Prof. Dr. Matthias Schrappe

Das Methodenpapier des IQTIG: keine Kursänderung in Sicht

Die ex post-Qualitätskontrolle bleibt die vorherrschende Doktrin

Zusammenfassung

Das nun in einem ersten Entwurf vorliegende Methodenpapier des Institutes für Qualität und Transparenz im Gesundheitswesen (IQTIG) ist ein wichtiges Dokument für die Diskussion des Themas in den nächsten Jahren. Allerdings bietet dieses Papier Anlaß zu ernsthafter Kritik auf zahlreichen Ebenen. In seinen Grundannahmen bleibt es einem traditionellen, auf die ex post-Qualitätskontrolle von Ergebnissen ausgerichteten Verständnis von Qualität verhaftet. Elemente der kontinuierlichen Qualitätsverbesserung mit Integration des Qualitätsgedankens in den Ablauf der Leistungserbringung auf institutioneller und Systemebene sind weitgehend in den Hintergrund gedrängt.

Schlüsselwörter

Methodenpapier IQTIG, ex post-Qualitätskontrolle, Qualitätsverständnis, Indikatorkonzept

Das nun in einem ersten Entwurf vorliegende Methodenpapier des Institutes für Qualität und Transparenz im Gesundheitswesen (IQTIG) ist ein wichtiges Dokument für die Diskussion des Themas in den nächsten Jahren. Allerdings bietet dieses Papier Anlaß zu ernsthafter Kritik auf zahlreichen Ebenen. In seinen Grundannahmen bleibt es einem traditionellen, auf die ex post-Qualitätskontrolle von Ergebnissen ausgerichteten Verständnis von Qualität verhaftet. Elemente der kontinuierlichen Qualitätsverbesserung mit Integration des Qualitätsgedankens in den Ablauf der Leistungserbringung auf institutioneller und Systemebene sind weitgehend in den Hintergrund gedrängt. Der Begriff der Patientenzentrierung wird in paternalistischer Tradition einseitig auf die Ergebnisqualität bezogen, die diskutierten Qualitätsdimensionen sind auf Fragen der Legitimität und statistischen Auswertbarkeit reduziert. Es wird das Konzept der linearen Messung von Qualität durch Qualitätsindikatoren vertreten, Indikatoren werden entgegen der international und in der Versorgungsforschung üblichen Auffassung nicht als hoch-sensitive Monitoring-Instrumente, sondern als klassische Messinstrumente eingesetzt, die gleichzeitig für die Erfassung von Qualitätsproblemen und für die Identifikation von „exzellenter" Qualität zuständig sind. Die Begriffe der Reliabilität und Validität von Indikatoren werden nicht zutreffend verwendet, es wird sogar die Behauptung aufgestellt, es gäbe keine sinnvolle Anwendung des Begriffes Validität auf Qualitätsindikatoren. Hinzu kommen weitere handwerkliche Fehler, so wird Indikatoren eine Bewertungsfunktion zugewiesen und der Begriff der Qualitätssicherung nicht vom Begriff der Qualitätsverbesserung und von Instrumenten wie Pay for Performance differenziert.

>> Das Methodenpapier gliedert sich in drei Teile: Grundlagen (Teil A), Entwicklung und Durchführung (Teil B) und methodische Elemente (Teil C). Es werden zahlreiche grundlegende Definitionen verwendet, die nicht dem gängigen wissenschaftlichen und praktischen Verständnis entsprechen (s. „Curriculum Qualitätsmanagement" der Bundesärztekammer, der kassenärztlichen Bundesvereinigung und der Arbeitsgemeinschaft der Medizinisch-wissenschaftlichen Fachgesellschaften (AWMF) (Bundesärztekammer 2007), die aktuelle Ausgabe der DIN-Vorschriften (s. DIN-TERM online 2017), die aktuellen Begriffsbestimmungen der gemeinsamen Arbeitsgruppe der Gesellschaft für Qualitätsmanagement in der Gesundheitsversorgung (GQMG) und der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie (GMDS) (das sog. „QM-Glossar" (Sens et al. 2007), derzeit in Überarbeitung) und die internationale Literatur (z.B. Nomenklatur der Agency of Healthcare Research and Quality (AHRQ 2017)).

Selbstverständlich darf nicht vergessen werden, dass das Institut an die Ausführungen des Gesetzgebers und die Anforderungen des Gemeinsamen Bundesausschusses (G-BA) gebunden ist. Trotzdem wäre zu fordern, dass das Institut die notwendige Fortentwicklung des statischen, auf die Sicherung von Qualität ausgerichteten Verständnisses zu einem international anschlussfähigen Konzept der Qualitätsverbesserung (quality improvement) nicht nur anspricht, sondern in den Mittelpunkt der perspektivischen Entwicklung stellt.

Qualitätsverständnis: Qualitätskontrolle aus der ex post-Perspektive

Das im Methodenpapier vertretene Qualitätsverständnis stellt anders als die führenden, international gängigen Ansätze (Donabedian 1990, Arah et al. 2006, AHRQ 2017) einseitig die Dimension der „Legitimität" in den Vordergrund, die als Meta-Kriterium konstruiert wird (*„Eine objektive, Vergleiche ermöglichende Darstellung von Versorgungsqualität setzt voraus, dass eine begrenzte Anzahl legitimer Anforderungen an die Versorgung identifiziert und explizit gemacht wird."* (S. 17)). Hierunter wird Patientenzentrierung, Beeinflussbarkeit durch die Leistungserbringer und Unbedenklichkeit verstanden. Diese Auswahl bedeutet eine schwerwiegende Verengung des Fokus und muss zusätzlich auch in ihrer Ausführung kritisch gesehen werden:

- **Patientenzentrierung:** Der auffallend häufig verwendete Begriff der Patientenzentrierung wird im Methodenpapier ausschließlich auf die Ergebnisqualität bezogen, obwohl Patienten in viel stärkerem Maße an Fragen der Strukturqualität (z.B. Ausbildung der Ärzte) und an Kooperation, Koordination und Information (also Prozessparametern) interessiert sind (aus der umfangreichen Literatur: Geraedts und de Cruppé 2011, Schoen et al. 2011). So gut Patientenzentrierung heutzutage klingt, es wäre zu erwarten gewesen, dass man neben diesen präferenz-sensitiven Aspekten mindestens noch die Dimensionen der Wirksamkeit und der Anbietersensitivität mit einschließt (Wennberg et al. 2002, s. auch Frosch et al. 2010). Die einseitige Betonung der Ergebnisqualität führt zu einer reinen ex post-Betrachtung im Sinne eines Qualitätskontroll-Ansatzes, statt eine Integration von Qualitätsaspekten in den Prozess der Leistungserbringung im Sinne eines kontinuierlichen Qualitätsverbesserungsansatzes (continuous quality improvement) zu fördern.
- **Beeinflussbarkeit durch die Leistungserbringer:** Dieses Kriterium ist genau genommen keine Qualitätsdimension, sondern stammt aus der Bewertung von Qualitätsindikatoren.
- **Die Dimension der Unbedenklichkeit:** Diese erscheint zwar auf den ersten Blick sinnvoll, ist aber international und in den in Deutschland gebräuchlichen Systematiken an keiner Stelle vorzufinden. Natürlich ist nichts dagegen zu sagen, dass *„die gestellten Anforderungen (…) keine für die Patientinnen und Patienten schädlichen Nebenwirkungen*

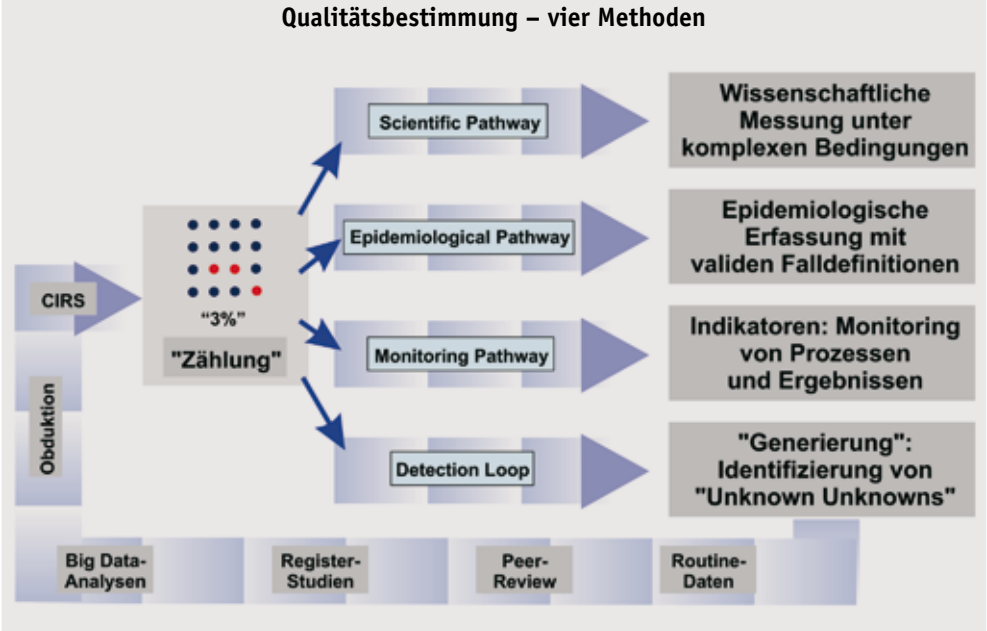


Abb. 1: Qualitätsbestimmung – vier Methoden (s. Text) (aus Schrappe 2016).

entfalten" dürfen. Es stellt sich allerdings die Frage, weshalb dann nicht Wirksamkeit oder Sicherheit im Sinne der gebräuchlichen Systematiken direkt genannt werden.

Wie sehr der Gesichtspunkt der juristischen bzw. politischen Absicherung im Vordergrund steht, wird auch durch die starke Betonung der statistischen Methodik klar. Gleich im Anschluss an die Ausführungen zur Legitimität heißt es hierzu (S. 19, Hervorh. der Verf.):

„Die Fokussierung auf definierte, legitime Anforderungen ist eines der zentralen Charakteristika der externen Qualitätssicherung. ***Ein zweites ist die primär statistische Herangehensweise an die Messung und Bewertung der Versorgungsqualität*** *der einzelnen Leistungserbringer im Vergleich zueinander."*

Statistische Fragen sind natürlich von großer Wichtigkeit, sie stellen aber gegenüber der Zielformulierung und der Beschreibung der inhärenten Merkmale keine primären Aspekte eines Qualitätssverständnisses dar, sondern sind in erster Linie technischer Natur.

Begriff Qualitätssicherung

Der Begriff der Qualitätssicherung wird nicht sachgerecht verwendet. Das IQTIG versteht darunter nicht den „Teil des Qualitätsmanagements, der auf das Erzeugen von Vertrauen darauf gerichtet ist, dass Qualitätsanforderungen erfüllt werden" (Sens et al. 2007 und DIN EN ISO 9000:2015-11), sondern bezieht die *„Weiterentwicklung der Versorgungsqualität"* mit ein (S. 15). Der Verbesserungsgedanke, so wichtig er im aktuellen gesundheitspolitischen Kontext erscheint, ist kein Bestandteil der Definition von Qualitätssicherung, sondern fällt unter den Begriff der Qualitätsverbesserung. Auf der institutionellen Ebene ist der Aspekt der Verbesserung im Konzept des Qualitätsmanagements aufgehoben, auf der Systemebene steht heute international der Begriff des Quality Improvement (QI) bzw. Improvement of Health Care Performance im Vordergrund (z.B. Berwick et al. 2003) und ist als Improvement Science ein wichtiger Bestandteil der Versorgungsforschung (Berwick 2008, Marshall et al. 2013). In der Nomenklatur der AHRQ kommt der dem Begriff der Qualitätssicherung entsprechende englischsprachige Begriff „Quality Assurance" gar nicht mehr vor, sondern ist vollständig durch den Begriff des Quality Improvement verdrängt (ARHQ 2017).

In einem wissenschaftlichen Methodenpapier muss auf die sachlich korrekte Verwendung der zentralen Begriffe Wert gelegt werden, zumindest durch eine differenzierte Darstellung und Interpretation. Um den Gedanken der Qualitätsverbesserung in den derzeit im deutschen Gesundheitssystem wirksamen bzw. geplanten Maßnahmen wie Public Reporting, qualitätsorientierte Vergütung und qualitätsorientierte Krankenhausplanung hervorzuheben, hat daher das DNVF das 5. DNVF-Forum im Mai 2017 unter den (englischsprachigen) Begriff des „Quality Improvement" gestellt.

Bestimmung und Erfassung von Qualität

Im Methodenpapier wird zugrunde gelegt, dass sich die Qualität der Gesundheitsversorgung ähnlich wie biomedizinische Marker durch lineare Messinstrumente quantifizieren lässt („Die Messung der Qualität medizinischer Versorgung erfolgt grundsätzlich mittels Qualitätsindikatoren" (S. 21)). Die Indikatoren können nach dem Konzept des IQTIG verschiedene Ausprägungen einnehmen, die verwendeten Begriffe zur Abstufung von Qualität reichen von „ungenügend", „nicht ausreichend", „durchschnittlich gut" über „gut" bis zu „außerordentlich gut" und sogar „Exzellenz" (S. 19).

Aus Sicht der Versorgungsforschung kann man dem Konzept einer linearen Qualitäts-„messung" jedoch nicht folgen. Wie bereits einleitend ausgeführt (s. Abb. 1), sind grundsätzlich bei der Quantifizierung des Grades, in dem inhärente Merkmale die Anforderungen erfüllen (so die DIN-Definition von Qualität),

- wissenschaftsgetriebene Verfahren,
- der Einsatz klinisch-epidemiologischer Falldefinitionen,
- der Einsatz von Indikatoren zu Zwecken des Monitoring und
- die generierenden Verfahren (z.B. CIRS)

zu unterscheiden. Nur die beiden erstgenannten Verfahren kommen einer Messung nahe. Der wissenschaftliche Zugang ist auf der Basis einer Modellbildung bemüht, ein Messinstrument mit gleichermaßén hoher Sensitivität und Spezifität einzusetzen, den Messvorgang selbst als (komplexe) Intervention zu verstehen und die Komplexität des Kontextes der Messung zu berücksichtigen (z.B. Messung der Händedesinfektionscompliance durch direkte Beobachtung). Der Aufwand einer solchen Messung ist sehr hoch, so dass solche Verfahren nur zur Etablierung einer Methode oder in besonders dringenden Situationen eingesetzt werden. Auch beim Einsatz klinisch-epidemiologischer Falldefinitionen (z.B. Erfassung von postoperativen Wundinfektionen) ist der Erhebungsaufwand so hoch, dass man den Einsatz dieser Instrumente auf definierte Ereignisgruppen (z.B. nosokomiale Infektionen) beschränkt. Allerdings sind diese klinisch-epidemiologischen Instrumente einfacher einzusetzen als die primär Wissenschafts-getriebenen Ansätze, man ist nicht bei jedem Einsatz darauf angewiesen, eine individuelle Modellbildung vorzunehmen und die doppelte Komplexität von Kontext und (Mess-)Intervention zu bewerten. Für den Routinebetrieb

verbleibt daher als einzig denkbare Lösung der Einsatz von Monitoring-Instrumenten, so wie sie durch Indikatoren gegeben sind (auf die generierenden Verfahren wird hier nicht näher eingegangen).

Das Verständnis von Qualitätsindikatoren, das im IQTIG-Methodenpapier vertreten wird, weicht allerdings in dreierlei Hinsicht vom gängigen Verständnis ab:

- Indikatoren werden als Messinstrumente verstanden,
- Indikatoren wird eine Bewertungsfunktion zugeschrieben, und
- der Begriff der Validität sei nicht auf Indikatoren anwendbar.

Alle drei Aspekte sind weit von dem Verständnis von Indikatoren entfernt, wie es z.B. im „QM-Curriculum" der Bundesärztekammer, KBV und AWMF (4. Auflage) in Anlehnung an die internationale Literatur vertreten wird (S. 73f, Hervorh. der Autor):

„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. ***Ein Indikator ist kein direktes Maß der Qualität.*** *Es ist mehr ein Werkzeug, das zur Leistungsbewertung benutzt werden kann, das Aufmerksamkeit auf potentielle Problembereiche lenken kann, die einer intensiven Überprüfung innerhalb einer Organisation bedürfen könnten. Qualitätsindikatoren bilden die Qualität einer Einheit durch Zahlen bzw. Zahlenverhältnisse indirekt ab. Man kann sie auch als qualitätsbezogene Kennzahlen („Qualitätskennzahlen") bezeichnen. (…) Darüber hinaus müssen Qualitätsindikatoren – je nach Anwendung – den Anforderungen der Validität, Reliabilität, Sensitivität und Spezifität genügen."*

Qualitätsindikatoren zur Qualitäts-„messung"?

Indikatoren müssen in erster Linie sensitiv eingestellt werden, denn bei einem Monitoring-Ansatz ist es notwendig, möglichst alle problematischen Fälle zu erfassen (Vermeidung von falsch-negativen Ergebnissen), auch auf die Gefahr hin, falsch-positive Ergebnisse zu erhalten (daher geht man Kompromisse bei der Spezifität ein). Patienten und Zuweiser müssen vor allem darauf vertrauen könne, dass bei fehlendem Ansprechen des Indikators möglichst keine Qualitätsprobleme vorliegen. Aus dieser Sicht weisen Indikatoren eine ganz andere Konfiguration als Messinstrumente und auch als diagnostische Verfahren in der Medizin auf. Messinstrumente werden im Allgemeinen auf einen möglichst optimalen Kompromiss von Sensitivität und Spezifität eingestellt (z.B. je 85%), und diagnostische Verfahren weisen in erster Linie eine hohe Spezifität auf (weswegen im medizinischen Bereich Indikatoren oft nicht zu vermitteln sind).

An diesem Punkt liegt eine der entscheidenden Schwächen im Indikatorenverständnis im Methodenpapier des IQTIG. Wenn man Qualitätsindikatoren, so wie oben ausgeführt, in einem Kontinuum von schlechter bis außerordentlich guter Qualität einsetzt, kommt man nicht umhin, ihnen ähnliche Eigenschaften wie klassischen Messinstrumente zuzuweisen. Da es aus grundlegenden statistischen Überlegungen nicht möglich ist, ein Messinstrument gleichzeitig auf eine Sensitivität und Spezifität von 100% einzustellen, muss man bei der Sensitivität Kompromisse machen, und in der Folge verlieren Indikatoren ihre Monitoring-Funktion.

An dieser Stelle kommt das Konzept der sog. Exzellenz-Indikatoren ins Spiel. Wenn Qualitätsindikatoren, die Qualitätsprobleme identifizieren sollen (Monitoring), eine hohe Sensitivität für diese Ereignisse besitzen müssen (man will möglichst keine Probleme übersehen), dann sollten Qualitätsindikatoren, die besonders hohe Qualität identifizieren sollen, auf eine besonders hohe Spezifität eingestellt sein, denn man will ja vermeiden, dass Exzellenz dort festgestellt wird, wo sie nicht vorliegt (man will falsch-positive Befunde vermeiden). In der Konsequenz heißt dies nichts anderes, als dass die gleichen Qualitätsindikatoren, die für die Identifikation von Qualitätsproblemen verwendet werden, nicht gleichzeitig als Indikatoren für besonders gute Qualität fungieren können. Es liegt auf der Hand, dass diese Ableitung nicht nur den Grundannahmen im Methodenpapier des IQTIG widerspricht, sondern auch die gesetzlichen Regelungen in Frage stellt, die folglich energisch hinterfragt werden sollten.

Qualitätsindikatoren mit „Bewertungsfunktion"?

Dem Methodenpapier des IQTIG kann man folgende Definition von Qualitätsindikatoren entnehmen:

„Die Messung der Qualität medizinischer Versorgung erfolgt grundsätzlich mittels Qualitätsindikatoren. Qualitätsindikatoren sind Konstrukte, die aus Versorgungsdaten (Input) ***nachvollziehbare Bewertungen der Versorgungsqualität*** *(Output) ableiten"* (S. 21, Hervorh. der Verf.).

Im Methodenpapier werden nachfolgend „drei Komponenten" erwähnt, die Indikatoren beinhalten sollen (S. 21):

- *„einem konkreten Ziel für die Versorgungsqualität (z.B. „die Häufigkeit von Komplikationen soll möglichst gering sein"),*
- *einem spezifischen Dokumentations- und Messverfahren für das Qualitätsmerkmal (z.B. Spezifikation, Dokumentation, Rechenregel)* sowie
- *einem Bewertungskonzept, das die Zielerreichung bewertet und einen Handlungsanschluss bietet (z.B. Referenzbereich einschließlich Klassifikationsverfahren)*."

Positiv ist hier zu vermerken, dass explizit die Zielorientierung von Qualitätsindikatoren hervorgehoben wird (vgl. Schrappe 2014, S. 68ff und 86ff), und auch die Forderung nach einer genauen Spezifikation ist selbstverständlich nicht zu kritisieren. Allerdings steht die Annahme, zum Begriff des Indikators würde auch ein „Bewertungskonzept" gehören, nicht im Einklang mit den gängigen Definitionen. Sie basiert auf der sachlich falschen Überzeugung, Indikatoren könnten direkt Qualität messen und insofern einen direkten „Handlungsanschluss" liefern. Monitoring-Instrumente sind dazu jedoch nicht in der Lage, was allerdings nicht ausschließen soll, dass man aus übergeordneten Gründen an die Ergebnisse der Indikatoren Handlungskonsequenzen knüpft (z.B. wie das Bußgeld bei Geschwindigkeitsüberschreitung in geschlossenen Ortschaften).

Der grundsätzliche Unterschied besteht darin, dass das Methodenpapier impliziert, die „Bewertung" bzw. der „Handlungsanschluss" sei aus dem Indikator heraus abzuleiten, während die gängige Position davon ausgeht, dass der Indikator lediglich als Ampel fungiert, die die Aufmerksamkeit auf einen Handlungsbedarf lenkt und eventuell Handlungsoptionen auslöst, die aber extern abgeleitet sind. An dieser Stelle schließt sich die wichtige Diskussion an, ob Indikatoren, die als Monitoring-Instrumente eingestellt sind, eigentlich zur Steuerung auf Systemebene z.B. in P4P-Programmen oder zur Krankenhausplanung verwendet werden können. Dies ist durchaus der Fall, obwohl dies zunächst kontraintuitiv erscheint (keine Messung, „nur" Monitoring, wie sollen daraus Entscheidungen abgeleitet werden?), denn in anderen Bereichen der Gesellschaft wird dies oft in dieser Weise gehandhabt (z.B. bei Ver- und Geboten). Die Handlungskonsequenzen (z.B. Abschläge) müssen jedoch extern gesetzt werden und können nicht aus

dem Indikator abgeleitet werden, die „Handlungsanschlüsse" stellen also keine genuine Indikatoreigenschaft dar.

Begriff der Validität nicht auf Indikatoren anwendbar?

Die Ansicht, Indikatoren wären als Instrumente der Qualitätsmessung anzusehen, und die daraus resultierenden Ungenauigkeiten in der Definition von Indikatoren gipfelt letztendlich in einem fragwürdigen und nicht korrekten Verständnis der Validität von Indikatoren. Wie in Abb. 2 im oberen Abschnitt ausgeführt, resultiert ein lineares Messverfahren, wie es im Methodenpapier des IQTIG für Indikatoren angenommen wird, in einem Verständnis von Validität, das sich darauf bezieht, ob das Messergebnis (z.B. Rate postoperativer Wundinfektionen oder Mortalität) richtig wiedergegeben wird (die Reliabilität gäbe in diesem Verständnis die Zuverlässigkeit des Messvorganges wieder).

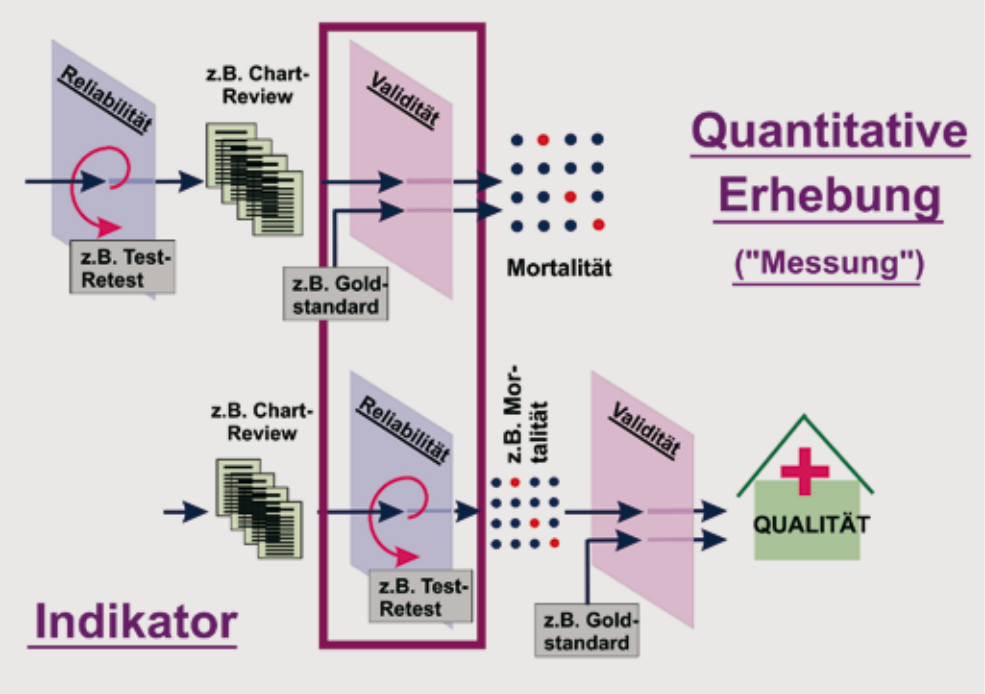


Abb. 2: Die Validität einer linearen Messung (quantitative Erhebung) von Qualitätsparametern entspricht der Reliabilität von Indikatoren (nähere Erklärung s. Text).

Verwendet man aber Indikatoren richtigerweise als Monitoring-Instrumente (Abb. 2 untere Hälfte), besteht deren Validität nicht darin, das Indikatorziel richtig zu messen, sondern Anlass für weitere Evaluationsschritte zu geben und damit die Qualität der Leistungserbringung als Einflussfaktor auf deren Varianz zu beschreiben. Die oben ausgeführte Notwendigkeit, Indikatoren hoch-sensitiv einzustellen, bezieht sich also auf deren Reliabilität, nämlich deren Fähigkeit, das Indikatorziel (z.B. Mortalität) richtig zu erkennen. Die Validität entspricht dagegen der Fähigkeit, diese Ereignisse mit der Qualität der Leistungserbringung in Zusammenhang zu bringen. Es ist dabei selbsterklärend, dass Indikatoren, die nicht reliabel sind (z.B. Abrechnungsdaten bei nosokomialen Infektionen), nie valide Indikatoren abgeben können (eine defekte Flinte (Reliabilität) kann nicht häufig ins Schwarze treffen (Validität), höchstens per Zufall). Auch lässt sich ableiten, dass Indikatoren, die risikoadjustiert werden müssen, weniger geeignet sind (z.B. weil die Adjustierung nie perfekt ist), Prozessparameter dagegen aber sehr gut mit direkten Konsequenzen verbunden werden können.

Literatur

Agency of Healthcare Research and Quality (AHRQ) (2017): The Six Dimensions of Health Care Quality, https://www.ahrq.gov/professionals/quality-patient-safety/talkingquality/create/sixdomains.html, letzter Zugriff 18.2.2017
Arah, O.A. (2006), Westert, G.P., Hurst, J., Klazinga, N.S.: A Conceptual Framework for the OECD Health Care Quality Indicators Project. Int. J. Qual. Health Care 18, 5-13
Berwick, D.M. (2003), James, B., Coye, M.J.: Connections Between Quality Measurement and Improvement. Med. Care 41, Suppl. 1: 30-8
Berwick, D.M. (2008): The Science of Improvement. JAMA 299, 1182-84
Bundesärztekammer (2007), Kassenärztlicher Bundesvereinigung, Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (Hrsg.), M. Schrappe, R. Klakow-Franck, W. Müller, B. Gibis, M. Brüggemann, M. Kelterbaum (Redaktion): Curriculum Ärztliches Qualitätsmanagement, 4. überarbeitete Auflage, ISSN 0945-1951
Campbell, S. (2010), Reeves, D., Kontopantelis, E., Middleton, E., Sibbald, B., Roland, M.: Quality of Primary Care in England with the Introduction of Pay for Performance. N. Engl. J. Med. 357, 181-190
DIN-TERM online, http://www.din.de/de/service-fuer-anwender/din-termonline, registrierungspflichtige Datenbank, letzter Zugriff 18.2.2017
Donabedian, A. (1990): The Seven Pillars of Quality. Arch. Path. Lab. 114, 1115-8
Frosch, D.L. (2010), Härter, M., Simon, D., Mulley, A.G.: Variation und Verteilungsgerechtigkeit. Patientenpräferenzen berücksichtigen. Dtsch. Aerztebl. 107, A 2100-4
Geraedts, M., de Cruppé, W. (2011): Wahrnehmung und Nutzung von Qualitätsinformationen durch Patienten. In: Klauber, J., Geraedts, M., Friedrich, J., Wasem, J. (Hrsg.): Krankenhausreport 2011. Schattauer, Stattgart, S. 93-104
Institute of Medicine (2001): Crossing the Quality Chasm: A New Health System for 21th Century. National Academy Press, Washington
JCAHO (1991): Primer on Indicator Development and Application, Joint Commission on Accreditation of Healthcare Organizations. One Renaissance Blvd, Oakbrook Terrace, Illinois 60181
Marshall, M.N., Pronovost, P., Dixon-Woods, M. (2013): Promotion of Improvement as a Science. Lancet 381, 419-21
Schoen, C. (2011), Osborn, R., Squires, D., Doty M., Pierson, R., Applebaum, S.: New 2011 Survey of Patients with Complex Care Needs in Eleven Countries Finds That Care is Often Poorly Coordinated. Health Aff. 20, 2437-42
Schrappe, M. (2014): Qualität 2030 – die umfassende Strategie für das Gesundheitswesen. Mit einem Geleitwort von Ulf Fink und Franz Dormann. Medizinisch Wissenschaftliche Verlagsgesellschaft, Berlin
Schrappe, M. (2016): Vortrag „Indikatorenkonzept und KHSG: neue Anforderungen und Weiterentwicklung", 15. Deutscher Kongress für Versorgungsforschung, Urania, Berlin, 6.10.2016
Schrappe, M. (2017): Planungsrelevante Qualitätsindikatoren. Impulsvortrag. Tagung „Qualität im Krankenhaus. Fluch und Wirklichkeit?!" Evangelischer Verbund Ruhr, 5. EVR-Forum, Bochum 15.2.2017
Sens, B. (2007), Fischer, B., Bastek, A., Eckardt, J., Kaczmarek, D., Paschen, U., Pietsch, B., Rath, S., Ruprecht, T., Thomeczek, C., Veit, C., Wenzlaff, P.: Begriffe und Konzepte des Qualitätsmanagements - 3. Auflage. GMS Med Inform Biom Epidemiol. 2007;3(1):Doc05
Wennberg, J.E. (2002), Fisher, E.S., Skinner, J.S.: Geography and the Debate over Medicare Reform. Health Aff. 21, Suppl. Web Exclusives, w96-114
Werner, R.M. (2011), Kolstad, J.T., Stuart, E:A., Polsky, D.: The Effect of Pay-For-Performance in Hospitals: Lessons for Quality Improvement. Health Aff. 30, 690-8

So ist ein niedriger Verbrauch von Händesdesinfektionsmitteln immer von einer mangelnden Händedesinfektion gefolgt, ohne dass eine Risikoadjustierung notwendig wäre.

Zusammenfassend kann also festgehalten werden, dass im Konzept des IQTIG, bedingt durch das Verständnis von Indikatoren als lineare Messinstrumente, die Reliabilität von Indikatoren als Validität fehlinterpretiert wird. Diese Vertauschung begegnet dem aufmerksamen Beobachter im Gesundheitswesen häufig: So wird beklagt, dass sich eine zu hohe Rate von Katheterinfektionen in der QS-Dokumentation bei der näheren Nachprüfung vor Ort als falsch herausgestellt hätte, denn einige Katheterinfektionen seien fehldokumentiert worden (falsch-positiv) – dieser Indikator sei eben „unspezifisch", also „nicht valide". Diese häufige Ansicht beschreibt wie unter dem Vergrößerungsglas die ganze Misere, denn niemand fragt nach, ob der Indikator vielleicht trotzdem alle tatsächlich vorliegenden Katheterinfektionen erkannt hat, also (wie verlangt) hoch-sensitiv ist, und dabei eben auch falsch-positive Befunde verursacht hat. Dazu müsste man allerdings hingehen, mit einer Goldstandard-Methode nachzusehen, wie viele Katheterinfektionen tatsächlich im fraglichen Zeitraum vorlagen, und mit den durch die QS-Dokumentation erfassten Zahlen vergleichen. Der Indikator wäre dann von hoher Reliabilität, wenn er alle Infektionen identifiziert hätte, und von hoher Validität, wenn sich z.B. durch ein sich anschließendes Peer Review-Verfahren herausgestellt hätte, dass diesen Infektionen Qualitätsmängel zugrundeliegen.

Diese grundlegende Unstimmigkeit zur Verwendung der Begriffe Reliabilität und Validität im Methodenpapier ist von schwerwiegenden Konsequenzen gefolgt, denn es erwächst die Schwierigkeit, den Terminus der Validität von Indikatoren positiv zu definieren. Das Methodenpapier versucht diese Situation so zu lösen, dass der Begriff der Validität ganz umgangen wird, und es versteigt sich zu der Ansicht, dass der Begriff der Validität im Gesundheitswesen nicht zu verwenden sei. Unter der Überschrift *„Validität und Korrektheit der Berechnungsspezifikation"* heißt es auf S. 95f (Hervorh. der Verf.):

„Unter der ***Validität einer Messung wird das Ausmaß verstanden, in dem der Indikator tatsächlich das abbildet, was er zu messen vorgibt*** *(…). Dieses Kriterium stammt aus der Psychometrie und bezieht sich auf latente, nicht direkt beobachtbare Messgrößen wie etwa Intelligenz oder Persönlichkeitsmerkmale.* ***Auf Qualitätsindikatoren, die direkt messbare Größen abbilden sollen, wie etwa Mortalität, Komplikationsraten oder ähnliches, lässt sich dieses Konzept demnach nicht übertragen.*** *Dies trifft vor allem auf Qualitätsindikatoren zu, die auf der Dokumentation durch die Leistungserbringer oder auf Sozialdaten bei den Krankenkassen basieren.* **Die Validität der Messung solcher Indikatoren wird im Wesentlichen dadurch bestimmt, dass die Berechnungsvorschrift für den Indikatorwert angemessen und korrekt spezifiziert ist** *(siehe Abb. 9). Beispielsweise ist der Indikator „Krankenhaussterblichkeit bei Behandlung ambulant erworbener Pneumonien" nur zu dem Ausmaß valide, zu dem korrekt festgelegt wurde, welche Daten das Merkmal „Sterblichkeit", den Zeitraum „Krankenhausaufenthalt" sowie die Nenner-Population des Indikators („Pneumonie" und „ambulant erworben") beschreiben."*

Mit anderen Worten, das IQTIG ist der Ansicht, dass der Begriff der Validität, der nach allgemeiner Übereinkunft zentral im Verständnis von Indikatoren ist, für Qualitätsindikatoren im Gesundheitswesen nicht anwendbar sei. Nur der Vollständigkeit halber soll darauf verwiesen werden, dass z.B. im QM-Curriculum der Bundesärztekammer, KBV und der AWMF die Validität von Indikatoren folgendermaßen definiert wird:

„Gültigkeit. Grad der Genauigkeit, mit dem ein Testverfahren das misst, was es messen soll. Die Validität ist das wichtigste, jedoch auch das am schwierigsten zu bestimmende Gütekriterium (siehe auch Reliabilität). Nach der Bestimmungsmethode werden verschiedene Gültigkeiten voneinander unterschieden (z.B. Konstruktvalidität, Augenscheinvalidität etc.)" (S. 81f).

Die Literatur zur Validität von Indikatoren ist unübersehbar groß. Es sei hier die Formulierung der Joint Commission herausgegriffen (JCAHO 1991): *„Degree to which indicators identify events that merit further review."*

Im Methodenpapier des IQTIG wird der Begriff der Validität stattdessen ersetzt durch die Spezifizierung des Messwertes. Man kann abschließend nur darauf hinweisen, dass diese hier zutage tretende Ansicht keine Basis für einen sinnvollen Umgang mit der Thematik der Qualitätsindikatoren und genenerell für die Arbeit mit dem Ziel der Qualitätsverbesserung darstellt. Wenn die Begriffe der Reliabilität und Validität von Indikatoren nicht zutreffend differenziert werden, und wenn die Ansicht vertreten wird, es gäbe keine sinnvolle Anwendung des Begriffes Validität auf die Arbeit mit Qualitätsindikatoren, verlieren wir uns im deutschen Gesundheitswesen in Bezug auf die Qualitätsfrage in einer Sackgasse. <<

The method-paper of IQTIG: no change of course in sight

The first version of the method paper, recently released by the German Institute for Quality and Transparency in Health Care (IQTIG), stays tied to the traditional concept of ex post-quality check and denies modern concepts of continuous quality improvement. Dimensions of quality are reduced to legitimacy and statistical questions, terms as quality assurance and quality indicators are not applied in a correct manner. Patient centeredness is focussed only on outcome, coordination and other process- and structural parameters are not considered. Quality indicators are regarded as direct measures of quality and not (as recommended in the utmost part of national and international documents) as an indirect measure of quality (monitoring), which directs attention on problematic areas for further review. As a striking and remarkable consequence, the term validity of indicators is seen as not applicable to quality issues. The paper needs profound revision.

Keywords

Methods paper IQTIG, quality dimensions, quality control, quality indicator

Prof. Dr. med. Matthias Schrappe

Internist, Infektiologe, 1996-2001 QM Univ.-Klinik Köln, 2002-05 Vorstandsvorsitzender Univ.-Klinik Marburg, danach Dekan/Wiss. GF Univ. Witten, Generalbevollmächtigter Univ.-Klinik Frankfurt. 2009 Ruf W3 Univ. Bonn, bis 2011 Direktor Inst. f. Patientensicherheit Univ. Bonn. Bis 2011 Mitglied/Stellv. Vors. Sachverständigenrat Gesundheit, 2001-7 Vorsitzender d. GQMG, 2005-9 Gründungsvorsitzender Aktionsbündnis Patientensicherheit, derzeit APL-Prof. Univ. Köln mit Lehrauftrag „Patientensicherheit und Risikomanagement", Vorstandsmitglied des Deutschen Netzwerk Versorgungsforschung (DNVF)
Kontakt: matthias@schrappe.com

