

Stellungnahme IQTIG-Methodenpapier

1. Fassung (26.2.2017)

1. Einführung

Das nun in einem ersten Entwurf vorliegende Methodenpapier des Institutes für Qualität und Transparenz im Gesundheitswesen (IQTIG) weist einige positive Aspekte auf (z.B. Zielorientierung des Qualitätsbegriffes, Qualitätsmodell als Voraussetzung für die Identifikation von Indikatoren), bleibt aber in seinen Grundannahmen einem traditionellen, auf die ex post-Qualitätskontrolle von Ergebnissen ausgerichteten Verständnis von Qualität verhaftet. In den folgenden Ausführungen darf selbstverständlich 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.

2. Allgemeines

Das Methodenpapier gliedert sich in drei Teile: Grundlagen (Teil A), Entwicklung und Durchführung (Teil B) und methodische Elemente (Teil C). Das Methodenpapier behandelt die *„wissenschaftlichen Methoden, die den Entwicklungen und Weiterentwicklungen von QS-Verfahren zugrunde liegen“* (S. 13), Andererseits wird hervorgehoben, es handle sich *„weder [um] ein Lehrbuch für Qualitätssicherung noch beinhaltet es eine Abhandlung der Geschichte der medizinischen Qualitätssicherung“* (S. 13).

In der Auseinandersetzung mit dem Methodenpapier, das für die weitere Entwicklung der Qualitätsdiskussion im deutschen Gesundheitswesen eine wichtige Rolle spielen wird, kann man von zwei unterschiedlichen Positionen ausgehen:

Aus der Perspektive einer **normativen Position** muss klar konstatiert werden, dass in dem vorliegenden Methodenpapier zahlreiche grundlegende Definitionen verwendet werden, die nicht dem gängigen wissenschaftlichen und praktischen Verständnis entsprechen. Grundlegende Quellen, die in Deutschland in den letzten 20 Jahren für die Qualitätsdiskussion und vor allem auch für die Lehre, die Aus- und die Fortbildung der Gesundheitsberufe in diesem Bereich Gültigkeit hatten, werden nicht berücksichtigt, insbesondere

- das von der Bundesärztekammer, der kassenärztlichen Bundesvereinigung und der Arbeitsgemeinschaft der Medizinisch-wissenschaftlichen Fachgesellschaften (AWMF) herausgegebene „Curriculum Qualitätsmanagement“ (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).

Die verwendeten Begriffsbestimmungen weichen in dem Methodenpapier durchaus nicht nur graduell von den in diesen (und anderen) Quellen verwendeten Definitionen ab.

Aus der **wissenschaftlichen Perspektive** ist festzuhalten, dass das Methodenpapier nicht der aktuellen Diskussion über den Qualitätsbegriff entspricht, wie sie z.B. im Rahmen der Versorgungsforschung (aber auch darüber hinaus) geführt wird. Die Versorgungsforschung geht davon aus, dass Begriffe und Konzepte im Gesundheitswesen nur im Rahmen komplexer Kontextbedingungen zu verstehen und quantitativ zu fassen sind. Gerade solche Konstrukte wie die Qualität der Gesundheitsversorgung basieren auf Aushandlungsprozessen über „Merkmale“ und „Anforderungen“, die dem Qualitätsbegriff zugrunde liegen und eine direkte, lineare Messung wie im biomedizinischen Kontext nur in Ausnahmefällen möglich machen. Zusätzlich ist zu bedenken, dass jeder Ansatz einer „Qualitätsmessung“ selbst als Intervention aufzufassen ist, die die Merkmale und Anforderungen sofort und nachhaltig verändert. Man muss sich daher über die Messmethodik (primär) und die Datenquellen (sekundär) sehr genau Rechenschaft ablegen; im geplanten breiten Einsatz von Indikatoren als Instrumente einer Qualitätsverbesserung auf Systemebene (*Public Reporting* oder *Pay for Performance*) können diese nur als Monitoring-Instrumente konzipiert werden, die hoch-sensitiv eingestellt werden (s. Abb. 1). Messmethoden wie bei wissenschaftlichen Studien oder begrenzten klinisch-epidemiologischen Erhebungen scheiden wegen des damit verbundenen Messaufwandes aus (s. Abschnitt 3.3.).

Monitoring-Instrumente müssen dabei durch wissenschaftliche Erkenntnisse angeleitet sein, können aber in den meisten Fällen *nicht vollständig durch wissenschaftliche Erkenntnisse begründet werden* (Beispiel: Mindestmengen als Qualitätsindikator – Problematik der Grenzwerte). Ebenso wie in anderen gesellschaftlichen Bereichen (z.B. Verkehrssicherheit: Geschwindigkeitsbegrenzung in geschlossenen Ortschaften) müssen sie wissenschaftlich abgeleitet, aber letztlich politisch gesetzt werden, um im Kontext weiterer Maßnahmen übergeordneten gesellschaftlichen Zielen gerecht zu werden. Nicht in jedem Fall ist es möglich, z.B. die Grenzwerte (z.B. 50 oder 55 km/h) valide abzuleiten oder in jedem Einzelfall eine konkrete Gefährdung nachzuweisen. Neben den genannten Mindestmengen sind daher Indikatoren über den Zugang zur Versorgung und zur regionalen Versorgung von hoher Bedeutung, da sie unmittelbar in den Bereich der politischen Handlungskompetenz fallen; allerdings ist dafür ein politisches Rahmenkonzept notwendig, wie es z.B. in den USA durch *Crossing the Quality Chasm* gegeben ist (IOM 2001). Im konkreten Zusammenhang mit der Qualitätsthematik heißt dies: die ex post-Qualitätskontrolle mittels Ergebnisparametern scheitert nicht nur an ihrer mangelnden Integration in den Prozess der Leistungserbringung, auch nicht an der (niemals vollständigen) Risikoadjustierung, sondern an der mangelnden biometrischen Spezifizierung (zu niedrige Sensitivität) und letztlich an der fehlenden übergeordneten Zielsetzung.

Dabei muss betont werden, dass das Methodenpapier des IQTIG sehr komplex und ehrgeizig ist und in Teilen (z.B. in den Ausführungen zum „Qualitätsmodell“) innovative Gedanken verfolgt, die durchaus diskussionswürdig erscheinen. Der Aufeinanderfolge von Qualitätsaspekten, Qualitätsmerkmalen und Qualitätsindikatoren kann man sicher einiges abgewinnen, eventuell ergibt sich hier die Möglichkeit, der Zielorientierung des Qualitätsbegriffes und von Indikatoren auch langfristig bessere Sichtbarkeit zu gewähren. Man muss aber intensiv über die Grundannahmen diskutieren, die dem Methodenpapier zugrunde liegen, insbesondere zum Verständnis von Qualität, zur Qualitätsdarstellung, der Bestimmung von Qualität (als Qualitätsmessung bezeichnet) und zum hier verwendeten Konzept der Qualitätsindikatoren sowie deren Validität.

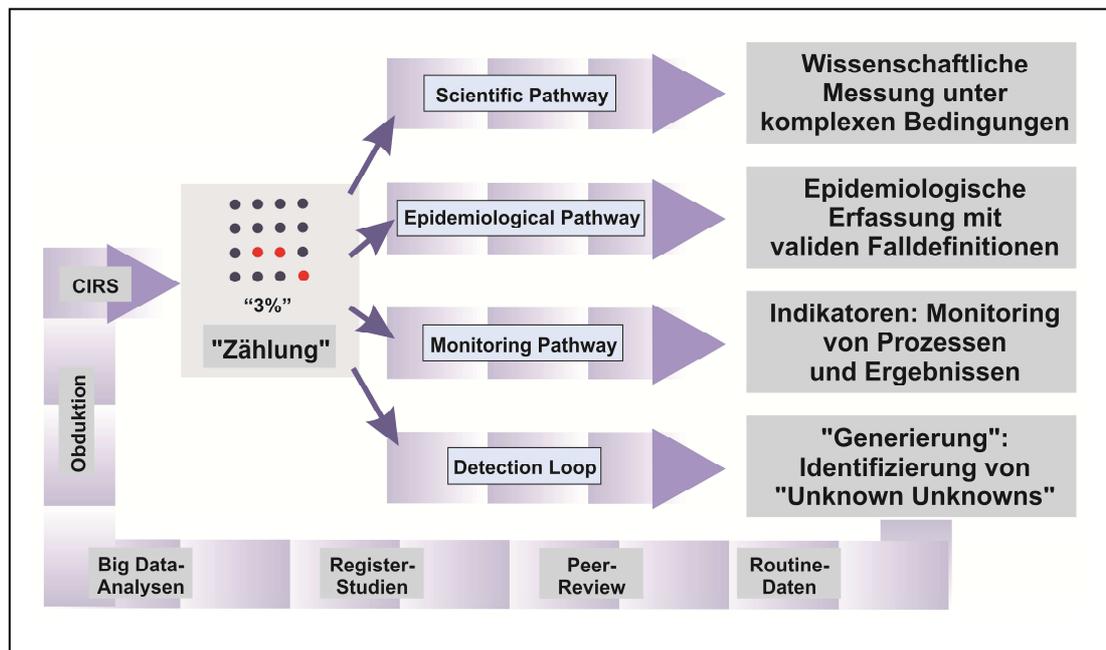


Abb. 1: Qualitätsbestimmung – vier Methoden. Bei einer wissenschaftlich intendierten Messung kommt die Messung einer komplexen Intervention gleich, eine vorherige Theoriebildung ist unumgänglich. Klinisch-epidemiologische Falldefinitionen werden z.B. in der klinischen Infektiologie verwendet und basieren auf genauen, nach epidemiologischen Gesichtspunkten gestalteten Kriterien. Qualitätsindikatoren bedienen sich eines Monitoring-Ansatzes und erlauben eine Beobachtung von großen Ereignismengen; sie sind mit hoher Sensitivität eingestellt. Die sog. generierenden Verfahren wie das CIRS verfolgen das Ziel, die *Unknown Unknowns* aufzudecken (aus Schrappe 2016).

3. Konzeption und Qualitätsverständnis

3.1. Qualitätsverständnis

Das dem Methodenpapier zugrunde liegende Qualitätsverständnis wird in den Kap. 2.1 und 2.2. ausgeführt (S. 16 ff). Ausgehend von einer sachlich richtigen Definition von Qualität, die auf den Begriffen „inhärente Merkmale“ und „Anforderungen“ basiert, wird die Dimension der Patientenzentrierung besonders hervorgehoben, flankiert durch „Systemqualität“ (ein unüblicher Begriff, den man aber prägen kann), den Zugang und „equity“ (Gleichheit) (S. 16). Diese Dimensionen können als relevant gelten, aus Sicht der Versorgungsforschung ist besonders die Betonung der Patientenzentrierung positiv zu werten. Die getroffene Auswahl ist jedoch vor dem Hintergrund der national und international geltenden Dimensionen und Rahmenkonzepte als unvollständig und in ihrer Auswahl nicht begründet zu werten. So sind im allgemein als grundlegend geltenden Konzept der „Seven Pillars“ von Donabedian (1990) die sechs Dimensionen *efficacy*, *effectiveness*, *efficiency*, *acceptability*, *optimality*, *legitimacy* und *equity* enthalten; das auch vom AQUA-Institut genutzte OECD Konzept umfasst *effectiveness*, *safety*, *timeliness* und *patient-centeredness* (Arah et al. 2006); das aktuelle AHRQ-Konzept umfasst die Attribute *safe*, *effective*, *patient-centered*, *timely*, *efficient* und *equitable* (AHRQ 2017).

In Kap. 2.2. kommt zu dieser wenig begründeten Auswahl noch der Umstand hinzu, dass aus der **Dimension „Legitimität“** eine Art 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)

Die Begründung wird darin gesehen, dass

„... diese Anforderungen der externen Qualitätssicherung Grundlage für Verbesserungsinitiativen und Steuerungsverfahren im Gesundheitswesen sein sollen“ (S. 17)

und sie daher *„in besonderer Weise legitimiert sein“* müssen. Die Kriterien für diesen **Begriff der Legitimität** werden genau genannt und lauten

- Patientenzentrierung,
- Beeinflussbarkeit durch die Leistungserbringer und
- Unbedenklichkeit.

Es wird also a priori eine Auswahl der in Frage kommenden Dimensionen getroffen. Diese Auswahl muss sowohl in ihrer Konstellation als auch in ihrer Ausführung kritisch gesehen werden:

- **Patientenzentrierung:** Der Begriff der Patientenzentrierung wird im Methodenpapier ausgesprochen häufig verwendet. Man wird den Verdacht nicht los, dass hier eine Art Augenscheinvalidität aufgebaut wird, da gegenwärtig Patientenzentrierung gerade im Rahmen der Versorgungsforschung ein so prominentes Thema darstellt. In der Umsetzung zeigt sich aber, dass Patientenzentrierung in einer sehr paternalistischen Weise verstanden wird, in dem sie ausschließlich auf die Ergebnisqualität bezogen wird. Dagegen zeigen entsprechende Untersuchungen, dass Patienten in viel stärkerem Maße an Fragen der Strukturqualität (z.B. Ausbildung der Ärzte) sowie an Kooperation, Koordination und Information (also eindeutigen Prozessparametern) interessiert sind (aus der umfangreichen Literatur: Geraedts und de Cruppé 2011, Schoen et al. 2011). Die einseitige Betonung der Ergebnisqualität erscheint auf den ersten Blick sinnvoll, führt aber zu einer reinen ex post-Betrachtung im Sinne eines Qualitätskontroll-Ansatzes, statt zu einer Integration von Qualitätsaspekten im Prozess der Leistungserbringung im Sinne eines kontinuierlichen Qualitätsverbesserungsansatzes (*continuous quality improvement*). Die Fokussierung auf die Patientenzentrierung erscheint auch vor dem Hintergrund der derzeitigen Diskussion um die Regionalität der Versorgung nicht als einziger Blickwinkel sinnvoll zu sein, denn neben den präferenz-sensitiven Aspekten, zu denen Patientenzentrierung zu rechnen ist, sollte man mindestens noch die Dimension der Wirksamkeit und Anbietersensitivität mit einschließen (Wennberg et al. 2002, s. auch Frosch et al. 2010).

- **Beeinflussbarkeit durch die Leistungserbringer:** Dieses Kriterium ist genau genommen keine Qualitätsdimension, sondern stammt aus der Bewertung von Qualitätsindikatoren (z.B. RUMBA-Regel). So wie unter der Überschrift „*Beeinflussbarkeit durch den Leistungserbringer*“ auf Seite 92 richtig dargestellt, ist dieser Aspekt bei der Auswahl von Indikatoren wichtig, für die Entwicklung von Qualitätsdimensionen jedoch nachgeordnet, denn wenn Qualitätsprobleme bestehen, die durch einen Leistungserbringer nicht beeinflussbar sind, dann bestehen sie dessen ungeachtet fort.

- **Unbedenklichkeit:** Diese Dimension 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 entfalten*“ dürfen. Es stellt sich allerdings die Frage, weshalb dann nicht Wirksamkeit oder Sicherheit im Sinne der gebräuchlichen Systematiken direkt genannt werden.

Die vorherrschende Stellung der „Legitimität“ ist der wissenschaftlichen und praktischen Betrachtung von Qualität sowohl im institutionellen und im Systemkontext weitgehend

fremd (eine der wenigen Nennungen in Steffen 1988) bzw. darf als selbstverständliche Grundbedingung gelten. Es scheint hier der Versuch vorzuliegen, eine juristische Absicherung der Aktivitäten vorzunehmen, allerdings auf Kosten anderer Dimensionen des Qualitätsbegriffes, die hier durchaus sinnvoll hätten aufgeführt und diskutiert werden können (*optimality, acceptability, safety* etc.). Klar wird dies in folgenden Ausführungen:

*„Versorgungsqualität im Rahmen der extern vergleichenden Qualitätssicherung ist definiert als der **Grad der Erfüllung der legitimen Anforderungen**, die durch das jeweilige QS-Verfahren normativ vorgegeben sind.“* (S. 20, Hervorh. der Verf.)

*„Die Qualität der Versorgung besteht in der Fähigkeit, bestimmte legitime Ziele zu erreichen (Steffen 1988). In der externen Qualitätssicherung werden diese Ziele als **legitime Anforderungen an die Leistungserbringer** formuliert und konkret über Qualitätsmerkmale dargestellt ...“* (S. 92, Hervorh. der Verf.).

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. Die statistische Analyse von Qualitätssicherungsdaten ist eine über den Einzelfall hinausgehende Betrachtungsweise und kann dadurch Aspekte der Qualität quantitativ darstellen, die in Einzelfallanalysen nicht sichtbar werden. Die Bewertung der Versorgungsqualität eines Leistungserbringers erfolgt anhand statistischer Daten der Versorgung von Patientinnen und Patienten, die dieser in einem bestimmten Zeitraum (meist während eines Jahres) behandelte. Die externe Qualitätssicherung hat damit einen anderen Schwerpunkt als etliche Instrumente der internen Qualitätssicherung, die mehr auf kasuistischen Einzelfallbewertungen vor Ort beruhen...“*

Abgesehen, dass der letzte Satz sachlich nicht zutreffend ist, sind statistische Fragen natürlich von großer Wichtigkeit, sie stellen aber gegenüber der Zielformulierung und der Beschreibung der inhärenten Merkmale ein sekundäres Problem dar.

Das Qualitätsverständnis des Methodenpapiers entspricht dem Ansatz der Qualitätskontrolle aus der ex post-Perspektive. Die diskutierten Qualitätsdimensionen sind auf Fragen der Legitimität und statistischen Auswertbarkeit reduziert, die juristische Absicherung des Verfahrens steht ganz im Vordergrund. Sehr häufig wird der Begriff der Patientenzentrierung genannt, jedoch in paternalistischer Tradition einseitig auf die Ergebnisqualität bezogen. Elemente der kontinuierlichen Qualitätsverbesserung mit Integration des Qualitätsgedankens in den Ablauf der Leistungserbringung auf institutioneller und Systemebene sind weitgehend in den Hintergrund gedrängt.

3.2. Begriff Qualitätssicherung

Der Begriff der Qualitätssicherung wird nicht sachgerecht verwendet. Das IQTIG beschreibt diesen Begriff folgendermaßen:

*„Unter Qualitätssicherung im Gesundheitswesen werden alle Aktivitäten verstanden, die als gezielte Maßnahmen das Einhalten bestimmter Qualitätsziele der Patientenversorgung gewährleisten sollen. Es kann dabei einerseits um die Einhaltung von Mindestanforderungen gehen oder aber andererseits um das Erreichen bestmöglicher Qualität und die Weiterentwicklung dessen, was als bestmögliche Qualität gelten kann. Insofern ist mit Qualitätssicherung immer auch **Weiterentwicklung der Versorgungsqualität** gemeint.“* (S. 15, Hervorh. der Verf.)
und

*„Die Aufgabe der (externen) Qualitätssicherung ist es daher, die Einhaltung von Mindestanforderungen zu gewährleisten und die **Unterschiede in der Versorgungsqualität in Richtung bestmöglicher Qualität zu reduzieren.**“* (S. 15, Hervorh. der Verf.)

Nach den allgemein gültigen Quellen und Standards ist Qualitätssicherung dagegen definiert als

„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)

Im Curriculum der Bundesärztekammer, KBV und AWMF heißt es zur Definition von Qualitätssicherung (S. 88):

„Teil des Qualitätsmanagements, der auf das Erzeugen von Vertrauen darauf gerichtet ist, dass Qualitätsanforderungen erfüllt werden. In der aktuellen Terminologie wird Qualitätssicherung – gemäß der wörtlichen Übersetzung – als „Qualitätszusicherung“ verstanden. Hierunter sind Aktivitäten zu verstehen, die bei Kunden und Partnern im Gesundheitswesen Vertrauen schaffen, dass eine Organisation alle festgelegten, üblicherweise vorausgesetzten und verpflichtenden Erfordernisse und Erwartungen erfüllt. In der Gesundheitsversorgung in Deutschland spielte der Begriff Qualitätssicherung bisher eine zentrale Rolle für verschiedenste Aktivitäten. Es wurde traditionell zwischen interner und externer Qualitätssicherung unterschieden. „Interne Qualitätssicherungsmaßnahmen“ umfassten durchaus auch Aspekte der Qualitätsverbesserung und des Qualitätsmanagements. Unter „externer Qualitätssicherung“ wurden insbesondere Qualitätssicherungsmaßnahmen mit externen Vergleichen verstanden.“

Wie aus den vorstehenden Definitionsansätzen zu erkennen, bezieht sich der Begriff der Qualitätssicherung auf die Darlegung von Qualität, gemäß der vorgenannten Anforderungen; er umfasst also weder Maßnahmen, die „das Einhalten bestimmter Qualitätsziele (...) gewährleisten“, noch solchen, die der „Einhaltung von Mindestanforderungen“, dem „Erreichen bestmöglicher Qualität“ und der „Weiterentwicklung der Versorgungsqualität“ dienen oder „Unterschiede in der Versorgungsqualität (...) reduzieren“. Der Verbesserungsgedanke, so wichtig er im aktuellen gesundheitspolitischen Kontext erscheint, ist kein Bestandteil der Definition von Qualitätssicherung, sondern fällt unter den Begriff der Qualitätsverbesserung (s.u.). Diese Differenzierung ist von großer Bedeutung, denn in der Auseinandersetzung um die Qualität der Gesundheitsversorgung kann nur mit klar definierten Begriffen eine sinnvolle Weiterentwicklung erreicht werden.

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 *Health Services Research* (Versorgungsforschung) (Berwick 2008, Marshall et al. 2013). In der Nomenklatur der *Agency of Healthcare Research and Quality* (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).

Auch im deutschsprachigen Raum ist der Begriff der Qualitätsverbesserung klar definiert: „Teil des Qualitätsmanagements, der auf die Erhöhung der Eignung zur Erfüllung der Qualitätsanforderungen gerichtet ist“ (Sens et al. 2007, DIN EN ISO 9000:2015-11).

Im Curriculum der Bundesärztekammer, KBV und AWMF heißt es zur Definition von Qualitätsverbesserung (S. 89):

„Teil des Qualitätsmanagement, der auf die Erhöhung der Fähigkeit zur Erfüllung der Qualitätsanforderungen gerichtet ist. Dies umfasst alle in einer Organisation ergriffenen Maßnahmen zur Erhöhung der Effektivität und Effizienz von Tätigkeiten und Prozessen, um zusätzlichen Nutzen sowohl für die Organisation als auch für ihre Kunden zu erzielen. Qualitätsverbesserung und Qualitätsplanung sollten zusammenwirken. Erkenntnisse aus der Qualitätsverbesserung können sich auf die Qualitätsplanung auswirken.“

Man kann natürlich argumentieren, dass dieser Begriff in Deutschland vor allem auf die organisatorische Perspektive beschränkt ist. So zitiert das IQTIG in seinem Methodenpapier zwar die QSKH- und die Qesü-Richtlinie (S. 15, Ende 2. Absatz), lässt aber die mit großem Aufwand erstellte Qualitätsmanagement-Richtlinie vom 15.9.2016, die die sektoralen Bestimmungen (Krankenhäuser, Kassenärztliche Vereinigungen, Kassenzahnärztliche Vereinigungen) zusammenfasst und besonders den Verbesserungsgedanken betont, völlig außer acht. Es wird so die Möglichkeit ausgelassen, auf den Unterschied von Qualitätsdarlegung (Qualitätssicherung) und den vielfältigen Möglichkeiten der Qualitätsverbesserung zu verweisen.

Natürlich ist das IQTIG, wie einleitend ausgeführt, an die Vorgaben des Gesetzgebers und des G-BA gebunden. In einem wissenschaftlichen Methodenpapier muss aber trotzdem auf die sachlich korrekte Verwendung der zentralen Begriffe Wert gelegt werden, zumindest durch eine differenzierte Darstellung und Interpretation. Geschieht dies nicht, sind Formulierungen wie die Folgende nicht zu vermeiden, die deutlich zeigt, wie sehr es dem Methodenpapier an einem fachlich adäquaten Verständnis des Begriffes Qualitätssicherung und seiner benachbarten Begriffe fehlt (S. 15):

„Internationale Longitudinaluntersuchungen zeigen, dass in sehr vielen Bereichen Versorgungsqualität auch ohne Maßnahmen der externen Qualitätssicherung kontinuierlich verbessert wird (Campbell et al. 2007, Werner et al. 2011).“

Die beiden zitierten Arbeiten sind prominent; die Arbeit von Campbell et al. (2007) ist eine der wichtigsten (und frühesten) Arbeiten zum sog. QOF-Projekt (*Quality Outcome Framework*), einem groß angelegten und gut evaluierten P4P-Projekt in der ambulanten Versorgung des NHS, und die Arbeit von Werner et al. (2011) berichtet über eine Subgruppe des *Premier Hospital Quality Incentive Demonstration Project* (HQIP), des großen P4P-Projektes der *Centers of Medicare and Medicaid Services* (CMS), das der Vorbereitung des *Value-Based Purchasing* Programms in den USA diene. Die Studie von Campbell et al. untersuchte in den Jahren 2003 bis 2005 ein Kollektiv von Praxen, das bereits seit 1998 in Studien zur Qualitätssicherung eingebunden war, und außerdem war in den Jahren vor Start des QOF-Projektes eine heftige öffentliche Diskussion zu den

Indikatoren im Gange. Zu behaupten, dass diese Studie eine „kontinuierliche“ Verbesserung auch ohne Qualitätssicherungsmaßnahmen nachweisen würde, ist zumindest gewagt, denn die langfristig untersuchten Praxen waren durch die wissenschaftliche Begleitung und die öffentliche Diskussion natürlich qualitätsfördernden Einflüssen ausgesetzt. Vor allem gilt dies aber für die Studie von Werner et al., denn tatsächlich waren alle Krankenhäuser in der Kontrollgruppe in das damals größte Qualitätssicherungsprojekt der USA im stationären Sektor, *Hospital Compare*, integriert, aus dem auch die gesamten Daten für diese wichtige Studie stammen¹.

Es sei also dringend dazu geraten, im Methodenpapier des IQTIG die Begriffe Qualitätssicherung und Qualitätsverbesserung differenziert darzustellen, um Missverständnissen zu begegnen und vor allem klare Botschaften an diejenigen zu senden, die vor Ort verantwortlich sind für Qualität: es geht um die Verbesserung, nicht um die Sicherung von Qualität. 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.

Das Methodenpapier verwendet den Begriff der Qualitätssicherung nicht zutreffend, sondern vermischt ihn mit dem Begriff der Qualitätsverbesserung. In der Folge werden Verfahren der Qualitätssicherung nicht sauber von Methoden der Qualitätsverbesserung auf Systemebene (z.B. P4P) getrennt.

¹ Es wäre wünschenswert, wenn diese Studie auch in den vorbereitenden Papieren des IQTIG zu P4P zur Geltung käme, denn es wird hier sehr gut dargestellt, dass die Einführung von P4P mit Indikatoren, die vorher schon zu Zwecken des Public Reportings eingesetzt werden sinnlos ist, vor allem wenn sie in hoch-kompetitiven Regionen eingesetzt werden, in denen Public Reporting besonders wirksam und .P4P nur wenig wirksam ist.

3.3. „Qualitätsmessung“ und Indikatorenkonzept

Das Methodenpapier des IQTIG geht davon aus, dass sich Qualität in der Gesundheitsversorgung linear messen lässt, und verwendet Qualitätsindikatoren als Messinstrumente, ähnlich wie sie im technischen Bereich oder in der Labormedizin eingesetzt werden. Dieser Ansicht ist aus der Perspektive der Qualitäts- und Versorgungsforschung deutlich zu widersprechen. In der vorliegenden Stellungnahme sind drei Punkte genauer auszuführen:

- Grundlage des Indikatorenverständnisses: Monitoring-Funktion
- Definition von Indikatoren (normativer Zugang)
- Unterscheidung von Reliabilität und Validität von Indikatoren

3.3.1. 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).

„Kern der Aufgaben des Instituts ist die Qualitätsmessung in der Gesundheitsversorgung mittels Qualitätsindikatoren (S. 87).

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“:

*„Die Qualitätsförderung (...) unterscheidet bislang lediglich qualitativ unauffällige und qualitativ auffällige Ergebnisse, also gute und nicht ausreichende Qualität. Das KHSG identifiziert dagegen beispielsweise im Rahmen von qualitätsbezogenen Zu- und Abschlägen **außerordentlich gute und am anderen Ende der Skala unzureichende Qualität und damit indirekt auch die dazwischen liegende, durchschnittlich gute Qualität**. Damit wird deutlich, dass zukünftig weitere, über die Differenzierung von guter und nicht ausreichender Qualität hinausgehende Differenzierungen notwendig werden. (...) Auch im Rahmen planungsrelevanter Qualitätsindikatoren findet sich eine feinere Differenzierung der Versorgungsqualität: Minimalanforderungen bei Planungsentscheidungen, Ausschlusskriterien bei Interventionsentscheidungen und Exzellenzaspekte bei Vergabeentscheidungen.“ (S. 19, Hervorh. der Verf.)*

Aus Sicht der Versorgungsforschung kann man dem Konzept einer linearen Qualitäts“messung“ 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 (s. Abb. 1). Nur die beiden erstgenannten Verfahren kommen einer Messung nahe. Der wissenschaftliche Zugang ist auf der Basis einer Modellbildung bemüht, ein Messinstrument mit gleichermaßen 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, so wie sie in der Hygiene (z.B. Erfassung von postoperativen Wundinfektionen) oder im Bereich der Patientensicherheit bei der Erfassung unerwünschter Ereignisse verwendet werden, ist der Erhebungsaufwand so hoch, dass man den Einsatz dieser Instrumente auf definierte Ereignisgruppen (z.B. nosokomiale Infektionen) beschränkt². 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 bewerten³.

Für den Routinebetrieb verbleibt daher als einzig denkbare Lösung der Einsatz von Monitoring-Instrumenten, so wie sie durch Indikatoren gegeben sind. 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 (Kompromisse bei der Spezifität). Im Gesundheitswesen ist die Erfassung von problematischen (negativen) Fällen bzw. Leistungserbringern von vorrangiger Wichtigkeit, denn sonst können die resultierenden Ergebnisse keine Steuerungswirkung entfalten:

² Allerdings lässt der Methodentransfer zwischen der Klinischen Infektiologie mit ihrem *Infection Control*-Ansatz und dem Qualitäts- und Patientensicherheitsbereich viele Wünsche offen, denn es wäre ein Leichtes, klinisch-epidemiologische Falldefinitionen z.B für postoperative Thrombosen zu entwickeln, denen nicht die mangelnde Sensitivität von Abrechnungsdaten anhaftet, und die – ähnlich wie bei nosokomialen Infektionen – so genau sein könnten, dass sie sogar internationale Vergleiche erlauben würden. Wichtig ist auch hier die Feststellung, dass es sich dabei nicht um klinische, d.h. für therapeutische Verfahren verwertbare Diagnosen handelt, sondern um primär epidemiologische Instrumente.

³ Die vierte Methode, die generierenden Verfahren, sind lediglich zur Behebung des „blinden Fleckes“ in Form der „Unknown Unknowns“ geeignet, weisen keine sinnvolle Sensitivität und Spezifität auf und stehen in dem hier diskutierten Zusammenhang nicht im Mittelpunkt.

Patienten und Zuweiser müssen vor allem darauf vertrauen können, dass bei fehlendem Ansprechen des Indikators *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 verstanden werden).

An diesem Punkt liegt einer der entscheidenden Schwächen im Indikatorenverständnis, so wie es im Methodenpapier des IQTIG und auch darüber hinaus oft im deutschen Gesundheitswesen vertreten wird. 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 klassische 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. Welche statistischen Eigenschaften muss(t)en Exzellenz-Indikatoren aufweisen? 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 an dieser Stelle hinterfragt werden sollten.

Das Methodenpapier des IQTIG vertritt das Konzept der linearen Messung von Qualität durch Qualitätsindikatoren. Eine differenzierte Betrachtung unterschiedlicher Erhebungsmethoden wird nicht vorgenommen. Entgegen der international und im deutschen Schrifttum üblichen Auffassung werden Indikatoren 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.

3.3.2. Verständnis von Qualitätsindikatoren

Nach diesen analytisch-inhaltlichen Ausführungen kommt man nicht umhin, wieder zu normativen Fragen zurückzukehren (die sich daraus ableiten lassen), bevor im nächsten Absatz das zweite inhaltliche Grundproblem der IQTIG-Position angesprochen wird (unklare Abgrenzung von Reliabilität und Validität von Indikatoren). Dem Methodenpapier des IQTIG kann man folgende Definitionen 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.).

*„Als Qualitätsindikatoren gelten leistungserbringerbezogene **Qualitätsmessungen**, die ein konkretes Qualitätsziel verfolgen und eine **Qualitätsbewertung** (...) ermöglichen“* (S. 52, Hervorh. der Verf.)

Wie oben bereits ausgeführt, geht das Methodenpapier davon aus, dass Qualitätsindikatoren in der Lage sind, Qualität zu „messen“, und dass sie außerdem der Bewertung von Qualität dienen können. Zum ersten Teil dieser Aussage wurde in Abschnitt 4.3.1. schon Stellung genommen, der zweite Teil („Bewertung“) ist jedoch nicht weniger unzutreffend. Es kann hier nicht die gesamte Literatur wiedergegeben werden, aber im „QM-Curriculum“ der Bundesärztekammer, KBV und AWMF (4. Auflage) heißt es beispielsweise zur Definition von Qualitätsindikatoren (S. 73f):

“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. Die Ausprägung eines Indikators kann mit guter bzw. schlechter Qualität in Verbindung gebracht werden. Hierzu verwendet man definierte Ausprägungen des Indikators, den sog. Referenzwert oder Referenzbereich. Qualitätsindikatoren sind struktur-, prozess- und/oder ergebnisbezogen. Darüber hinaus müssen Qualitätsindikatoren – je nach Anwendung – den Anforderungen der Validität, Reliabilität, Sensitivität und Spezifität genügen.“

Diese Definition liegt mit den international gebräuchlichen Definitionen z.B. der Joint Commission auf einer Linie, wie sie auch im „QM-Glossar“ der GMDS und GQMG vertreten werden:

*“An indicator is a quantitative measure that can be used to monitor and evaluate the quality of important governance, management, clinical, and support functions that affect patient outcomes. **An indicator is not a direct measure of quality.** Rather, it is a tool that can be used to assess performance that can direct attention to potential performance issues that may require more intense review within an organisation”* (JCAHO 1991, 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 daran aus übergeordneten Gründen Handlungskonsequenzen knüpft (z.B. wie das Bußgeld bei Geschwindigkeitsüberschreitung in geschlossenen Ortschaften). Der grundsätzliche Unterschied besteht hier darin, dass das Methodenpapier impliziert, die „Bewertung“ bzw. der „Handlungsanschluss“ sei aus dem Indikator heraus abzuleiten, während die richtige 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 in P4P-Programmen oder zur Krankenhausplanung sinnvoll verwendet werden können⁴.

⁴ Sie können tatsächlich eingesetzt werden, obwohl dies zunächst kontraintuitiv erscheint (keine Messung, „nur“ Monitoring, wie sollen daraus Entscheidungen abgeleitet werden?). In anderen Bereichen der Gesellschaft wird dies aber durchaus oft in dieser Weise gehandhabt (z.B. Ver- und Gebote). Die Handlungskonsequenzen (z.B. Abschlüge) müssen jedoch extern gesetzt werden und können nicht aus

Die im Methodenpapier des IQTIG verwendeten Definitionsansätze von Qualitätsindikatoren entsprechen nicht den gebräuchlichen Standards, insbesondere indem sie diesem Begriff ein Bewertungsfunktion zuweisen.

3.3.3. Validität von Indikatoren

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 gibt die Zuverlässigkeit des Messvorganges wieder). 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 als Einflussfaktor auf die Varianz der Leistungserbringung zu beschreiben⁵. Die oben ausgeführte Notwendigkeit, Indikatoren hoch-sensitiv einzustellen, bezieht sich also auf deren Reliabilität, die Fähigkeit, Qualitätsprobleme richtig zu erkennen. 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 immer bzw. häufig treffen (Validität), höchstens per Zufall).

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

⁵ Es folgt hieraus zwangsläufig, dass Indikatoren, die risikoadjustiert werden müssen, weniger geeignet sind, Prozessparameter dagegen direkte Konsequenzen haben können. So ist ein niedriger Verbrauch von Händedesinfektionsmitteln immer von einer mangelnden Händedesinfektion gefolgt, ohne dass eine Risikoadjustierung notwendig ist.

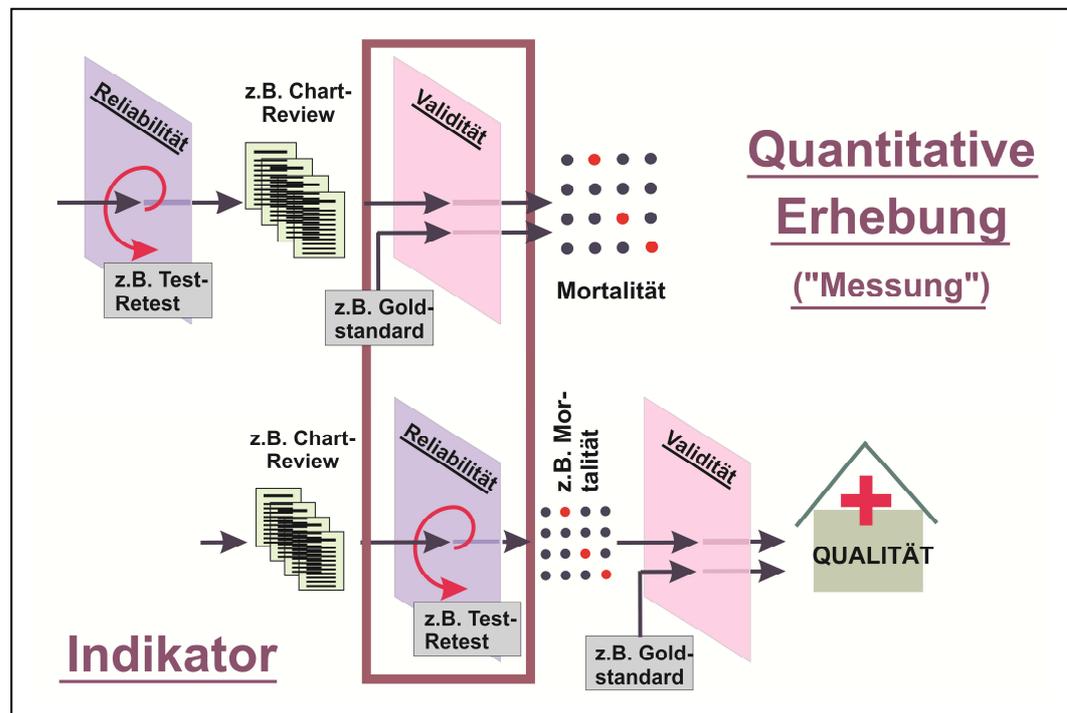


Abb. 2: Die Validität einer linearen Messung (quantitative Erhebung) von Qualitätsparametern entspricht der Reliabilität von Indikatoren (nähere Erklärung s. Text).

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 Zuhörer im Gesundheitswesen häufig: so wird beklagt, dass sich eine zu hohe Rate von Katheterinfektionen in der §137a-Dokumentation bei der näheren Nachprüfung vor Ort als falsch herausgestellt hätte, denn einige Katheterinfektionen seien fehdokumentiert 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, oder der Indikator 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 also von hoher Validität, wenn er alle Infektionen identifiziert hätte.

Das Methodenpapier des IQTIG lässt indirekt erkennen, dass diese grundlegende Unstimmigkeit durchaus nicht ganz ohne Aufmerksamkeit geblieben ist. Aus der nicht zutreffenden Verwendung der Begriffe Reliabilität und Validität erwächst nämlich 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 Abbildung 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 ist. 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 die hier zutage tretende Ansicht nur aus dem im Methodenpapier vertretenen Qualitäts- und Indikatorenverständnis zu interpretieren ist. Eine Basis für einen sinnvollen Umgang mit der Thematik der Qualitätsindikatoren stellt sie nicht dar.

Zusammenfassend muss festgehalten werden, dass das Methodenpapier des IQTIG die Begriffe der Reliabilität und Validität von Indikatoren nicht zutreffend differenziert. Es wird im vorliegenden Methodenpapier im Gegensatz zur allgemein (national und international) vertretenen Ansicht die Behauptung aufgestellt, es gäbe keine sinnvolle Anwendung des Begriffes Validität auf Qualitätsindikatoren. Dieser Ansicht kann man nicht beipflichten.

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, Stuttgart, 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

Steffen, G.E. (1988): Quality Medical Care. A Definition. JAMA 260, 56-61

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