



Institut für Qualitätssicherung und
Transparenz im Gesundheitswesen

Methodische Grundlagen V1.0s

Würdigung der Stellungnahmen

Stand: 14. März 2018

Impressum

Thema:

Methodische Grundlagen V1.0s. Würdigung der Stellungnahmen

Ansprechpartner:

Dr. Dennis Boywitt, Dr. Sven Bungard

Stand:

14. März 2018

Zitierhinweis:

IQTIG (2018): Methodische Grundlagen V1.0s. Würdigung der Stellungnahmen. Stand: 14. März 2018. Berlin: Institut für Qualitätssicherung und Transparenz im Gesundheitswesen. URL: https://iqtig.org/downloads/berichte/2018/IQTIG_Methodische-Grundlagen-V1-0s_Wuerdigung-der-Stellungnahmen_2018-03-14.pdf

Herausgeber:

IQTIG – Institut für Qualitätssicherung
und Transparenz im Gesundheitswesen

Katharina-Heinroth-Ufer 1
10787 Berlin

Telefon: (030) 58 58 26-0
Telefax: (030) 58 58 26-999

methoden@iqtig.org
info@iqtig.org
<https://www.iqtig.org>

Inhaltsverzeichnis

Abkürzungsverzeichnis.....	4
Stellungnahmeverzeichnis	6
1 Stellungnahmeverfahren zu Version 1.0s	8
2 Würdigung allgemeiner, themenübergreifender Stellungnahmen	9
3 Würdigung der Stellungnahmen zu einzelnen Kapiteln	13
3.1 Neuentwicklung von Qualitätssicherungsverfahren.....	13
3.1.1 Konzeptstudie für Qualitätssicherungsverfahren	14
3.1.2 Dokumentations- und sozialdatenbasierte Neuentwicklung von Qualitätssicherungsverfahren	14
3.1.3 Machbarkeitsprüfung für Qualitätssicherungsverfahren.....	15
3.1.4 Neuentwicklung von Patientenbefragungen im Rahmen von Qualitätssicherungsverfahren	16
3.2 Grundlagen der systematischen Literatur- und Informationsbeschaffung	19
3.3 Eignung von Qualitätsindikatoren.....	20
3.4 Veröffentlichung von Indikatorergebnissen	32
3.5 Bewertungskonzepte und Referenzbereiche.....	35
3.6 Datenvalidierung von QS-Dokumentationsdaten und Strukturierter Dialog.....	37
3.7 Methoden der Datenanalyse	39
Literatur.....	50

Abkürzungsverzeichnis

Abkürzung	Erklärung
An-Institut	Wissenschaftlicher Beirat des An-Instituts für Qualitätssicherung in der operativen Medizin gGmbH
AOLG	Arbeitsgemeinschaft der Obersten Landesgesundheitsbehörden
AUC	Area under the curve
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V.
BÄK	Bundesärztekammer
BPtK	Bundespsychotherapeutenkammer
BZÄK	Bundeszahnärztekammer
DDG	Deutsche Diabetes Gesellschaft
DGIM	Deutsche Gesellschaft für Innere Medizin e. V.
DGK	Deutsche Gesellschaft für Kardiologie – Herz- und Kreislaufforschung e. V.
DKG	Deutsche Krankenhausgesellschaft e. V.
DNVF	Deutsches Netzwerk Versorgungsforschung e. V.
DPR	Deutscher Pflegerat e. V.
DRV	Deutsche Rentenversicherung Bund
G-BA	Gemeinsamer Bundesausschuss
GKV-SV	Spitzenverband Bund der Krankenkassen (Gesetzliche Krankenversicherung)
GMDS	Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V.
GQMG	Gesellschaft für Qualitätsmanagement in der Gesundheitsversorgung e. V.
IBS-DR	Deutsche Region der Internationalen Biometrischen Gesellschaft
IQTIG	Institut für Qualitätssicherung und Transparenz im Gesundheitswesen
KBV	Kassenärztliche Bundesvereinigung
KZBV	Kassenzahnärztliche Bundesvereinigung
O/E	Verhältnis aus der beobachteten und der erwarteten Rate (<i>observed to expected ratio</i>)

Abkürzung	Erklärung
PatV	Patientenvertretung: maßgebliche Patientenorganisationen nach § 140f SGB V
PatV-Kuratorium	Patientenvertretung im Kuratorium des IQTIG
PID	patientenidentifizierende Daten
plan. QI-RL	Richtlinie zu planungsrelevanten Qualitätsindikatoren
PSN	Patientenpseudonym
Qesü-RL	Richtlinie zur einrichtungs- und sektorenübergreifenden Qualitätssicherung
QFR-RL	Qualitätssicherungs-Richtlinie Früh- und Reifgeborene
QIDB	Qualitätsindikatorendatenbank
QS	Qualitätssicherung
QS PCI	Qualitätssicherungsverfahren <i>Perkutane Koronarintervention (PCI) und Koronarangiographie</i>
QS WI	Qualitätssicherungsverfahren <i>Vermeidung nosokomialer Infektionen – postoperative Wundinfektionen</i>
QSKH-RL	Richtlinie über Maßnahmen der Qualitätssicherung in Krankenhäusern
SGB	Sozialgesetzbuch
SMR	standardisierte Mortalitätsrate (<i>standardized mortality ratio</i>)
SOP	Standard Operating Procedure
WINEG	Wissenschaftliches Institut der TK für Nutzen und Effizienz im Gesundheitswesen

Stellungnahmeverzeichnis

Arbeitsgemeinschaft der Obersten Landesgesundheitsbehörden (AOLG)

Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. (AWMF)

Bundesärztekammer (BÄK)

Bundespsychotherapeutenkammer (BPtK)

Bundesverband der Vertragspsychotherapeuten e. V. (bvvp)

Bundeszahnärztekammer (BZÄK)

Prof. Dr. Andreas Büscher

Deutsche Diabetes Gesellschaft (DDG)

Deutsche Gesellschaft für Gastroenterologie, Verdauungs- und Stoffwechselkrankheiten (DGVS)

Deutsche Gesellschaft für Hämatologie und medizinische Onkologie e. V. (DGHO)

Deutsche Gesellschaft für Innere Medizin e. V. (DGIM)

Deutsche Gesellschaft für Kardiologie – Herz- und Kreislaufforschung e. V. (DGK)

Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e. V. (GMDS)
und Deutsche Region der Internationalen Biometrischen Gesellschaft (IBS-DR)

Deutsche Gesellschaft für Neuroradiologie e. V. (DGNR)

Deutsche Gesellschaft für Orthopädie und Unfallchirurgie e. V. (DGOU), Deutsche Gesellschaft
für Unfallchirurgie e. V. (DGU) und Deutsche Gesellschaft für Orthopädie und Orthopädische
Chirurgie e. V. (DGOOC)

Deutsche Gesellschaft für Palliativmedizin e. V. (DGP)

Deutsche Gesellschaft für Rehabilitationswissenschaften e. V. (DGRW)

Deutsche Krankenhausgesellschaft e. V. (DKG)

Deutsche Rentenversicherung Bund (DRV)

Deutscher Pflegerat e. V. (DPR)

Deutsches Netzwerk Versorgungsforschung e. V. (DNVF)

Prof. Dr. Astrid Elsbernd

Prof. Dr. Max Geraedts und Prof. Dr. Eva Bitzer

Gesellschaft für Qualitätsmanagement in der Gesundheitsversorgung e. V. (GQMG)

GKV-Spitzenverband (GKV-SV)

Kassenärztliche Bundesvereinigung (KBV)

Kassenärztliche Vereinigung Rheinland-Pfalz (KVRLP)

Kassenzahnärztliche Bundesvereinigung (KZBV)

Prof. Dr. David Klemperer und Prof. Dr. Dr. Martin Härter

Marburger Bund Bundesverband e. V.

Dr. Ulrike Ossig

Patientenvertretung im Kuratorium des IQTIG

Patientenvertretung: maßgebliche Organisationen nach § 140f SGB V

Paul-Ehrlich-Gesellschaft für Chemotherapie e. V. (PEG)

Picker Institut Deutschland gGmbH

Prof. Dr. Matthias Schrappe

Dr. Joachim Steller

Prof. Dr. Stephanie Stock

Wissenschaftlicher Beirat des An-Instituts für Qualitätssicherung in der operativen Medizin
gGmbH

Wissenschaftliches Institut der TK für Nutzen und Effizienz im Gesundheitswesen (WINEG)

1 Stellungnahmeverfahren zu Version 1.0s

Am 1. Februar 2017 veröffentlichte das IQTIG die Version 1.0s seiner „Methodischen Grundlagen“ und leitete das Stellungnahmeverfahren dazu ein. Die „Methodischen Grundlagen“ stellen die wissenschaftlichen Arbeitsgrundlagen des IQTIG als fachlich unabhängiges, wissenschaftliches Institut nach § 137a SGB V dar. Sie umfassen die Methoden der Entwicklung, Weiterentwicklung und Durchführung von Qualitätssicherungsverfahren im Auftrag des Gemeinsamen Bundesausschusses (G-BA). Bis zum 31. März 2017 hatten die Institutionen des Gesundheitswesens, wissenschaftliche Institutionen, Patientinnen und Patienten und deren Vertreterinnen und Vertreter sowie weitere fachkundige Einzelpersonen die Gelegenheit, schriftliche Stellungnahmen zu den „Methodischen Grundlagen V1.0s“ einzureichen. Insgesamt gingen 40 Stellungnahmen von Institutionen, Fachgesellschaften, den Trägerorganisationen des G-BA, Berufsverbänden und Einzelpersonen fristgerecht ein. Die eingegangenen Stellungnahmen wurden gesichtet, gewürdigt und die „Methodischen Grundlagen“ wurden auf Basis der Stellungnahmen überarbeitet. Am 15. September 2017 wurde die überarbeitete Version der „Methodischen Grundlagen V1.0“ auf der Website des Instituts veröffentlicht.

Das Institut bedankt sich ausdrücklich bei allen Stellungnehmenden für die konstruktive Kritik an den „Methodischen Grundlagen V1.0s“. Im vorliegenden Dokument werden die zentralen Aspekte aus den Stellungnahmen zusammengefasst und es wird erläutert, wie das Institut mit den vorgebrachten Hinweisen und der Kritik umgegangen ist. Die Hinweise und die Kritik aus den Stellungnahmen sind nach der Reihenfolge der Kapitel, wie sie in den „Methodischen Grundlagen“ vorkommen, geordnet.

2 Würdigung allgemeiner, themenübergreifender Stellungnahmen

Einige Stellungnahmen kritisieren das Fehlen von Definitionen für verschiedene Begriffe oder die gewählten Definitionen für diese Begriffe, wie etwa für „Qualitätssicherung“, „Qualitätsmodell“, „Qualitätsmerkmal“ und „Qualitätsaspekt“ (Stock, S. 2; Schrappe, S. 8 ff.; DNVF, S. 2; DRV, S. 1; KBV, S. 9 f.).

IQTIG: Gerade der Begriff der „Qualitätssicherung“ wird in unterschiedlichen Kontexten unterschiedlich verwendet. So wird „Qualitätssicherung“ in der DIN EN ISO 9000:2015 definiert als der „Teil des Qualitätsmanagements [...], der auf das Erzeugen von Vertrauen darauf gerichtet ist, dass Qualitätsanforderungen [...] erfüllt werden“ (DIN, NQSZ 2015). Von dieser Definition weicht das IQTIG jedoch in Bezug auf sein Tätigkeitsfeld im Gesundheitswesen ab. Aufgabe des IQTIG als Einrichtung der Selbstverwaltung im Gesundheitswesen ist vor allem die unabhängige Messung von Versorgungsqualität in Kooperation mit den verschiedenen Partnern im Gesundheitswesen sowie die transparente Darstellung der Ergebnisse. Gerade weil das IQTIG nicht primär beauftragt ist, Vertrauen zu schaffen, sondern die Versorgungsqualität kritisch zu prüfen, kann es dem vertrauensbildenden Ziel der Qualitätssicherungsmaßnahmen insgesamt dienlich sein. Nicht das IQTIG soll Vertrauen schaffen, sondern die Ergebnisse der Qualitätssicherungsmaßnahmen im Gesundheitswesen sollen das bewirken. Damit unterscheidet sich der Qualitätssicherungsbegriff der externen Qualitätssicherung in der konkreten Umsetzung von dem der DIN EN ISO 9000:2015 und folgt dem Gebrauch in den gesetzlichen Regelungen (SGB V). Allerdings soll natürlich der Einsatz der Qualitätssicherung als politisches Element der Steuerung des Gesundheitswesens das Vertrauen der Bevölkerung in die gesundheitliche Versorgung wachsen lassen. In diesem übergeordneten Rahmen trifft die Definition der DIN EN ISO 9000:2015 zu. Es kommt darauf an, auf welchen Rahmen oder auf welches Handlungsfeld diese Definition angewandt wird. In der überarbeiteten Version der „Methodischen Grundlagen 1.0“ wurde der Begriff der Qualitätssicherung ausführlicher erläutert und eingeordnet.

Auch die für die Methodik des IQTIG wichtigen Begriffe „Qualitätsmodell“, „Qualitätsmerkmal“ und „Qualitätsaspekt“ werden in der überarbeiteten Version der „Methodischen Grundlagen“ ausführlicher erläutert. Qualitätsaspekte sind konkrete Themen der Versorgungspraxis, wie z. B. Indikationsstellung zur Herzkatheteruntersuchung oder Komplikationen bei Hüftoperationen, anhand derer sich die Qualität der Gesundheitsversorgung beschreiben lässt. Zu den Qualitätsaspekten werden Qualitätsmerkmale ermittelt, d. h. konkrete Eigenschaften der Versorgung, die mit Anforderungen verknüpft sind. Die Ausprägung der Qualitätsmerkmale in der Versorgungspraxis wird dann durch Qualitätsindikatoren gemessen und bewertet. Die Summe der für

ein Qualitätssicherungsverfahren ausgewählten Qualitätsaspekte wird als Qualitätsmodell für dieses Verfahren bezeichnet.

In manchen Stellungnahmen wird eine klarere Erläuterung bzw. Herleitung der definierten Zwecke der Qualitätssicherung „Qualitätsförderung“, „Information“ und „Regulierung“ gefordert (DNVF, S. 2; KBV, S. 5 f.; GKV-SV, S. 3 und 7 ff., PatV S. 8).

IQTIG: Die externe Qualitätssicherung nach den Richtlinien des G-BA war in der Vergangenheit im stationären Bereich auf die Maßnahmen der Richtlinie über Maßnahmen der Qualitätssicherung in Krankenhäusern (QSKH-RL) beschränkt. Qualitätssicherungsverfahren im Rahmen dieser Richtlinie haben den Zweck, die Qualität der Versorgung zu fördern, indem durch das Instrument des Strukturierten Dialogs etwaige Qualitätsprobleme bei den Leistungserbringern aufgedeckt und dann gegebenenfalls behoben werden sollen. Durch verschiedene Novellierungen des SGB V wurde das Spektrum der gesetzlichen Qualitätssicherung mittlerweile jedoch deutlich erweitert. So ist nun die Berücksichtigung der Versorgungsqualität bei der Krankenhausplanung (planungsrelevante Qualitätsindikatoren, § 136c SGB V), bei der Vergütung (qualitätsbezogene Zu- und Abschläge; § 136b SGB V) und beim Abschluss von Selektivverträgen (Qualitätsverträge; § 110a SGB V) zwischen Krankenkassen und Leistungserbringern durch den Gesetzgeber vorgesehen. Im Gegensatz zur Qualitätsförderung durch den Strukturierten Dialog greifen diese Maßnahmen nun regulierend, beispielsweise durch Vergütungsanreize, in die Qualitätsentwicklung ein.

Darüber hinaus kann ein Zweck von Qualitätssicherungsverfahren die gezielte Information der Öffentlichkeit sein. Dies ist beispielsweise gesetzlicher Auftrag des IQTIG nach § 137a Abs. 3 SGB V. Solch eine gezielte Information der Öffentlichkeit geht über den reinen Anspruch der Transparenz hinaus, da dafür unter anderem Informationen zielgruppenspezifisch aufbereitet und dargestellt werden sollen.

Diese unterschiedlichen Zwecke, die Qualitätssicherungsverfahren verfolgen können, stellen mitunter unterschiedliche Anforderungen an die Messung und Darstellung der Qualität. Daher werden diese Verfahrenszwecke der Qualitätssicherung in der überarbeiteten Version der „Methodischen Grundlagen“ spezifisch und ausführlicher erläutert.

In einigen Stellungnahmen wird kritisiert, dass die Beschreibung der Schritte, nach denen Qualitätssicherungsverfahren entwickelt werden, das Vorgehen nicht reproduzierbar mache (DNVF, S. 1; BpTK, S. 3; GKV-SV, S. 6).

IQTIG: Das IQTIG teilt nicht die Auffassung, dass die „Methodischen Grundlagen“ die Entwicklungsschritte so beschreiben könnten, dass sie es Dritten erlaubten, zu den gleichen Entwicklungsprodukten zu kommen wie das Institut. Die „Methodischen Grundlagen“ stellen vielmehr einen wissenschaftlichen

Rahmen dar, innerhalb dessen Qualitätssicherungsverfahren entwickelt werden. Darin sind viele methodische Grundsatzentscheidungen festgehalten, wie etwa die faktenorientierte Ausrichtung bei der Entwicklung von Patientbefragungen, der Einsatz von Fokusgruppen zur Abbildung der Patientensperspektive und die Annahme von Stochastizität bei der Messung von Qualitätsindikatoren. Durch die transparente Darlegung und wissenschaftliche Begründung der „Methodischen Grundlagen“ und der Entwicklungsberichte stellt das Institut die Wissenschaftlichkeit seiner Entwicklungsprodukte sicher. Eine rezeptartige Beschreibung der Entwicklungsmethodik, die, genau befolgt, zu dem gleichen Entwicklungsprodukt führt wie das des Instituts, ist im Rahmen der „Methodischen Grundlagen“ weder sinnvoll noch möglich. An vielen Stellen im Entwicklungsprozess müssen Entscheidungen getroffen werden, die die weitere Entwicklung beeinflussen, wie etwa die Zusammensetzung von Fokusgruppen, die Entwicklung von MeSH-Terms für die Literaturrecherche oder die Entscheidung für eine Methode der Veränderungsmessung. Solche Entscheidungen können nicht vollumfänglich in den „Methodischen Grundlagen“ antizipiert und als algorithmische Regeln beschrieben werden. Sie sind immer vor dem Hintergrund einer konkreten Beauftragung und dem methodischen und gesetzlichen Rahmen, innerhalb dessen sich das Institut bewegt, abzuwägen und zu begründen. Daher legt das Institut in seinen Abschlussberichten immer in ausführlichen Begründungen dar, wie und weshalb das Institut zu den vorgelegten Ergebnissen kam.

In einer Stellungnahme wird eine allgemeinverständlichere Darstellung gefordert, damit auch Patientinnen und Patienten sowie Gesundheitsfachkräfte der Routineversorgung das Dokument verstehen könnten (PatV-Kuratorium, S. 4).

IQTIG: Die „Methodischen Grundlagen“ haben das Ziel, durch transparente Darstellung der wissenschaftlichen Grundsätze, nach denen das IQTIG Qualitätssicherungsverfahren entwickelt, diese nachvollziehbar und damit auch kritisierbar zu machen. Um das zu erreichen, müssen sich die Beschreibungen bisweilen einer spezifischen wissenschaftlichen Sprache bedienen, die nicht immer allgemeinverständlich sein kann. Damit richten sich die „Methodischen Grundlagen“ in erster Linie an Personen, die in der Qualitätssicherung Erfahrung haben und die in die Thematik eingearbeitet sind, wie z. B. die Vertreterinnen und Vertreter der Trägerorganisationen des G-BA. Während also die „Methodischen Grundlagen“ nicht den Anspruch haben können, in allen Einzelheiten laienverständlich zu sein, müssen es die Ergebnisse der Qualitätssicherung umso mehr sein. Daher hat sich das IQTIG das Ziel gesetzt, die Ergebnisse der Qualitätssicherung zukünftig sowohl für die Gesundheitsfachkräfte als auch für die Patientinnen und Patienten besser verständlich und vor allem hilfreich darzustellen. Selbstverständlich ist das IQTIG auch bemüht, seine „Methodischen Grundlagen“ stets so verständlich wie möglich zu gestalten. Zur Verständlichkeit der Ergebnisse zählt eben auch die Verständlichkeit, wie diese hergeleitet sind. Hierzu wird ein optimaler Kompromiss zwischen Verständlichkeit und wissenschaftlicher Präzision in der Beschreibung angestrebt.

Einige Stellungnahmen betonen das noch vorhandene Entwicklungspotenzial der „Methodischen Grundlagen“ im Allgemeinen (KBV, S. 6) oder dass wichtige Elemente noch nicht dargestellt seien, wie z. B. Ausführungen zu planungsrelevanten Qualitätsindikatoren und qualitätsorientierten Zu- und Abschlägen (GKV-SV, S. 7; PatV-Kuratorium, S. 11; BÄK S. 22; AOLG, S. 3 f.).

IQTIG: Die „Methodischen Grundlagen 1.0“ haben nicht den Anspruch, alle aktuellen Herausforderungen der gesetzlichen Qualitätssicherung abzubilden. Die Version 1.0 deckt bisher nur die Entwicklung von Qualitätssicherungsverfahren ab, die mit dem Strukturierten Dialog eine Förderung der Qualität in den Einrichtungen verfolgen. Die Methoden der Entwicklung neuer Qualitätssicherungsverfahren, die eine qualitätsbezogene Krankenhausplanung (planungsrelevante Qualitätsindikatoren) oder eine qualitätsabhängige Vergütung (Zu- und Abschläge) verfolgen, sind explizit nicht Bestandteil der vorliegenden Version 1.0, da bei ihrer Veröffentlichung diese Methoden noch nicht abschließend entwickelt waren. Die „Methodischen Grundlagen“ werden jedoch regelmäßig weiterentwickelt, aktualisiert und überarbeitet, sodass auch diese Themen mit den kommenden Versionen abgedeckt werden sollen. An anderen Stellen der Version 1.0 besteht ebenfalls Weiterentwicklungsbedarf, wie etwa bei der Risikoadjustierung von Qualitätsindikatoren oder bei der Darstellung des konkreten Vorgehens bei der Eignungsprüfung von Qualitätsindikatoren. Diese und weitere Themen sollen im Zuge der Überarbeitung in den kommenden Versionen der „Methodischen Grundlagen“ detaillierter Berücksichtigung finden.

In einigen Stellungnahmen wird kritisiert, dass Patientinnen und Patienten in die Verfahrensentwicklung nur mangelnd eingebunden seien (DNVF, S. 2) und dass die Patientenorientierung der Qualitätssicherungsverfahren fehle oder nur vage skizziert sei (PatV-Kuratorium, S. 5; Schrappe, S. 6; Klemperer, S. 1).

IQTIG: Die frühzeitige Einbindung der Patientenperspektive in die Entwicklung von Qualitätssicherungsverfahren ist für das IQTIG unabdingbar. Dazu werden zum einen im Rahmen von Konzeptstudien Fokusgruppen durchgeführt und zum anderen Patientinnen und Patienten in die beratenden Expertengremien berufen, die die Entwicklung von Qualitätssicherungsverfahren begleiten. Darüber hinaus können qualitative Studien aus der Literaturrecherche eine wichtige Rolle bei der Ermittlung und Abbildung der Patientenperspektive hinsichtlich Erfahrungen, Einstellungen, Haltungen und Akzeptanz im Rahmen der Entwicklung von Qualitätssicherungsverfahren spielen. Somit wird die Auswahl der Qualitätsaspekte für ein Qualitätssicherungsverfahren auch dadurch geleitet, inwiefern sie am Patientennutzen und an den Interessen und Bedürfnissen der Patientinnen und Patienten orientiert sind. Dies ist den Abschnitten der „Methodischen Grundlagen“ zur Verfahrensentwicklung geschildert. Aufgrund der hohen Bedeutung der Patientenorientierung wurde zudem ein Abschnitt zur Patientenorientierung ergänzt (Abschnitt 2.1.3).

3 Würdigung der Stellungnahmen zu einzelnen Kapiteln

3.1 Neuentwicklung von Qualitätssicherungsverfahren

Wiederholt wird in den Stellungnahmen danach gefragt, welche Modifikationen sich in den Entwicklungsprojekten ergeben, wenn ein anderer Verfahrenszweck als die Qualitätsförderung gesetzt sei. Es wird zudem vorgeschlagen, zunächst nur von einem „vorläufigen Verfahrenszweck“ zu sprechen und erst nach der Konzeptstudie einen „definitiven Verfahrenszweck“ zu formulieren (PatV-Kuratorium, S. 16; KZBV, S. 6).

IQTIG: In der Version 1.0 der „Methodischen Grundlagen“ wird noch nicht auf andere Verfahrenszwecke als die Qualitätsförderung eingegangen. Dies ist ab Version 2.0 nach Abschluss der ersten Projekte mit einem anderen Verfahrenszweck geplant.

Das IQTIG kann sich keinen „vorläufigen“ Verfahrenszweck für einen Projektauftrag vorstellen. Qualitätsziel und Messmethodik als zwei der drei Hauptelemente eines Qualitätsindikators können unabhängig vom Verfahrenszweck definiert werden. Bewertungsregeln für einen Qualitätsindikator müssen jedoch immer auf dem Verfahrenszweck aufbauen. Über den Verfahrenszweck sollte nicht erst auf Basis der verfahrensspezifischen Konzeptstudie, sondern bereits auf Basis des zuvor vom G-BA erstellten Kriterienkatalogs von diesem entschieden werden. Wenn vom G-BA gewünscht, kann das IQTIG hierbei beratend unterstützen.

Es wird in einer Stellungnahme angeregt, die Qualitätsziele in den Entwicklungsschritten explizit zu formulieren, da diese als „Anker“ für das Qualitätsmodell dienen könnten (DKG, S. 30).

IQTIG: Es wurden Erläuterungen zur Ableitung der konkreten Qualitätsziele in den Abschnitten 4.1 und 4.2 aufgenommen.

Es wird in einer Stellungnahme nachgefragt, wer entscheide, welche Expertinnen und Experten in die Entwicklungsschritte eingebunden werden sollen (KZBV, S. 8).

IQTIG: Das IQTIG als wissenschaftlich unabhängiges Institut entscheidet anhand transparenter Kriterien darüber, welche Expertinnen und Experten das IQTIG in einem bestimmten Fachbereich beraten sollen. Die ausgewählten Expertinnen und Experten, ihre Qualifikationen und potenziellen Interessenkonflikte werden in den Abschlussberichten an den G-BA und durch die nachfolgende Veröffentlichung der Berichte transparent gemacht. Beim Umgang mit Interessenkonflikten folgt das IQTIG den entsprechenden, vom Vorstand genehmigten Regelungen, die Teil der Geschäftsordnung des IQTIG sind.

Es wird darum gebeten auszuführen, für welche Berichte Stellungnahmeverfahren vorgesehen seien (KZBV, S. 6).

IQTIG: Im Text wurde ergänzt, dass Stellungnahmeverfahren sowohl in der dokumentations- und sozialdatenbasierten Neuentwicklung von Qualitätsindikatoren als auch in der Entwicklung von Patientenbefragungen vorgesehen sind.

3.1.1 Konzeptstudie für Qualitätssicherungsverfahren

Es wird in einer Stellungnahme bemängelt, dass die Darstellung der Versorgungspraxis in der Konzeptstudie zu sehr medizinisch ausgerichtet sei und berufsgruppenübergreifende Anstrengungen nicht ausreichend berücksichtige (DPR, S. 2).

IQTIG: Im Text wurden mehrere Ergänzungen im Sinne der „ärztlichen, pflegerischen und therapeutischen Versorgungspraxis“ vorgenommen.

Es wird von einer stellungnehmenden Organisation gefordert, dass im Rahmen der Aufwand-Nutzen-Abschätzung eines Qualitätssicherungsverfahrens in einer Konzeptstudie eine Kosten-Nutzen-Analyse durchgeführt werden solle (KBV, S. 12).

IQTIG: Eine Kosten-Nutzen-Analyse durch das IQTIG ist an keiner Stelle vorgesehen. Es kann lediglich eine grobe Abschätzung des Aufwands-Nutzen-Verhältnisses erfolgen.

3.1.2 Dokumentations- und sozialdatenbasierte Neuentwicklung von Qualitätssicherungsverfahren

In einer Stellungnahme wird kritisiert, dass es unklar sei, in welchem Verhältnis während der Neuentwicklung eines Qualitätssicherungsverfahrens etablierte Qualitätsindikatoren (z. B. aus der Literatur) und De-novo-Entwicklungen stehen werden/sollen (DKG, S. 12).

IQTIG: Aus der Literatur zu entnehmende Qualitätsindikatoren stellen lediglich Hinweise und Vorschläge für mögliche Operationalisierungen von Qualitätsmerkmalen dar, die den vorgegebenen Prüf- und Entwicklungsschritten des IQTIG unterliegen. Sie gehen nie unmittelbar in ein Indikatorenset ein, das zur Anwendung empfohlen wird.

Im Falle der Lieferung von Sozialdaten der Krankenkassen sollte eine frühzeitige Einbindung aller kooperierenden Kassen erfolgen, um eine einheitliche Qualität der zu liefernden Indikatoren sicherzustellen. (WINEG, S. 5)

IQTIG: Die Absprachen mit allen Krankenkassen werden über den GKV-SV, z. T. auch über den Fachausschuss QS-IT und Spezifikation des G-BA durchgeführt. Das IQTIG entwickelt eine Sozialdatenspezifikation auf Basis der vom G-BA beschlossenen Allgemeinen Sozialdatenspezifikation. Die Krankenkassen liefern Daten an eine eigene Datenannahmestelle der Krankenkassen, die die Daten

leistungserbringerbezogen pseudonymisiert an die Vertrauensstelle weiterleitet, die die Daten patientenbezogen pseudonymisiert. Erst dann gelangen die Daten zur Bundesauswertungsstelle (IQTIG). Bislang erfolgte noch keine Lieferung von Sozialdaten bei den Krankenkassen im Regelbetrieb, sodass noch keine diesbezüglichen Erfahrungen vorliegen.

3.1.3 Machbarkeitsprüfung für Qualitätssicherungsverfahren

Es wird in einer Stellungnahme bemängelt, dass es unklar bliebe, ob Machbarkeitsprüfungen nach Vorstellung des IQTIG standardmäßig oder nur, wenn explizit vom G-BA beauftragt durchgeführt werden sollen (GKV-SV, S. 15).

IQTIG: Der Text wurde dahingehend ergänzt, dass regelhaft in der Entwicklung eines Qualitätssicherungsverfahrens eine Machbarkeitsprüfung beauftragt werden sollte.

In mehreren Stellungnahmen wird hinterfragt, warum keine Probebetriebe zur Prüfung der Datenflüsse mehr vorgesehen werden. Dann stelle sich aber die Frage, wer hierfür die Software zur Verfügung stelle und die Kosten übernehme (PatV-Kuratorium, S. 16; KZBV, S. 12).

IQTIG: Die Erfahrung der früheren Jahre hat gelehrt, dass kein Unternehmen eine kostenlose Probesoftware entwickeln wird, Gelder hierfür aber vom G-BA nicht bereitgestellt werden können. Aus diesem Grunde wurde der Entwicklungsschritt eines Probebetriebs vom G-BA aufgegeben. Die Datenflüsse selbst können in der Machbarkeitsprüfung somit nicht getestet werden. Vielmehr wurde vom G-BA eine Erprobungsphase im Regelbetrieb (verpflichtende Teilnahme, Software des Regelbetriebs) eingeführt. Da diese durch die themenspezifischen Bestimmungen der Richtlinie zur einrichtungs- und sektorenübergreifenden Qualitätssicherung (Qesü-RL) festgelegt wird, ist sie nicht Teil der in den Abschnitten 4.1 bis 4.3 dargestellten Entwicklungsschritte bis zum Regelbetrieb.

Für die Machbarkeitsprüfung in der vorgeschlagenen Form bedarf es keiner Filter- oder Dokumentationssoftware, die wenigen Fälle werden durch Auswahl-/Filterschritte im Krankenhausinformationssystem (KIS) bzw. in der Praxissoftware identifiziert und händisch dokumentiert.

Es wird von einer stellungnehmenden Organisation angeregt, dass den Expertinnen und Experten, die an der Entwicklung des Verfahrens beteiligt waren, eine abschließende fachliche Einschätzung des finalen Indikatorensets ermöglicht wird (BÄK, S. 11).

IQTIG: Der Vorschlag wurde aufgegriffen (Abschnitt 4.3.4).

3.1.4 Neuentwicklung von Patientenbefragungen im Rahmen von Qualitätssicherungsverfahren

Aussagekraft von Patientenbefragungen

In einzelnen Stellungnahmen wird die Aussagekraft von Patientenbefragungen zu Zwecken der Qualitätssicherung oder im Allgemeinen bezweifelt (DGK, S. 2). So wird beispielsweise die Vermutung geäußert, Patientinnen und Patienten könnten grundsätzlich wenig bis nichts über Behandlungsabläufe oder den Behandlungserfolg berichten (DKG, S. 13; DGK, S. 2). Als Gründe werden unter anderem fehlendes Fachwissen der Patientinnen und Patienten und vermutete Subjektivität der Antworten angeführt (Steller, S. 8).

IQTIG: International haben sich Patientenbefragungen zur Beurteilung der Behandlungsqualität etabliert (siehe z. B. National Health Service in Großbritannien, Centers for Medicare and Medicaid Services in den USA). Da die Patientinnen und Patienten aus erster Hand über den Versorgungsprozess berichten können, ist die Einbindung ihrer Erfahrungen in die Qualitätsbeurteilung als eine zusätzliche Informationsquelle unabdingbar. Dabei nehmen die Patientinnen und Patienten jedoch selbst keine Beurteilung der Behandlungsqualität vor, sondern berichten lediglich über ihre Erfahrungen im Rahmen der Behandlung oder über mögliche Veränderungen in ihrem Gesundheitszustand (z. B. Symptome, Fähigkeiten). So können Patientinnen und Patienten beispielsweise angeben, ob sie über alternative Behandlungsmöglichkeiten informiert wurden, ob ihnen respektvoll begegnet wurde oder mit welchen Schmerzen sie konfrontiert waren. Patientenzentrierung heißt, dass genau diese und ähnliche Aspekte bei der Beurteilung der Versorgungsqualität berücksichtigt werden müssen. Dagegen werden Patientinnen und Patienten nicht zu medizinischen Details befragt, für deren Verständnis Fachwissen erforderlich ist.

Sowohl in der Versorgungsforschung als auch in der Qualitätsmessung wurden wissenschaftliche Verfahren etabliert, mit denen valide Instrumente zur Befragung von Patientinnen und Patienten entwickelt werden können, deren Ergebnisse zu belastbaren Aussagen zur Versorgungsqualität führen. Dabei wird natürlich darauf geachtet, dass die Fragen sich auf die Themen beziehen, zu denen Patientinnen und Patienten relevante Aussagen machen können. Nach § 137a Abs. 3 SGB V ist die Entwicklung von Patientenbefragungen zur Messung und Darstellung der Versorgungsqualität eine gesetzlich verankerte Aufgabe des IQTIG. Es wird erwartet, dass die bisherige Qualitätsbewertung mit diesen Befragungen um wichtige, nützliche Aspekte erweitert wird.

Vom Fragebogenitem zum Qualitätsindikator

Für eine stellungnehmende Organisation erschien die Darstellung des Wegs vom Fragebogenitem zum Qualitätsindikator nicht hinreichend nachvollziehbar (DNVF, S. 2).

IQTIG: In der überarbeiteten Fassung der „Methodischen Grundlagen“ wurde dieser Entwicklungsschritt nun deutlicher dargestellt. Grundsätzlich werden neben Items zur Erfassung der soziodemographischen Daten der Befragten und Items zur Risikoadjustierung nur solche Items in den Fragebögen aufgenommen, die sich als Qualitätsindikatoren entweder für sich genommen oder zusammengefasst in Form eines Index eignen. Als Qualitätsindikatoren gelten grundsätzlich Qualitätsmessungen, die ein konkretes Qualitätsziel verfolgen und eine Qualitätsbewertung ermöglichen (d. h. einen Referenzbereich aufweisen).

Details zu Methoden und Herausforderungen von Patientenbefragungen

In einigen Stellungnahmen werden weitergehende Details zu den Methoden und Herausforderungen der Patientenbefragung gefordert, wie etwa Veränderungsmessungen, Befragungszeitpunkt, Erinnerungsfähigkeit, Non-Response und Stichprobenziehung (Picker Institut, S. 3; KBV, S. 14 f.; GKV-SV, S. 22; WINEG, S. 5; BZÄK, S. 5; DKG, S. 14).

IQTIG: Der Wunsch nach Informationen zu diesen zentralen Aspekten von Patientenbefragungen ist nachvollziehbar, diesem kann allerdings nicht im Rahmen der „Methodischen Grundlagen“ entsprochen werden. Da das IQTIG keinen generischen Fragebogen, sondern für jedes Befragungsthema ein spezifisches Instrument entwickelt, muss auch für jede Befragung individuell entschieden werden, wie diese methodischen Fragen adressiert werden. Beispielsweise gibt es keine pauschal richtige Antwort auf die Frage, welche Methode der Veränderungsmessung geeignet ist. Dies kann zwischen verschiedenen Befragungsthemen variieren. Gleiches gilt für die Beurteilung der Erinnerungsfähigkeit und den Befragungszeitpunkt, der in Abhängigkeit des Verfahrens jedes Mal neu ermittelt wird. Handelt es sich beispielsweise um eine diagnosebezogene Befragung, ist der optimale Befragungszeitpunkt anders zu ermitteln als etwa für eine prozedurbezogene Befragung. Fragen zu konkreten Ereignissen in der Behandlung wie z. B. Aufklärungsinhalte sollten eher einen kürzeren Erinnerungszeitraum haben, dagegen zeichnen sich manche Behandlungseffekte erst zu einem späteren Zeitpunkt ab. Ebenso wird für jede zu entwickelnde Befragung erneut geprüft, welches Verfahren zur Erhöhung der Rücklaufquote am geeignetsten erscheint. Diese Fragen werden für jede zu entwickelnde Befragung individuell beantwortet und entziehen sich damit einer pauschalen Beantwortung. Konkrete Antworten auf diese Fragen werden in den Abschlussberichten zu den jeweiligen Entwicklungen von Patientenbefragungen gegeben.

Details zum Ablauf von Patientenbefragungen

In einigen Stellungnahmen wird der Wunsch nach weiteren Informationen zum konkreten Ablauf von Patientenbefragungen wie etwa zum Modus der Dateneingabe (von Hand, maschinell), zum Datenschutz, zum Datenfluss und zu der Fragebogenlogistik, geäußert (KBV, S. 15 f.; PatV, S. 18).

IQTIG: Derartige verfahrenstechnische Aspekte sind Bestandteil aktueller Entwicklungsarbeiten für Patientenbefragungen. Die verfahrensübergreifenden Vorgehensweisen, die mit der ersten Patientenbefragung entwickelt werden, müssen in der Praxis etabliert werden, um im Regelbetrieb Anwendung zu finden. Die Berücksichtigung von datenschutzrechtlichen Aspekten ist dabei unabdingbar. Die „Methodischen Grundlagen“ beinhalten jedoch die wissenschaftlichen Grundlagen für die Entwicklung und Durchführung von Qualitätssicherungsverfahren im Allgemeinen und von Patientenbefragungen im Speziellen. Technische Aspekte wie etwa die Fragebogenlogistik sind nicht Teil der wissenschaftlichen Entwicklungsmethoden.

Validität von Patientenbefragungen

In einigen Stellungnahmen werden weitergehende Darstellungen zu den Methoden zur Sicherstellung der Validität der Befragungsergebnisse, die über die Sicherstellung der Inhaltsvalidität hinausgehen, gefordert (DNVF, S. 2; WINEG, S. 5; QQMG, S. 3).

IQTIG: Das IQTIG setzt bei der Entwicklung von Patientenbefragungen schwerpunktmäßig auf einen faktenorientierten Befragungsansatz, bei dem konkrete Ereignisse und Erfahrungen der Patientinnen und Patienten sowie patientenbezogene Ergebnisse im Mittelpunkt der Befragung stehen. Dies sei am Beispiel folgender, für diese Befragungen typischen Frage erläutert: „Eine Ärztin oder ein Arzt hat mir erklärt, wie ich meine Medikamente einnehmen soll“ (Antwortmöglichkeiten: „Ja“, „Nein“, „Weiß nicht“). Die einzige, eher hypothetische Möglichkeit zu überprüfen, ob die Antwort der Patientin tatsächlich dem Sachverhalt entspricht, wäre, die Situation des Aufklärungsgesprächs zu protokollieren und das Protokoll mit der Antwort der Patientin abzugleichen. Methoden zur Prüfung der Konstruktvalidität helfen nicht, die Frage zu beantworten, ob die Antwort der Patientinnen und Patienten den Ereignishergang wahrheitsgemäß wiedergibt. Auch Faktoranalysen mit weiteren ähnlichen Fragen, die häufig zur Prüfung der sogenannten faktoriellen Validität herangezogen werden, helfen nicht festzustellen, ob die Antwort der Patientinnen und Patienten den Ereignishergang wahrheitsgemäß wiedergibt. Viel wichtiger als mögliche Korrelationen zwischen den Antworten auf verschiedene Fragen ist, ob die Fragen relevante und aussagekräftige Qualitätsmerkmale der Behandlung abbilden. Anders formuliert: Ist es ein Merkmal guter Behandlungsqualität, wenn möglichst viele Patientinnen und Patienten angeben, darüber aufgeklärt worden zu sein, wie sie ihre Medikamente einnehmen sollen?

Daher erfolgt die Sicherstellung der Validität der Befragungsergebnisse von Patientenbefragungen auf zwei Ebenen: Zum einen wird durch den in den

„Methodischen Grundlagen“ beschriebenen Prozess der Entwicklung von Qualitätsmerkmalen und Items sichergestellt, dass das Set an Items in dem endgültigen Fragebogen tatsächlich die qualitätsrelevanten Themen abdeckt – dieser Teil der Validierungsstrategie betrifft die Inhaltsvalidität der Befragung. Zum anderen muss sichergestellt werden, dass die qualitätsbezogenen Aussagen, die aus den späteren Befragungsergebnissen abgeleitet werden sollen, tatsächlich gerechtfertigt und belastbar sind. Dies kann durch verschiedene Faktoren beeinflusst werden wie etwa Non-Response, mangelnde Erinnerungsfähigkeit oder eine nicht adäquate Methode der Veränderungsmessung. Diese diversen und von Befragungsthema zu Befragungsthema teilweise variierenden Einflussfaktoren werden vom IQTIG in der Entwicklung aufgegriffen, z. B. im Rahmen der Pretestung der Fragebögen, und im Befragungskonzept entsprechend adressiert.

3.2 Grundlagen der systematischen Literatur- und Informationsbeschaffung

Auswahl der Instrumente bei der kritischen Bewertung relevanter Publikationen

In einigen Stellungnahmen wird die Auswahl der Bewertungsinstrumente von Publikationen angesprochen (u. a. BZÄK, S. 6; Stock, S. 2; KZBV, S. 18).

IQTIG: Bei der kritischen Bewertung von Publikationen orientiert sich das Institut an international etablierten standardisierten Bewertungsinstrumenten. Unter Berücksichtigung der aktuellen wissenschaftlichen Methoden und Diskussionen wurden einige Instrumente hierfür beispielhaft erwähnt. In der überarbeiteten Fassung der „Methodischen Grundlagen“ wurde dies ergänzt.

Zur Leitlinienbewertung wendet das Institut das auf Deutsch übersetzte sowie in einer überarbeiteten Version vorliegende AGREE-II-Instrument (Appraisal of Guidelines of Research & Evaluation) an (Shea et al. 2007), da das Deutsche Leitlinien-Bewertungsinstrument (DELBI, modifiziert nach AGREE) derzeit noch nicht aktualisiert vorliegt.

Die in einer Stellungnahme beispielhaft genannten Standards stellen Leitlinien für Forschungsberichte dar, die dazu gedacht sind, als Orientierungshilfe Mängel der Berichterstattung zu verhindern bzw. die Berichtsqualität zu verbessern. Sie dürfen nicht gleichgesetzt werden mit Bewertungsinstrumenten, die versuchen, das Verzerrungspotenzial einer Studie einzuschätzen.

Bewertung qualitativer Studien

In einigen Stellungnahmen wird auch für qualitative Studien eine Bewertung der Studienqualität vorgeschlagen (GKV-SV, S. 17; DPR, S. 4; BZÄK, S. 6).

IQTIG: Dieser Abschnitt (6.3.4) wurde diesbezüglich entsprechend ergänzt: „Ausgehend von der Form der Evidenzaufbereitung werden eingeschlossene qualitative Studien von mindestens zwei Personen bzgl. ihrer methodischen Qualität mithilfe eines Bewertungsinstrumentes (z. B. der Checkliste des Critical Appraisal Skills Programme/CASP) bewertet.“

3.3 Eignung von Qualitätsindikatoren

Allgemeine Themen

Herleitung der Eignungskriterien

Einige Stellungnahmen merken an, dass die Eignungskriterien im Text der „Methodischen Grundlagen“ ausführlicher begründet und der Bezug und die Unterschiede zu anderen Kriterienkatalogen detaillierter dargestellt werden sollten (KBV, S. 21; BÄK, S. 13; GQMG, S. 4; GKV-SV, S. 15).

IQTIG: Die Anregung wurde aufgegriffen und der Text an einigen Stellen ausführlicher gestaltet, um die Verständlichkeit zu erhöhen und Missverständnissen besonders bei solchen Begriffen und Konzepten vorzubeugen, die in der Vergangenheit oder in anderen Zusammenhängen anders verwendet wurden. Die Funktion einer Systematik von Eignungskriterien besteht allerdings darin, dass bei einer Entwicklung, Weiterentwicklung oder Überprüfung von Qualitätssicherungsverfahren und Qualitätsindikatoren alle wichtigen Fragestellungen im Blick bleiben und berücksichtigt werden, es geht nicht um die Systematik an sich. Auf eine Diskussion, wie die Eignungskriterien im Unterschied zu anderen Vorschlägen in der Literatur kategorisiert und sortiert werden, wird daher weiterhin verzichtet.

Konkrete Nutzung der Eignungskriterien

In mehreren Stellungnahmen wird beschrieben, dass die praktische Anwendung der Eignungskriterien nicht deutlich werde. Die Anmerkungen zielen dabei zum einen auf die Darstellung des Ablaufs bzw. der Zeitpunkte der Eignungsprüfung und zum anderen auf den Detailgrad der Darstellung wie z. B. die Angabe von Grenzwerten, Informationsgrundlagen oder einen Algorithmus für die Ermittlung einer Gesamtbeurteilung der Indikatoreignung (DGIM, S. 4; KBV, S. 21; GKV-SV, S. 13 f. und 19; DKG, S. 25). Eine der Stellungnahmen fordert mit Bezug auf die Ausführungen zur Erfüllung von Eignungskriterien, dass es für die Eignung von Qualitätsindikatoren positive Hinweise im Sinne einer „offensichtlichen Eignung“ geben müsse (DKG, S. 21). Offen bleibe zudem die Frage, wie ermittelt werde, welche Indikatoren geeigneter als andere sind (BÄK, S. 13), und wie Transparenz über die Anwendung der Eignungskriterien geschaffen werde (BPtK, S. 11), insbesondere vor dem Hintergrund, dass in der Regel keine quantitative Bewertung der Kriterien erfolge (GKV-SV, S. 19).

IQTIG: Ziel des Kapitels „Eignungskriterien“ ist die Darstellung der grundsätzlichen Überlegungen und Anforderungen bei der Entwicklung und der Beurteilung der Eignung von Qualitätsmessungen. Der konkrete Ablauf der Verfahrensentwicklung und die Einbettung der Beurteilung der Eignungskriterien in diesen Ablauf werden in den entsprechenden Kapiteln beschrieben.

Wie in den einleitenden Abschnitten von Kapitel 8 ausgeführt, können die Eignungskriterien und das Ausmaß, zu dem diese von den Indikatorenkandidaten für ein Qualitätssicherungsverfahren „erfüllt“ werden, immer nur als Informationsgrundlage dienen. Viele der Eignungskriterien sind qualitativer Natur und einer Quantifizierung nicht zugänglich. So ist beispielsweise das Ausmaß, zu dem Leistungserbringer Einfluss auf ihren Indikatorwert haben, nicht messbar, sondern es kann nur eine Einschätzung vorgenommen werden, ob ein Indikator so weit durch Leistungserbringer beeinflussbar ist, dass er sich zur Qualitätsmessung eignet. Auf der Basis möglichst reichhaltiger Informationen (z. B. Literatur, Expertengespräche) muss die Auswahl eines Indikators durch Zusammenschau der Informationen und Abwägung von Vor- und Nachteilen erfolgen und ist nicht sinnvoll durch einen Algorithmus zu treffen (Meyer 2004: 40). Für die einzelnen Eignungskriterien ist aus diesem Grund eine Angabe fester Grenzwerte, jenseits derer ein Indikator grundsätzlich nicht mehr in Betracht gezogen wird, nicht sinnvoll. Die Informationsgrundlage für die Prüfung der einzelnen Eignungskriterien und Kriterien zu deren Beurteilung wurden im Text der „Methodischen Grundlagen“ soweit möglich detaillierter dargestellt. In den Berichten zur Verfahrens- bzw. Indikatorenentwicklung werden die Beurteilungen der Indikatoren hinsichtlich der Eignungskriterien als Indikatorprofil dargestellt, z. B. in Form von Bewertungstabellen. Eine entsprechende Erläuterung wurde in den Text aufgenommen. Die Transparenz über das entwickelte Verfahren entsteht somit nicht durch die Quantifizierung kaum quantifizierbarer Kriterien, sondern durch die nachvollziehbare Darlegung der Gründe, weshalb das IQTIG beispielsweise einen Indikator für ausreichend durch die Leistungserbringer beeinflussbar hält oder, allgemein, weshalb die Informationen als Hinweise auf eine angemessene oder auf eine unzureichende Erfüllung der Eignungskriterien angesehen werden. Diesen Gründen kann dann in einem Stellungnahmeverfahren widersprochen werden und es können Gegenargumente angeführt werden. Eine Zahlengabe würde eine Präzision suggerieren, die dieser Einschätzung nicht zugrunde liegt, und wird hier daher als wenig hilfreich angesehen.

Bezug der Eignungskriterien auf Qualitätsaspekte und Indikatorensets

In einigen Stellungnahmen wird angemerkt, dass nicht klar werde, ob sich die im Kapitel „Eignung von Qualitätsindikatoren“ genannten Kriterien auch auf die Prüfung der Qualitätsaspekte, der Qualitätsmerkmale und des Sets der Qualitätsindikatoren eines Qualitätssicherungsverfahrens beziehen (GKV-SV, S. 14 und 19; ähnlich DKG S. 21).

IQTIG: Bei der Darstellung der Eignungskriterien in den „Methodischen Grundlagen“ wurden mit dem Ziel einer leichteren Lesbarkeit die Kriterien in Bezug auf Qualitätsindikatoren dargestellt. Da Qualitätsindikatoren letztendlich die Operationalisierung der Qualitätsaspekte und Qualitätsmerkmale darstellen, betrifft eine Beurteilung der Indikatoreignung auch diese zugrunde liegenden theoretischen Konzepte. Wenn es also nicht darum geht, bereits etablierte Qualitätsindikatoren ex post zu beurteilen, sondern der Prozess der Verfahrensentwicklung unter Zuhilfenahme der Eignungskriterien gesteuert werden soll, kommen diese Kriterien bei verschiedenen Schritten zur Anwendung, ggf. auch mehrfach. Dieser Bezug auf mehrere Ebenen im Rahmen der Verfahrensentwicklung wurde bei der Bearbeitung des Textes nun deutlicher herausgestellt.

Bezug zum Verfahrenszweck

Es wird von einer stellungnehmenden Organisation angemerkt, dass der Bezug der Eignungskriterien zu den Verfahrenszwecken der Qualitätssicherungsverfahren nicht konkret genug dargestellt werde (DKG, S. 22).

IQTIG: Wie an verschiedenen Stellen im Text angemerkt, geht das IQTIG davon aus, dass die Eignung von Qualitätsindikatoren in unterschiedlichen Kontexten auch unterschiedlich ausfallen kann. Allerdings trifft dies nicht auf alle Eignungskriterien gleichermaßen zu. So kann davon ausgegangen werden, dass z. B. die Bedeutung eines Qualitätsaspektes bzw. Qualitätsmerkmals für die Patientinnen und Patienten oder z. B. die Evidenzgrundlage für mittelbar patientenrelevante Qualitätsziele unabhängig vom Verfahrenszweck beurteilt werden können, während etwa die Brauchbarkeit der Indikatoregebnisse für verschiedene Adressatenkreise unterschiedlich sein kann. Je nach Regelungskontext könnten beispielsweise auch unterschiedliche Anforderungen an die statistische Klassifikationsgüte gestellt werden. Auch die (qualitative) Gewichtung der Eignungskriterien untereinander kann sich je nach Verfahrenszweck unterscheiden. Eine generelle Festlegung a priori, in welchem Maße ein Eignungskriterium bei welchem Verfahrenszweck zur Geltung kommt, ist daher nicht vorgesehen. Überdies sind die grundlegenden Konzepte für die über die Qualitätsförderung hinausgehenden Verfahrenszwecke (siehe dazu Kapitel 2 der „Methodischen Grundlagen“) noch in Entwicklung und daher noch nicht Bestandteil der Version 1.0.

Anwendung auf bestehende Qualitätsindikatoren

In den Stellungnahmen wird die Frage aufgeworfen, ob und wie die Eignung der in den bestehenden Qualitätssicherungsverfahren verwendeten Qualitätsindikatoren überprüft werde (GKV-SV, S. 19; DKG, S. 17).

IQTIG: Das in den „Methodischen Grundlagen“ vorgestellte Konzept der Eignungskriterien beschreibt die grundsätzlichen Anforderungen, die an Messungen und Bewertungen mittels Qualitätsindikatoren zu stellen sind. Bei der Entwicklung neuer Qualitätssicherungsverfahren werden diese Anforderungen

von Anfang an berücksichtigt, um möglichst gut geeignete Qualitätsindikatoren nutzen zu können. Für in Gebrauch befindliche Qualitätsindikatoren gelten aus methodischer Sicht die gleichen Anforderungen, eine Prüfung aller Indikatoren bedeutet aber einen hohen Aufwand, der vorerst nicht geleistet werden kann. Möglichkeiten, wie die Eignungskriterien mit vertretbarem Aufwand für diese Indikatoren angewendet werden bzw. mit welcher Priorität bestimmte Eignungskriterien bzw. bestimmte Indikatoren behandelt werden, sind derzeit in Entwicklung.

Einzelne Eignungskriterien

Bedeutung für die Patientinnen und Patienten; Potenzial zur Verbesserung

In verschiedenen Stellungnahmen werden diese beiden thematisch verwandten Eignungskriterien und ihr Bezug zueinander diskutiert. So wird gefragt, wie Patientenrelevanz operationalisiert werde (Stock, S. 3), und gefordert, diese systematisch zu definieren (DDG, S. 3) oder auf extern konsentiertere Versorgungsziele wie die Core Outcome Sets zu beziehen (AWMF, S. 5; QMG, S. 2). Einige Stellungnahmen fordern, Aspekte wie hohe Kosten und öffentliches Interesse nicht als Hinweis auf die Patientenrelevanz eines Indikators zu verstehen (DGIM, S. 4), während andere Stellungnahmen diese auf das Gesundheitssystem als Ganzes bezogenen Aspekte diesem Eignungskriterium zurechnen (BÄK, S. 14 f.) bzw. der Bedeutung eines Indikators (DKG, S. 21) zurechnen. In diesem Zusammenhang wird auch vorgeschlagen, das „Potenzial zur Verbesserung“ wie im QUALIFY-Instrument als Teilaspekt der Bedeutung des Indikators zu verstehen und das Kriterium wie in QUALIFY zu benennen (BÄK, S. 14).

IQTIG: Bei der Beurteilung, ob für ein bestimmtes Qualitätsmerkmal der Versorgung eine Qualitätsmessung sinnvoll ist, sollte aus Sicht des IQTIG zwischen zwei Facetten von „Bedeutung“ unterschieden werden:

- a) Handelt es sich um ein Merkmal, das für Patientinnen und Patienten persönlich in der Versorgung wichtig ist (Patientenrelevanz).
- b) Handelt es sich um ein Merkmal, bei dem durch dessen Messung und die sich aus der Messung ergebenden korrektiven Maßnahmen überhaupt positive Effekte auf die Gesundheitsversorgung erwartbar sind (Verbesserungspotenzial).

Dabei soll Patientenrelevanz jedoch nicht theoretisch und allgemein definiert werden, sondern für jedes Thema konkret und spezifisch ermittelt werden (ähnlich: Mühlhauser und Müller 2009). In den Entwicklungsprojekten des IQTIG erfolgt dies insbesondere mittels Fokusgruppen mit Patientinnen und Patienten und durch Einbindung von Patientenvertreterinnen und -vertretern in die beratenden Expertengruppen sowie ergänzend durch Hinweise aus Literatur und von Expertinnen und Experten. Das Konzept der Core Outcome Sets weist zwar mit der Entwicklung und Konsentierung von Patient-Reported Outcomes Parallelen zur Entwicklung von Qualitätsindikatoren auf, hat seinen Fokus aber auf der Etablierung vergleichbarer Berichtsendpunkte in Studien (Prinsen et al. 2016, Macefield et al. 2014). Die Auswahl von Messzielen und

Messinstrumenten erfolgt nicht unter der Perspektive der Qualitätssicherung und ist damit auf diese nicht ohne weiteres übertragbar; sie kann jedoch als Informationsquelle bei der Themenerschließung genutzt werden.

Für die Bedeutung einer Qualitätsmessung im Gesundheitswesen können im Rahmen der Indikatorenentwicklung nur Hinweise gewonnen werden. Eine angemessene Bewertung der Fragestellung, ob sich positive Effekte auf die Gesundheitsversorgung insgesamt ergeben, könnte allenfalls im Rahmen einer Evaluation des Qualitätssicherungsverfahrens vorgenommen werden. Dennoch sind bereits bei der Indikatorenentwicklung Hinweise zu suchen, ob zumindest ein Potenzial für Verbesserungen besteht. Diesem Zweck dient die Berücksichtigung des gleichnamigen Eignungskriteriums, das daher – anders als bei QUALIFY – unabhängig von der Patientenrelevanz beurteilt werden soll. Die Trennung der Kriterien soll bei der Indikatorenentwicklung den Blick dafür schärfen, dass es Qualitätsmerkmale geben kann, die zwar eine hohe Bedeutung für die betroffenen Patientinnen und Patienten haben, aber bei denen wegen durchgängig guter Versorgung eine Messung nicht erforderlich ist. Umgekehrt kann Verbesserungspotenzial bei Merkmalen bestehen, die aus Patientensicht als wenig bedeutsam erlebt oder eingestuft werden. Richtig ist die Anmerkung, dass die auf das Gesundheitssystem als Ganzes bezogenen Teilaspekte nicht zum Kriterium Patientenrelevanz gehören. Der Text wurde entsprechend korrigiert und überarbeitet, um die beabsichtigte Trennung klarer herauszustellen.

Übereinstimmung mit dem zugehörigen Qualitätsaspekt

In einer Stellungnahme wird dieses Eignungskriterium infrage gestellt und geltend gemacht, dass es für einen hochwertigen Indikator von untergeordneter Bedeutung sei, ob er in das theoretische Qualitätsmodell genau passe oder nicht. Hochwertige Indikatoren sollten nicht wegen mangelnder Passgenauigkeit verworfen werden (BÄK, S. 14).

IQTIG: Sofern die Stellungnahme mit „Passgenauigkeit“ die Zuordnung eines Indikators zu einem von mehreren infrage kommenden Qualitätsaspekten meint, ist ihr insoweit zuzustimmen, als es bei der Entwicklung der Qualitätsindikatoren weniger auf eine exakte Zuordnung, sondern vielmehr auf die Legitimation des Qualitätsziels eines Indikators ankommt. Der Begriff des „hochwertigen Indikators“ impliziert ja bereits, dass er einen relevanten Aspekt valide misst. Der Nutzen der Prüfung des Bezugs eines Qualitätsindikators zu den im Qualitätsmodell zusammengefassten Aspekten besteht darin, dass

- a) die Entwicklung von Indikatoren vermieden wird, die weniger bedeutende Aspekte (die daher nicht im Qualitätsmodell enthalten sind) abbilden,
- b) die Vollständigkeit oder Unvollständigkeit der Abbildung des Qualitätsmodells im Qualitätssicherungsverfahren beurteilt werden kann (dies entspricht der Inhaltsvalidität des Indikatorensets) und

- c) geprüft werden kann, ob mehrere Indikatoren das Gleiche abbilden und ob eine solche mehrfache Abbildung erforderlich ist (Inhaltsvalidität des Indikators versus Sparsamkeit bei der Indikatornutzung).

Der Text wurde dahingehend angepasst.

Zusammenhang mit patientenrelevantem Ziel

Es wird in einer Stellungnahme angemerkt, dieses Kriterium wie in QUALIFY mit dem Begriff „Indikatorevidenz“ zu bezeichnen, und die Notwendigkeit hinterfragt, zwischen unmittelbaren und mittelbaren Indikatorzielen zu unterscheiden (BÄK, S. 15). Außerdem wird in einer anderen Stellungnahme die Frage gestellt, wie hoch die Anforderungen seien, die an den Nachweis des Zusammenhangs gestellt werden (GKV-SV, S. 20).

IQTIG: Prinzipiell könnte die Bezeichnung als „Indikatorevidenz“ für dieses Eignungskriterium auch gewählt werden, da die Evidenz dafür geprüft wird, ob ein Struktur-, Prozess- oder (Surrogat-)Ergebnis-Parameter mit einem unmittelbar patientenrelevanten Ziel stringent zusammenhängt. Für Ergebnisparameter, deren gewähltes Outcome für die Patientinnen und Patienten von unmittelbarer Relevanz ist, liegt dieser Nachweis per se bereits vor. Beispielhaft seien das Erreichen des Therapieziels (z. B. Gehfähigkeit nach Gelenkersatz) und das Fehlen von Komplikationen genannt. Hier bedarf es keiner weiteren Argumentation zur Patientenrelevanz.

QUALIFY fordert für die Indikatorevidenz bei patientenrelevanten Ergebnisindikatoren jedoch die Beeinflussbarkeit durch die Leistungserbringer und die Korrektheit des Messzeitpunkts (Reiter et al. 2007). Dies sind aber Kriterien, die das IQTIG nicht unter Patientenrelevanz subsumiert. Die Beeinflussbarkeit ist ein Kriterium für sich, weil es dabei nicht direkt um die Patientenrelevanz eines Outcome, sondern um die Verfügbarkeit wirksamer Interventionsmöglichkeiten geht. Nun können künftig im Rahmen der Entwicklung neuer medizinischer Verfahren bestimmte patientenrelevante Aspekte der Versorgung beeinflussbar werden, bei denen das heute nur in geringerem Maße oder gar nicht möglich ist. Ein entsprechender Ergebnisindikator ist immer patientenrelevant, aber die Möglichkeiten, seine Ziele zu erreichen, verbessern sich ggf. mit der Zeit, sodass er selbst an Bedeutung gewinnt. Das sollte aber bei Ergebnisindikatoren nicht unter Patientenrelevanz subsumiert werden, sondern in einem getrennten Kriterium der Beeinflussbarkeit. Die Korrektheit des Messzeitpunkts ist andererseits ein Aspekt der Validität des Messinstrumentariums und nicht primär ein Aspekt der Patientenrelevanz. Die im QUALIFY-Konzept für Ergebnisindikatoren zusätzlich geforderten Kriterien werden also in den „Methodischen Grundlagen“ an anderer Stelle berücksichtigt. Daher konnte der Begriff der Indikatorevidenz nicht übernommen werden.

Ein fester Grenzwert, ab welcher Qualität und Konsistenz der recherchierten Evidenz der Zusammenhang mit einem unmittelbar patientenrelevanten Ziel als ausreichend belegt anzusehen ist, wird nicht definiert (s. a. Abschnitt „Konkrete Nutzung der Eignungskriterien“). Die Informationsgrundlage aller vom

IQTIG beurteilten Indikatoren, in diesem Fall die Bewertung der Evidenzbasis, wird aber in den jeweiligen Entwicklungsberichten transparent gemacht.

Beeinflussbarkeit durch die Leistungserbringer

Zur Beeinflussbarkeit des Qualitätsmerkmals und damit des Indikatorwertes durch den Leistungserbringer wird angemerkt, dass diese auch Faktoren unterliege, die nicht unter der Kontrolle der Leistungserbringer stehen (KZBV, S. 2; DKG, S. 22; An-Institut, S. 1; DGK, S. 4), u. a. auch von den zur Verfügung stehenden Ressourcen. Es wurden Klarstellungen zur Abbildung in diesem Abschnitt gefordert.

IQTIG: Die Benennung der Beeinflussbarkeit als eines der zentralen Eignungskriterien für eine Qualitätsmessung soll genau den in den Stellungnahmen vorgebrachten Punkt deutlich machen, nämlich, dass diese Beeinflussbarkeit bei verschiedenen Qualitätsmerkmalen in unterschiedlichem Maße gegeben sein kann. In den „Methodischen Grundlagen“ wird dabei unterschieden zwischen der grundsätzlichen Beeinflussbarkeit einerseits und andererseits der Frage, welche konkreten Anforderungen unter den bestehenden Rahmenbedingungen – z. B. im Gesundheitssystem verfügbare Ressourcen – gestellt werden können, d. h., wie die „Erreichbarkeit des Referenzbereichs“ (siehe gleichnamiges Kriterium) beurteilt wird. Die Abbildung im Text visualisiert das mögliche Spektrum, ohne die genannten Beispiele mit konkreten Zahlen zu belegen. Um Missverständnisse zu vermeiden, wurde die Erläuterung zur Abbildung überarbeitet.

Datenvalidität und Schritte der Datenerhebung

In einer Stellungnahme wird angemerkt, dass die Datenvalidität ein Teilaspekt der Objektivität der Messung sei und entsprechend eingeordnet werden müsste. Zudem sei der Bezug der Eignungskriterien auf die Schritte bei der Datenerhebung nicht klar genug dargestellt (DKG, S. 22).

IQTIG: Es ist richtig, dass die Objektivität der Messung insgesamt auch davon abhängt, inwieweit einzelne am Messprozess beteiligte Personen einen verzerrenden Einfluss auf die verwendeten Daten haben. In der externen Qualitätssicherung besteht – wie in den „Methodischen Grundlagen“ und in der Stellungnahme beschrieben – allerdings die Besonderheit, dass bei Nutzung von QS-Dokumentationsdaten und von Sozialdaten bei den Krankenkassen die Messung formal sehr deutlich in zwei Schritte unterteilt ist: zum einen die primäre Datenerhebung bei den Patienten und Patientinnen im Rahmen der Behandlung und zum anderen die Erfassung und Weiterleitung dieser Daten für die externe Qualitätssicherung, entweder direkt (QS-Dokumentationsdaten) oder indirekt (Abrechnungsdaten der Krankenkassen). Das IQTIG hält es daher für sinnvoll, bei der Prüfung der Messeigenschaften von Qualitätsindikatoren zwischen diesen Ebenen zu unterscheiden und die Datenvalidität als separates Kriterium zu benennen, da dieser Aspekt der Objektivität in der Verantwortung der Dokumentierenden liegt und erst anhand einer empirischen Überprüfung beurteilt werden kann. Hingegen kann der Einfluss von

korrekter Operationalisierung und Spezifikation des Indikators und des Erhebungsverfahrens auf die Objektivität der Messung bereits bei der Indikatorenentwicklung mittels sachlogischer Überlegungen zumindest vorläufig eingeschätzt werden. Der Text zu den Eignungskriterien wurde nochmals überarbeitet, um die Ebenen der Datenerhebung und den Bezug der Eignungskriterien deutlicher zu machen.

Brauchbarkeit

Es wird von einer stellungnehmenden Organisation angemerkt, dass unter dem Kriterium Brauchbarkeit sehr unterschiedliche Aspekte zusammengefasst wurden, und angeregt, diese als getrennte Kriterien aufzuführen. Zudem beziehe sich der unter Brauchbarkeit subsummierte Aspekt der Ausgewogenheit nicht auf einzelne Indikatoren, sondern auf ein Indikatorenset (BÄK, S. 16).

IQTIG: Der gemeinsame Nenner der vom IQTIG genannten Teilaspekte der Brauchbarkeit ist, wie gut sich die ermittelten Indikatorwerte nutzen lassen, wenn man eine angemessene Zielsetzung, Messung und Bewertung unterstellt. Dies entspricht in der internationalen Literatur weitgehend dem Begriff der *usability* (im Gegensatz zur *feasibility*: Praktikabilität des Verfahrens) (NQF 2015, McGlynn 2003). Die Kritik in der Stellungnahme wird so verstanden, dass befürchtet wird, durch die Zusammenfassung mehrerer Teilaspekte zu einem übergeordneten Kriterium würden diese Aspekte weniger sorgfältig beurteilt als bei getrennter Darstellung. Ein unterschiedliches Gewicht für Kriterien, die als eigene Eignungskriterien aufgeführt werden, und Kriterien, die als Teilaspekte aufgeführt werden, ist jedoch im IQTIG-Konzept der Eignungsprüfung nicht vorgesehen. Die Entscheidung für oder gegen Aufnahme eines Indikators in das empfohlene Indikatorenset erfolgt nicht auf Basis eines Algorithmus, sondern durch qualitative Abwägung von Vor- und Nachteilen (siehe Ausführungen zur Nutzung der Eignungskriterien).

Für den Teilaspekt der Ausgewogenheit könnte auch eine alternative Einsortierung in die Systematik der Eignungskriterien vorgenommen werden. Wie richtig angemerkt, bezieht sich die Ausgewogenheit auf ein Set von Indikatoren und geht damit zwei Fragen nach. Erstens, ob das Konzept, das das Indikatorenset abbilden soll, repräsentativ erfasst ist. Dieser Aspekt beschreibt die Inhaltsvalidität des Indikatorensets bei der Darstellung des Sachverhalts (Hartig et al. 2012: 148). Dabei muss je Teilaspekt (Dimension) des Konzepts mindestens ein Indikator vorhanden sein (Meyer 2004: 46). Zweitens, ob keine überflüssigen Indikatoren im Set vorhanden sind, weil derselbe Aspekt mehrfach abgebildet wird. Es sollten also für jeden Teilaspekt (Dimension) des abzubildenden Konzepts nur so viele Indikatoren wie nötig verwendet werden (Meyer 2004: 46). Nicht in jedem Qualitätssicherungsverfahren ist die vollständige Abbildung eines Konzepts das Ziel: So wurde z. B. für den ersten Schritt des Verfahrens zu planungsrelevanten Indikatoren auf solche Qualitätsindikatoren zurückgegriffen, die jeder für sich Hinweise auf eine Patientengefährdung geben, ohne dabei den Anspruch zu haben, die Versorgungsqualität eines

Fachbereichs als Ganzes abzubilden. Aufgrund dieser Abhängigkeit des Kriteriums Ausgewogenheit vom Verfahrenszweck wurde dieses Kriterium dem Eignungskriterium Brauchbarkeit zugeordnet.

Zuschreibbarkeit

Es wird in einer Stellungnahme angemerkt, dass die Frage der Zuschreibbarkeit zukünftig z. B. durch Langzeit-Follow-up-Verfahren eine zunehmende Bedeutung bekomme und dass nicht deutlich werde, wie die Zuschreibbarkeit beurteilt werden und wie bei unklarer Zuschreibbarkeit zu verfahren sei (GKV-SV, S. 20). Zudem, so eine andere Stellungnahme, könne für manche Verfahren eine Zuschreibbarkeit von untergeordneter Bedeutung sein, etwa wenn Indikatoren für das interne Qualitätsmanagement oder als Area-Indikatoren genutzt würden (KBV, S. 6).

IQTIG: Grundlage für die Beurteilung der Zuschreibbarkeit sind die Informationen zum Versorgungsgeschehen aus Literatur, Expertengesprächen und Gruppendiskussionen sowie ggf. den daraus abgeleiteten Versorgungspfaden. Wenn mehrere Leistungserbringer am Zustandekommen eines Indikatorwerts beteiligt sind, kann der jeweilige Anteil nicht quantifiziert, sondern nur orientierend eingeschätzt werden (s. a. die Ausführungen zur konkreten Nutzung der Eignungskriterien). Indikatoren, die keine eindeutige Zuschreibbarkeit zu einem Leistungserbringer aufweisen, sind für die weitere Verwendung wie richtig angemerkt erst geeignet, wenn für die Nutzung der Indikatorergebnisse ergänzende Konzepte vorliegen. Diese sind noch zu entwickeln und daher in den „Methodischen Grundlagen“ nicht enthalten.

Validität

In einer Reihe von Stellungnahmen wird thematisiert, dass der Stellenwert der Validität von Qualitätsindikatoren in den „Methodischen Grundlagen“ nicht angemessen behandelt werde. So wird gefordert klarzustellen, dass der Validitätsbegriff auf Qualitätsindikatoren anzuwenden sei (Schrappe, S. 20), Aussagen zu weiteren Validitätskonzepten über die Inhaltsvalidität hinaus zu treffen (DNVF, S. 19; WINEG, S. 6) und deutlich zu machen, dass die Qualitätsmessung latente Konstrukte adressiere (DKG, S. 22; GKV-SV, S. 19). In den Stellungnahmen finden sich dabei verschiedene Verwendungen des Validitätsbegriffs. Des Weiteren wird in einer Stellungnahme die Ansicht vertreten, dass die Anforderungen an die Validität von QS-Aussagen stets und unabhängig vom Verfahrenszweck hoch sein sollten (PatV, S. 8 f.).

IQTIG: Es ist richtig, dass es sich bei Qualität um ein theoretisches Konzept handelt, nämlich den Grad, in dem Anforderungen erfüllt werden. Um Aussagen zu einem solchen theoretischen Konzept treffen zu können, ist eine Operationalisierung in Form von Indikatoren erforderlich, die den theoretischen Sachverhalt durch beobachtbare Sachverhalte messbar machen (Meyer 2004: 28 und 51, Döring und Bortz 2016: 223 und 228, Schnell et al. 2013: 117 und 121). Für die Qualität der Gesundheitsversorgung geschieht dies z. B. mittels Qualitätsindikatoren. Nach der gängigen wissenschaftlichen Definition von

Messung als Zuordnung von Zahlen zu Objekten (Schnell et al. 2013: 128) handelt es sich bei der quantitativen Beschreibung der Erfüllung von Anforderungen anhand von Qualitätsindikatoren damit um eine Qualitätsmessung. Die in einer Stellungnahme geäußerte Auffassung, Indikatoren unterschieden sich grundsätzlich von anderen Messinstrumenten (Schrappe, S. 12 ff.), kann aus der sozial- und evaluationswissenschaftlichen Literatur (s. o.) nicht nachvollzogen werden und widerspricht auch der in der Literatur zur Qualitätssicherung in der Gesundheitsversorgung üblichen Bezeichnung als „Maße“ oder „measures“ (Sens et al. 2007, NQF 2015). Richtig ist, dass die Messung von Qualitätsaspekten insoweit als „indirekt“ bezeichnet werden kann, als es sich bei Qualitätsaspekten um theoretische Konzepte handelt. Im Kontext der externen Qualitätssicherung erhalten die Qualitätsaspekte durch die Qualitätsindikatoren operationale Definitionen, die über den Beschluss des G-BA normativ festgeschrieben werden und dann für die Anwendung im Qualitätssicherungsverfahren maßgeblich sind.

Validität bezeichnet das Ausmaß, in dem die Angemessenheit von Interpretationen und von Maßnahmen auf Basis von Messergebnissen empirisch und theoretisch gestützt ist (vgl. Hartig et al. 2012: 144). Interpretation des Erhebungsergebnisses kann dabei Verschiedenes bedeuten, wie z. B. die Bewertung des Ergebnisses, seine Verallgemeinerung oder das Fällen von Entscheidungen (Kane 2001). In der externen Qualitätssicherung kann sich Validität dementsprechend auf verschiedene Aspekte und auf verschiedene Ebenen beziehen, wie z. B. die folgenden Fragestellungen:

- Bildet der Indikatorwert das Qualitätsmerkmal ab, das der Indikator operationalisieren soll?
- Stimmt diese Klassifikation mit dem Ergebnis anderer Erhebungsverfahren (Strukturierter Dialog) überein?
- Lässt die Bewertung eines Indikatorwertes anhand des Bewertungskonzeptes für diesen Indikator (Referenzbereich) eine angemessene Klassifikation zu (in Bezug auf anzuschließende Handlungen)?
- Werden die Qualitätsaspekte durch die ausgewählten Indikatoren vollständig adressiert?
- Umfasst das Qualitätsmodell alle Aspekte, die für das dem Verfahren zugrunde liegende theoretische Konzept wichtig sind?

Wegen dieser vielen möglichen Facetten von Validität in der externen Qualitätssicherung würde es ein hohes Risiko für Missverständnisse bergen, von der Validität eines Qualitätsindikators zu sprechen. Dies wurde daher in den „Methodischen Grundlagen“ bewusst vermieden. Die zugrunde liegenden Fragestellungen sind aber im Konzept der Eignungsprüfung aufgegriffen. Um dies nachvollziehbarer zu machen, wurde der Bezug zum Validitätsbegriff an einigen Stellen im Text deutlicher dargestellt.

Reliabilität

In einer Stellungnahme wird angemerkt, die Verwendung des Begriffs Reliabilität sei nicht zutreffend, und Reliabilität stattdessen als Fähigkeit, Qualitätsprobleme richtig zu erkennen, definiert (Schrappe, S. 17).

IQTIG: Die in der Stellungnahme vorgeschlagene Definition von Reliabilität entspricht nicht der Verwendung in der wissenschaftlichen Literatur. Die Reliabilität eines Messinstruments wird übereinstimmend definiert als Genauigkeit einer Messung, verstanden als Anteil der Varianz in den Messwerten, der durch tatsächliche Unterschiede verursacht ist und nicht auf zufällige Messfehler zurückgeht (Schermelleh-Engel und Werner 2012: 120, Meyer 2004: 28, O'Brien 1990, z. B. Bühner 2011: 60, Raykov und Marcoulides 2011: 138). Eine abweichende Definition von Reliabilität für Messinstrumente, die als Indikatoren verwendet werden, ist in den Sozial- und Evaluationswissenschaften nicht etabliert. Auch im Bereich der medizinischen Qualitätssicherung wird international Reliabilität in diesem Sinne verstanden (Adams et al. 2010, Campbell et al. 2003, Reiter et al. 2007, Mainz et al. 2003, NQF 2015).

Mathematisch wird Reliabilität als ein Varianzverhältnis definiert, das ausdrückt, wie viel der Varianz in den beobachteten Messwerten auf Varianz in den wahren Werten zurückgeht (Raykov und Marcoulides 2011: 138). Diese Definition beinhaltet den für das Verständnis von Reliabilität unverzichtbaren Verweis auf einen wahren Wert. Verfahren zur Bestimmung von Reliabilitätskoeffizienten (z. B. Test-Retest) basieren auf der Annahme, dass der Messung ein wahrer, nicht beobachtbarer Wert zugrunde liegt. Der wahre Wert ist definiert als der Erwartungswert unendlich vieler Messungen des gleichen Objekts unter den gleichen Umständen (Raykov und Marcoulides 2011: 118). Ist diese Trennung von wahren und beobachtetem Wert für eine Messung nicht plausibel, ist die Bestimmung der Reliabilität nicht sinnvoll. Dies betrifft v. a. die Messung direkt beobachtbarer Merkmale einzelner Patientinnen und Patienten, wie z. B., ob ein Patient verstorben ist. Nicht zu verwechseln ist der wahre Wert mit Werten aus anderen Datenquellen, wie beispielsweise der Patientenakte. Stimmt etwa der für die Qualitätssicherung dokumentierte Wert nicht mit dem Wert aus einer Patientenakte überein, ist das zwar unerwünscht, aber für die Messung auf Ebene einzelner Patientinnen und Patienten kein Reliabilitätsproblem im eigentlichen Sinn, sondern ein Dokumentationsproblem, das durch das Eignungskriterium Datenvalidität adressiert wird.

Die Annahme eines zugrunde liegenden wahren Werts bei der Messung der Qualität medizinischer Versorgung kann jedoch auf anderer Ebene hilfreich sein: Qualitätsindikatoren fassen Messwerte einer definierten Gruppe behandelte Patientinnen und Patienten zusammen, um die Qualität der Versorgung einzelner Leistungserbringer zu beschreiben. Es liegt eine hierarchische Struktur mit zwei Ebenen vor, wobei es sich auf der unteren Ebene um die Beschreibung einzelner Patientinnen und Patienten und auf der oberen Ebene um die Beschreibung der Leistungserbringer handelt. Für diese Messung auf Ebene der Leistungserbringer kann in Analogie zu der obigen Definition der wahre

Wert eines Leistungserbringers definiert werden als der Erwartungswert der Messungen, wenn unendlich viele Patientinnen und Patienten durch diesen Leistungserbringer behandelt worden wären und somit zufallsbedingte Einflüsse ausgeglichen wären. Möglichkeiten, wie die Reliabilität auf Leistungserbringerebene bestimmt werden kann, sind Gegenstand zukünftiger Entwicklungen.

Weitere Themen

Nicht als Eignungskriterien genannte Begriffe

In einigen Stellungnahmen werden weitere Anforderungen an Qualitätsindikatoren als Eignungskriterien benannt: „Robustheit gegen Manipulation“ und „Änderungssensitivität“ (DKG, S. 30), „Verständlichkeit für Patienten“ (BÄK, S. 13).

IQTIG: Die Benennung von Änderungssensitivität als eigenständige Anforderung an Messinstrumente ist umstritten. Ein Messinstrument zeigt in dem Maße Änderungen im gemessenen Merkmal an, in dem es zutreffend und präzise dieses Merkmal misst. Änderungssensitivität folgt also aus Validität und Reliabilität des Messinstruments und umgekehrt kann sie als Hinweis auf eine valide Messung betrachtet werden (Hays und Hadorn 1992). Für den Vergleich eines Indikatorwerts mit einem Referenzbereich spielt die Änderungssensitivität keine Rolle.

Robustheit gegen Manipulation zielt auf die Qualität der Datengrundlage für die Qualitätsindikatoren. Dieser Aspekt ist unter dem Eignungskriterium Datenvalidität berücksichtigt, soweit es um die Übereinstimmung zwischen Spezifikation, Primärdokumentation (z. B. in der Patientenakte) und den für die Qualitätssicherung erfassten Daten geht. Darüber hinaus wird Manipulationssicherheit berücksichtigt, wenn die Korrektheit der Berechnungsspezifikation eines Indikators geprüft wird.

Verständlichkeit eines Indikators für Patientinnen und Patienten und Öffentlichkeit ist grundsätzlich anzustreben. Verständlichkeit ist aber keine Eigenschaft eines Indikators, sondern eine Vermittlungsaufgabe in Bezug auf die Indikatorinhalte und die daraus ableitbaren Aussagen. Dies darf nicht verwechselt werden, wenn zusätzlich zu den Ergebnissen der Qualitätsindikatoren den Patientinnen und Patienten und der Öffentlichkeit weitere Informationen zur Verfügung gestellt werden sollen, die per se keine Qualitätsindikatoren sind und dennoch dazu geeignet sind, Wahlentscheidungen von Patientinnen und Patienten zu unterstützen. Hier wird es künftig eine Entwicklung von Methoden zur Bereitstellung von zusätzlichen Versorgungsinformationen geben, die noch nicht Inhalt der derzeitigen ersten Version der „Methodischen Grundlagen“ sind.

Treffsicherheit, Sensitivität, Spezifität

In einer Stellungnahme wird gefordert, Maße der Klassifikationsgenauigkeit wie Sensitivität, Spezifität sowie positiv und negativ prädiktiven Wert in den „Methodischen Grundlagen“ zu diskutieren (DKG, S. 21). Dabei bezieht sich die Stellungnahme auf Unterschiede zwischen „realen Qualitätsproblemen“ und „Messungenauigkeiten“, z. B. durch „abweichendes Patientengut“.

IQTIG: Die Stellungnahme bezieht die genannten Maße auf ein Außenkriterium, mit dem die Indikatorergebnisse verglichen werden müssten, um eine Aussage zur Klassifikationsgenauigkeit treffen zu können. Damit spricht dieser Aspekt die Kriteriumsvalidität des Indikators an (siehe Ausführungen zum Begriff Validität). Allerdings ist ein solches Außenkriterium nicht verfügbar, da ein Goldstandard zur Beschreibung von Qualität nicht vorliegt (vgl. die Ausführungen zu QUALIFY, das Sensitivität und Spezifität im gleichen Sinne wie die Stellungnahme definiert: Reiter et al. 2007). Das Konzept des IQTIG zur Eignungsbeurteilung sieht daher eine Globalbeurteilung der Angemessenheit der Klassifikation nicht vor, sondern unterscheidet stattdessen zwischen den verschiedenen Gesichtspunkten, die Einfluss auf die Klassifikation nehmen, anhand der Eignungskriterien. Dementsprechend wird das Eignungskriterium der Klassifikationsgüte in den „Methodischen Grundlagen“ nur in einem eng definierten Sinn verwendet und bezieht sich auf die Klassifikation des Indikatorwerts im Vergleich zu einem Referenzwert, wenn stochastische Effekte auf das Messergebnis zu berücksichtigen sind (vgl. Zuordnungsregeln im Kapitel „Bewertungskonzepte und Referenzbereiche“ in den „Methodischen Grundlagen“). Dies entspricht dem Konzept der Teststärkeanalyse (vgl. Döring und Bortz 2016: 807 ff.).

3.4 Veröffentlichung von Indikatorergebnissen

Verständlichkeit von Patienteninformationen

In mehreren Stellungnahmen wird geltend gemacht, dass die Verständlichkeit veröffentlichter Indikatorergebnisse essenziell sei (PatV-Kuratorium, S. 4) und geprüft werden solle (BÄK, S. 13).

IQTIG: Gemeint ist, dass zwar die Verständlichkeit aller veröffentlichten Qualitätsinformationen möglichst hoch sein sollte und eine entsprechende Darstellung angestrebt werden sollte, dass aber eine Einschätzung der Verständlichkeit nicht zur Bedingung gemacht werden soll für die Empfehlung, Indikatorergebnisse zu veröffentlichen. Diese offensichtlich missverständliche Formulierung wurde überarbeitet. Auf die erforderliche Trennung von valider Messung und verständlicher Vermittlung von indikatorbasierten Qualitätsergebnissen wurde bereits hingewiesen.

Abgrenzung zwischen Transparenz und Verfahrenszweck Information

In mehreren Stellungnahmen wird um Erläuterung gebeten, wie sich die im Kapitel „Veröffentlichung von Indikatorergebnissen“ thematisierte Transparenz von Informationen für Patientinnen und Patienten abgrenze (KBV, S. 22; GKV-SV,

S. 19 ff.; BÄK, S. 16; DKG, S. 11), mit welchen Darstellungsformen und Berichtsformaten Informationen für verschiedene Zielgruppen bereitgestellt werden sollen (GKV-SV, S. 21; KBV, S. 17) und auf welche Weise allgemeinverständliche Übersichten über die Qualität der Versorgung gemäß § 137a SGB V Abs. 3 Nr. 5 veröffentlicht werden sollen (PatV, S. 9 f.; BÄK, S. 16).

IQTIG: In Abschnitt 2.4 der „Methodischen Grundlagen“ wurden Erläuterungen zur Abgrenzung von Transparenz und der Bereitstellung gezielter Patienteninformationen eingefügt. Ein Konzept, wie und auf Grundlage welcher Daten gezielte, d. h. bedarfs- und nutzerorientierte, Informationen über die Qualität der Einrichtungen dargestellt werden können, muss zukünftig entwickelt werden. Es ist daher noch nicht Bestandteil der „Methodischen Grundlagen“.

Wirkungen von Public Reporting

Einige Stellungnahmen thematisieren die möglichen Wirkungen der Veröffentlichung einrichtungsbezogener Qualitätsinformationen (Public Reporting) und werfen die Frage auf, ob die erwünschten Wirkungen die unerwünschten überwiegen (Steller, S. 9 ff. und 17; DKG, S. 23 f.; DGK, S. 4 f.). Dies wird vom IQTIG als Forderung verstanden, dass die leistungserbringerbezogene Veröffentlichung von Indikatorwerten von einer Abschätzung von Nutzen und Risiken dieser Veröffentlichung abhängig gemacht werden sollte.

IQTIG: Eine vertiefte Literaturanalyse und Diskussion über erwünschte und unerwünschte Wirkungen von Public Reporting wird nicht angestrebt, da das IQTIG im Grundsatz eine Veröffentlichung aller Indikatorergebnisse empfiehlt. Diese Transparenz wird als ethisches und soziales Erfordernis verstanden. Daher rührt auch die gesetzlich festgelegte Pflicht zur Transparenz. Die uneinheitliche Beurteilung der Wirkungen von Public Reporting in der Literatur – soweit eine zusammenfassende Beurteilung angesichts der Unterschiedlichkeit in Art und Inhalt der bisherigen Veröffentlichungen überhaupt sinnvoll ist – wird vom IQTIG weder als Nutzenbeleg noch als Widerlegung eines Nutzens interpretiert. Der Verweis in den „Methodischen Grundlagen“ auf wichtige systematische Übersichtsarbeiten zu diesem Thema dient vielmehr dazu, deutlich zu machen, dass sich Argumente gegen eine Veröffentlichung nicht schlüssig aus der Literatur ableiten lassen.

Veröffentlichung weiterer Kennzahlen

In einer Stellungnahme wird gefragt, wie hinsichtlich der Veröffentlichung mit Kennzahlen umgegangen werden soll, die für die Interpretation von Indikatorergebnissen benötigt werden (GKV-SV, S. 20).

IQTIG: Das Kapitel 9 der „Methodischen Grundlagen“ erläutert, welche Bedingungen aus Sicht des IQTIG an die Empfehlung zur Veröffentlichung von Indikatorwerten gestellt werden. Es ist richtig, dass die ergänzende Darstellung weiterer Kennzahlen sinnvoll sein kann, um die Interpretation der Indikator-

werte zu erleichtern, beispielsweise durch Angabe stratenzspezifischer Ergebnisse oder Deskription der zugrunde liegenden Fälle. Ein entsprechendes Konzept wurde in der Zwischenzeit entwickelt und soll in die nächste Version der „Methodischen Grundlagen“ integriert werden.

Prüfung auf Veröffentlichungsfähigkeit

Einige Stellungnahmen merken an, dass die Veröffentlichung von Indikatoregebnissen mit Leistungserbringerbezug nur nach Prüfung der Veröffentlichungsfähigkeit der Indikatoren erfolgen sollte. Für diese Prüfung werden die Eignungskriterien der Indikatoren nach Kapitel 8 der „Methodischen Grundlagen“ (KZBV, S. 22) vorgeschlagen oder eine Weiterentwicklung der vom AQUA-Institut entwickelten Kriterien (KBV, S. 22). Ergebnisse von Indikatoren, die Schwächen in einzelnen Eignungskriterien aufweisen, sollten nicht veröffentlicht werden (KZBV, S. 22).

IQTIG: Dieser Auffassung schließt sich das IQTIG nicht an. Wie in den „Methodischen Grundlagen“ erläutert, hält das IQTIG das Transparenzgebot über die Ergebnisse der externen Qualitätssicherung für maßgeblich im Hinblick auf seine Empfehlungen zur Veröffentlichung von Indikatorwerten. Ergebnisse von Qualitätsindikatoren, deren Eignung gut genug für eine Nutzung in Qualitätssicherungsverfahren ist, sollen auch veröffentlicht werden.

Risikoadjustierung

Eine Stellungnahme führt an, dass der gesetzliche Auftrag zur Veröffentlichung vergleichender risikoadjustierter Übersichten impliziere, dass nur adjustierte Ergebnisse veröffentlicht werden dürften (KBV, S. 22). Eine weitere Stellungnahme fordert, dass Ergebnisindikatoren mit Ausnahme von Sentinel-Event-Indikatoren für eine Veröffentlichung risikoadjustiert sein müssten (BÄK, S. 17).

IQTIG: Das IQTIG schließt sich der Auffassung an, dass für Qualitätsindikatoren grundsätzlich zu prüfen ist, ob und für welche Faktoren eine Risikoadjustierung erforderlich ist. Dies ist besonders bei Ergebnisindikatoren regelmäßig der Fall. Allerdings wird dabei davon ausgegangen, dass Qualitätsindikatoren, deren Risikoadjustierung bereits als angemessen beurteilt wurde, um auf Qualitätsdefizite und -potenziale hinzuweisen, keiner erneuten Prüfung der Angemessenheit der Risikoadjustierung vor einer Veröffentlichung bedürfen. Der Interpretation des Gesetzestextes von § 137a SGB V Abs. 3 Nr. 5 dahingehend, dass eine Risikoadjustierung nicht nur für die noch zu entwickelnden Übersichten über die Qualität, sondern generell für alle Indikatorergebnisse (z. B. auch alle Prozessindikatoren) verpflichtend ist, wird nicht zugestimmt. Dies würde auch nicht berücksichtigen, dass es gute Prozessindikatoren gibt, die sehr eindeutig beschreiben, wann welche Prozessschritte durchzuführen sind, und die auch Ausnahmefälle ausreichend berücksichtigen, ohne dass ein Risikoadjustierungsmodell erstellt werden muss.

3.5 Bewertungskonzepte und Referenzbereiche

Qualitätsdarstellung nicht nur zum Vergleich der Leistungserbringer untereinander

In mehreren Stellungnahmen wird kritisiert, dass Bewertungskonzepte ohne feste Referenzbereiche nur einen Vergleich der Einrichtungen untereinander erlaubten und damit keine Aussage zum Qualitätsniveau in den einzelnen Einrichtungen trafen (An-Institut, S. 1; GKV-SV, S. 12 und 21; PatV, S. 18). Dadurch werde hohe oder niedrige Qualität nicht gewürdigt, wenn sie beispielsweise bei allen Einrichtungen hoch bzw. niedrig sei. In ähnlichem Sinne merkt eine Stellungnahme an, dass im System liegende Qualitätsdefizite durch Querschnittsvergleiche nicht identifiziert werden könnten, sondern dass longitudinale Vergleiche oder Vergleiche mit einem definierten Qualitätsstandard durchgeführt werden müssten (PatV-Kuratorium, S. 1).

IQTIG: Das IQTIG sieht ebenfalls die Schwierigkeiten, die sich durch Nutzung relativer Vergleichswerte wie z. B. Perzentilen als Referenzbereichsgrenzen ergeben. Aus diesem Grunde wird bereits in der Stellungnahmeversion der „Methodischen Grundlagen“ zwischen verteilungsunabhängigen und verteilungsabhängigen Referenzbereichen unterschieden und wird die Möglichkeit des Vergleichs mit Vorwerten erwähnt. Es sei insbesondere auf die Erläuterungen im Abschnitt 10.3.2 (Spezifikation des Referenzbereichs) verwiesen.

Qualitätsdarstellung bei kleinen Fallzahlen oder Ereignisraten

Eine Stellungnahme thematisiert die Problematik niedriger Fall- oder Ereigniszahlen bei der Qualitätsmessung mittels Indikatoren. Niedrige Fallzahlen könnten dazu führen, dass Qualitätsindikatoren in diesem Fall nicht oder beispielsweise nur durch Zusammenfassung längerer Betrachtungszeiträume zur Darstellung von Qualität geeignet seien (Steller, S. 3).

IQTIG: Die Schwierigkeit, bei geringen Fallzahlen oder seltenen Ereignissen zu sicheren Aussagen bezüglich der Versorgungsqualität zu kommen, wird in der Stellungnahme zu Recht angesprochen. Aus diesen Gründen wurde in den „Methodischen Grundlagen“ an mehreren Stellen auf die Notwendigkeit hingewiesen, die statistische Unsicherheit bezüglich der Lage der Indikatorwerte darzustellen oder, sofern sinnvoll, Indizes oder Sentinel Events zu nutzen.

Berücksichtigung des Verfahrenszwecks

In einigen Stellungnahmen wird gefragt, wie die unterschiedlichen Zwecke der Qualitätssicherungsverfahren (wie z. B. planungsrelevante Indikatoren oder Qualitätszu- und -abschläge) bei der Entwicklung von Bewertungskonzepten für Qualitätsindikatoren konkret berücksichtigt werden (GKV-SV, S. 21 f.; KBV, S. 11).

IQTIG: Die Entwicklung der grundlegenden methodischen Konzepte für Qualitätssicherungsverfahren, deren Zweck über die bisherige Qualitätsförderung hinausgeht, ist derzeit noch nicht abgeschlossen. Die bisher vom G-BA beauftragten ersten Schritte für die Verfahren der planungsrelevanten Indikatoren und der Qualitätszu- und -abschläge wurden explizit lediglich als Erweiterung auf Basis der bestehenden Qualitätsindikatoren der externen vergleichenden

Qualitätssicherung beauftragt, nicht als grundlegende Neuentwicklung. In den sich jetzt anschließenden weiteren Stufen wird der Horizont der bisherigen Indikatoren überschritten. Eine detaillierte Darstellung der Klassifikationen und Bewertungskonzepte für diese neuen Verfahrenszwecke kann daher erst in zukünftigen Versionen der „Methodischen Grundlagen“ erfolgen.

Begriff der Auffälligkeit

In mehreren Stellungnahmen wird angemerkt, dass der Begriff der Auffälligkeit und „rechnerisch auffällige“ Ergebnisse von Indikatoren nicht klar oder uneinheitlich verwendet werden (KBV, S. 23; DKG, S. 25). Wichtig sei auch, zwischen „rechnerischer Auffälligkeit“ und „auffälliger Qualität“ zu unterscheiden, da beides nicht gleichzusetzen sei (DKG, S. 25 f.)

IQTIG: Der Begriff der rechnerischen Auffälligkeit wird in der QSKH-RL verwendet und meint dort das Vorliegen von Indikatorwerten (einer Einrichtung) außerhalb des definierten Referenzbereichs. In der QSKH-RL wird zudem festgelegt, dass Fallzahlen und Konfidenzintervalle dabei nicht berücksichtigt werden sollen. Die Qesü-RL spricht demgegenüber von „Auffälligkeiten“, ohne diese näher zu spezifizieren. Da Indikatorwerte immer errechnete Werte sind, wird in den „Methodischen Grundlagen“ der sprachlichen Einfachheit halber auch die verkürzte Form „Auffälligkeit“ oder „auffälliger Indikatorwert“ verwendet. Auffälligkeiten können in verschiedenen Verfahren auf unterschiedliche Weise definiert sein. Indikatorwerte, die nicht nur eine Referenzbereichsgrenze überschreiten, sondern statistisch signifikant von dieser abweichen, werden als „statistisch auffällig“ bezeichnet. Das Glossar wurde entsprechend ergänzt.

Eine Gleichsetzung von auffälligen Indikatorwerten und auffälliger Qualität erfolgt nicht. Allerdings ist es für die Festlegung, welche Bereiche auf der Werteskala eines Indikators Anlass für weitere Handlungen geben, erforderlich, die jeweils zu repräsentierende Qualität zu benennen (z. B. „unzureichende Qualität“ oder „außerordentlich gute Qualität“). Das bedeutet, der Referenzbereich und der außerhalb des Referenzbereichs liegende Bereich der Indikatorkala operationalisieren die verschiedenen Erfüllungsgrade von Qualität. Liegt der Indikatorwert einer Einrichtung außerhalb des Referenzbereichs, so weist das Instrument Qualitätsindikator zunächst auf eine Nicht-Erfüllung der über den Indikator und sein Bewertungskonzept operationalisierten Qualitätsanforderungen hin. Das Instrument Strukturierter Dialog kann nachfolgend zur gleichen oder zu einer anderen Einschätzung bezüglich der Erfüllung der Qualitätsanforderungen des Indikators kommen.

3.6 Datenvalidierung von QS-Dokumentationsdaten und Strukturierter Dialog

Einige Stellungnahmen kritisieren die Bewertung der Datenvalidität beim Stichprobenverfahren mit Datenabgleich, bei dem die für die QS-Dokumentation übermittelten Daten mit den Angaben aus der Patientenakte für zufällig ausgewählte Leistungserbringer abgeglichen werden. Es sei unverständlich, wieso eine Übereinstimmung von nur 80 % noch als verbesserungsbedürftige Datenqualität toleriert werden könne (GKV-SV, S. 22; PatV, S. 9).

IQTIG: Der Satz zur Bewertung der 80%igen Übereinstimmungsrate als verbesserungsbedürftige Datenvalidität wurde gestrichen und durch eine Beschreibung der für die Bewertung der Datenvalidität und Ableitung des Handlungsbedarfs relevanten Faktoren ersetzt. Das IQTIG entwickelt derzeit eine neue Bewertungsmethodik. Diese wird künftig in den „Methodischen Grundlagen“ beschrieben.

Einige Stellungnahmen merken kritisch an, dass die „Methodischen Grundlagen“ zur Datenvalidierung sich lediglich auf die fallbezogene QS-Dokumentation beziehen – und nicht auf weitere, in der externen Qualitätssicherung (zukünftig) verwendete Datenquellen (z. B. Sozialdaten der Krankenkassen, Patientenbefragung, einrichtungsbezogene QS-Dokumentation) eingegangen wird (DKG, S. 17; PatV, S. 19).

IQTIG: Das IQTIG entwickelt derzeit eine Methodik zur Validierung weiterer Datenquellen – insbesondere für die Verfahren der sektorenübergreifenden Qualitätssicherung nach der Qesü-RL. Die Methodik wird in den zukünftigen Versionen der „Methodischen Grundlagen“ dargestellt.

Eine Stellungnahme kritisiert, dass eine methodische Beschreibung des Datenvalidierungsverfahrens nach der Qualitätssicherungs-Richtlinie Früh- und Reifgeborene (QFR-RL) fehle (DKG, S. 26).

IQTIG: Das methodische Vorgehen bei der Datenvalidierung nach der QFR-RL wurde in Abschnitt 12.3 der „Methodischen Grundlagen“ ergänzt.

Eine Stellungnahme kritisiert die fehlende Präzisierung, ob und in welcher Form das IQTIG externe Expertise bei der Datenvalidierung einbinden möchte (GKV-SV, S. 239).

IQTIG: Die Einbindung von Expertinnen und Experten ist an mehreren Stellen zu einzelnen Schritten des Datenvalidierungsverfahrens erwähnt. Allgemeine methodische Hinweise zur Einbindung von Expertinnen und Experten durch das IQTIG sind in Kapitel 7 der „Methodischen Grundlagen“ erläutert.

Einige Stellungnahmen kritisieren die fehlende Differenzierung der Methodik der Datenvalidierung und des Strukturierten Dialogs nach verschiedenen Verfahrenszwecken (Qualitätssicherung und -förderung, qualitätsorientierte Bedarfsplanung und Vergütung; GKV-SV, S. 23; PatV, S. 9).

IQTIG: Die neuen regulativen Verfahren (planungsrelevante Qualitätsindikatoren, Qualitätszu- und -abschläge) sind nicht Gegenstand der Version 1.0 der „Methodischen Grundlagen“.

In einer Stellungnahme wird gefragt, was unter weichen und harten Plausibilitäts- und Vollständigkeitsregeln zu verstehen sei (KBV, S. 25).

IQTIG: Dies ist in Abschnitt 12.1 erläutert: Die Verletzung weicher Plausibilitätsregeln in der Spezifikation der QS-Dokumentation führt zu einer Warnmeldung bei der Dateneingabe durch den Leistungserbringer; der Dokumentationsbogen kann dennoch abgeschlossen werden. Weiche Plausibilitätsregeln kommen dann zum Einsatz, wenn es nachvollziehbare Sonderfälle geben kann, bei denen die Plausibilitätsregeln nicht anzuwenden sind. Für diese Fälle sollte ein Abschluss des Dokumentationsbogens möglich sein. Nur wenn es keine Ausnahmefälle gibt, werden harte Plausibilitätsregeln eingesetzt, da hier ein Abschluss des Bogens ohne eine bestimmte Angabe durch den Leistungserbringer nicht mehr möglich ist.

Eine Stellungnahme kritisiert die fehlende Darstellung von Weiterentwicklungsbedarfen (z. B. gezielter Datenabgleich bei Leistungserbringern mit Dokumentationsmängeln) in Bezug auf das Datenvalidierungsverfahren (PatV, S. 19).

IQTIG: Vorschläge zur Weiterentwicklung sind nicht Bestandteil der „Methodischen Grundlagen“, sondern werden künftig in Diskussionspapieren des IQTIG formuliert und den zuständigen Fachexpertinnen und -experten und Gremien zur Diskussion gegeben.

3.7 Methoden der Datenanalyse

Allgemeine Themen

IQTIG: Das IQTIG bedankt sich für die vielen konstruktiven Vorschläge zur Verbesserung des Kapitels und die reichhaltigen Literaturangaben zu alternativen Methoden und Ansätzen bei der Risikoadjustierung, wie z. B. LASSO, Random-Forest und Propensity-Score-Matching. Die Vorschläge werden als Teil des kontinuierlichen Entwicklungsprozesses der Verfahren am IQTIG mit aufgenommen und entsprechend evaluiert. Wir möchten an dieser Stelle noch einmal den vorläufigen Charakter des Kapitels betonen. Der Fokus der Arbeit am IQTIG in der Aufbauphase lag in erster Linie auf der Umsetzung der Qualitätssicherungsverfahren im Regelbetrieb und den konkreten, vom G-BA beauftragten Entwicklungsprojekten. Zukünftige Versionen der „Methodischen Grundlagen“ werden dann entsprechend die jetzigen Stellungnahmen berücksichtigen. Insbesondere sollen die „Methodischen Grundlagen“ nur Verfahren und Methoden beinhalten, die tatsächlich angewandt werden.

Spezielle Themen

Mehrere der Stellungnahmen bitten um weitere Beispiele für stochastische Einflüsse und weisen speziell darauf hin, dass der Leistungserbringer beim Thema Medizinprodukt in der Verantwortung stehe (z. B. GKV-SV, S. 23; BÄK, S. 20). Speziell sollte deshalb auch auf den Begriff „fallbezogene Einflüsse bei Prozessindikatoren“ nochmal eingegangen werden.

IQTIG: In der überarbeiteten Version der „Methodischen Grundlagen“ hat das IQTIG die Liste an Beispielen für „weitere Einflüsse“ überarbeitet. Bei der erwähnten Art „weitere Einflüsse“ geht es vor allem um Einflussfaktoren, die weder als patientenseitig noch (bzw. nur sehr bedingt) als vom Leistungserbringer zu beeinflussen angesehen werden können. In seltenen Extremsituationen, wie z. B. globalen Krankheitsausbrüchen oder Naturkatastrophen, ist eine Zuordnung zur Kategorie „weitere Einflüsse“ weitestgehend unstrittig. In anderen Situationen wird die Trennung zwischen einem „weiterem Einfluss“ und der Leistungserbringerverantwortung jedoch schnell komplex und strittig, da der Leistungserbringer beispielsweise für die Auswahl des Medizinprodukts oder des Spenderorgans die Verantwortung trägt. Jedoch kann es auch hier zu speziellen Konstellationen kommen, die für den Leistungserbringer im Vorfeld nicht vorherzusehen und somit schwer oder überhaupt nicht zu beeinflussen sind. Bei Medizinprodukten sind die Verantwortlichkeiten tatsächlich komplex, sollten deshalb aber erwähnt bleiben. Ein Beispiel aus der Transplantationsmedizin sind nicht entdeckte Tumore bei der Spenderin oder dem Spender oder eine extern verursachte verlängerte Transportzeit des Spenderorgans. Auch wenn diese Konstellationen nur selten vorkommen, sollten sie bedacht und erwähnt werden, da viele seltene Konstellationen in der Summe doch eine gewisse Relevanz erreichen.

Bei Ergebnisindikatoren lassen sich, wie in Abschnitt 13.2 der „Methodischen Grundlagen“ beschrieben, zahlreiche Beispiele für patientenseitige Einflüsse finden. Wie jedoch in Abschnitt 13.1 beschrieben, kann hingegen bei Prozessindikatoren nicht pauschal beantwortet werden, ob es patienten- bzw. fallbezogene Einflüsse gibt und wie diese berücksichtigt werden sollen. Für Prozessindikatoren muss die Diskussion für jeden Indikator einzeln geführt werden. Im Folgenden wird die Thematik „fallbezogener Einflüsse“ am Beispiel des Prozessindikators „Perioperative Antibiotikaphylaxe bei Kaiserschnittentbindung“ (QI-ID 50045) aus dem Qualitätssicherungsverfahren *Geburtshilfe* illustriert. Das Qualitätsziel dieses Indikators ist eine möglichst hohe Rate an perioperativer Antibiotikaphylaxe bei Kaiserschnittentbindung. Es existiert jedoch eine Reihe von möglichen Gründen, weshalb berechtigterweise auf eine Antibiotikagabe verzichtet wird. Dazu zählen beispielsweise Antibiotikaunverträglichkeiten sowie Kreuzallergien bei der Patientin. Auch erfordert die Antibiotikagabe eine vorherige Aufklärung der Patientin, für die in manchen Fällen die Zeit wegen der Dringlichkeit des Kaiserschnitts fehlt. Letztendlich können die Patientinnen auch die Prophylaxe ablehnen. Obige Beispiele sind alles legitime Gründe, auf eine Antibiotikagabe zu verzichten, Gründe, die jedoch nicht in den Daten erfasst werden. Aus diesem Grund können die genannten fallbezogenen Einflüsse nicht explizit berücksichtigt werden. Stattdessen werden sie implizit berücksichtigt, indem der Referenzbereich des Indikators auf $\geq 90\%$ anstelle von $= 100\%$ festgelegt wird.

Anmerkung: Es ist bedauerlich, dass das Methodenpapier kein Auswertungskonzept für Follow-up-Indikatoren beinhaltet, da diese in der sektorenübergreifenden Qualitätssicherung und in einigen Leistungsbereichen der externen stationären Qualitätssicherung zunehmend von Bedeutung sein werden. (BÄK, S. 20)

IQTIG: Ein solches biometrisches Methodenkonzept für die mit Patientenpseudonym (PSN) gestützten Follow-up-Indikatoren in der QSKH-RL in den Qualitätssicherungsverfahren *Hüftendoprothesenversorgung*, *Knieendoprothesenversorgung* und *Herzschrittmacherversorgung* steht seit Mai 2017 unter <https://iqtig.org/das-iqtig/grundlagen/biometrische-methoden/?L=0> zu Verfügung. Bei dem Dokument handelt es sich um eine umfassende biometrische Darstellung der Auswertungsmethodik im Rahmen der statistischen Ereigniszeitanalyse. Das Dokument fokussiert dabei auf eine mathematisch präzise und im biometrischen Sinne transparenzschaffende Darstellung des Verfahrens.

Dabei ist wichtig zu beachten, dass der Begriff Follow-up-Indikatoren derzeit im QS-Sprachgebrauch für sehr unterschiedliche Sachverhalte verwendet wird:

1. Follow-up-Indikatoren gemäß Anlage 2 in der QSKH-RL bei den Qualitätssicherungsverfahren *Hüftendoprothesenversorgung*, *Knieendoprothesenversorgung* und *Herzschrittmacherversorgung*. Bei diesem Verfahren werden QS-Dokumentationsbögen von Ersteingriffen und Folgeeingriffen anhand von patientenidentifizierenden Daten (PID-Felder) mittels PSN

verknüpft. Zu Todesfällen liegen bei diesem Verfahren keine Informationen vor, weshalb diese die Analysen für längere Follow-up-Perioden verzerren.

2. Follow-up-Indikatoren als 1-, 2- und 3-Jahres-Indikatoren in den transplantationsmedizinischen Qualitätssicherungsverfahren der QSKH-RL, welche durch ein aktives Follow-up-Verfahren erhoben werden und zurzeit jeweils als separate Anteilsindikatoren ausgewertet werden
3. sozialdatengestützte Follow-up-Indikatoren bei Qualitätssicherungsverfahren nach der Qesü-RL, z. B. bei *Perkutane Koronarintervention (PCI) und Koronarangiographie (QS PCI)* und *Vermeidung nosokomialer Infektionen – postoperative Wundinfektionen (QS WI)*

Das oben erwähnte methodische Konzept zur Auswertung von Follow-up-Indikatoren wurde für Indikatoren aus der ersten Gruppe entwickelt. Das Konzept könnte sich mit einigen Modifikationen ggf. auf die Indikatoren der zweiten Gruppe übertragen lassen. Für Indikatoren der dritten Gruppe ist zum jetzigen Zeitpunkt unklar, ob und inwieweit die zeitlichen Verzögerungen der Lieferung von Sozialdaten bei den Krankenkassen einen abweichenden Analyserahmen erfordern werden. Um dies in adäquater Weise empirisch erproben zu können, muss zunächst eine geregelte und funktionierende Übermittlung der Sozialdaten erfolgen.

Die Wahl adäquater biometrischer Auswertungsmethodik beinhaltet eine Evaluation anhand simulierter Daten und realer, im Rahmen des Verfahrens erhobener Daten. Eine solche datengestützte Entwicklung ist aufwendig und kann nur schwierig und unvollständig vor dem Erhalt der Real-Daten erstellt werden. Somit ist das Auswertungskonzept bei Entwicklungsaufträgen dementsprechend an geeigneter Stelle einzuplanen. In vielen Situationen ist es daher auch angemessen, zunächst nur eine vorläufige Auswertungsmethodik zu präsentieren, die dann im Rahmen des laufenden Verfahrens verfeinert wird.

In Zukunft wird das IQTIG im Rahmen der Weiterentwicklungsaufträge mit Auswertungskonzepten immer erst eine Darstellung der statistischen Methoden anstreben. In einer solchen Darstellung wird die Auswertungsmethodik klar festgelegt und durch ihre Veröffentlichung auch transparent dargestellt. Ein Beispiel für ein solches sogenanntes biometrisches Methodendokument ist die erwähnte Auswertungsmethodik bei den PID-gestützten Follow-up-Indikatoren. Erst nach einer gewissen Stabilität eines Verfahrens werden die biometrischen Methoden dann in einem zweiten Schritt vereinfacht in den „Methodischen Grundlagen“ dargestellt.

Mehrere Stellungnahmen beziehen sich auf eine fehlende Beschreibung der konkreten Verfahren zur Validierung von Risikoadjustierungsmodellen (GMDS/IBS-DR, S. 2 f.; KBV, S. 25; WINEG, S. 11; DKG, S. 27 f.).

IQTIG: Aktuell wird am IQTIG im Rahmen des IQTIG-internen Projekts „Weiterentwicklung Risikoadjustierung“ an einer standardisierten Vorgehensweise

(SOP) gearbeitet, die den Entwicklungs- und Validierungsprozess von Risikoadjustierungsmodellen vereinheitlichen soll. In diesem Zusammenhang wird darüber hinaus erarbeitet, welche Informationen die Veröffentlichung von Risikoadjustierungsmodellen umfassen soll.

Wichtig ist, dass bei der Verwendung von inferenzstatistischer Modellierung zwischen den Zielen Erklärung und Prädiktion unterschieden werden muss (Shmueli 2010):

- Beim Ziel Erklärung handelt es sich um Analysen mit dem Ziel des epidemiologischen Erkenntnisgewinnes, z. B. die Untersuchung, welche Risikofaktoren im Zusammenhang mit der betrachteten Zielvariable stehen. Oft ist die eigentliche Fragestellung die Aufdeckung kausaler Zusammenhänge.
- Beim Ziel Prädiktion handelt es sich um die bestmögliche Vorhersage des Outcomes bei neuen bzw. zukünftigen Beobachtungen. Die Vorhersagen können mittels Punktschätzer, Prädiktionsintervalle oder prädiktiver Verteilungen geschehen.

Das Ziel des Risikoadjustierungsmodells, bei der jetzigen Berechnungsart der indirekten Standardisierung durch eine Referenzpopulation, ist eher prädiktiver Natur. Jedoch handelt es sich tatsächlich nur um *partielle* prädiktive Modelle, denn es soll nur für messbare Faktoren, die nicht in der Verantwortung des Leistungserbringers stehen, adjustiert werden.

Weitere Ausführungen zu relevanten Aspekten der Risikoadjustierung auf Basis von Regressionsmodellen, die unter dem Begriff der „Gütekriterien“ zusammengefasst werden können, werden nicht gemacht. Dies ist erstaunlich, da gerade hierzu Publikationen vorliegen, die die Bedeutung dieser Kriterien – auch als Grundlage der öffentlichen Berichterstattung – ausführlich beschreiben (so zum Beispiel Becker 2014a; Becker 2014b; Becker et al. 2016). Richtigerweise erkennt das IQTiG die Insuffizienz der aktuellen Ausführungen und verweist auf die aktuelle Entwicklungsarbeit: „In der vorliegenden Fassung der „Methodischen Grundlagen“ werden lediglich Grundbegriffe der zum Zeitpunkt der Erstellung eingesetzten Methoden der Risikoadjustierung eingesetzt. Wichtige Aspekte der gegenwärtig bereits verwendeten Verfahren werden derzeit durch das Institut wissenschaftlich weiterentwickelt.“ (Abschnitt 13.2, Seite 131) Auch wenn derzeit an den Verfahren gearbeitet wird, so fehlt die schon heute mögliche Darlegung der bei der Entwicklung und Validierung zu berücksichtigenden Gütekriterien. Hierzu ist keine „Neuerfindung“ erforderlich, schon in dieser Version des Methodenpapiers hätte das IQTiG seine Qualitätsziele und auch Qualitätskriterien bei der Entwicklung und Validierung geeigneter Modelle im Sinne der von den Leistungserbringern geforderten Qualitätstransparenz darlegen können. Da dies nicht geschehen ist, kann die folgende Aussage des IQTiG nur als Absichtsbekundung gewertet werden: „Umso wichtiger ist es, z. B. die Definition der Risikofaktoren oder die Berechnungsmodelle zur Risikoadjustierung öffentlich zur Verfügung zu stellen und die entsprechende Transparenz beim gewählten Vorgehen herzustellen.“ (Abschnitt 13.2, Seite 131) Einschränkend führt das IQTiG aber auch (vorsorglich?) aus: „Insgesamt

kann die Wahl eines adäquaten Risikoadjustierungsmodells zwar zu einem gewissen Grad standardisiert werden, aber eine auf jede Fragestellung oder Situation passende, allgemeingültige Beurteilung der Modellgüte erscheint nicht sinnvoll.“ (Abschnitt 13.2, Seite 131)) Warum eine allgemeingültige Beurteilung der Modellgüte, die auf jede Fragestellung und jede Situation (also: Risikoadjustierungsmodell?) passt, als „nicht sinnvoll“ erscheint, erschließt sich nicht im Ansatz. Die entsprechende Feststellung des IQTiG ist aus Sicht des Autors dieser Stellungnahme mit dem internationalen Standard der Wissenschaften zu diesem Thema nicht vereinbar. (DKG, S. 27 f.)

IQTiG: Das IQTiG bedankt sich für die Literaturangaben. Diese werden zurzeit zusammen mit internationaler wissenschaftlicher Literatur bei der Erstellung einer SOP zur Risikoadjustierung begutachtet. Jedoch vertritt das IQTiG die Auffassung, dass eine „Neuerfindung“ tatsächlich nötig ist, da aus Sicht des Instituts keine einzelne gängige Modellgüte-Metrik die besonderen Bedürfnisse der Risikoadjustierung in Qualitätssicherungsverfahren umfassend berücksichtigt. Aus der obigen Stellungnahme wird speziell auch nicht klar, was genau unter dem Begriff Modellgüte subsumiert wird. Aus Sicht des IQTiG ist die Modellgüte keine rein statistische Größe, u. a. auch weil –wie weiter unten beschrieben wird– nicht alle Fragestellungen rein datengestützt beantwortet werden können. Dies bedeutet insbesondere auch, dass sich die Modellgüte grundsätzlich nicht in einer (einzigen) Zahl zusammenfassen lässt. In den Stellungnahmen werden einige Gütemaße aus der Literatur zitiert, die aus Sicht der Stellungnehmenden Anwendung finden sollten. Jedoch bleibt offen, wie aus dieser Palette an Metriken eine eindeutige Entscheidung für ein Modell generiert werden soll. Auch sind viele der vorgeschlagenen Maße relativ, d. h., sie können nur zum Vergleich von Modellen im gleichen Datensetting genommen werden. Neben der Tatsache, dass die in der Stellungnahme erwähnten Maße ausschließlich zur Bewertung des rein statistischen Teils der Modellgüte beitragen können, führen sie zu folgenden Schwierigkeiten:

- Der Hosmer-Lemeshow-Test liefert schnell ein signifikantes Testergebnis, welches jedoch bei den großen Datenmengen mancher Risikoadjustierungsmodelle nicht gleichbedeutend einer relevanten Abweichung der Modellannahmen ist.
- Die Area under the curve (AUC) eines Modells hängt von folgenden Faktoren ab:
 - Heterogenität der Risikofaktoren in den verwendeten Daten bezüglich der Outcome-Variable (Beispiel: Im Qualitätssicherungsverfahren *Hüftendoprothesenversorgung* sind leicht hohe AUC-Werte bei den Risikoadjustierungsmodellen zu erzielen, weil hier die Trennung zwischen elektiven Eingriffen und Femurfrakturen bezüglich der Gehunfähigkeit groß ist.)
 - Häufigkeit eines Risikofaktors (Wird im untenstehenden fiktiven Datensatz bestehend aus 10.010 Fällen für den binären Risikofaktor adjustiert, hat das entsprechende logistische Regressionsmodell einen

AUC-Wert von 0.504, obwohl der Risikofaktor das Outcome so gut wie determiniert.)

		Risikofaktor	
		0	1
Outcome	0	9000	1
	1	1000	9

- Der Brier-Score hängt von der Prävalenz des Outcomes ab.

Obige Überlegungen zeigen, dass Neuentwicklungen im Kontext der Risikoadjustierung nötig sind – auch wenn sie im Endeffekt nur aus der Kombination verschiedener bereits existierender Gütemaße bestehen. Erschwerend kommt hinzu, dass die Modellwahl auch eine Vielzahl an Aspekten zu berücksichtigen hat, die nicht statistischer Natur sind. Beispiele:

- inhaltliche Entscheidungen wie z. B., ob Leistungserbringer, die ASA-5-Patientinnen und -Patienten behandelt haben, nun einen „Risikozuschlag“ bekommen sollten (wenn sich das aus den Daten ergibt) oder ob man diese Patientinnen und Patienten eigentlich überhaupt nicht hätte behandeln sollen
- Ausschluss von Dokumentationsfeldern aufgrund der Vermutung, dass die Felder nicht richtig interpretiert wurden (Datenvalidität), auch wenn die Felder inhaltlich sinnvoll wären und ggf. sogar die Prädiktion verbessern
- Nebenwirkungen, z. B. aufgrund möglicher Fehlanreize durch Patientenselektion

Somit besteht das Thema Validität von Risikoadjustierung aus vielen Dimensionen (Iezzoni 2013), die *alle* gemeinsam zu berücksichtigen sind:

- Inhaltsvalidität: medizinisch plausible und möglichst vollständige Berücksichtigung
- Datenvalidität: z. B. bei Verwendung von Routinedaten: beim Vorhandensein einer Diagnose ist zwischen „Present on Admission“ und „Present at Discharge“ zu unterscheiden.
- Modellvalidität:
 - wie gehen die Risikofaktoren in das Modell ein, z. B. durch die Gruppierung von kategorialen Kovariablen bzw. bei der Wahl der funktionalen Form bei kontinuierlichen Kovariablen
 - wie gut schneidet das Modell nach bestimmten Gütemaßen ab, z. B.
 - bei der Distanz zwischen vorhergesagter Wahrscheinlichkeit und tatsächlich beobachteten Werten (Kalibrierung)
 - bei der Fähigkeit des Modells, zwischen Fällen mit und ohne Ereignis zu unterscheiden (Diskriminationsfähigkeit)

- **Methodenvalidität:** Ist zum Beispiel die standardisierte Mortalitätsrate (*standardized mortality ratio, SMR*)¹ für den gedachten Verfahrenszweck das Richtige? Die Antwort fällt hier differenzierter aus, wenn das Ziel die Patienteninformation ist und nicht der Vergleich der Leistungserbringer mit einem normativen Standard.

Zusammenfassend lässt sich festhalten, dass Risikoadjustierungsmodelle bei der jetzigen Berechnungsmethodik durch O/E-Indikatoren zwar mit dem Ziel der Prädiktion entwickelt werden, jedoch keine rein datengetriebenen Vorhersagemodelle sind, da auch medizinische Plausibilität und normative Festlegungen eine wichtige Rolle spielen. Aus diesen Gründen bleibt das IQTiG bei der Aussage, dass das Vorgehen der Risikoadjustierung nie komplett standardisiert werden kann. Für das IQTiG ist deshalb die Behauptung, dass eine solche Aussage nicht mit dem wissenschaftlichen Standard vereinbar ist, nicht nachvollziehbar. Trotz alledem arbeitet das IQTiG aktuell daran die Teile, die im Prozess standardisiert werden können, durch eine SOP zu vereinheitlichen.

Von Seiten der Leistungserbringer und auch der Wissenschaft ist zu fordern, dass nicht nur die Berechnungsmodelle zur Risikoadjustierung öffentlich zur Verfügung gestellt werden, sondern auch die zur Entwicklung und Validierung des jeweiligen Risikoadjustierungsmodells verwendeten Daten. Nur so wird die auch vom IQTiG propagierte Transparenz vollständig gewährleistet und das Vertrauen in die Qualität der Qualitätsmessung gefördert. (DKG, S. 28)

IQTiG: Das IQTiG setzt sich für Transparenz bei den Risikoadjustierungsmodellen ein. Dabei ist das Spannungsfeld zwischen kompletter Veröffentlichung von den dazu notwendigen Daten und dem Datenschutz der Falldaten von Patientinnen und Patienten zu berücksichtigen. Somit ist es unkomplizierter, einen hohen Grad an Transparenz durch gute Dokumentation der Risikoadjustierungsmodelle zu bekommen.

Konkrete Vorstellungen des IQTiG bezüglich der Dokumentation von Risikoadjustierungsmodellen beinhalten eine Beschreibung des Datenpools, welcher für die Schätzung der Regressionsparameter verwendet wurde, z. B. durch Angabe des Erfassungsjahres des Datenpools und der Anzahl an Fällen im Datenpool. Stimmt der Datenpool aus inhaltlichen Überlegungen nicht direkt mit der nach der Indikatordefinition vorgegebenen Grundgesamtheit überein, ist dies speziell festzuhalten. Die Frage, mit der sich das IQTiG aktuell auseinandersetzt, ist, wie und an welcher Stelle diese Informationen geeignet kommuniziert werden können. Das eigentliche Ergebnismodell der Risikoadjustierung wird bereits jetzt transparent in der Qualitätsindikatorendatenbank (QIDB) und in den Bundesauswertungen dargestellt:

- Die Koeffizienten der Modelle werden in tabellarischer Form im Rahmen der QIDB veröffentlicht. Somit ist schon immer klar, welche Risikofaktoren auf welche Art berücksichtigt werden und welche nicht.

¹ D. h. die zurzeit verwendete Methodik für risikoadjustierte Indikatoren durch indirekte Standardisierung, welche durch die Berechnung des O/E-Verhältnisses (= der SMR) operationalisiert wird.

- Seit dem Erfassungsjahr 2016 werden die Modelle auch in Rechenregel-Skriptform veröffentlicht. Somit ist transparent, wie genau gerechnet wird.
- Die Darstellung der Risikoadjustierungsmodelle in den Bundesauswertungen des Erfassungsjahres 2016 wurde dahingehend überarbeitet, dass die risikoadjustierten Ergebnisse verständlicher dargestellt werden. Es werden nun das Grundrisiko, die Odds-Ratios für die einzelnen Risikofaktoren sowie Vertrauensintervalle für die Odds-Ratio dargestellt. Konkrete Rechenvorschriften sind in den Dokumenten der QIDB zu finden.

Die in der Stellungnahme geforderte komplette Transparenz bezüglich der Daten, aus denen Risikoadjustierungsmodelle entwickelt wurden, würde einen direkten Zugriff der Öffentlichkeit auf die anonymisierten QS-Daten der Patientinnen und Patienten bedeuten, welcher aus Datenschutzgründen nicht möglich ist. Im Rahmen eines sekundären Datennutzungsantrags besteht allerdings die Möglichkeit, die QS-Daten unter Einhaltung der entsprechenden Geschäftsordnung und Datenschutzaufgaben der sekundären Datennutzung auswerten zu lassen und somit extern die Risikoadjustierungsmodelle zu validieren. Somit ist die Möglichkeit externer Kontrolle prinzipiell bereits gegeben. Zu diskutieren ist dagegen, wie genau Transparenz bezüglich des Auswahlprozesses der Risikofaktoren geschaffen werden kann, denn, wie bereits in Abschnitt 13.2 der „Methodischen Grundlagen“ geschrieben, handelt es sich hier um einen iterativen Prozess, der sowohl daten- als auch expertengestützt ist und sich deshalb nicht vollkommen standardisieren lässt. In dem Rahmen, in dem sich Vorgänge jedoch vereinheitlichen lassen, werden zurzeit konkrete Vorschläge vom IQTIG erarbeitet.

[...] Es wäre sinnvoll, für die Erstellung der Referenzbereiche Regressionsverfahren (etwa die Quantilenregression [Koenker, 2005]) zu verwenden. Solche Verfahren oder Verfahrensklassen werden im Papier nicht genannt. (GMDS/IBS-DR, S. 2)

IQTIG: Prinzipiell gibt es zwei Möglichkeiten, die Risikoadjustierung für unterschiedliche Patientenkollektive durchzuführen: Entweder man adjustiert das Ergebnis oder man adjustiert den Referenzbereich. Das Standardvorgehen des IQTIG, um die unterschiedliche Zusammensetzung von Patientenprofilen über die Leistungserbringer zu berücksichtigen, ist derzeit eine indirekte Standardisierung der Leistungserbringerergebnisse anhand von SMRs. Die empirischen Quantile werden dann auf den realisierten SMR-Ergebnissen der Leistungserbringer berechnet. Durch die Festlegung der Grenze des Referenzbereichs beispielsweise als 95%-Perzentil erreicht man, dass ungefähr 5 % der Leistungserbringer als rechnerisch auffällig eingestuft werden. Dabei ist zu beachten, dass diese quantitative Einstufung der Leistungserbringer dann in einem zweiten Schritt (im Rahmen des Strukturierten Dialogs) qualitativ verifiziert und bewertet werden. Somit stellt der quantitative Schritt in der Qualitätsförderung oft eine Vorselektion der zu untersuchenden Leistungserbringer für den Strukturierten Dialog, d. h. der qualitativen Bewertung, dar. Die prinzipielle Frage für zukünftige Auswertungen ist, inwieweit diese Schritte parallel

oder sequentiell geschehen werden und wie die Gewichtung zwischen den beiden Befunden bei der Bewertung stattfindet.

In den jetzigen Verfahren der Qualitätssicherung ist zurzeit ein für alle Leistungserbringer einheitlich geltender Referenzbereich festgelegt. Das hat u. a. den Vorteil, dass dieser externe Standard klar kommuniziert werden kann. Bei den sektorenübergreifenden Verfahren wird die Frage, ob ein einheitlicher Referenzbereich sektorenübergreifend gelten soll, zurzeit unterschiedlich bei QS PCI und QS WI beantwortet. Insgesamt sieht das IQTIG perzentilbasierte Referenzbereiche als nützliches Instrument, um für neue Indikatoren Referenzbereiche festzulegen; längerfristig sollten solche relativen Referenzbereiche jedoch durch feste Referenzbereiche ausgetauscht werden (vgl. Abschnitt 10.3.2). Eine Möglichkeit wäre hier z. B., für einen risikoadjustierten Indikator die Setzung vorzunehmen, dass eine Sterblichkeit nur um maximal 25 % oder 50 % höher als das durch die Risikoadjustierung berechnete ‚E‘ sein darf.

Das IQTIG interpretiert obige Stellungnahme so, dass Quantilsregression verwendet werden könnte, sodass Einflüsse von Risikofaktoren auf die Perzentile abgebildet werden und die Quantilsregression somit als alternative Art der Risikoadjustierung verwendet werden könnte. Aus Sicht des IQTIG sind die Ergebnisse einer solchen Risikoadjustierung jedoch schwieriger vermittelbar, weil nicht das Ergebnis des Leistungserbringers, sondern der Referenzbereich adjustiert wird. Somit würde es keinen einheitlichen bundesweiten Referenzbereich mehr geben. Auch würde eine solche Verwendung der Quantilsregression nicht berücksichtigen, dass die Leistungserbringerergebnisse wegen der unterschiedlichen Fallzahlgrößen auch entsprechend unterschiedliche Variabilität haben.

... Generell liefern Vielebenenmodelle [R.2] methodische Ansätze, die es erlauben, über hierarchisierte Zufallseffekte die Qualität einer Institution adjustiert für quantifizierte institutionsspezifische Faktoren durch institutionsbezogene Zufallseffekt zu quantifizieren und gleichzeitig patientenspezifische gemessene Risikofaktoren (sowie nicht gemessene Risikofaktoren über individuelle Zufallsfaktoren) zu berücksichtigen. Im Rahmen Bayesianischer Strategien [Henschel et al, 2009] wurden solche Ansätze erprobt und besitzen ein großes Potential, um im Methodenspektrum des IQTIG wertvoll zu werden. (GMDS/IBS-DR, S. 4)

IQTIG: Das IQTIG sieht dies ähnlich. Jedoch muss bei der Durchführung der sogenannten indirekten Verfahren im Rahmen der QSKH-RL berücksichtigt werden, dass die konkrete Berechnung, die zur rechnerischen Auffälligkeitseinstufung der Leistungserbringer führt, von der auf Landesebene beauftragen Stelle durchgeführt wird. Aus diesem Grund werden in der jetzigen Vorgehensweise einheitliche Rechenvorschriften vom IQTIG festgelegt, die über die QIDB kommuniziert werden. Speziell für risikoadjustierte Indikatoren mittels logistischer

Regression bzw. des multiplikativen Hazardmodells beinhaltet die Rechenvorschrift ein Skript zur Berechnung der erwarteten Anzahl an Ereignissen für die vorliegenden Patientenprofile.²

Im Regelfall werden die Risikoadjustierungsmodelle zurzeit auf Grundlage aller Fälle, die in die Berechnung des Qualitätsindikators im Vorjahr eingingen, erstellt. Des Weiteren wird ein einheitlicher perzentilbasierter Referenzbereich anhand der Leistungserbringerergebnisse aus dem aktuellen Erfassungsjahr festgelegt. Danach kann die Berechnung unabhängig auf Länderebene durchgeführt werden. Hierarchische Modelle, bei denen Risikofaktoren und Leistungserbringereffekt im gleichen Modell integriert sind, würden eine solche zweistufige Rechnung nicht erlauben, da sie jederzeit komplett vom aktuellen Datenpool abhängig sind. Als Beispiel: Neue Daten für einen einzelnen Leistungserbringer können zur Verschiebung aller Ergebnisse und Auffälligkeitseinstufungen führen. Die Transparenz dieser Verfahren ist somit für den Leistungserbringer geringer und es ist nicht möglich, vorab, z. B. als Teil der internen Qualitätssicherung, unterjährig mit den gültigen Rechenvorschriften des Erfassungsjahres zu rechnen. Somit ist die Entscheidung, welches Verfahren für die Risikoadjustierung zu wählen ist, eine Abwägung zwischen sehr vielen Dimensionen, z. B. statistischer Effizienz, Klassifikationsgüte, Reproduzierbarkeit und Kommunizierbarkeit.

Hierarchische Modelle kommen jedoch bereits intern am IQTIG zum Einsatz, z. B. bei Sonderanalysen im Rahmen von Indikatorentwicklungen.

Das Methodenpapier des IQTIG beschreibt keine Methoden, wie Ergebnisse der Qualitätsanalysen präsentiert und kommuniziert werden. Im ersten Abschnitt zum Kapitel 13 wird ein zweistufiges Verfahren beschrieben: Bereitstellung quantitativer Information zur Qualität einer Institution, dann im zweiten Schritt gegebenenfalls eine fachlich-qualitative Bewertung. Diese Methoden sind einerseits verfahrensspezifisch. Es sollte dennoch klare methodische Vorgehensweise formuliert werden, wie Ergebnisse numerisch und/oder graphisch darzustellen sind und wie eine objektive Interpretation dieser Zahlen ermöglicht wird. Hierzu sollten reporting guidelines vorgelegt und der Einsatz graphischer Methoden beschrieben werden, um etwa Benchmarking möglich zu machen

So stellt etwa der Funnelplot ein zentrales Instrument beim Benchmarking risikoadjustierter Indikatoren dar. Dieses Instrument ist im Methodenpapier nicht erwähnt worden. Methodisch verweisen wir hier auf Arbeiten wie [Spiegelhalter, 2005]. (GMDS/IBS-DR, S. 4)

IQTIG: Das IQTIG dankt für diesen Kommentar und schließt sich der Meinung an, dass eine gute und klare Darstellung und Kommunikation in der Berichterstattung erstrebenswert ist. Dieses Ziel ist jedoch schwierig zu erreichen, denn die Berichte bedienen sehr viele unterschiedliche Interessen.

² D. h. der Wert von „E“ bei der Berechnung von O/E-Indikatoren. Diese Berechnungsvorschrift entspricht für die einzelne Patientin / den einzelnen Patienten der vorhergesagten Wahrscheinlichkeit für das Eintreten eines Ereignisses bei Patientinnen und Patienten mit diesem Risikoprofil.

Mit der Vielzahl an zu entwickelnden Berichten, die sich aus den zahlreichen neuen Richtlinien und Verfahren ergeben, liegt der Fokus zunächst auf der Entwicklung einer passenden *Auswertungsmethodik* durch adäquate statistische Methoden (siehe z. B. Follow-up-Methodik bei PID-gestützten Indikatoren der QSKH-RL sowie die statistische Auffälligkeitseinstufung bei der Richtlinie zu planungsrelevanten Qualitätsindikatoren, plan. QI-RL) und einer *fristgerechten* Berichterstattung.

Funnel-Plots (auch Trichter-Plots) werden vom IQTIG bereits an folgenden Stellen eingesetzt:

- bei der Visualisierung der Ergebnisse im Abschlussbericht zu den planungsrelevanten Indikatoren (siehe Abschnitt 4.3 von IQTIG 2016)
- in den Jahresberichten für die Leistungserbringer nach der plan. QI-RL (ab Frühjahr 2018)
- bei der übersichtlichen 2-Jahres-Zusammenfassung für die Bundesfachgruppen von Indikatorergebnissen und dazugehörigem Strukturierten Dialog (sogenannte Kommentierungstabellen)

Dabei sind die Funnel-Plots nicht nur für risikoadjustierte Indikatoren interessant, sondern bei allen Indikatorarten, denn durch die Darstellung differenziert nach Fallzahl wird die Variabilität der Ergebnisse in Abhängigkeit von der Fallzahl anschaulich visualisiert.

Literatur

- Adams, JL; Mehrotra, A; McGlynn, EA (2010): Estimating Reliability and Misclassification in Physician Profiling [Technical Report]. Santa Monica, US-CA: RAND. URL: https://www.rand.org/content/dam/rand/pubs/technical_reports/2010/RAND_TR863.pdf (abgerufen am: 18.08.2017).
- Bühner, M (2011): Einführung in die Test- und Fragebogenkonstruktion. 3., aktualisierte und erweiterte Auflage. München [u. a.]: Pearson. ISBN: 978-3-86894-033-6.
- Campbell, SM; Braspenning, J; Hutchinson, A; Marshall, MN (2003): Research methods used in developing and applying quality indicators in primary care. *BMJ – British Medical Journal* 326(7393): 816-819. DOI: 10.1136/bmj.326.7393.816.
- DIN [Normenausschuss Qualitätsmanagement], NQSZ [Statistik und Zertifizierungsgrundlagen] (2015): Qualitätssicherung. Kapitel 3.3.6. In: DIN [Normenausschuss Qualitätsmanagement], NQSZ [Statistik und Zertifizierungsgrundlagen]: *Qualitätsmanagementsysteme – Grundlagen und Begriffe (ISO 9000:2015); Deutsche und Englische Fassung EN ISO 9000:2015*. [Stand:] November 2015. Berlin: Beuth, 31.
- Döring, N; Bortz, J (2016): Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. 5. vollständig überarbeitete, aktualisierte und erweiterte Auflage. Berlin: Springer. ISBN: 978-3-642-41088-8.
- Hartig, J; Frey, A; Jude, N (2012): Validität. Kapitel 7. In: Moosbrugger, H; Kelava, A; Hrsg.: *Testtheorie und Fragebogenkonstruktion*. 2., aktualisierte und überarbeitete Auflage. Berlin [u. a.]: Springer, 143-171. ISBN: 3-642-20071-0.
- Hays, RD; Hadorn, D (1992): Responsiveness to change: an aspect of validity, not a separate dimension. *Quality of Life Research* 1(1): 73-75. DOI: 10.1007/bf00435438.
- Iezzoni, LI (2013): Reasons for Risk Adjustment. Chapter 1. In: Iezzoni, LI; Hrsg.: *Risk Adjustment for Measuring Healthcare Outcome*. 4th edition. Chicago, US-IL: Health Administration Press, 1-14. ISBN: 978-1-56793-437-3.
- IQTIG [Institut für Qualitätssicherung und Transparenz im Gesundheitswesen] (2016): Planungsrelevante Qualitätsindikatoren [Abschlussbericht zur Auswahl und Umsetzung]. Stand: 31.08.2016. Berlin: IQTIG. URL: https://iqtig.org/downloads/berichte/2016/IQTIG_Planungsrelevante-Qualitätsindikatoren_Abschlussbericht.pdf (abgerufen am: 25.08.2017).
- Kane, MT (2001): Current Concerns in Validity Theory. *JEM – Journal of Educational Measurement* 38(4): 319-342. DOI: 10.1111/j.1745-3984.2001.tb01130.x.
- Macefield, RC; Jacobs, M; Korfage, IJ; Nicklin, J; Whistance, RN; Brookes, ST; et al. (2014): Developing core outcomes sets: methods for identifying and including patient-reported outcomes (PROs). *Trials* 15: 49. DOI: 10.1186/1745-6215-15-49.

- Mainz, J (2003): Developing evidence-based clinical indicators: a state of the art methods primer. *International Journal for Quality in Health Care* 15(Suppl. 1): i5-i11. DOI: 10.1093/intqhc/mzg084.
- McGlynn, EA (2003): Selecting Common Measures of Quality and System Performance. *Medical Care* 41(1): I-39-I-47.
- Meyer, W (2004): Indikatorenentwicklung: eine praxisorientierte Einführung. 2. Auflage. (CEval-Arbeitspapier, Nr. 10). Saarbrücken: Universität des Saarlandes, Fak. 05 Empirische Humanwissenschaften, CEval [Centrum für Evaluation]. URL: <http://www.ssoar.info/ssoar/bitstream/handle/document/11124/ssoar-2004-meyer-indikatorenentwicklung.pdf?sequence=1> (abgerufen am: 28.06.2017).
- Mühlhauser, I; Müller, H (2009): Patientenrelevante Endpunkte und patient-reported outcomes in klinischer Forschung und medizinischer Praxis. In: Klusen, N; Fließgarten, A; Nebling, T: *Informiert und selbstbestimmt. Der mündige Bürger als mündiger Patient.* (Beiträge zum Gesundheitsmanagement, 24). Baden-Baden: Nomos, 34-65. ISBN: 978-3-8329-4474-2.
- NQF [National Quality Forum] (2015): Measure Evaluation Criteria and Guidance for Evaluating Measures for Endorsement. Effective April 2015. Washington, D. C.: NQF. URL: <http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=79434> (abgerufen am: 30.01.2017).
- O'Brien, RM (1990): Estimating the Reliability of Aggregate-Level Variables Based on Individual-Level Characteristics. *Sociological Methods and Research* 18(4): 473-504. DOI: 10.1177/0049124190018004004
- Prinsen, CAC; Vohra, S; Rose, MR; Boers, M; Tugwell, P; Clarke, M; et al. (2016): How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" – a practical guideline. *Trials* 17(1): 449. DOI: 10.1186/s13063-016-1555-2.
- Raykov, T; Marcoulides, GA (2011): Introduction to Psychometric Theory. London: Routledge. ISBN: 978-0-415-87822-7.
- Reiter, A; Fischer, B; Kötting, J; Geraedts, M; Jäckel, WH (2007): QUALIFY: Ein Instrument zur Bewertung von Qualitätsindikatoren. Düsseldorf: BQS [Bundesgeschäftsstelle Qualitätssicherung].
- Schermelleh-Engel, K; Werner, CS (2012): Methoden der Reliabilitätsbestimmung. Kapitel 6. In: Moosbrugger, H; Kelava, A; Hrsg.: *Testtheorie und Fragebogenkonstruktion.* 2., aktualisierte und überarbeitete Auflage. Berlin [u. a.]: Springer, 119-143. ISBN: 3-642-20071-0.
- Schnell, R; Hill, PB; Esser, E (2013): Methoden der empirischen Sozialforschung. 10., überarbeitete Auflage. München: Oldenbourg. ISBN: 978-3-486-72899-6.

Sens, B; Fischer, B; Bastek, A; Eckardt, J; Kaczmarek, D; Paschen, U; et al. (2007): Begriffe und Konzepte des Qualitätsmanagements – 3. Auflage. *GMS Medizinische Informatik, Biometrie und Epidemiologie* 3(1). Doc05. URL: <http://www.egms.de/tools/download.jsp?path=journals/mibe/2007-3/mibe000053.1a.pdf&mime=application/pdf&name=GMS-QM-Glossar-Auflage3.pdf> (abgerufen am: 30.01.2017).

Shea, BJ; Grimshaw, JM; Wells, GA; Boers, M; Andersson, N; Hamel, C; et al. (2007): Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC: Medical Research Methodology* 7(10): 1-7. DOI: 10.1186/1471-2288-7-10.

Shmueli, G (2010): To Explain or to Predict? *Statistical Science* 25(3): 289-2310. DOI: 10.1214/10-STS330.