissen: Welches dieser Krankenhäuser kann mir mit 95-prozentiger Wahrscheinlichkeit nachweisen, dass die Komplikationsrate bei einem Eingriff, etwa der Implantation einer künstlichen Hüfte, niedrig ist. Nehmen Sie die Zahl der Revisionsoperationen binnen eines Jahres. Dafür will ich nicht wissen, wo der Punktschätzer SMR liegt, sondern ich möchte wissen, mit welcher Sicherheit ein beliebiges Haus auch wirklich einen niedrigen SMR-Wert hat. Wir betrachten dafür den oberen Grenzwert des Konfidenzintervalls. Je niedriger der liegt, desto besser. Die 20 % Kliniken, die hier die niedrigsten oberen Grenzwerte zeigen, kommen in die Kategorie Plus.

? Ein Beispiel?

Angenommen, für ein Haus findet sich bei ungeplanten Folgeoperationen binnen 365 Tagen nach dem Eingriff ein SMR von 0,8 und ein 95 %-Konfidenzintervall von 0,6–0,9. Damit zeigt sich für dieses Haus, dass das Ergebnis mit hoher Wahrscheinlichkeit unter 0,9 liegt, damit besser ist als der Durchschnitt aller Kliniken ist.

Und angenommen bei einem weiteren Haus mit dem gleichen SMR von 0,9 ginge das Intervall von 0,6–1,3. Da würde ich als Patient doch in jedem Fall in das Haus mit dem niedrigeren oberen Wert für das Konfidenzintervall von 0,9 gehen. Und eher nicht in das Haus mit dem Wert 1,3, das mir also mit 95-prozentiger Wahrscheinlichkeit nur sagen kann, das Risiko liegt unter 1,3. Der SMR könnte hier also auch bis zu 30 % über dem Durchschnitt aller Häuser liegen. Das sagt der Grenzwert von 1,3 ja aus.

? Und was gilt für jene 20 %, die in der Kategorie Minus landen?

Dafür schauen wir von der anderen Seite auf das Intervall. Wir nehmen die Häuser mit den höchsten unteren Grenzwerten für das 95 %-Konfidenzintervall. Sie haben die höchsten Risiken auf Komplikationen.

? Harald Englisch und Burkhard Fischer werfen Ihnen vor, dass mit dieser Arithmetik durchaus gute Häuser durch den Rost fallen könnten und anders herum. Dahinter stecke ein methodischer Fehler Ihres Verfahrens, das eben von den Fallzahlen abhänge. Das rechnen Sie anhand der Zahlen für die Sterblichkeit binnen 90 Tagen nach Implantation einer Hüftendoprothese nach hüftgelenksnahem Bruch vor.

Ja, und eine Seite später rechnen uns die Autoren dann vor, dass die Fallzahl bei den von ihnen ausgewerteten QSR-Daten zur Implantation eines künstlichen Hüftgelenks nach hüftgelenksnahem Bruch gar nicht mit dem Abschneiden beim Parameter Sterblichkeit binnen 90 Tagen in den Kliniken korreliert. Der Grund für diese Verwirrung, ich muss es so hart sagen: Die Arbeit der Autoren ist methodisch so nicht haltbar. Wir haben das nachgerechnet.

So möchten die Autoren Aussagen zu den AOK-Lebensbäumen machen, schauen sich dann aber nur Sterblichkeiten an. Das kann man aber gar nicht vergleichen – denn in die Bewertung nach Lebensbäumen gehen ja bei der gewählten Indikation 3 Einzelindikatoren ein. Das sind Chirurgische Komplikationen innerhalb von 90 Tagen, die Sterblichkeit innerhalb von 90 Tagen nach dem Eingriff, und ungeplante Folge-Operation bis zu 365 Tage nach dem Eingriff.

? Also hat Ihre Berechnungsweise bei QSR gar keine Makel?

Wir kämpfen mit einem grundlegenden statistischem Problem, das jeden trifft, der derart Daten für eine Qualitätsbewertung aufarbeiten möchte. Eine verlässliche Bewertung wird bei höherer Fallzahl einfacher möglich. Das drückt sich darin aus, dass das Konfidenzintervall um den SMR kleiner wird. Und deswegen haben Häuser mit einer höheren Fallzahl auch bei den Lebensbäumen eine höhere Wahrscheinlichkeit entweder in die 1er oder 3er-Kategorie einsortiert zu werden, eben weil die Qualitätseinstufung leichter möglich wird.

? Und Häuser mit niedrigeren Fallzahlen?

Da haben Sie immer das Problem, dass in diesen Häusern die Konfidenzintervalle tendenziell breiter sind. Das macht bei Häusern mit geringer Fallzahl aufgrund des Einflusses zufälliger Effekte die Auswahl schwieriger.

? Mit der Folge, dass sie die am Ende eher in die mittlere Kategorie sortieren müssen – da Sie es nicht genauer wissen?

Richtig. Und das ist gar nicht erstaunlich – dieser Punkt ist ein statistisches Problem, vor dem alle stehen, die vergleichende Qualitätsbetrachtungen machen möchten. Was man zugestehen muss, das ist aber eine statistische Trivialität, ist, dass die Konfidenzintervalle schmaler sind bei hoher Fallzahl.

? Also tun Sie Häusern mit hohen Fallzahlen eventuell doch Unrecht?

Nein. Gerade bei Kliniken mit hoher Fallzahl ist aufgrund der schmalen Konfidenzintervalle eine Qualitätseinstufung gut

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

möglich. Bei Kliniken mit geringer Fallzahl ist es hingegen aus statistischen Gründen viel schwieriger gute oder schlechte Behandlungsqualität zu erkennen. Wenn Kliniken weniger als 30 Patienten in einer Indikation in 3 Jahren behandeln, veröffentlichen wir diese Daten im Übrigen nicht. Unter 30 sind die Werte zu volatil.

? Fällt damit nicht gerade die Gruppe der Häuser durchs Raster, die AOK wie auch viele Orthopäden aufgrund kleiner Fallzahlen besonders kritisch beim Thema Qualität sehen?

Nein. Die Kliniken mit sehr kleinen Fallzahlen muss man methodisch anders auswerten. In Volume-Outcome-Analysen fasst man sie üblicherweise zu einer Gruppe zusammen. So konnten wir auch für Deutschland einen inversen Zusammenhang von Fallzahl und Komplikationsrate beim Gelenkersatz feststellen, also weniger Komplikationen bei höherer OP-Zahl [2].

? Warum nehmen Sie nicht einfach die SMR-Werte für Ihre Einstufung? Das ist eine Forderung Ihrer Kritiker.

Nein, das reicht eben nicht. Wie dargelegt ist das SMR als Punktschätzer alleine nicht ausreichend. Die statistische Sicherheit muss in Form der Konfidenzintervalle mit berücksichtigt werden.

? Wie ist die Zukunft? QSR nimmt für sich in Anspruch, Langzeitdaten zu erheben – und das über die Grenzen von stationär und ambulant hinweg. Kommt so etwas jetzt auch in der gesetzlichen Qualitätssicherung?

Das wünschen wir uns. Schon das AQUA-Institut hat ja im Bereich Herzkatheter, PCI, ein Verfahren vorgeschlagen, dass erstmals sektorübergreifend Daten erheben soll, einschließlich von Routinedaten. Das geht nach Beschluss des G-BA in diesem Jahr los [3]. Bei Hüftgelenk- und Knieersatz könnte der G-BA ebenfalls längst damit loslegen, leider geht das aber nach unserem Eindruck nicht voran.

? Was meinen Sie?

Der G-BA hatte AQUA zuletzt am 22.01.2015 beauftragt, zu prüfen, wie man in der Endoprothetik ein längerfristiges Follow-Up nach einem stationären Aufenthalt gestalten kann. In seinem seit 29.01.2016 vorliegenden Abschlussbericht hat das AQUA-Institut auch Verfahren dabei, die auf unsere QSR-Indikatoren und Konzepte aufsatteln. Als da sind „Sterblichkeit innerhalb von 90 Tagen“, Revisionsraten, Folge-OP (innerhalb von 1,2 und 5 Jahren) [4]. Und mehr noch: AQUA hatte derartige Indikatoren bereits 2012 und 2013 vorgelegt – die eine hohe Übereinstimmung mit den damaligen QSR-Definitionen hatten. Der neue Bericht ist da nur eine Konkretisierung.

? Und jetzt?

Es geht nicht voran. Der G-BA muss beim Institut für Qualitätssicherung und Transparenz im Gesundheitswesen, beim IQTiG, die Umsetzung beauftragen. Da kämen dann übrigens die Routinedaten aller Kassen zum Tragen, was wir ebenfalls begrüßen würden.

? Man hört schon mal Befürchtungen, Sie wollten mit QSR einen Endlosprozess anstoßen. Man nehme die 20 % schlechtesten Häuser raus aus der Versorgung, werte danach die verbleibenden 80 % neu aus, und lege dann wieder die 20 % schlechtesten still?

Mit Verlaub: Das ist ein theoretisches Gedankenspiel. Die QSR-Daten informieren Ärzte und Patienten über potentielle Komplikationen im Behandlungsverlauf bis zu einem Jahr. Sie geben Orientierung im Endlosprozess, Qualität zu sichern und zu verbessern. Die Unterschiede sind heute noch erheblich. Man muss sich vergegenwärtigen, dass die Revisionsraten zwischen dem Viertel der Kliniken mit den niedrigsten Raten und dem Viertel mit den höchsten um den Faktor 2 differieren. Wenn Komplikationen reduziert werden und sich überall auf ähnlichem Niveau bewegen, dann ist das Ziel erreicht und die QSR-Einstufung hat sich überflüssig gemacht. Das ist meine Wunschvorstellung.

Literatur

1 http://www.qualitaetssicherung-mit-routinedaten.de/entwicklung/experten/index_12615.html
2 Jeschke E, Günster C. Zum Zusammenhang zwischen Behandlungshäufigkeit und -ergebnis in der Hüftendoprothetik. In: Klauber J, Geraedts M, Friedrich J, Wasem J, Hrsg. Krankenhaus-Report 2014. Schwerpunkt: Patientensicherheit. Stuttgart 2014: Schattauer, 227-240
3 https://www.g-ba.de/institution/presse/pressemitteilungen/570/
4 https://sqg.aqua-institut.de/aktuelles/pressemitteilungen/komplikation_follow-up_knie-huefte.html

Das Interview führte Bernhard Epping

Weitere Informationen

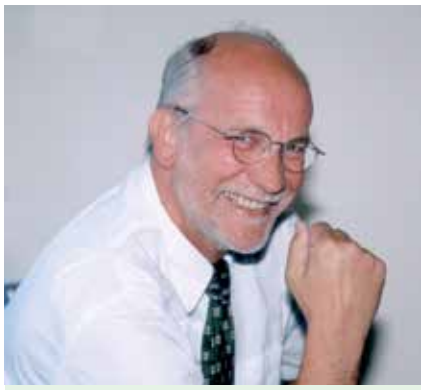
http://www.qualitaetssicherung-mit-routinedaten.de/

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

Validierung

„QSR-Verfahren muss validiert werden“

Professor Hans-Konrad Selbmann moniert den zu schleppenden Prozess einer Validierung von Systemen zur Qualitätsmessung – auch und gerade beim System der AOK, der Qualitätssicherung mit Routinedaten (QSR).



Prof. Hans-Konrad Selbmann (Jahrgang 1941) war bis zu seiner Emeritierung 2007 Leiter des Instituts für Medizinische Informationsverarbeitung der Universität Tübingen. Selbmann, aktuell als Schatzmeister und Stellungnahmebeauftragter im Präsidium der AWMF, ist einer der Protagonisten für statistische Verfahren zur Qualitätsbewertung in der Medizin, seine Gruppe publizierte bereits Anfang der 1990er Jahre Grundzüge eines Erhebungssystems zur Qualitätsbewertung – zunächst in der Perinatalmedizin und der Chirurgie.

? Eine neue Publikation in der Zeitschrift Gesundheitsökonomie & Qualitätsmanagement kritisiert die Methodik bei QSR. Ein Hauptstreitpunkt scheint die Frage, ob man die 95 %-Konfindenzintervalle für die Einteilung der Kliniken in 3 Qualitätsgruppen verwenden darf oder nicht. Was sagen Sie dazu?

Ich darf mal mit meinen Interessenskonflikten beginnen. Ich bin unter anderem Mitherausgeber der Zeitschrift Gesundheitsökonomie & Qualitätsmanagement, aus der die neue Arbeit stammt. Die Autoren sind mir bekannt. Ich habe aber am Reviewing des Manuskripts nicht mitgewirkt.

? Danke, wir befragen Sie als unabhängigen Experten – wie bewerten Sie nun den Konflikt?

Ich hatte von Anfang an Probleme mit den Qualitätsindikatoren von QSR. Faszinierend an den Routinedaten ist, dass diese Daten eh schon da sind. Andererseits ist die Validität der Qualitätsindikatoren aus Routinedaten trotz mehrfacher Aufforderung bis heute nicht belegt. Das gilt sowohl für die Daten selbst als auch für ihre Interpretation – vor allem die Ursachenzuweisung von postoperativen Ereignissen! Meines Wissens legen zudem die verschiedenen Kassenarten immer noch nicht ihre Daten zusammen. Das wäre aber wichtig, damit wir wegkommen von einer reinen AOK-Aus- und Bewertung.

? Was meinen Sie mit „Validität nicht belegt“?

Das statistische Problem der Qualitätsindikatoren bei QSR ist die unterschiedliche Fallzahl von Haus zu Haus. Bei kleinen Fallzahlen sind die Konfidenzintervalle – wenn man überhaupt annimmt, dass es sich um Stichproben handelt – so groß, dass man keine Unterschiede erkennen kann. Bei großen Fallzahlen sind sie wiederum so klein, dass auch kleine Unterschiede auffällig werden. Deshalb versucht QSR die Grenzen der Konfidenzintervalle ins Spiel zu bringen. Dies geschieht bei QSR eher pragmatisch als anhand einer statistischen Theorie. Das gilt auch für die Kombination von mehreren Qualitätsindikatoren zu einem Gesamtqualitätsindikator in den Lebensbäumen. Ich kenne keine Arbeit, die einmal die Sicherheit und die Richtigkeit der so entstanden Indikatoren untersucht hat. Die neue Arbeit von Harald Englisch und Burkhard Fischer zeigt nun, dass das pragmatische Verfahren des AOK-Navigators bestimmte Fallzahlklassen bevorzugt. Dies sollte bei einem Qualitätsindikator natürlich nicht sein.

? Das WIdO argumentiert allerdings damit, eine Darstellung der Qualität rein über die SMR-Werte würde zu viel größeren Fehlern führen – der Punktschätzer SMR könne schließlich ganz falsch liegen.

Was die Standard Mortalitäts- oder Morbiditäts-Ratio, alias SMR, anbelangt, finde ich zunächst mal eine andere Erkenntnis von Englisch und Fischer richtig – die unterschiedliche Zusammensetzung der Patientenkollektive der Kliniken kann auch zu unterschiedlichen Ausprägungen der Qualitätsindikatoren führen. Denn man braucht zur Standardisierung, quasi zur statistischen Gleichmachung der Zusammensetzung der Kollektive, Daten, die man in der Regel eben nicht in den Routinedaten findet. Das WIdO verweist zwar auf die Durchführung einer Regressionsanalyse – dabei bleibt aber die Frage offen, welche und wie die Daten zur Standardisierung dabei eingesetzt werden

? Sie meinen die Risikoadjustierung bei QSR ist unklar?

Ja.

? Das WIdO erklärt, Dinge wie Übergewicht, Diabetes etc. würden genau adjustiert – das Verfahren ist auch nachzulesen [1].

Ja aber ich bin sicher, dass in den Routinedaten etliche Angaben, die für die Vergleichbarkeit der Kollektive wichtig sind, nicht vorliegen. Die Alternative wäre ein Erfassungssystem – wie wir es schon 1992 diskutiert haben, mit der Perinatalstudie [2]. Man kommt nicht umhin, besondere Daten zu den Risiken erheben. Und, ja – ich höre schon die Gegenargumente – das schafft wiederum Erhebungsaufwand, den die Kassen ungern bezahlen möchten und deswegen beharren sie auf den Routinedaten. Natürlich können Sie den Vitalstatus eines Patienten mit Routinedaten nachverfolgen, lebt der Patient noch oder nicht? Aber den Kausalzusammenhang kriegt man damit nicht heraus. Wenn der Patient 1 Jahr nach einer OP nicht mehr lebt, kann man nicht einfach behaupten, dass sein Tod auf die schlechte Operation zurückzuführen ist. Er konnte ja auch einen Unfall gehabt haben.

? Aber in einem Erfassungssystem – wenn Sie da alle Risikofaktoren kriegen wollen, um über Kausalzusammenhänge nachzuden-

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.

ken – dann sind Sie am Ende wieder auf Studienniveau, oder?

Absolut. Und daher sehe ich da nur eine Lösung: Bei wem alle Daten zusammenkommen, ist der Patient. Man muss daher den Patienten mit einbeziehen und darf nicht hergehen und ungeprüft aus Routinedaten irgendwas ablesen. Der Patient kann Ihnen sagen, in welchen stationären Einrichtungen er war, wie er wo mit welchem Erfolg versorgt wurde.

? Sollte QSR Ihrer Ansicht nach mangels Validität besser runter vom Netz, keine Veröffentlichung im AOK-Krankenhausnavigator?

Nein, aber man sollte mehr auf mögliche Fallstricke hinweisen. Mit dem Baumschema habe ich meine Schwierigkeiten.

? Internetnutzer wollen aber eine klare Botschaft, zum Beispiel mit 1-3 Bäumen. Nehme ich dann als Patient ein Haus mit den 3 Bäumen, habe ich ein Gefühl von Sicherheit.

Das kann ich nachvollziehen, da wollen wir alle hin. Nur müssen diese Bäume auch stimmen. Sie könnten für viele Krankenhäusern durchaus stimmen, aber eben nicht für alle.

? Das heißt im Umkehrschluss, bei den meisten Häusern kann man davon ausgehen, dass das QSR-Ergebnis in Ordnung ist?

So würde ich das sagen. Patienten rate ich, über ihre Suche auch mit dem Arzt zu reden. Ich finde es eindrucksvoll, mit welcher Sicherheit Ärzte wissen, wer in welchem Krankenhaus gut arbeitet und wer vielleicht etwas weniger gut.

Es ist aber ganz wichtig, QSR endlich zu validieren – ich appelliere an Forscherinnen und Forscher dafür jetzt Anträge beim Innovationsfonds zu stellen.

? Was meinen Sie?

Der Innovationsfonds beim G-BA schüttet jedes Jahr 300 Millionen Euro aus. 225 Millionen für neue Versorgungsformen. und 75 Millionen für Versorgungsforschung. Und unter der Rubrik Versorgungsforschung lassen sich auch Anträge mit dem Thema Routinedaten stellen. 75 Millionen sind unglaublich viel Geld – liebe Wissenschaftler stellt Anträge. Den Link zur Ausschreibung finden Sie auch über die AWMF-Homepage unter AWMF aktuell zu finden [3].

? Was für Studien müssen das sein? Es gibt eine neue Studie, in der immerhin gezeigt wird, dass Routinedaten in der Endoprothetik auch die Daten aus der Gesetzlichen Qualitätssicherung gut ergänzen könnten [4].

Das ist eine statistische Aussage, die nicht für den Einzelfall gilt. Validierung heißt unter anderem, jede Patientenakte zu prüfen, ob ein Haus, das bei QSR schlecht abschnitt, auch tatsächlich schlechtere Qualität geliefert hat.

? Das AQUA-Institut hat ein Verfahren zur Qualitätssicherung bei Knie- und Hüftgelenken präsentiert, nebst Verwendung von Routinedaten der Kassen [5]. Ist damit nicht ein Verfahren nach QSR offiziell gebilligt, um den Patienten auch für längere Zeit nachzuverfolgen und Daten zur Qualität der Versorgung auch noch bis zu 5 Jahre nach der Operation zu gewinnen? Trotzdem macht das bisher nur die AOK?

Aus meiner Sicht schwelt dahinter von Anfang ein nach wie vor existierender Streit: Routinedaten gegen Spezialerfassung. Die AOK war vor über 10 Jahren die erste Kassenart, die Qualitätsindikatoren aus Routinedaten forderte und dies zu ihrem Alleinstellungsmerkmal machte. Es kam in der Folge zu kassenpolitischen Streitgesprächen, was man auch daran sieht, dass die anderen Kassenarten nicht mitgezogen haben.

? Wieso eigentlich nicht?

Naja, mit diesen Daten wird auch ein Vergleich von Kassenarten möglich. Damit könnten sich die Kassen gegenseitig vorrechnen, wie gut sie sind?

? Es steht längst im Gesetz, dem SGBV, dass diese Sozialdaten der Kassen zusammen geführt werden sollen. Passiert ist das nicht.

Da kann die Politik lange reden, wenn die Selbstverwaltung es wegen Qualitätsbedenken nicht umsetzt. Aktuell beschäftigt uns aus diesem Themenkreis ein neuer Auftrag, den der G-BA soeben an das IQTiG, gegeben hat. Es soll aus den bereits vorhandenen Daten Indikatoren entwickeln, die für die Krankenhausplanung der Länder verwendet werden können. Das Institut arbeitet seit Mitte März daran. Auch wir von der AWMF sollen zu diesen Indikatoren Stellung nehmen. Ich bin Stellungnahmebeauftragter in der AWMF und erlebe gerade, wie sehr da aufs Tempo gedrückt wird. Das IQTiG hat bereits jetzt zu diesem Auftrag ein Stellungnahmeverfahren gestartet. Dabei ist das eigentliche Produkt noch gar nicht fertig. Die größte Sorge bereitet mir dabei die Beschränkung auf die vorhandenen Daten.

Literatur

1 http://www.qualitaetssicherung-mit-routinedaten.de/methoden/
2 Pietsch-Breitfeld B, Selbmann HK. Qualitätssicherung am Beispiel der Perinatologie und Chirurgie. Z Orthop Unfall 1992; 130: 352–356
3 http://www.awmf.org/die-awmf/awmf-aktuell/detail/news/antraege-fuer-projekte-aus-innovationsfonds.html
4 Wessling M, Gravius S, Gebert C et al. Ergebnisqualität in der Revisionsendoprothetik: Eine Analyse von Routinedaten mit dem Vergleich zur externen Qualitätssicherung. Z Orthop Unfall 2016; 154: 63–71
5 https://www.aqua-institut.de/de/aktuelles/komplikation_follow-up_knie-huefte.html

Das Interview führte Bernhard Epping

Dieses Dokument wurde zum persönlichen Gebrauch heruntergeladen. Vervielfältigung nur mit Zustimmung des Verlages.