

Eignung von Qualitätsindikatorensets in der Gesundheitsversorgung für verschiedene Einsatzgebiete – Forschungs- und Handlungsbedarf

Döbler K, Schrappe M, Kuske S, Schmitt J, Sens B, Boywitt D, Misselwitz B, Nothacker M, Geraedts M

Stand: 16.6.2019

1. Hintergrund und Zielsetzung

Indikatoren-gestützte Verfahren zur Qualitätsmessung und -bewertung in der Gesundheitsversorgung sind in Deutschland primär aus Initiativen wissenschaftlicher Fachgesellschaften entstanden und wurden über Jahrzehnte vor allem mit den Zielen der Forschung und der Qualitätsverbesserung eingesetzt [1]. Spätestens seit Einführung der verpflichtenden Qualitätsberichterstattung für Krankenhäuser im Jahr 2006 werden Qualitätsindikatoren zunehmend auch mit dem Ziel der Versorgungssteuerung verwendet [2].

Mit dem Krankenhausstrukturgesetz aus dem Jahr 2016 sollen Ergebnisse von Qualitätsindikatoren nunmehr noch weitgehender für regulative und steuernde Interventionen im Krankenhausbereich verwendet werden, so beispielsweise für Vergütungszu- und -abschläge sowie für Entscheidungen von Krankenhausplanungsbehörden [3].

Unter einem Qualitätsindikator versteht man nach weithin akzeptierter Definition der JCAHO „ein quantitatives Maß, welches zum Monitoring und zur Bewertung der Qualität wichtiger Leitungs-, Management-, klinischer und unterstützender Funktionen genutzt werden kann, die sich auf das Behandlungsergebnis beim Patienten auswirken.“ [4]. Als wesentliche Einsatzgebiete von Qualitätsindikatoren im Gesundheitssystem benennt die US-amerikanische Agency for Healthcare Research and Quality (AHRQ) „Qualitätsverbesserung, Steuerung und Forschung“ [5].

Aus der dargestellten Entwicklung ergibt sich die Notwendigkeit, den Einsatz von einzelnen Qualitätsindikatoren und insbesondere auch von Indikatoren-Sets in ein Rahmenkonzept einzubetten, auf dessen Grundlage die weitere Entwicklung geplant und später auch evaluiert werden kann. Die spätere Evaluation erscheint von besonderer Bedeutung, da internationale Ergebnisse aus einem breiten Einsatz von Qualitätsinformationen bekannt sind, die teilweise einen positiven Einfluss im Sinne einer Qualitätsverbesserung nicht belegen konnten. Ein wichtiges aktuelles Beispiel ist das National Surgical Quality Improvement Program des American College of Surgeons (ACS-NSQIP), dessen Auswertung auf insgesamt 1,6 Millionen Patienten mit elektiven Eingriffen basiert [6, 7, 8].

Mit dem Koalitionsvertrag für die 19. Legislaturperiode haben die Regierungsparteien nun erneut die Schwerpunktsetzung der sektorenübergreifenden und regionalen Versorgung in den Mittelpunkt gestellt [9]. Auch hieraus können sich wichtige Konsequenzen für den Einsatz von Qualitätsindikatoren ergeben: es können neben der bislang im Mittelpunkt stehenden Beobachtungs- und Handlungsebene einzelner Einrichtungen weitere Beobachtungsebenen relevanter werden (wie beispielsweise Versorgungsverträge oder Regionen) und das Einsatzgebiet „Forschung“ bekommt größere Bedeu-

tung, da der Einsatz von Qualitätsindikatoren sich vom Vergleich einzelner Einrichtungen auf die Ebene der Systembetrachtung ausweiten kann.

Neue gesetzliche Anforderungen, aber auch die Entwicklung der realen Versorgungsstrukturen und -prozesse stellen neue – höhere - Anforderungen an den Einsatz von Qualitätsindikatoren und Qualitätsindikatorensets in der Gesundheitsversorgung.

Diese veränderte Ausgangssituation führt dazu, dass in Zukunft nicht nur die Qualitätsförderung einzelner Einrichtungen, sondern auch regional-sektorenübergreifende, qualitätssteuernde Zwecke zu verfolgen sein werden und bedingt erheblichen und teilweise dringlichen Forschungs- und Handlungsbedarf.

Die an den bisherigen Anforderungen ausgerichteten Konzepte zur Entwicklung und zum Einsatz von Qualitätsindikatoren und Qualitätsindikatorensets scheinen für diese neuen Anforderungen teilweise nur noch bedingt geeignet.

Das vorliegende Positionspapier stellt dar, welcher prioritäre Forschungs- und Handlungsbedarf sich aus Sicht der Qualitäts- und Patientensicherheitsforschung abzeichnet.

2. Die Situation in Deutschland

Vor diesem Hintergrund ergeben sich in Deutschland weitergehende Anforderungen an die Qualitätsmessung, -bewertung, und -darlegung. Diese muss zunehmend auch für regional-sektorenübergreifende Beobachtungs- und Handlungsebenen gestaltet werden und zudem Rahmenbedingungen und Grenzen für regulative Verfahren wie beispielsweise Zu- und Abschläge bzw. einen Versorgungsausschluss aufzeigen [10].

Parallel zu der Entwicklung, Qualitätsindikatoren für Steuerungszwecke zu verwenden, findet eine intensive Diskussion um Kriterien zur Bewertung methodischer Eigenschaften („Gütekriterien“) von Indikatoren statt [11, 12, 13]. Aufbauend auf dem QUALIFY-Bewertungsinstrument für Qualitätsindikatoren der BQS haben das AQUA-Institut und das Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG) hierzu jeweils eigene Konzepte in ihren jeweiligen Methodenpapieren formuliert [14,15].

Festzustellen ist hierbei, dass sich unterdessen für die Kriterien zur Darstellung methodischer Eigenschaften von Qualitätsindikatoren recht umfangreiche Literatur findet, während für die Bedeutung der Ausprägung dieser Eigenschaften für bestimmte Einsatzgebiete („Eignung“) noch wenig Klarheit zu bestehen scheint [11, 16].

Zudem steht eine Diskussion um die Eigenschaften, Ausrichtung und Ausgewogenheit von Indikatorensets, d.h. einer Zusammenstellung einzelner Qualitätsindikatoren zu einem Thema, hinter der Diskussion um Eigenschaften einzelner Indikatoren noch deutlich zurück [17].

Beide Ebenen (einzelner Indikator, Indikatorenset) spielen eine wichtige Rolle für die Bewertung der Eignung von Verfahren zur Qualitätsdarstellung, da sowohl der einzelne Indikator als auch das Indikatorenset jeweils Konstrukte der Qualität darstellen, deren Eigenschaften sich unterscheiden können und damit für bestimmte Einsatzgebiete mehr oder weniger gut geeignet sein können [18].

Um dies beispielhaft zu illustrieren:

1. Auf Ebene einzelner Indikatoren können für ein Verfahren mit dem Ziel einer Qualitätsverbesserung in einem anonymen Rahmen unter sorgfältiger Abwägung von Nutzen und Risiken ggf. Kompromisse an der Güte der Risikoadjustierung hingenommen werden, wenn dadurch gleichzeitig ein geringerer Erfassungsaufwand resultiert.

Für einen Einsatz mit möglichen einschneidenden Konsequenzen ist dagegen eine möglichst optimale Risikoadjustierung zu fordern, da hierfür ein ggf. erhöhter Aufwand den Nutzen rechtfertigt.

2. Auf Ebene des Indikatorensets kann für das Ziel einer Qualitätsverbesserung in Bezug auf ein erkanntes Qualitätsdefizit (z.B. Implementierung einzelner Leitlinienempfehlungen) ein sehr schmales, auf eben diesen Ausschnitt (z.B. einzelner Aspekt aus der Dimension Sicherheit) fokussiertes Indikatorenset eingesetzt werden oder aber die Implementierung selber wird zum Gegenstand gemacht und deren Ergebnisse werden gemessen [19].

Wenn Indikatoren genutzt werden sollen, um Patienten bei der Auswahl von Krankenhäusern zu einem bestimmten Thema zu unterstützen (z.B. in der Geburtshilfe), ist hingegen ein ausgewogeneres, umfassenderes Indikatorenset erforderlich, das Informationen zu mehreren Dimensionen, wie etwa Sicherheit (Vermeidung unerwünschter Ereignisse), Effektivität (werden die Behandlungsziele erreicht?) und zur Patientenzentrierung liefert.

Mit der Einführung einer sektorenübergreifenden Qualitätssicherung und zunehmendem Fokus auf eine bedarfsgerechte, sektorenübergreifende Steuerung der Gesundheitsversorgung werden Qualitätsindikatoren in Zukunft über die Beobachtungs- und Handlungsebene der einzelnen Einrichtung hinaus stärker für weitere Beobachtungs- und Handlungsebenen (z.B. für regionale Betrachtungen) zu modellieren sein [10,20]. Es zeigt sich, dass für die Qualitätsbetrachtung auf einer sektorenübergreifenden „Systemebene“ die aktuellen Instrumente derzeit nur eingeschränkt einsetzbar sind [21].

3. Spezifische Fragestellungen zur Bewertung von Indikatorensets

Die Gesundheitsversorgung ist ein multidimensionales, komplexes Geschehen. Gut ausgewählte und valide Indikatoren können zwar Aussagen zur Versorgungsqualität auch für größere Bereiche der Gesundheitsversorgung ermöglichen, trotzdem bleiben sie immer selektiv. Zudem wird es nie möglich sein, alle Aspekte der Gesundheitsversorgung abzubilden.

In einigen Fällen ist dies hinreichend, wenn z.B. nur ein bestimmter, spezifischer Aspekt der Gesundheitsversorgung beobachtet werden soll. Soll jedoch die Qualität der Gesundheitsversorgung in einem ganzen Leistungsbereich (z. B. „Geburtshilfe“) und dann auch noch möglichst sektorenübergreifend (z. B. ambulante Schwangerschaftsvorsorge, Geburt, Nachsorge) bewertet werden, dann sind Indikatorensets (d.h. Gruppen mehrerer einzelner Qualitätsindikatoren zu einem Thema) nötig, die unterschiedliche Teilaspekte bewerten. Dies gilt zumindest für Bewertungen, aus denen eine direkte Handlungsrelevanz abgeleitet werden soll, die also beispielsweise die Leitlinienkonformität der Behandlung einzelner Leistungserbringer beurteilen. Auf den Sonderfall der aggregierten Auswertung mehrerer Einzelindikatoren („Composite Measures“) wird in Abschnitt 4.5 eingegangen.

Indikatorensets der Gesundheitsberichterstattung, die ganze Gesundheitssysteme bewerten, sind hinlänglich bekannt (z. B. World Health Report 2000, OECD - Health Care Quality Indicators). Auch Indikatorensets, die die Behandlung durch einzelne Leistungserbringer in definierten Regionen bewerten, sind international bekannt. Eines der ältesten und am besten erprobten derartigen Indikatorensets wurde von der Health Care Financing Administration (HCFA) bzw. später den Centers for Medicare & Medicaid Services (CMS) entwickelt und für den Vergleich der Versorgungsqualität der Bundesstaaten der USA verwendet [22, 23]. Dieses Indikatorenset basiert auf sechs Erkrankungen (akuter Myokardinfarkt, chronische Herzinsuffizienz, Schlaganfall, ambulant erworbene Pneumonie, Brustkrebs und Diabetes) mit insgesamt 24 Indikatoren. Das Instrument ergibt über die Zeit hochgradig korrelierende Ergebnisse. Die Auswahl der Indikatoren geschah nach transparent ausgewählten Kriterien:

- disease is a major source of morbidity or mortality
- certain processes of care are known to improve outcomes
- measurement of these processes is feasible
- offers substantial scope for improvement in performance
- managerial intervention can potentially improve performance.

International werden Kriterien angewendet, die allgemein zur Priorisierung von Indikatoren für die Verwendung in den jeweiligen Indikatorensets genutzt werden. Orientiert an den jeweiligen Zielsetzungen und Einsatzgebieten werden die Kriterien „Wissenschaftlichkeit“, „Schweregrad“, „Häufigkeit“, „Kostenfaktoren“, „Interesse der Öffentlichkeit“, „Machbarkeit“ und „Vulnerabilität“ herangezogen [24]. Je nach Fragestellung können weitere Kriterien wie „Ausgewogenheit der Settings“ sowie „Ausgewogenheit der Qualitätsdimensionen“ sowie „Repräsentativität“ herangezogen werden. Dies sind jedoch Kriterien, die nur von einzelnen Institutionen explizit genannt werden (z.B. The Joint Commission, Organisation for Economic Co-operation and Development, Agency for Healthcare Research and Quality, National Quality Forum).

Bei der Entwicklung von Indikatorensets finden letztlich Auswahlentscheidungen auf mehreren Ebenen statt:

- Auswahl prioritärer Themen, die eine große Bedeutung für die Gesundheit der Bevölkerung haben (mögliche Kriterien: Häufigkeit, Schwere der Erkrankungen);
- Auswahl von Indikatoren zu Versorgungsaspekten, bei denen Qualitätsunterschiede wie Über-, Unter-, oder Fehlversorgung unter Berücksichtigung von Bedarfen festzustellen sind;
- Auswahl von Indikatoren mit verbliebenem Verbesserungspotenzial, für die evidenzbasierte Behandlungsempfehlungen bestehen, die wiederum unter Kontrolle der Leistungserbringer sind;

In Erweiterung dieser Auswahlkriterien sind bei Indikatorensets weitere Aspekte zu berücksichtigen:

- Auswahl der einzelnen Indikatoren unter dem Gesichtspunkt ihrer methodischen Eigenschaften / methodischen Güte;
- Auswahl messbarer, also praktikabler Indikatoren.
- Auswahl der einzelnen Indikatoren unter dem Gesichtspunkt der ausgewogenen Abbildung des Versorgungsaspekts (Eigenschaften des Indikatorensets).

Es ergibt sich aus dem in den Abschnitten 2 und 3 Beschriebenen, dass bei der Bewertung von Indikatorensets primär zwei Ebenen zu berücksichtigen sind:

- 1) Die Ebene der Eigenschaften („Gütekriterien“)
 - a) der einzelnen Indikatoren
 - b) der Indikatorensets;
- 2) Die Ebene der Eignung für bestimmte Zielsetzungen und Einsatzgebiete
 - a) der einzelnen Indikatoren
 - b) der Indikatorensets.

zu 1) Bewertung auf der Ebene der Eigenschaften („Gütekriterien“)

Eine Bewertung auf Ebene der Eigenschaften liefert eine Bewertung der Güte des Indikators, jedoch keine Bewertung der Eignung eines Indikators für einen bestimmten Zweck im Sinne von „gut für“ oder „nicht gut für“, sondern nur im Sinne von „ausgeprägt“ oder „weniger ausgeprägt“, bestenfalls im Sinne von „gut“ oder „weniger gut“ bzw. „besser“. So kann beispielsweise festgestellt werden, dass ein Indikator eine mehr oder weniger differenzierte und damit spezifische Risikoadjustierung aufweist. Dies lässt aber zunächst noch keine Aussage zu, ob der Indikator damit für eine bestimmte Zielsetzung mehr oder weniger gut geeignet ist.

Somit sind für eine Bewertung auf der Ebene der Eigenschaften Kriterien erforderlich, die die Eigenschaften eines Indikators und eines Indikatorensets möglichst umfassend beschreiben und somit ein Profil dieser Eigenschaften liefern.

Für Kriterien zur Beschreibung der Eigenschaften einzelner Indikatoren erscheint der Begriff „Gütekriterien“ unterdessen geläufig und es liegen eine Reihe von Empfehlungen für solche Gütekriterien vor [11-15, 25-27].

Für Kriterien zur Beschreibung der Eigenschaften von Indikatorensets liegen derzeit keine klaren Empfehlungen vor [17].

zu 2) Bewertung auf Ebene der Eignung für bestimmte Zielsetzungen und Einsatzgebiete

Eine Eignung kann nur in Bezug auf die beabsichtigte Anwendung beurteilt werden: Eignung wofür? Im Vordergrund steht dabei die Zielbestimmung des Einsatzes der Indikatoren.

Berücksichtigt werden müssen dabei die strukturellen Umgebungsbedingungen, mögliche erwünschte sowie nicht erwünschte strukturelle Folgen des Indikatoreneinsatzes sowie die Beobachtungs- und Handlungsperspektive. Der letztgenannte Aspekt ist besonders wichtig, denn speziell in Deutschland werden Indikatoren vielfach unter dem Blickwinkel des Institutionen- bzw. Anbieterbezugs genutzt (Beobachtungs- und Handlungsebene: einzelne Einrichtung). In Zukunft erscheint zusätzlich besonders ein Populations- bzw. regionaler Bezug von Bedeutung (sektorenübergreifende und patientenbezogene Perspektive).

Immanent mit der Frage der Zielsetzung verbunden ist die Frage nach den möglichen Konsequenzen, d.h. praktisch die Frage nach den Interventionen, die auf Grundlage der Indikatorergebnisse erfolgen sollen.

Dies betrifft die Grundfragen nach freiwilligen oder verpflichtenden Datenerhebungen und Maßnahmen, externer oder interner Nutzung der Ergebnisse sowie nach den sich ergebenden Konsequenzen, z.B. Veröffentlichung der Ergebnisse, Einsatz qualitätsfördernder oder steuernder Maßnahmen, Benchmarking-Projekte der beteiligten Einrichtungen.

Somit müssen zunächst mögliche Ziele und Einsatzgebiete der Indikatoren sowie mögliche Konsequenzen der ermittelten Ergebnisse beschrieben werden.

Im nächsten Schritt sind dann Kriterien festzulegen, mit Hilfe derer die Eignung von Indikatoren und Indikatorensets mit bestimmten Eigenschaften für bestimmte Ziele und Einsatzgebiete bewertet werden kann.

4. Prioritärer Forschungs- und Handlungsbedarf

Im Folgenden soll der prioritäre Forschungs- und Handlungsbedarf auf den mehrfach genannten zwei Ebenen beschrieben werden:

Ebene Qualitätsindikatoren

1. Bewertung der Eigenschaften (Gütekriterien) von Qualitätsindikatoren
2. Bewertung der Eignung von Qualitätsindikatoren für bestimmte Zielsetzungen und Einsatzgebiete

Ebene Qualitätsindikatorensets

1. Bewertung der Eigenschaften (Gütekriterien) von Qualitätsindikatorensets
2. Bewertung der Eignung von Qualitätsindikatorensets für bestimmte Zielsetzungen und Einsatzgebiete

4.1 Bewertung der Eigenschaften von Qualitätsindikatoren

Vielfach verwendet werden Gütekriterien aus der vom RAND-Institut vorgeschlagenen „RAND-Appropriateness-Methode“ und des Qualify-Instruments. Für beide Instrumente gilt jedoch, dass die Kriterien meist in modifizierter Form verwendet werden, d.h. dass nur ausgewählte Kriterien verwendet werden, und/oder die Kriterien modifiziert wurden und/oder das Vorgehen zur Bewertung der Kriterien modifiziert wurde [12, 28-30].

Der Einsatz und die Ausgestaltung dieser Instrumente hängt wesentlich davon ab, ob die Kriterien zur Bewertung der Eigenschaften von Qualitätsindikatoren während der Entwicklung von Qualitätsindikatoren oder zur Bewertung „fertiger“ Indikatoren im Routinebetrieb verwendet werden.

Keines der in Deutschland eingesetzten Verfahren (Qualify durch BQS, modifiziertes Qualify durch NVL, modifiziertes RAND durch AQUA, Kriterien des IQTIG) wurde bislang systematisch empirisch validiert; das Qualify-Instrument wurde zumindest interviewbasiert evaluiert [31]. Dabei erscheint es sinnvoll, zwischen den eigentlichen Gütekriterien (z.B. Diskriminationsfähigkeit, Risikoadjustierung, etc.) und deren Operationalisierung, etwa mittels eines Instru-

ments wie QUALIFY oder modifiziertes RAND, zu unterscheiden. So bedienen sich beispielsweise unterschiedliche Instrumente ähnlicher Gütekriterien, operationalisieren diese jedoch teilweise auf unterschiedliche Weise.

- ***Forschungsbedarf: systematische Validierung eines oder mehrerer der eingesetzten Instrumente zur Bewertung der methodischen Güte von Indikatoren***

4.2 Bewertung der Eignung von Qualitätsindikatoren für bestimmte Zielsetzungen und Einsatzgebiete

Die BQS hat das Qualify-Instrument genutzt, um die Eignung von im Routinebetrieb eingesetzten Indikatoren für ein einrichtungsbezogenes Public Reporting zu bewerten. Für diese spezifische Zielsetzung wurden 14 der 20 Qualify-Kriterien ausgewählt und es wurden Annahmeschwellen definiert, ab wann ein Indikator unter methodischen Gesichtspunkten als geeignet für Public Reporting anzusehen ist [12]. Das AQUA-Institut hat dieses Vorgehen modifiziert [32]. In beiden Fällen wurden ausgewählte Gütekriterien verwendet und mittels eines individuell festgelegten Bewertungsschemas unter Einbeziehung von Expertenpanels bewertet.

Das IQTIG spricht in seinem Methodenpapier von „Eignungskriterien“ für verschiedene „Verwendungszwecke“. Das Konzept des IQTIG differenziert daher nicht primär zwischen „Eigenschaften“ und „Eignung“, sondern verwendet letztlich einen integrierten Bewertungsansatz [15].

- ***Forschungsbedarf: systematische Validierung eines oder mehrerer der eingesetzten Konzepte zur Eignungsprüfung für unterschiedliche Verwendungszwecke***

Insbesondere erscheint es relevant, Vor- und Nachteile des integrierten Bewertungsansatzes des IQTIG mit einer differenzierten Bewertung von „Eigenschaften“ und „Eignung“ zu vergleichen und zu bewerten.

Darüber hinaus erscheint es empfehlenswert, einen einheitlichen Sprachgebrauch darüber herzustellen, was als „Einsatzgebiet“ bzw. „Verwendungszweck“ von Qualitätsindikatoren anzusehen ist. Die von der AHRQ vorgeschlagenen und vom IQTIG grundsätzlich aufgenommenen drei Kategorien (Qualitätsverbesserung, Steuerung, Forschung) sind sicherlich als akzeptierte Konvention anzusehen, erscheinen allerdings für den praktischen Einsatz und für Forschungsaktivitäten nicht ausreichend differenziert. So liegt beispielsweise nahe, dass unterschiedliche Eignungskriterien für (möglicherweise unterschiedliche Formen von) Pay-for-Performance und für Public Reporting sachgerecht sind, obwohl beide Einsatzgebiete der Kategorie „Steuerung“ zuzuordnen sind.

Das IQTIG differenziert in seinem aktuellen Methodenpapier für den „Kontext der externen Qualitätssicherung nach SGB V“ zwischen „Qualitätsförderung, Information, selektivvertraglichen Anreizen, Planung der Versorgungsstrukturen und Qualitätszu- und –abschlägen“ [15].

Es erscheint empfehlenswert, auch über die Vorgaben des SGB V hinaus einen Rahmen zu definieren, der für Einsatzgebiete, Handlungskonsequenzen (Interventionen) und Beobachtungsebenen eine Grundlage für einheitlich verwendbare Begriffe schafft

- **Entwicklungsbedarf: Erarbeitung eines Rahmenkonzepts zu verschiedenen Zielsetzungen, Beobachtungs- und Handlungsebenen von Qualitätsindikatoren und Qualitätsindikatorensets**

4.3 Bewertung der Eigenschaften (Ausrichtung und Ausgewogenheit) von Indikatorensets

Konsentiertere oder anerkannte Kriterien zur Bewertung der Güte von Indikatorensets stehen derzeit nicht zur Verfügung. Verbreitet sind allerdings Zuordnungen von Qualitätsindikatoren zu sogenannten Qualitätsdimensionen oder Domänen [14, 15, 17, 33-35].

Wohl am bekanntesten und gebräuchlichsten ist die Einteilung nach Donabedian in die sog. „Dimensionen“ Struktur-, Prozess- und Ergebnisqualität [33]. Diese Einteilung erlaubt aber keine Aussage zur Ausgewogenheit in Bezug auf die inhaltliche Ausrichtung von Indikatorensets, da beispielsweise durch ein Indikatorensatz allein die Dimension der Sicherheit (nach dem Rahmenkonzept der OECD) angesprochen werden kann – aber möglicherweise eben mit je einem Indikator zur Strukturqualität, einem zur Prozess- und einem zur Ergebnisqualität.

Um die inhaltliche Ausgewogenheit von Indikatorensets zu beurteilen, könnten die von einer Arbeitsgruppe der OECD konsentierten Qualitätsdimensionen genutzt werden [34, 35].

Damit wird eine Zuordnung zu folgenden zentralen Dimensionen („core dimensions“) möglich:

- Effektivität (werden die Behandlungsziele erreicht?),
- Sicherheit (werden unerwünschte Ereignisse vermieden?),
- Indikationsstellung und
- Patientenzentrierung

Aktuell wurde das Modell der Qualitätsdimensionen der OECD als methodischer Ansatz zur Bewertung der inhaltlichen Schwerpunktsetzung und Ausgewogenheit von Indikatorensets genutzt und an den Indikatorensatz der stationären Qualitätssicherung nach §136 SGB V überprüft [17]. Grundsätzlich erscheint der genutzte Ansatz geeignet, eine orientierende Aussage zur Ausrichtung von Indikatorensatz vorzunehmen. Es können auch andere Modelle genutzt werden, auf die die im Folgenden formulierten Fragestellungen vom Grundsatz her analog anzuwenden sind [36-38]. Auch das IQTIG hat in der aktuellen Version seiner Methodischen Grundlagen hierzu ein eigenes Modell vorgeschlagen [15].

Sowohl für die Dimension der Effektivität als auch die Dimension der Sicherheit können Core Outcome Sets ein hilfreiches Instrument darstellen, um festzustellen, wie umfassend die jeweiligen Dimensionen abgebildet werden [26, 27]. Die Bewertung der Abbildung der jeweiligen Dimensionen ist zunächst unabhängig davon, ob Prozess- oder Ergebnisindikatoren eingesetzt werden. Die Evidenzbasierung von Prozessindikatoren drückt sich darin aus, dass jeder Indikator sich auf mindestens einen patientenrelevanten Endpunkt bezieht (dies gilt auch für Indikatoren zur Indikationsstellung). Somit können sowohl Prozess- als auch Ergebnisindikatoren anhand der adressierten patientenrelevanten Endpunkte den einzelnen Dimensionen zugeordnet werden. Die Abbildung der Dimension der Sicherheit kann somit beispielsweise nicht nur durch die Erfassung der „Abwesenheit von unerwünschten Ereignissen“ son-

dern auch durch die Erfassung der „Eigenschaften von Organisationen und Systemen“ erfolgen [36].

Damit erscheint die Diskussion der Vor- und Nachteile von Prozess- und Ergebnisindikatoren für diese Betrachtung als zunächst nachgelagert – auch wenn sie grundsätzlich von hoher Bedeutung ist (insbesondere vor dem Hintergrund der Grenzen der Adjustierbarkeit von Ergebnisindikatoren).

Auf der Basis der ersten Erprobungsergebnisse des OECD-Modells zur Beurteilung der Ausgewogenheit von Indikatorensets stellen sich folgende weitergehende Forschungsfragen:

- Welche Dimensionen können prinzipiell als geeigneter Rahmen für eine Bewertung der Ausgewogenheit und Ausrichtung von Indikatorensets angesehen werden?
- Welcher Dimension wird die Sterblichkeit zugeordnet?
 - Wenn die Dimensionen der OECD genutzt werden, muss die Zuordnung der Mortalität geklärt werden. Dieser Endpunkt kann abhängig vom Zeitpunkt der Erfassung, dem Gesundheitszustand des Patienten und von der Dringlichkeit eines Eingriffs der Dimension „Vermeidung unerwünschter Ereignisse“ oder der Dimension „Erreichen primärer Behandlungsziele“ zugeordnet werden. Für diagnosebezogene Verfahren kann die Sterblichkeit abhängig von der betrachteten Behandlungsepisode verschiedenen Dimensionen zugeordnet werden. Möglicherweise ist es vor diesem Hintergrund eine Option, solche Indikatoren als „Sterblichkeit“ zu bezeichnen, wenn sie unerwünschte Ereignisse abbilden (Zuordnung zur Dimension „Sicherheit“) und „Überleben“, wenn sie das Erreichen von Behandlungszielen erfassen (Zuordnung zur Dimension „Effektivität“).
 - Alternativ kann auch eine Mehrfachzuordnung von Indikatoren vorgenommen werden (ein Indikator deckt mehrere Dimensionen ab).
- In welcher Tiefe werden die einzelnen Dimensionen durch ein Indikatorenset abgebildet?
 - Ein Indikatorenset kann ggf. einen Aspekt innerhalb einer Dimension mit einem Indikator adressieren (z.B. eine Komplikation, zugeordnet zur Dimension Sicherheit) und somit die Dimension formal abdecken. Möglicherweise erfasst ein Indikator jedoch nur einen marginalen Ausschnitt dieser Dimension (z.B. ausschließliche Verwendung eines einzelnen Indikators zu einem Teilaspekt der Antibiotikaprophylaxe bei einem bestimmten Eingriff für die Dimension der Sicherheit)
 - Es bedarf daher neben der Erfassung der Anzahl der Indikatoren für die einzelnen Dimensionen auch einer Bewertung des Umfangs und der Qualität der Abdeckung durch die Indikatoren.
 - Hierzu ist theoretisch ein Abgleich mit einem Goldstandard, einem „idealen“ oder „umfassenden“ Indikatorenset sinnvoll.

Hierbei können „Core outcome sets“, z. B. die Sets des International Consortium for Health Outcomes (ICHOM) oder der COMET Initiative [37] und evidenzbasierte starke Leitlinienempfehlungen eine wichtige Rolle spielen, um ein solches „ideales Indikatorenset“ zu entwerfen.

4.4 **Bewertung: Eignung von Indikatorensets für bestimmte Zielsetzungen und Einsatzgebiete**

Indikatorensets, die alle Dimensionen adressieren, liefern einen Blick auf ein prinzipiell umfassendes und ausgewogenes Konstrukt der Qualität zu einem bestimmten Thema. Indikatorensets, die sich ggf. auf eine Dimension konzentrieren, sind möglicherweise für eine definierte Zielsetzung der Qualitätsdarstellung ebenfalls sachgerecht und geeignet, wenn diese Dimension eine hohe Priorität darstellt.

Auch bei „umfassenden“ Indikatorensets werden zwangsläufig niemals alle Aspekte der Gesundheitsversorgung erfasst werden können. Ziel sollte jedoch sein, ein Instrumentarium zu schaffen, mit dem die Stärken und Schwächen von Indikatorensets systematisch transparent gemacht werden können.

Um bei der Bewertung der Eignung von Indikatorensets für bestimmte Zielsetzungen und Einsatzgebiete voran zu kommen, sollte daher folgender Entwicklungs- bzw. Forschungsbedarf in den Fokus genommen werden:

- Rahmen für einheitliche Begriffsverwendung (siehe 4.2):
 - Abstimmung einer Konvention zur differenzierten Benennung der jeweiligen Einsatzgebiete
 - Abstimmung einer Konvention zur differenzierten Benennung der jeweiligen Beobachtungs- / Handlungsebenen;
- Entwicklung von Kriterien zur Bewertung der Eignung von Indikatorensets für bestimmte Einsatzgebiete (z.B. Public Reporting, P4P);

4.5 **„Composite Measures“**

Weiterhin besteht Forschungsbedarf zur Eignung von Composite Indicators / Composite Measures / Scores / Indices (im Folgenden: „Composite Measures“)..

Während Indikatorensets mehrere Indikatoren (mit jeweils einzelnen Ergebnissen) zu einem Thema nebeneinander darstellen, werden in „Composite Measures“ die Ergebnisse mehrerer einzelner Indikatoren / Kennzahlen zu einem Gesamtergebnis aggregiert.

Auch hier besteht Bedarf, ein einheitliches Begriffsverständnis herzustellen.

Die bereits langjährige Diskussion der Thematik hat international zuletzt bezüglich des PSI90-Indikatorsets im Rahmen des Value Based Purchasing (P4P) erheblich an inhaltlicher Tiefe und Relevanz gewonnen [40, 41].

Bereits einzelne Qualitätsindikatoren bilden komplexe Konstrukte der Gesundheitsversorgung „in einer Zahl“ ab (beispielsweise „Indikationsstellung zu einem Eingriff“). „Composite Measures“ nehmen eine noch stärkere Aggregation vor. Dies erscheint auf der einen Seite „verlockend“, da es zu einer verbesserten Übersichtlichkeit und scheinbar einfacheren Verständlichkeit führt, birgt allerdings relevante Risiken, vor allem in Bezug auf eine zu starke Aggregation im Sinne einer nicht angemessenen Vereinfachung [36,42].

Forschungs- und Entwicklungsbedarf besteht daher bei der Bildung von solchen aggregierten Kennzahlen vor allem in Bezug auf die Möglichkeiten, Vor- und Nachteile der Aggregation verschiedener Dimensionen / Domänen (z.B. von Indikatoren zur Prozess- und Ergebnisquali-

tät oder von Indikatoren zur Indikationsstellung und der Ergebnisqualität), in Bezug auf die Interpretierbarkeit und den Einsatz für regulierende Interventionen, wobei ein besonderes Augenmerk auf die Problematik der Gewichtung der Indikatoren innerhalb der Composite Measures gelegt werden sollte [42].

Zusammenfassend erfordert die aktuelle Situation – Weiterentwicklung in Richtung auf eine sektorübergreifende, populationsorientierte Qualitätsbewertung mit teilweise verstärkt steuernder Intention – eine Verstärkung der Forschung zur Güte und Eignung von Indikatorensets und den darin verwendeten einzelnen Indikatoren. Die Forschung sollte eingebettet sein in ein methodisches Rahmenkonzept, in dem Konventionen zu „Einsatzgebieten“ und „Handlungs- und Beobachtungsebenen“ von Qualitätsindikatoren sowie zu „Instrumenten / Interventionen“ (d.h. zu Steuerungsinstrumenten) festgelegt werden.

Literatur

1. Beck T. 10 Jahre Qualitätssicherung in der stationären Versorgung – ein institutioneller Rückblick. In Rebscher H (Hrsg.). Gesundheitsökonomie und Gesundheitspolitik im Spannungsfeld zwischen Wissenschaft und Politikberatung. Heidelberg, München, Landsberg, Berlin: Economica Verlag 2006: 659-675.
2. Bundesanzeiger. Bekanntmachung eines Beschlusses des Gemeinsamen Bundesausschusses über eine Änderung der Vereinbarung gemäß § 137 Abs. 1 Satz 3 Nr. 6 des Fünften Buches Sozialgesetzbuch (SGB V) über Inhalt und Umfang eines strukturierten Qualitätsberichts für nach § 108 SGB V zugelassene Krankenhäuser – Darstellung der Ergebnisse des BQS-Verfahrens. BAnz. Nr. 202a (Beilage S. 71) vom 27.10.2007.
3. Bundesgesetzblatt. Gesetz zur Reform der Strukturen der Krankenhausversorgung (Krankenhausstrukturgesetz – KHSVG) vom 10. Dezember 2015, Bundesgesetzblatt Jahrgang 2015 Teil I Nr. 51, ausgegeben zu Bonn am 17. Dezember 2015. Im Internet: https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl115s2229.pdf#bgbl_%2F%2F%5B%40attr_id%3D%27bgbl115s2229.pdf%27%5D_1498725224998 (Zugriff 19.8.2018).
4. Sens B, Pietsch B, Fischer B et al. Begriffe und Konzepte des Qualitätsmanagements - 4. Auflage. GMS Med Inform Biom Epidemiol 2018;14(1):Doc04.
5. Agency for Healthcare Research and Quality. Uses of Quality Measures. Im Internet (Page Last Reviewed or Updated: July 15, 2016) https://www.qualityindicators.ahrq.gov/FAQs_Support/FAQ_QI_Overview.aspx# (Zugriff 22.4.2019).
6. Etzioni DA, Wasif N, Dueck AC, Cima RR, Hohmann SF, Naessens JM, Mathur AK, Habermann EB. Association of Hospital Participation in a Surgical Outcomes Monitoring Program with Rates of Inpatient Complications. JAMA 2015; 313: 505-511.
7. Osborne NH, Nicholas LH, Ryan AM, Thumma JR, Dimick JB. Association of Hospital Participation in a Quality Reporting Program With Surgical Outcomes and Expenditures for Medicare Beneficiaries. JAMA 2015; 313: 496-504.
8. Berwick DM. Measuring Surgical Outcomes for Improvement. Was Codman Wrong? JAMA 2015; 313: 469-70.
9. Koalitionsvertrag zwischen CDU, CSU und SPD - Ein neuer Aufbruch für Europa. Eine neue Dynamik für Deutschland. Ein neuer Zusammenhalt für unser Land. 19. Legislaturperiode. Im Internet: https://www.cdu.de/system/tdf/media/dokumente/koalitionsvertrag_2018.pdf?file=1 (Zugriff 19.8.2018).
10. Expertengruppe der Friedrich-Ebert-Stiftung. Qualität in einem sektorübergreifenden Gesundheitswesen. Über die Bedeutung der Qualität für eine gute gesundheitliche Versorgung. Im Internet: <http://library.fes.de/pdf-files/wiso/14885.pdf> (Zugriff 4.4.2019).
11. Rode S, Ries V, Petzold T, Buch U, Untersweg F, Fischer B. Anforderungen und Anwendungshinweise für den Einsatz von Qualitätsindikatoren in der medizinischen Versorgung: Ergebnisse eines systematischen Reviews. Gesundheitswesen 2016; 78(S 01): e110-e119.

12. Reiter A, Fischer B, Kötting J, Geraedts M, Jäckel WH, Döbler K. QUALIFY--a tool for assessing quality indicator. Z Arztl Fortbild Qualitatssich. 2007;101(10): 683-688.
13. De Koning J. Development and validation of a measurement instrument for appraising indicator quality: Appraisal of Indicators through Research and Evaluation (AIRE) instrument (06.09.2007). www.egms.de/de/meetings/gmds2007/07gmds798.shtml (Zugriff 22.4.2019)
14. AQUA-Institut. Allgemeine Methoden im Rahmen der sektorenübergreifenden Qualitätssicherung im Gesundheitswesen nach §137a SGB V Version 4.0. Stand: 17. Februar 2015. Im Internet: https://www.aqua-institut.de/fileadmin/aqua_de/Projekte/248_Methodenpapier/Methodenpapier_4.0.pdf (Zugriff 22.4.2019).
15. Institut für Qualitätssicherung und Transparenz im Gesundheitswesen (IQTIG). Methodische Grundlagen V1.1. Stand: 15. April 2019. Im Internet: https://iqtig.org/dateien/dasiqtig/grundlagen/IQTIG_Methodische-Grundlagen-V1.1_2019-04-15.pdf (Zugriff 22.4.2019).
16. Jones P, Sheperd M, Wells S, Le Fevre J, Ameratunga S. Review Article: what makes a good healthcare quality indicator? A systematic review and validation study. Emerg Med Australas 2014; 26: 113 – 124.
17. Doeblner K, Geraedts M. Ausgewogenheit der Qualitätsindikatorensets der externen Qualitätssicherung nach §136 SGB V. Z Evid Fortbild Qual Gesundhwes 2018; 134: 9–17.
18. NHS Institute for Innovation and Improvement. The good indicators guide. January 2008. Im Internet: <http://webarchive.nationalarchives.gov.uk/20170106081009/http://www.apho.org.uk/resource/item.aspx?RID=44584> (Zugriff 19.8.2018).
19. Gutt A-K, Hoben M, Roes M, Willmeroth T, Wesselborg B, Kuske S. Systematische Übersetzung und Cross-Validierung definierter Implementierungsergebnisse der Gesundheitsversorgung. Z Evid Fortbild Qual Gesundhwes 2018;135-136: 72–80.
20. Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen. Bedarfsgerechte Steuerung der Gesundheitsversorgung. Gutachten 2018 https://www.svr-gesundheit.de/fileadmin/user_upload/Gutachten/2018/SVR-Gutachten_2018_WEBSEITE.pdf (Zugriff 8.6.2019)
21. IQTIG – Institut für Qualitätssicherung und Transparenz im Gesundheitswesen. Aktualisierung und Erweiterung des QS-Verfahrens Versorgung von volljährigen Patienten und Patientinnen mit Schizophrenie, schizotypen und wahnhaften Störungen. Abschlussbericht. Stand: 22. Dezember 2017 https://iqtig.org/downloads/berichte/2017/IQTIG_QS-Verfahren-Schizophrenie_Abschlussbericht_2017-12-22.pdf (Zugriff 8.6.2019)
22. Jencks SF, Cuerdon T, Burwen DR, Fleming B, Houck PM, Kussmaul AE, Nilasena DS, Ordian DL, Arday DR. Quality of Medical Care Delivered to Medicare Beneficiaries A Profile at State and National Levels. JAMA 2000; 284: 1670-1676.
23. Jencks SF, Huff ED, Cuerdon T. Change in the Quality of Care Delivered to Medicare Beneficiaries, 1998-1999 to 2000-2001. JAMA 2003; 289: 305-312.

24. Kuske S, Maass C, Weingärter V, Pöhlmann S, Schrappe M. Patient-safety indicators. A systematic review, criteria-based characterization and prioritization. *J Public Health* 2013;21(2): 201–214.
25. Geraedts M, Drösler SE, Döbler K, Eberlein-Gonska M, Heller G, Kuske S, Manser T, Sens B, Stausberg J, Schrappe M. DNVF-Memorandum III „Methoden für die Versorgungsforschung“, Teil 3: Methoden der Qualitäts- und Patientensicherheitsforschung. *Gesundheitswesen* 2017; 79: e95-e124.
26. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT Filter for Outcome Measures in Rheumatology. *J Rheumatol* 1998; 25(2): 198-199.
27. Schmitt J, Petzold T, Eberlein-Gonska M, Neugebauer EA. Anforderungsprofil an Qualitätsindikatoren. Relevanz aktueller Entwicklungen der Outcomes Forschung für das Qualitätsmanagement. *Z Evid Fortbild Qual Gesundheitswes* 2013;107(8): 516-22.
28. Fitch, K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, van het Loo M, McDonnell J, Vader JP, Kahan JP. The RAND/UCLA Appropriateness Method User's Manual. Santa Monica: RAND 2001.
29. Kerr EA, Asch SM, Hamilton EG, McGlynn EA (eds.). *Quality of Care for General Medical Conditions: A Review of the Literature and Quality Indicators*. Santa Monica, CA: RAND Corporation, 2000. Im Internet: http://www.rand.org/pubs/monograph_reports/MR1280 (Zugriff 5.4.2019).
30. Programm für Nationale VersorgungsLeitlinien. Methodenreport. 5. Auflage, 2017. Im Internet: <https://www.leitlinien.de/mdb/downloads/nvl/methodik/nvl-methodenreport-5aufl-vers1.pdf> (Zugriff 22.4.2019).
31. Lungen M, Rath T. Analyse und Evaluierung des QUALIFY Instruments zur Bewertung von Qualitätsindikatoren anhand eines strukturierten qualitativen Interviews. *Z Evid Fortbild Qual Gesundheitswes* 2011; 105/1: 38–43.
32. AQUA-Institut. Bericht zur Prüfung und Bewertung der Indikatoren der externen stationären Qualitätssicherung hinsichtlich ihrer Eignung für die einrichtungsbezogene öffentliche Berichterstattung. Stand: 17. April 2015. Im Internet: <https://www.aqua-institut.de/projekte/bericht-zur-qi-pruefung/> (Zugriff 19.8.2018).
33. Donabedian A. Evaluating the quality of medical care. *Milbank Q* 2005; 83(4): 691-729.
34. Arah OA, Westert GP, Hurst J, Klazinga NS. A conceptual framework for the OECD Health Care Quality Indicators Project. *Int J Qual Health Care* 2006; 18 Suppl 1: 5-13.
35. Carinci F, van Gool K, Mainz J, Veillard J, Pichora EC, Januel JM, Arispe I, Kim SM, Klazinga NS on behalf of the OECD Health Care Quality Indicators Expert Group. Towards actionable international comparisons of health system performance: expert revision of the OECD framework and quality indicators. *Int J Qual Health Care* 2015; 27(2): 137-146.
36. Schrappe M. APS-Weißbuch Patientensicherheit. Sicherheit in der Gesundheitsversorgung: neu denken, gezielt verbessern. Herausgegeben vom Aktionsbündnis Patientensicherheit e.V. (APS) 2018
https://www.aps-ev.de/wp-content/uploads/2018/08/APS-Weissbuch_2018.pdf
(Zugriff 8.6.2019)
37. Schrappe M. *Qualität 2030*. MWV Medizinisch Wissenschaftliche Verlagsgesellschaft Berlin, 2015

http://matthias.schrapppe.com/index_htm_files/Schrapppe_Qualitaet_2030_165x240_Download.pdf (Zugriff 8.6.2019)

38. Agency for Health Care Research and Quality. Organizing Measures by Quality Domain. Last reviewed March 2016. Im Internet: <https://www.ahrq.gov/professionals/quality-patient-safety/talkingquality/create/organize/qualitydomain.html> (Zugriff 19.8.2018).
39. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, Clarke M, Gargon E, Gorst S, Harman N, Kirkham JJ, McNair A, Prinsen CAC, Schmitt J, Terwee CB, Young B. The COMET Handbook: version 1.0. *Trials* 2017; 18(Suppl 3):280.
40. Rajaram R., Barnard C, Bilimoria KY. Concerns About Using the Patient Safety Indicator-90 Composite in Pay-for-Performance Programs. *JAMA* 2015; 313: 897-900.
41. Goddard M, Jacobs R. Using Composite Indicators to Measure Performance in Health Care. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S (eds.): *Performance Measurement for Health System Improvement. Experiences, Challenges and Prospects*. Cambridge: Cambridge University Press 2010: 339-368.
42. Friebel R, Steventon A. Composite measures of healthcare quality: sensible in theory, problematic in practice. *BMJ Qual Saf* 2019;28:85-88.