



Stellungnahme der Bundesärztekammer

zum IQTIG-Bericht

Weiterentwicklung des Strukturierten Dialogs mit Krankenhäusern

Vorbericht zu Stufe 1 und Stufe 2

Berlin, 21.11.2019

Korrespondenzadresse:

Bundesärztekammer
Herbert-Lewin-Platz 1
10623 Berlin

1. Vorbemerkung

Der Bericht beinhaltet gemäß Auftrag des G-BA vom 18.02.2018 Weiterentwicklungsvorschläge zum Strukturierten Dialog. Da mittelfristig alle Leistungsbereiche der externen stationären Qualitätssicherung (ESQS) aus der QSKH-RL in die DeQS-RL übergehen werden, sollen die Empfehlungen des Berichts trotz des anderslautenden Titels auch bzw. vor allem für zukünftige Stellungnahmeverfahren in der DeQS-RL gelten. Demzufolge gelten auch die Anmerkungen der Bundesärztekammer sowohl für den Strukturierten Dialog als auch für das Stellungnahmeverfahren.

2. Inhalt des Berichts

In einer Bestandsaufnahme wird die Ausgangssituation beim Strukturierten Dialog gemäß QSKH-RL analysiert (Kapitel 2). Die Heterogenität der Vorgehensweise der Landesgeschäftsstellen Qualitätssicherung (LQS) wird beleuchtet, z. B. bezüglich der Verwendung alleiniger Hinweise bei rechnerischen Auffälligkeiten. Betrachtungen zur Effizienz des Verfahrens zeigen, dass sich vor allem bei Einrichtungen mit niedrigen Fallzahlen aus einer großen Zahl von rechnerischen Auffälligkeiten häufig nur eine kleine Zahl von tatsächlichen qualitativen Auffälligkeiten ergibt.

In Kapitel 3 wird der gegenwärtig in den LQS durchgeführte Strukturierte Dialog unter methodischen Überlegungen zur Güte von Bewertungsprozessen betrachtet. Das Institut kritisiert einerseits, dass die auf Landesebene eingebundenen Expertinnen und Experten Entscheidungen zur qualitativen Auffälligkeit lediglich nach impliziten Bewertungsregeln vornehmen würden. Andererseits werde versucht, nicht nur die Qualitätsindikatorergebnisse, sondern auch die Behandlungsprozesse zu bewerten. Dies könne aber nicht die Aufgabe des Stellungnahmeverfahrens sein (Seite 56). Würden Qualitätsindikatoren lediglich als Aufgreifkriterium genutzt, sei zur Bewertung der Gesamtqualität ein separates von der indikatorbasierten Qualitätsmessung unabhängiges Verfahren notwendig (Seite 57). Zudem wird kritisiert, dass Elemente der Qualitätsbewertung und der Qualitätsförderung im Strukturierten Dialog vermischt würden.

Es wird konstatiert, dass nach Erkenntnissen der wissenschaftlichen Literatur zur Urteilsbildung die sogenannte klinische oder intuitive Urteilsbildung einer statistischen oder mechanischen Urteilsbildung unterlegen sei. Durch explizite Vorgaben an klinische Expertinnen und Experten sowie standardisierte Informationsgrundlagen könne die Güte von Entscheidungen erhöht werden (Seite 58). Vorrangige Ziele des Instituts seien daher die Formulierung expliziter Bewertungsregeln, die Standardisierung des Stellungnahmeverfahrens sowie die alleinige Beschränkung auf die Analyse der Indikatorergebnisse. Dadurch würden Stellungnahmeverfahren einerseits weniger heterogen und andererseits effizienter. Qualitätsbewertung und Qualitätsförderung seien zukünftig in Form von Modulen klar zu trennen (Kapitel 4).

Die Vorschläge zur statistischen Auswertungsmethodik nehmen einen Schwerpunkt des Berichts ein (Kapitel 5). Besonderes Augenmerk wird der sogenannten analytischen Herangehensweise an die statistische Datenauswertung gewidmet. Es wird postuliert, dass sich mit geeigneter mathematisch-statistischer Methodik über die Qualitätsindikatorergebnisse die ihnen zugrunde liegende „Kompetenz der Leistungserbringer“ abschätzen lasse. Damit seien nicht nur Rückschlüsse auf die Qualität der Behandlung bei Fällen eines zurückliegenden Erfassungsjahrs möglich, sondern auch „auf Fälle über das Erfassungsjahr hinaus“ (Seite 73).

Der Bericht enthält den Vorschlag, Stellungnahmeverfahren im Sinne eines Moduls „Qualitätsbewertung“ in einem zweischrittigen Verfahren neu zu konzeptionieren: ein erster

Schritt solle eine optimierte quantitative Auffälligkeitseinstufung beinhalten, sowohl für einen Ein-Jahreszeitraum als auch für einen Zwei-Jahreszeitraum, ein zweiter (qualitativer) Schritt soll einen neuen standardisierten Bewertungsprozess umfassen.

Als theoretische Basis dient dem Institut ein entscheidungstheoretisches Modell mittels bayesianischem Netzwerk. In diesem Modell hängen über Indikator gemessene Behandlungsergebnisse von einem „latenten Kompetenzparameter“ der Leistungserbringer ab. Indikatorergebnis, Fallzahl und Referenzbereich sind weitere Variablen, die in dem Modell zur Entscheidung über eine quantitative Auffälligkeit führen. Die Diskussion wird im Bericht mit dem Ziel geführt, eine möglichst effiziente, d. h. mit geringeren Aufwänden behaftete Entscheidung herbeizuführen. Es werden neben der aus der ESQS bekannten „rechnerischen Auffälligkeit“ und der aus dem PlanQI-Verfahren bekannten „statistisch signifikanten Auffälligkeit“ auch eine „statistisch relevante Auffälligkeit“ unterschieden, bei den die außerhalb des Referenzbereich liegenden maximal tolerierte Anzahl von Fällen im Fokus steht. Vor- und Nachteile dieser Auffälligkeitseinstufungen werden hinsichtlich unterschiedlicher Parameter wie z. B. Fallzahlabhängigkeit, Sensitivität und Spezifität miteinander verglichen. Im Weiteren werden verschiedene statistische Methoden zur Berechnung von Zwei-Jahresauffälligkeiten diskutiert.

Zusammenfassend schlägt das IQTIG vor, zukünftig bei der Berechnung und Bewertung der Indikatorergebnisse die statistische Unsicherheit zu berücksichtigen. Konkret wird die Anwendung der bayesianischen Version der statistisch signifikanten Auffälligkeit empfohlen. Für den Fall, dass viele Leistungserbringer nur kleine Fallzahlen aufweisen, sollten die Berechnungen auf den Daten von zwei Erfassungsjahren beruhen.

In Kapitel 6 werden Vorschläge für Schritt 2 des Stellungnahmeverfahrens gemacht. Liefert Schritt 1 bei einem Indikator einen „hinreichenden Hinweis auf ein Qualitätsdefizit“, so sollen die Expertinnen und Experten nach einem festen Bewertungsschema entscheiden, ob der Hinweis bestätigt wird oder nicht. Im Fokus des schriftlichen Stellungnahmeverfahrens soll stehen, ob in Schritt 1 bei einem Indikator auffällig gewordene Leistungserbringer besondere Einflussfaktoren geltend machen, die bei ihren Patientinnen und Patienten nicht berücksichtigt wurden. Dies soll für jeden einzelnen Fall eines Leistungserbringers erfolgen.

Die Aufgabe der Expertinnen und Experten soll es dann sein, die Angemessenheit dieser Einflussfaktoren zu prüfen. Werden Faktoren anerkannt, für die der Leistungserbringer nicht verantwortlich gemacht werden kann, so soll eine partielle Neuberechnung des Indikatorergebnisses erfolgen. Dabei können sowohl einzelne Fälle aus der Grundgesamtheit (Nenner) genommen werden, wenn sie zuvor fälschlicherweise eingeschlossen wurden, als auch aus dem Zähler, wenn fälschlicherweise ein Ereignis (z. B. Komplikation) gezählt wurde. Dokumentationsfehler („Datenfehler“) sollen dabei gesondert bewertet werden.

Das Ergebnis der Neuberechnung soll ein so genannter „informativer Indikatorwert“ sein. Auch der Referenzbereich ist neu zu berechnen, d. h. ein Anteil an „gerichteten Einflussfaktoren“ wird herausgerechnet. Liegt der informative Indikatorwert innerhalb des bereinigten Referenzbereichs, so ist für die Berichterstattung zum Stellungnahmeverfahren genau zu unterscheiden, ob dies aufgrund inhaltlicher Fehler, Datenfehler oder einer Mischung aus beiden geschehen ist. Liegt der informative Indikatorwert weiterhin außerhalb des bereinigten Referenzbereichs, so haben die Expertinnen und Experten zu entscheiden, ob berechtigterweise geltend gemachte Einflussfaktoren nicht berücksichtigt wurden. Dabei können sie entscheiden (so genannte heuristische Beurteilung), ob der „Hinweis auf ein Qualitätsdefizit“ bei einem Indikator trotzdem entkräftet werden kann. Andernfalls soll gemäß einem Bewertungsschema festgestellt werden, dass ein „Qualitätsdefizit“ vorliege.

Im Weiteren werden Vorschläge zu Mindestanforderungen des schriftlichen Stellungnahmeverfahrens und zur Zusammensetzung der Fachkommission gemacht. Das von der Qualitätsbewertung nach Auffassung des Instituts zu trennende „Modul“ der Qualitätsförderung wird kurz umrissen. So wird auch ein Peer Review als mögliche Maßnahme beschrieben. Für Zielvereinbarungen werden Mindestanforderungen dargelegt (Kapitel 7)

In Kapitel 8 wird eine zeitliche Verkürzung des Stellungnahmeverfahrens vorgeschlagen. Durch strikte Verfahrensvorgaben soll eine Verkürzung des neuen Moduls „Qualitätsbewertung“ gegenüber den Fristen gemäß DeQS-RL von vier auf drei Monate möglich sein.

Die Berichterstattung (Kapitel 9) soll nach Vorschlag des IQTIG zukünftig in getrennten Berichten zur Versorgungsqualität (Ergebnisse des Moduls Qualitätsbewertung) und zu Maßnahmen der Qualitätsförderung erfolgen.

Kapitel 10 enthält Empfehlungen zur Evaluation des neuen Konzepts. Neben einer formativen Evaluation während der Einführung des neuen Verfahrens wird eine summative Evaluation empfohlen, in der u. a. die Entwicklung der Heterogenität der Stellungnahmen und die Effizienz der Stellungnahmen überprüft werden sollen. Anhand der Daten der Erfassungsjahre 2016 und 2017 werden Beispielsberechnungen präsentiert.

Abschließend (Kapitel 11) werden die Empfehlungen zusammengefasst. Gemäß G-BA-Auftrag sollte das IQTIG außerdem prüfen, bei welchen Qualitätsindikatoren eine quantitative mit einer qualitativen Auffälligkeit gleichgesetzt werden kann. Das Institut kommt zu dem Schluss, dass trotz der Erhöhung der Spezifität durch neue statistische Verfahren vorerst nicht auf das Stellungnahmeverfahren verzichtet werden könne. In Beantwortung weiterer Einzelfragen des Auftrags spricht sich das Institut u.a. gegen den Einsatz von Indizes als Aufgreifkriterien aus und diskutiert den Einsatz leistungsbereichsübergreifender Qualitätsindikatoren.

3. Wesentliche Kritikpunkte

Konstrukt der Kompetenzkomponente der Leistungserbringer

Mit der Feststellung „Bei der analytischen Herangehensweise liegt das Interesse am zugrunde liegenden Prozess. Im Rahmen der externen Qualitätssicherung bedeutet dies, dass man sich für die zugrunde liegende Kompetenz des Leistungserbringers bezüglich des im Indikator abgebildeten Qualitätsziels interessiert“ (Seite 73) wird das theoretische Konstrukt der „Kompetenz der Leistungserbringer“ postuliert.

Im Weiteren werden umfangreiche mathematische Modellierungen auf Basis dieses Konstrukts vorgenommen unter der Maßgabe, dass mit Qualitätsindikatoren die zugrunde liegende Kompetenz erfasst werde. Der Satz „Dies bedeutet, dass...bei der analytischen Herangehensweise nicht nur Schlüsse über die behandelten Fälle im betrachteten Erfassungsjahr mit den entsprechenden Ereignissen gezogen werden sollen, sondern das Interesse in der zugrunde liegenden Kompetenz des Leistungserbringers liegt, die sich auch auf die Fälle über das Ende des Erfassungsjahrs hinaus auswirkt“ (Seite 73) kommt der Behauptung gleich, dass sich mit Qualitätsindikatoren nicht nur vergangene, sondern auch zukünftig zu erwartende Qualität messen lasse.

Es handelt sich nach Ansicht der Bundesärztekammer um eine Neudefinition der Qualität bzw. der Funktion eines Qualitätsindikators, die durch keine Ableitung aus der wissenschaftlichen Literatur gedeckt ist. Die Befassung mit und die Messung von Kompetenzen ist der Domäne der Bildung und Weiterbildung vorbehalten. Instrumente zur Messung von Kompetenzen sind nicht mit den Indikatoren der externen Qualitätssicherung

vergleichbar. Unabhängig davon ist auch der Begriff des Leistungserbringers in diesem Kontext unbestimmt. Das Spektrum der denkbaren Konstellationen ist breit, vom fest zusammengesetzten oder wechselnden interdisziplinären Team bis hin zum einzelnen Arzt.

Die Bundesärztekammer rät von der Verwendung des Begriffs der „Kompetenz“ in diesem Zusammenhang dringend ab, insbesondere deshalb, weil er einen zentralen Bestandteil in der ärztlichen Weiterbildung darstellt und die Gefahr einer inhaltlichen und sprachlichen Uneindeutigkeit gegeben ist.

Fachliches Bewertungsverfahren

Zentrales Element der vorgeschlagenen fachlichen Bewertung in Schritt 2 des Stellungnahmeverfahrens soll die Ergebnisbetrachtung aller Indikatoren auf Einzelfallebene sein. Das Ziel ist dabei nicht eine tatsächliche fachliche Erörterung des Einzelfalls mit den klinischen Experten. Vielmehr geht es um die fallweise Abklärung, ob sich der Indikatorwert und der Referenzbereich weiter „bereinigen“ lassen, indem vorher nicht berücksichtigte Risikofaktoren nachträglich berücksichtigt werden. Die Validität des Indikators soll hier quasi ex post rechnerisch optimiert werden.

Dieses Grundprinzip steht im Widerspruch zur Forderung an anderer Stelle im Bericht, dass „die Stellungnahme der Leistungserbringer möglichst auf eine Analyse von übergeordneten Gründen fokussieren soll“ (Seite 165). Dort heißt es (zu Recht) weiter: „Eine am Einzelfall orientierte Stellungnahme kann daher den Blick auf systemische Ursachen und notwendige systemische Verbesserungsansätze, wie z. B. die Verbesserung in der Einrichtungsorganisation oder der Patientensicherheitskultur, versperren“.

Der Aufwand, für jeden einzelnen Fall eines Erfassungsjahrs bei einem quantitativ auffälligen Indikator im Nachhinein mögliche Risikofaktoren zusammenzutragen, dürfte erheblich sein. Bei großen Fachabteilungen kann es notwendig werden, u.U. mehrere hundert Patientenakten für einen Indikator zu analysieren. Auch für die Expertinnen und Experten in den Fachgremien dürfte die Überprüfung der Angaben der Leistungserbringer aufwändig, wenn nicht sogar unpraktikabel sein.

Als Extrembeispiel sei im Leistungsbereich Hüftendoprothetik der QI 53004 (Präoperative Verweildauer bei endoprothetischer Versorgung einer hüftgelenknahen Femurfraktur) genannt: Im Jahre 2018 wurden 431 von 1.159 Krankenhäuser auffällig. Die Anzahl der auffälligen Fälle betrug 8.078 von 58.741 Fällen. Alleine für diesen Indikator müssten jedoch vermutlich mehrere Tausend Fälle betrachtet werden.

Das vorgeschlagene fachliche Verfahren will erklärtermaßen den Bewertungsspielraum der Expertinnen und Experten durch explizite Regeln einengen. „Es soll ein Maximum an Standardisierung und damit Objektivität der fachlichen Bewertung erreicht“ werden (Seite 163). Ein Dialog mit dem Leistungserbringer, wie er namensgebend für den Strukturierten Dialog war, ist nicht mehr vorgesehen.

Das Institut räumt selbst ein, dass durch die vorgeschlagenen expliziten Regeln das Verfahren deutlich komplexer werde und zudem für die klinischen Expertinnen und Experten ungewohnt sei (Seite 60). Tatsächlich dürfte das Verfahren inklusive der Neuberechnung der Indikatorwerte und Referenzbereiche für alle Beteiligten, also für die auf Landesebene beauftragten Stellen, die Expertinnen und Experten und die betroffenen Krankenhäuser und Vertragsärzte deutlich aufwändiger werden. Der im IQTIG-Konzept gewonnene Vorteil, dass durch die neue, auf statistischer Signifikanz beruhende Auswertungsmethodik weniger Auffälligkeiten im Stellungnahmeverfahren generiert werden, droht durch die Komplexität des nachfolgenden Bewertungsverfahrens

konterkariert zu werden. Es ist zudem zu befürchten, dass auch für die Adressaten der Qualitätsberichterstattung, die Versicherten und Patienten, die Ergebnisse des Stellungnahmeverfahrens nicht transparenter, sondern weniger anschaulich werden.

Im Bericht werden Komplikationsraten als Beispiele für das neue vorgeschlagene Bewertungsverfahren genannt. Das Spektrum der Qualitätsindikatoren in der ESQS reicht mit Prozessindikatoren, Indikatoren zur Indikation und komplexen Index-Indikatoren aber deutlich weiter. Hier sollte an weiteren Beispielen die Praktikabilität des neuen Verfahrens überprüft werden.

Die Ausführungen im Bericht beziehen sich zudem ausschließlich auf „negativ gepolte“ Indikatoren, d. h. Indikatoren, die ein negatives Ereignis, wie z. B. eine Komplikation, erfassen. Positiv gepolte Indikatoren, wie z. B. im Leistungsbereich Hüftendoprothetik QI 54002 (Indikation zum Hüftendoprothesen-Wechsel bzw. –Komponentenwechsel) sind umgekehrt zu behandeln. In diesen Fällen wird ein Referenzbereich bei Auffälligkeit nicht über- sondern unterschritten. Hier müsste von den betroffenen Krankenhäusern im Stellungnahmeverfahren für jeden nicht im Zähler mitgezählten (also primär nicht indizierten) Fall ex post nachgewiesen werden, dass er doch indiziert war. Diese Indikatorvarianten sollten im Bericht mit diskutiert werden.

Das Ergebnis des Bewertungsverfahrens soll ein neues Bewertungsschema sein. Statt dem bis dato verwendeten Begriff der „qualitativen Auffälligkeit“ soll zukünftig von einem „Qualitätsdefizit“ gesprochen werden. Es wird festgestellt, dass ersterer Begriff zu Missverständnissen führe, während letzterer „allgemeinverständlich das Nichterreichen der Anforderungen“ (Seite 144) repräsentiere. Tatsächlich ist dieses Bewertungsschema für die Außendarstellung von zentraler Bedeutung, da es die Bewertung der QS-Ergebnisse im Qualitätsbericht der Krankenhäuser wiedergibt. Es suggeriert jedoch die Sicherheit einer Tatbestandsfeststellung, die keineswegs gegeben ist. Dies wird schon alleine deshalb deutlich, da auch die Nichtwahrnehmung einer Möglichkeit zur Stellungnahme durch den Leistungserbringer und „teilweise durch Leistungserbringer zu verantwortende Datenfehler“ automatisch als „Qualitätsdefizit“ klassifiziert werden sollen.

Nachberechnung der Indikatorergebnisse

Im PlanQI-Verfahren wird gegenwärtig eine Neuberechnung der Indikatorergebnisse auf Basis von stichprobenhaften Datenvalidierungen durchgeführt. Das IQTIG hat in keinem seiner früheren Berichte wissenschaftlich dargelegt, dass die Vorgehensweise, das Ergebnis einer Vollerhebung auf Basis einer Stichprobenziehung zu korrigieren, methodisch angemessen bzw. statthaft ist.

Das Konzept des vorliegenden Berichts sieht nun vor, für bei einem Indikator auffällige Leistungserbringer partielle Nachberechnungen auf Basis der Rückmeldungen der Leistungserbringer im Stellungnahmeverfahren durchzuführen. Bei Ratenindikatoren können sich Zähler (interessierende Ereignisse, z. B. Komplikation) oder Nenner (eingeschlossene Grundgesamtheit des Indikators) ändern. Ergebnis soll ein so genannter informativer Indikatorwert sein. Da gleichzeitig auch der Referenzbereich individuell für jeden Leistungserbringer auf Basis der Ergebnisse „bereinigt“ werden soll, dürfte nicht selten das Ergebnis nach der „Bereinigung“ für das Krankenhaus noch ungünstiger sein als vorher (nachzuvollziehen anhand des fiktiven Beispiels in Abbildung 33 auf Seite 157). Die statistischen Auswertungen nach Neuberechnung eines Indikators und seines Referenzbereichs sind komplex und erfolgen unter Zeitdruck im Bewertungsverfahren. Bei komplexeren Indikatoren als einfachen Raten dürfte die Unanschaulichkeit der Ergebnisse zunehmen. Viele praktische Fragen bleiben offen. Kommt beispielsweise bei einem Indikator

mit einer statistischen Auffälligkeit über zwei Jahre auch eine Neuberechnung über zwei Jahre hinweg zur Anwendung?

In jedem Fall wäre für die Neuberechnung ein geeignetes Softwaretool unabdingbar, das den auf Landesebene beauftragten Stellen vom IQTIG zur Verfügung gestellt werden müsste.

Peer Review

An mehreren Stellen wird im Bericht auf das Ärztliche Peer Review Bezug genommen. Erfolgsfaktoren für dieses Instrument des kollegialen Dialogs wie Freiwilligkeit, Sanktionsfreiheit, fachliche Kompetenz der Peers und einheitliches Vorgehen werden zu Recht hervorgehoben (Kapitel 3.6). Allerdings ergeben sich Hinweise, dass bezüglich der Funktion eines Peer Reviews Missverständnisse bestehen. So wird auf Seite 23 eine Ähnlichkeit mit dem Strukturierten Dialog angedeutet, „auffällige Ergebnisse“ mit dem Leistungserbringer zu klären. Das Ärztliche Peer Review ist aber keinesfalls ein Qualitätsbewertungsinstrument, sondern ein Instrument des Qualitätsmanagements bzw. der Qualitätsförderung. Indikatorergebnisse können als Aufgreifkriterien dienen, müssen es aber nicht. Im Zentrum des kollegialen Dialogs steht nicht die Erörterung der Ergebnisse klinischer Einzelfälle, sondern die Betrachtung der Prozesse, die der Behandlung der Patientinnen und Patienten zugrunde liegt.

Auf Seite 174 wird vorgeschlagen, „diejenigen Leistungserbringer, bei denen mit hoher Sicherheit davon ausgegangen werden kann, dass sie die Qualitätsanforderungen erfüllt haben, zu identifizieren und als Peers für qualitätsfördernde Maßnahmen bei anderen Leistungserbringern zu gewinnen.“ Auch diese Überlegung, dass im Sinn eines „Lernen von den Besseren“ die besuchenden Peers „die Besseren“ seien, entspricht nicht dem Grundgedanken des Ärztlichen Peer Reviews. „Peers“ sind schon dem Wort nach Kolleginnen und Kolleginnen, die sich auf Augenhöhe begegnen. Im Peer Review kann auch der besuchende Peer vom besuchten Peer lernen. Zudem kann er oder sie selbst bei einem anderen Peer Review zum besuchten Peer werden.

4. Fazit

Der Auftrag des G-BA an das IQTIG vom 18.02.2018 zur Weiterentwicklung des Strukturierten Dialogs mit Krankenhäusern enthielt einen umfangreichen Katalog von Aufgaben zur Optimierung der Einheitlichkeit der Vorgehensweise, der Transparenz und Nachvollziehbarkeit der Entscheidungsfindung und der Effizienz des Verfahrens. Zudem sollte ein Rahmenkonzept für ein „Verfahren zur Feststellung von Auffälligkeiten in Krankenhäusern, zu deren Bewertung sowie zur Initiierung von Maßnahmen“ entwickelt werden. Die meisten dieser Aufgaben wurden im vorliegenden Bericht abgearbeitet, einige allerdings nur oberflächlich gestreift. Der Schwerpunkt des Berichts liegt auf der Qualitätsbewertung. Zur Qualitätsförderung hätte man sich substantziellere Vorschläge gewünscht.

Die Analyse des IQTIG zum Ist-Zustand des Strukturierten Dialogs gemäß QSKH-RL ist weitgehend nachvollziehbar. Betont werden die Heterogenität der Verfahren in den einzelnen Bundesländern und der Aufwand für die beteiligten Krankenhäuser und LQS. Letzterer ist durch eine niedrige Spezifität des Verfahrens bedingt.

Zur Erhöhung der Spezifität wird vom Institut eine statistische Methodik abgeleitet, nach der nicht mehr nur rechnerische, sondern statistisch signifikante Auffälligkeiten ausgewiesen werden sollen, bei Bedarf über zwei Erfassungsjahre statt über nur ein Erfassungsjahr. Die Ableitung der statistischen Methodik nimmt im Bericht einen dominanten Teil ein. Hier wäre die Auslagerung der mathematisch vertiefenden Ausführungen in eine Anlage zielführender gewesen. Die vorgeschlagene Berechnungsmethode erscheint - zumindest für die Beispielindikatoren und- auf theoretischer Ebene - plausibel. Eine Erhöhung der Spezifität der Indikatorergebnisse würde eine Reduzierung des Aufwands für alle Beteiligten bedeuten. Allerdings wird die vom Institut postulierte Grundannahme für die mathematische Modellierung, mit Qualitätsindikatoren ließe sich die „Kompetenz“ der Leistungserbringer messen, von der Bundesärztekammer nicht geteilt. Es existiert auch keine Grundlage in der Literatur dafür.

Da im Bericht nur wenige Beispiele für die neue Methodik der Indikatoreauswertung benannt werden, lässt sich noch keine Aussage darüber treffen, ob wirklich die gesamte datengestützte Qualitätssicherung von den neuen Methoden profitieren würde. Die Anschaulichkeit und Nachvollziehbarkeit der ermittelten Werte sollte Priorität haben.

Als methodische Grundlage für eine Neukonzeption des Strukturierten Dialogs gemäß QSKH-RL bzw. des Stellungnahmeverfahrens gemäß DeQS-RL sollen Forschungsergebnisse zur ärztlichen Urteilsbildung dienen. Eine statistische bzw. mechanistische Urteilsbildung sei der intuitiven Urteilsbildung hinsichtlich ihrer Treffgenauigkeit überlegen. Ob sich die zitierten Forschungsergebnisse wirklich 1:1 auf die spezielle Situation des Strukturierten Dialogs übertragen lassen, muss jedoch kritisch hinterfragt werden.

Das bis heute gültige Konzept, dass Qualitätsindikatoren zunächst als Aufgreifkriterien dienen und erst im kollegialen, strukturierten Dialog der Expertinnen und Experten mit den Leistungserbringern die wahren Qualitätsprobleme identifiziert werden, wird im Bericht quasi als obsolet erklärt. Stellungnahmeverfahren sollen nach Vorstellung des Instituts zukünftig nach expliziten, d. h. detailliert vorgegebenen Regeln erfolgen. Inhalt des Stellungnahmeverfahrens sollen alleine die Indikatorergebnisse sein. Dabei soll der Kontext, in dem diese Ergebnisse entstanden sind, nicht mehr in die Bewertung einfließen. Die Leistungserbringer sollen für jeden ihrer behandelten Fälle angeben, ob individuelle Risikofaktoren vorgelegen haben. Die Expertinnen und Experten sollen dann nach detaillierten Vorgaben für jeden Indikator und jede Vorgangsnummer überprüfen, ob eine Neuberechnung des Indikators gerechtfertigt ist. Lediglich für eine Restkategorie nicht in

Rechenregeln umsetzbarer Begründungen wird den Expertinnen und Experten ein gewisser heuristischer Beurteilungsspielraum eingeräumt.

Dieser Paradigmenwechsel im Strukturierten Dialog bzw. im Stellungnahmeverfahren wird von der Bundesärztekammer –zumindest in diesem Ausmaß – abgelehnt. Ein solcher mechanistischer Ansatz würde das Bewertungsverfahren auf einen abzuarbeitenden Algorithmus reduzieren. Ein eigentlicher Dialog mit den betroffenen Einrichtungen soll nicht mehr stattfinden. Zum Ausdruck kommt ein erschreckend reduktionistisches Qualitätsverständnis, bei dem nur noch Indikatorergebnisse, aber nicht mehr die zugrunde liegenden Prozesse und Strukturen betrachtet werden sollen. Letzteres angemessen zu berücksichtigen, ist jedoch genau die Stärke des Strukturierten Dialoges und sollte auch nicht mit „Heterogenität“ verwechselt werden, die es um jeden Preis zu beseitigen gelte.

Die vorgeschlagene Trennung des Strukturierten Dialogs in zwei unterschiedliche, nacheinander abzuarbeitende Module Qualitätsbewertung und Qualitätsförderung scheint vordergründig im Sinne einer klaren Aufgabentrennung vorteilhaft zu sein. Da dadurch allerdings in der Qualitätsbewertung keinerlei Raum mehr für einen kollegialen Dialog vor der abschließenden Bewertung gelassen wird, wird diese strikte Trennung von der Bundesärztekammer abgelehnt.

Ein weiteres und dabei grundlegendes Problem der Anwendbarkeit des neuen Konzepts besteht durch die Historie des Auftrags. Das Konzept ist formuliert in Anlehnung an den Strukturierten Dialog für den Krankenhausbereich. Inwieweit die Vorschläge für Bewertungsverfahren auch im vertragsärztlichen Bereich gelten können, wird gar nicht thematisiert.

Zusammenfassend enthält der IQTIG-Vorbericht einige bedenkenswerte Vorschläge v. a. zur methodischen Weiterentwicklung des Auswertungsverfahrens. Ein erster Schritt sollte allerdings die Überprüfung der bestehenden Indikatoren auf ihre Güte und ggf. die Streichung von Indikatoren bzw. die Neuentwicklung geeigneter valider Indikatoren sein. Das im Bericht vorgelegte Konzept betont jedoch einseitig die nachgelagerten Schritte der Auswertung und Bewertung mit dem offensichtlichen Ziel, mehr Validität ex post zu erreichen. Folgt man den vorgeschlagenen Änderungen im Bewertungsverfahren, würde am Ende das Kernstück des Strukturierten Dialogs – nämlich der Dialog – weitgehend abgeschafft sein.

Die Bundesärztekammer rät von einer Umsetzung der Änderungsvorschläge zum Bewertungsverfahren aufgrund der vielen offenen Fragen derzeit ab.