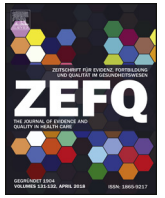




Contents lists available at ScienceDirect

Z. Evid. Fortbild. Qual. Gesundh. wesen (ZEFQ)

journal homepage: <http://www.elsevier.com/locate/zefq>

Qualität und Sicherheit in der Gesundheitsversorgung / Quality and Safety in Health Care

Methodischer Standard für die Entwicklung von Qualitätsindikatoren im Rahmen von S3-Leitlinien – Ergebnisse einer strukturierten Konsensfindung



Methodological Standard for the Development of Quality Indicators within Clinical Practice Guidelines – Results of a structured consensus process

Stefanie Deckert^{a,*}, Katrin Arnold^a, Monika Becker^b, Max Geraedts^c, Marie Brombach^a, Jessica Breuing^b, Marie Bolster^d, Cornelia Assion^e, Norbert Birkner^f, Eva Buchholz^{g,h}, Ernst-Günther Carlⁱ, Franziska Diel^j, Klaus Döbler^k, Markus Follmann^l, Timo Harfst^m, Monika Klinkhammer-Schalkeⁿ, Ina Kopp^d, Burkhard Lebert^o, Dagmar Lühmann^p, Claudia Meiling^q, Tim Niehues^r, Thomas Petzold^s, Susanne Schorr^t, Reina Tholen^u, Simone Wesselmann^v, Karen Voigt^{w,x}, Gerald Willms^y, Edmund Neugebauer^{b,z}, Dawid Pieper^b, Monika Nothacker^{d,1}, Jochen Schmitt^{a,1}

^a Zentrum für Evidenzbasierte Gesundheitsversorgung (ZEGV), Universitätsklinikum und Medizinische Fakultät Carl Gustav Carus an der Technischen Universität Dresden, Dresden, Deutschland

^b IFOM – Institut für Forschung in der Operativen Medizin, Department für Humanmedizin Universität Witten/Herdecke, Köln, Deutschland

^c Institut für Versorgungsforschung und Klinische Epidemiologie, Fachbereich Medizin, Philipps-Universität Marburg, Marburg, Deutschland

^d AWMF-Institut für Medizinisches Wissensmanagement (AWMF-IMWi), c/o Philipps-Universität, Marburg, Deutschland

^e Bundesministerium für Gesundheit (BMG), Referat 214 – Qualitätssicherung, Evidenzbasierte Medizin, Berlin, Deutschland

^f BQS Institut für Qualität & Patientensicherheit, Hamburg, Deutschland

^g Interessenvertretung Selbstbestimmt Leben in Deutschland e.V. (ISL), Berlin, Deutschland

^h Zentrum für Versorgungsforschung Brandenburg (ZVF-BB), Medizinische Hochschule Brandenburg Theodor Fontane, c/o Immanuel Klinik Rüdersdorf, Rüdersdorf, Deutschland

ⁱ Haus der Krebs-Selbsthilfe – Bundesverband e.V., Bonn, Deutschland

^j Kassenärztliche Bundesvereinigung (KBV), Dezernat Versorgungsqualität, Berlin, Deutschland

^k Kompetenzzentrum Qualitätssicherung / Qualitätsmanagement (KCQ), MDK Baden-Württemberg, Stuttgart, Deutschland

^l Deutsche Krebsgesellschaft e.V., Leitlinienprogramm Onkologie, Berlin, Deutschland

^m Bundespsychotherapeutenkammer, Berlin, Deutschland

ⁿ Arbeitsgemeinschaft Deutscher Tumorzentren e.V., Berlin, Deutschland

^o Frauenselbsthilfe Krebs – Bundesverband e.V., Bonn, Deutschland

^p Institut und Poliklinik für Allgemeinmedizin, Zentrum für Psychosoziale Medizin, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Deutschland

^q Deutscher Verband der Ergotherapeuten e.V., Referat Standards und Qualität, Karlsbad, Deutschland

^r Helios Klinikum Krefeld, Zentrum für Kinder- und Jugendmedizin, Krefeld, Deutschland

^s Gesellschaft für Qualitätsmanagement in der Gesundheitsversorgung e.V. (GQMG), Köln, Deutschland

^t Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Berlin, Deutschland

^u Deutscher Verband für Physiotherapie (ZVK) e.V., Köln, Deutschland

^v Deutsche Krebsgesellschaft e.V., Zertifizierung, Berlin, Deutschland

^w Bereich Allgemeinmedizin/MK3, Medizinische Fakultät Carl Gustav Carus, Technische Universität Dresden, Dresden, Deutschland

^x Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM), Berlin, Deutschland

^y aQua – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen, Göttingen, Deutschland

^z Medizinische Hochschule Brandenburg – Theodor Fontane, Neuruppin, Deutschland

* Korrespondenzadresse: Stefanie Deckert, MPH, Zentrum für Evidenzbasierte Gesundheitsversorgung (ZEGV), Universitätsklinikum Carl Gustav Carus an der Technischen Universität Dresden Fetscherstraße 74, 01307 Dresden, Deutschland.

E-mail: stefanie.deckert@uniklinikum-dresden.de (S. Deckert).

¹ geteilte Letztautorenschaft.

ARTIKEL INFO

Artikel-Historie:

Eingegangen: 27. August 2020
 Revision eingegangen: 11. November 2020
 Akzeptiert: 23. November 2020
 Online gestellt: 20. Januar 2021

Schlüsselwörter:

Versorgungsqualität
 Qualitätsindikator
 Leitlinie
 Konsens
 Evidenzbasiert

ZUSAMMENFASSUNG

Hintergrund: Empfehlungen aus S3-Leitlinien (S3-LL) stellen eine wichtige Grundlage für die Entwicklung von Qualitätsindikatoren (QI) dar. Ein einheitliches methodisches Vorgehen zur Entwicklung von QI im Kontext von S3-LL existiert in Deutschland bisher nicht. Aus diesem Grund wurde ein methodischer Standard für die leitlinienbasierte Entwicklung von QI (QI-Standard) anhand eines strukturierten Konsensprozesses unter Einbezug der relevanten Experten- und Interessengruppen erarbeitet.

Methoden: Die Inhalte des QI-Standards wurden evidenzgeleitet zusammengestellt, basierend auf den Ergebnissen von Übersichtsarbeiten und qualitativen Studien sowie unter Berücksichtigung der bereits existierenden Manuale zur leitlinienbasierten QI-Entwicklung. Für die Abstimmung des QI-Standards wurde ein multiperspektivisches Panel mit breiter Repräsentanz von Experten- und Interessengruppen aus dem deutschen Gesundheitssystem und Expertise in Leitlinienarbeit und/oder Qualitätsmanagement nominiert. Das iterative, strukturierte Konsensverfahren beinhaltete eine zweistufige Onlinebefragung angelehnt an das Delphi-Verfahren („vorläufige Abstimmung“) sowie eine moderierte Abschlusskonferenz, bei der diejenigen Empfehlungen in den QI-Standard aufgenommen wurden, die eine Zustimmung von >75% (Konsens-Kriterium) des Expertenpanels erhielten.

Ergebnisse: Der erarbeitete QI-Standard sieht als QI-Entwicklungsschritte zuerst die kriteriengestützte Auswahl „potenzieller QI“ vor, die nach ihrer Verabschiedung als „vorläufige QI“ in der S3-LL veröffentlicht werden und schließlich nach erfolgreicher Testung den Status „finale QI“ erlangen. Der QI-Standard setzt sich aus insgesamt 30 Empfehlungen zusammen, die sich auf sechs Bereiche verteilen: A) Vorbereitende Arbeitsschritte für die leitlinienbasierte Empfehlung von QI, B) Zusammensetzung der QI-Entwicklergruppe und Zusammenarbeit mit der LL-Gruppe, C) Entwicklung der potenziellen QI, D) Bewertung der potenziellen QI, E) Verabschiedung und Publikation sowie F) Pilotierung/Testung der vorläufigen und Überführung in finale QI.

Diskussion: Bevor der QI-Standard bei künftigen S3-LL-Erstellungen oder Aktualisierungen zur Implementierung empfohlen werden kann, sollte dieser in ausgewählten S3-Leitlinienprojekten erfolgreich erprobt worden sein. Neben methodischen Anforderungen an die QI-Entwicklung muss sichergestellt sein, dass den LL-Gruppen adäquate Ressourcen für die Umsetzung des QI-Standards zur Verfügung stehen.

Schlussfolgerung: Vom Einsatz des QI-Standards durch Leitliniengruppen können wissenschaftlich fundierte und für die Versorgung relevante QI erwartet werden.

ARTICLE INFO

Article History:

Received: 27 August 2020
 Received in revised form:
 11 November 2020
 Accepted: 23 November 2020
 Available online: 20 January 2021

Keywords:

Quality of healthcare
 Quality assurance
 Quality indicator
 Clinical practice guideline
 Consensus
 Evidence-based

ABSTRACT

Background: Recommendations of evidence- and formally consensus-based clinical practice guidelines (CPGs) represent a valuable source of quality indicators (QIs). Nevertheless, a standardized methodological procedure for developing QIs in the context of CPGs does not yet exist in Germany for all CPGs. For this reason, a methodological standard for the guideline-based development of QIs (QI Standard) was developed based on a structured consensus process involving multiple key stakeholders.

Methods: The proposed content of the QI Standard was derived from evidence, drawing upon results of reviews and qualitative studies, and considered German manuals for guideline-based QI development of two guideline programs. A multi-perspective consensus panel, broadly representing key stakeholders from the German healthcare system with expertise in CPGs and/or quality management, was nominated to vote on recommendations for guideline-based development of QIs. The iterative, structured consensus process included a two-stage online survey based on the Delphi method (“preliminary voting”) and a moderated final stakeholder conference where all those recommendations were definitely included in the QI Standard that received approval of more than 75 % (consensus criterion) of the consensus panel.

Results: Based on the agreed QI Standard, the QI development process starts with a criteria-based selection of „potential“ QIs which – in case of adoption – are published in CPGs as „preliminary“ QIs and can achieve the status “final” after successful testing. The QI Standard is composed of a total of 30 recommendations, which are allocated to six areas: A) preparatory work steps for the guideline-based recommendation of QIs, B) QI development group and cooperation with the CPG group, C) development of potential QIs, D) critical appraisal of potential QIs, E) formal adoption and publication as well as F) piloting/testing of preliminary QIs and conversion into final QIs.

Discussion: Before the QI Standard can be recommended for implementation in future CPGs, it should have been successfully tested in selected German CPG projects. In addition to methodological requirements for the QI development, it must be ensured that guideline groups have adequate resources for the implementation of the QI Standard.

Conclusion: By using the QI Standard, scientifically sound and healthcare-relevant QIs can be expected.

Einleitung

Mit Blick auf eine qualitativ hochwertige Gesundheitsversorgung wird Leitlinien (LL) eine große Bedeutung beigemessen [1]. Aufgrund der evidenz- und strukturiert konsensbasierten Entwicklungsmethodik sind S3-LL als methodischer Goldstandard etabliert und stellen eine wichtige Grundlage für die Entwicklung

von Qualitätsindikatoren (QI) dar. Aus LL-Empfehlungen abgeleitete QI, sog. leitlinienbasierte QI, ermöglichen als quantitatives Maß die Beurteilung des Umsetzungsgrades von Empfehlungen der S3-Leitliniengruppen und schaffen damit eine Basis für die Bewertung des Verbesserungspotenzials der medizinischen und pflegerischen Versorgungsqualität. Aufgrund dieser zentralen Rolle müssen solche QI hohe Anforderungen bezüglich ihrer

Entwicklungsmethodik inklusive der Prüfung ihrer methodischen Eigenschaften erfüllen [2,3].

Trotz der hohen Relevanz von S3-LL für die medizinische Versorgung spielen QI im Kontext von S3-LL bisher nur in ausgewählten Bereichen eine Rolle. Für Leitlinienautor*innen der geförderten Programme Nationale VersorgungsLeitlinien (NVL-Programm) [4] und Leitlinienprogramm Onkologie (OL-Programm) [5] ist die Ableitung und Verabschiedung leitlinienbasierter QI seit 2005 bzw. 2008 obligatorischer Bestandteil der LL-Erstellung bzw. -Aktualisierung. In der Planung und Umsetzung von S3-LL außerhalb der geförderten Programme (LL von Fachgesellschaften (FG) der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF): FG-Programm) besteht diese Verpflichtung nicht. Eine vergleichende Gegenüberstellung der derzeitigen QI-Entwicklungsprozesse in Abhängigkeit des jeweiligen Leitlinienprogramms ist in [Abbildung 1](#) dargestellt und verdeutlicht, dass ein für Deutschland einheitlich geltendes methodisches Vorgehen zur Entwicklung von leitlinienbasierten QI bisher nicht vorliegt.

Auch im internationalen Vergleich existiert bisher kein methodischer Goldstandard zur spezifischen Entwicklung von leitlinienbasierten QI [6]. Von der Performance Measures Working Group des Guidelines International Network (G-I-N) wurde 2016 in einem ersten Schritt ein Standard für die Berichtsqualität von leitlinienbasierten QI publiziert, der implizit methodische Anforderungskriterien an den QI-Entwicklungsprozess beinhaltet [7].

Ziel des von der Deutschen Forschungsgemeinschaft (DFG) geförderten Forschungsprojektes war es, einen evidenzgeleiteten und konsensbasierten Standard für die Entwicklung von leitlinienbasierten QI (QI-Standard) anhand eines strukturierten Konsensprozesses unter Einbezug der relevanten Experten- und Interessengruppen zu verabschieden.

Der strukturierten Konsensfindung ging ein für die Erarbeitung der Evidenzbasis des QI-Standards konzipiertes Mixed-Methods-Forschungsprojekt voraus, indem Elemente einer systematischen Evidenzsynthese (drei systematische Reviews) mit denen eines qualitativen Forschungsansatzes (zwei qualitative Studien) kombiniert wurden. Anhand dieser fünf Teilstudien wurde sowohl auf nationaler als auch auf internationaler Ebene untersucht, wie leitlinienbasierte QI entwickelt werden, welche Faktoren in diesem Zusammenhang förderlich bzw. hinderlich sind und welche methodischen Anforderungen generell an QI zu stellen sind.

Eine Übersicht der methodischen Vorgehensweisen sowie der Kernergebnisse der fünf Teilstudien sind im [Appendix A, Online Anlage 1](#) beschrieben. Eine detailliertere Darstellung findet sich in den Veröffentlichungen [8–11]. Mit Blick auf Deutschland stellte sich heraus, dass der Anteil der S3-LL mit QI-Empfehlungen sowie die Berichtsqualität der für die QI-Entwicklung angewandten Methodik zwischen den Leitlinienprogrammen NVL/OL und FG stark variiert [8] und damit eine generelle Überprüfbarkeit der leitlinienkonformen Versorgung inklusive des Zielerreichungsgrades derzeit stark vom Leitlinienthema abhängig ist. Neben einer personellen und finanziellen Unterstützung sind aus Sicht von S3-Leitlinienautor*innen vor allem die methodische Begleitung und die Implementierung (bspw. in etablierte Qualitätssicherungsverfahren) der erarbeiteten QI-Vorschläge wichtig, um deren Entwicklung und Einsatz zu fördern [9]. Auch international zeigte sich anhand einer Interviewstudie, dass der QI-Entwicklungsprozess durch eine methodische Begleitung und die Berücksichtigung von bereits bestehenden Qualitätsverbesserungsinitiativen unterstützt werden kann [10]. Der systematische Vergleich von S3-LL und inhaltlich vergleichbaren internationalen LL mit QI machte deutlich, dass die abgeleiteten QI national und international vielfach nicht vergleichbar sind. Aufgrund der ungenügenden Berichtsqualität konnten keine eindeutigen Gründe hierfür identifiziert werden [11]. Mit Hilfe eines systematischen

Reviews zu Instrumenten der methodischen Bewertung von QI konnte zudem festgestellt werden, dass international kein „Best Practice Verfahren“ zur QI-Bewertung gefunden wurde, welches im Leitlinienkontext genutzt werden könnte, da unterschiedliche Bewertungskriterien angewendet werden und aussagekräftige Evaluationsstudien fehlen.

Der sich auf Grundlage der Evidenzbasis anschließende Entwicklungsprozess mit Fokus auf die strukturierte Konsensfindung sowie die Darstellung der verabschiedeten Inhalte des QI-Standards sind Gegenstand des vorliegenden Beitrages.

Methoden

Die im Rahmen des Mixed-Methods-Forschungsprojektes erarbeitete Evidenzbasis [8–11] wurde bei einer Initialkonferenz unter Beteiligung eines multiperspektivischen Expertenpanels präsentiert und diskutiert. Das Studienteam erstellte im Anschluss eine Evidenzsynthese aus den Ergebnissen dieser Teilstudien und den in Deutschland bereits existierenden Manualen zur leitlinienbasierten QI-Entwicklung [4,5] und leitete daraus Vorschläge für die leitlinienbasierte Entwicklung von QI ab. Mit dem Ziel, diese national breit abzustimmen, wurde unter erneuter Beteiligung des Expertenpanels für das anschließende strukturierte, iterative Konsensverfahren eine zweistufige Onlinebefragung angelehnt an das Delphi-Verfahren („vorläufige Abstimmung“) sowie eine moderierte Abschlusskonferenz (finale Abstimmung) durchgeführt.

[Abbildung 2](#) stellt den gesamten Entwicklungsprozess dar. Im Folgenden werden die Auswahlkriterien für die Rekrutierung des Expertenpanels, die methodische Vorgehensweise bei der Erarbeitung des QI-Standards sowie der Konsentierungsprozess beschrieben.

Auswahl und Rekrutierung des Expertenpanels

Um unterschiedliche Perspektiven und Expertisen bei der Erarbeitung und Abstimmung des QI-Standards berücksichtigen zu können, wurde ein multiperspektivisches Expertenpanel rekrutiert. Dieses sollte sich aus Repräsentant*innen relevanter Institutionen aus dem deutschen Gesundheitssystem mit Fokus auf LL und Qualitätsmanagement zusammensetzen. Dabei beteiligt werden sollten

- Kliniker*innen verschiedener Fachdisziplinen repräsentiert durch Medizinische Wissenschaftliche Fachgesellschaften,
- Methodiker*innen aus dem Bereich Qualitätsmessung,
- Vertreter*innen der Versorgungs- und Qualitätsforschung sowie der evidenzbasierten Medizin,
- nicht-ärztlich tätige Leistungserbringer*innen (Einschlusskriterium: Verbände etc., die bei der Erstellung von > 10 aktuell gültigen S3-LL beteiligt waren (Stand April 2017)),
- Vertreter*innen von Patientenvereinigungen und
- gesundheitspolitische Repräsentant*innen.

[Tabelle 1](#) gibt einen Überblick über die eingeladenen 29 Fachgesellschaften und Institutionen. Die Nominierung der Mandatsträger*innen erfolgte intern durch die jeweiligen Fachgesellschaften und Institutionen.

Evidenzsynthese

In Vorbereitung auf die zu erstellende Evidenzsynthese fand eine moderierte Initialkonferenz statt, mit dem Ziel, dem Expertenpanel die Studienziele, das vorgesehene methodische Vorgehen gemäß der Projektskizze und die bis dato vorliegenden Studienergebnisse zu präsentieren und im Dialog zentrale Diskussionspunkte und Anregungen aufzugreifen. Die Teilnehmerbeiträge

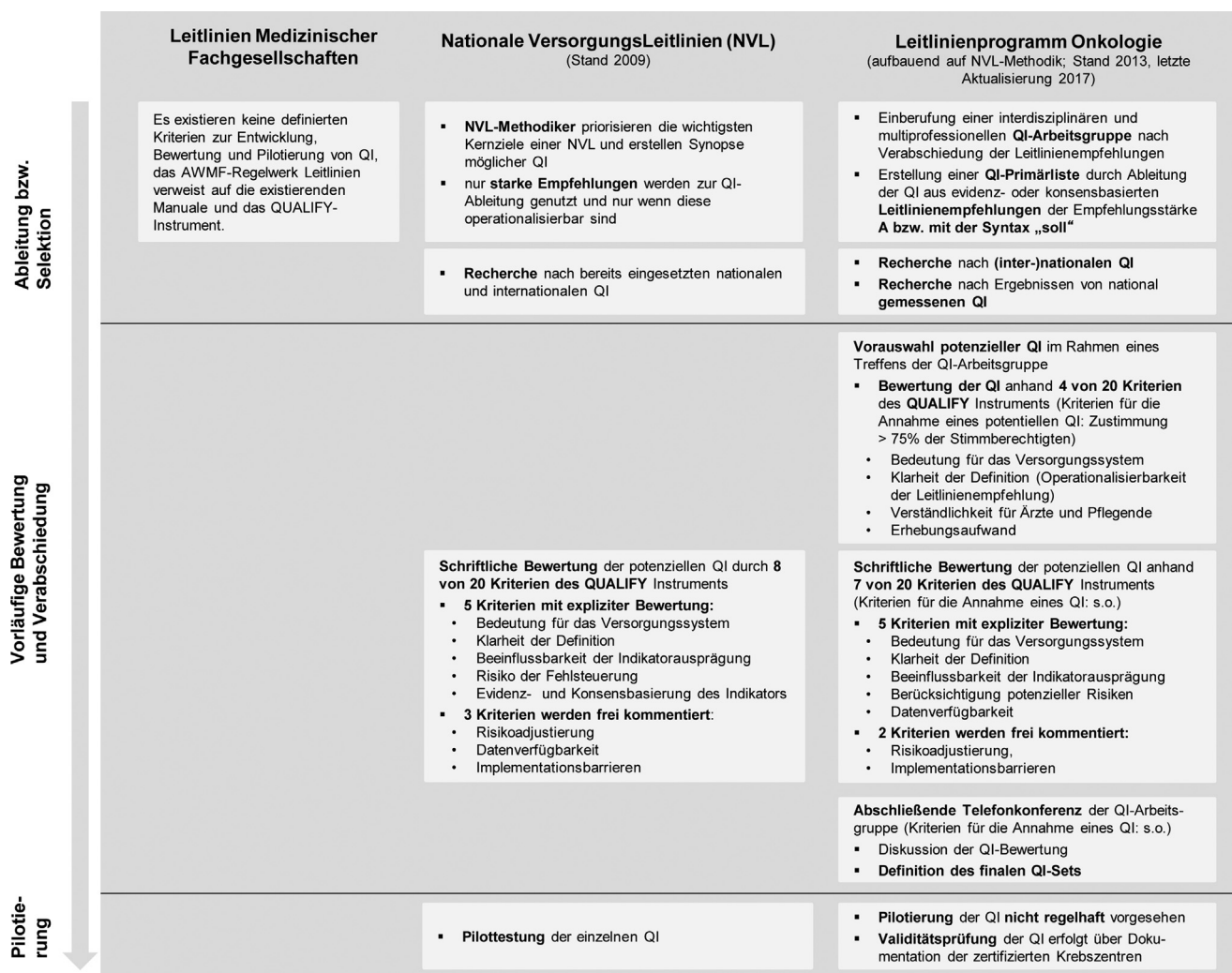


Abbildung 1. Vergleichende Gegenüberstellung des methodischen Vorgehens der QI-Entwicklung im Kontext von S3-LL in Abhängigkeit des jeweiligen Leitlinienprogramms (erweitert nach [8]). QI: Qualitätsindikator; QUALIFY: Instrument zur Bewertung von Qualitätsindikatoren.

wurden an einer Pinnwand mit Moderationskarten festgehalten und thematisch strukturiert. Die Initialkonferenz diente zudem der Verabschiedung von Verfahrens- (u.a. Stimmberechtigung) und Abstimmungsregeln (u.a. Konsens-Kriterium).

Im Nachgang der Initialkonferenz wurde anhand der existierenden Manuale für die leitlinienbasierte QI-Entwicklung [4,5] sowie der Auswertungen der Diskussionen mit dem Expertenpanel vom Studienteam zunächst ein aus drei zeitlichen Phasen (vor, während und nach der LL-Erstellung) und sechs Arbeitsschritten (Bereiche A – F) bestehendes Rahmenkonzept für den QI-Entwicklungsprozess im Kontext von S3-LL erarbeitet (Abb. 3).

Durch das Studienteam erfolgte im Anschluss eine Zusammenfassung der Ergebnisse der fünf Teilstudien und die Ableitung von möglichen Empfehlungen zur leitlinienbasierten QI-Entwicklung im Kontext des Rahmenkonzepts (Abb. 3) unter Berücksichtigung der vorliegenden QI-Manuale der Leitlinienprogramme. Bei inkonsistenten Studienergebnissen, wie z.B. den anzuwendenden QI-Bewertungskriterien (Appendix A, Online Anlage 1) wurden zunächst ausschließlich die bereits erprobten und bewährten Empfehlungen aus den QI-Manualen herangezogen. Wenn sich aus den Ergebnissen der Teilstudien keine konkreten Handlungsempfehlungen ableiten ließen, diese aber vom Studienteam und Expertenpanel im QI-Entwicklungsprozess als essenziell angesehen wurden, erfolgten entsprechend gekennzeichnete Vorschläge.

Der so entwickelte vorläufige leitlinienbasierte QI-Entwicklungsprozess umfasste die in Abb. 3 und Appendix A, Online Anlage 2 dargestellten jeweiligen Arbeitsschritte (Bereiche A bis F). Wesentliche Änderungen und zum Teil Erweiterungen verglichen mit den QI-Manualen des NVL- und OL-Programms (vgl. Abb. 1) werden im Folgenden pro Bereich kurz skizziert. Der konkrete Wortlaut der initial vorgeschlagenen Empfehlungen ist im Appendix A, Online Anlage 2 einsehbar.

Bereich A: Vorbereitende Arbeitsschritte für die leitlinienbasierte Empfehlung von QI

Damit LL-Gruppen eine informierte Entscheidung (Wissen über Konzept, Verwendungszwecke und generelle Entwicklungsschritte von QI) für bzw. gegen die Entwicklung von QI zu Beginn einer LL-Erstellung bzw. -Aktualisierung treffen können, wurde

- das Hinzuziehen einer Person mit Erfahrungen im Bereich Qualitätsmessung, die zu Beginn der LL-Erstellung die Leitliniengruppe über QI informieren soll (Empfehlung A1), vorgeschlagen.

Bereich B: Zusammensetzung der QI-Entwicklergruppe und Zusammenarbeit mit der LL-Gruppe

- Für die Schaffung eines gemeinsamen Grundverständnisses über die QI-Entwicklungsschritte wurde das Anbieten einer Schulung

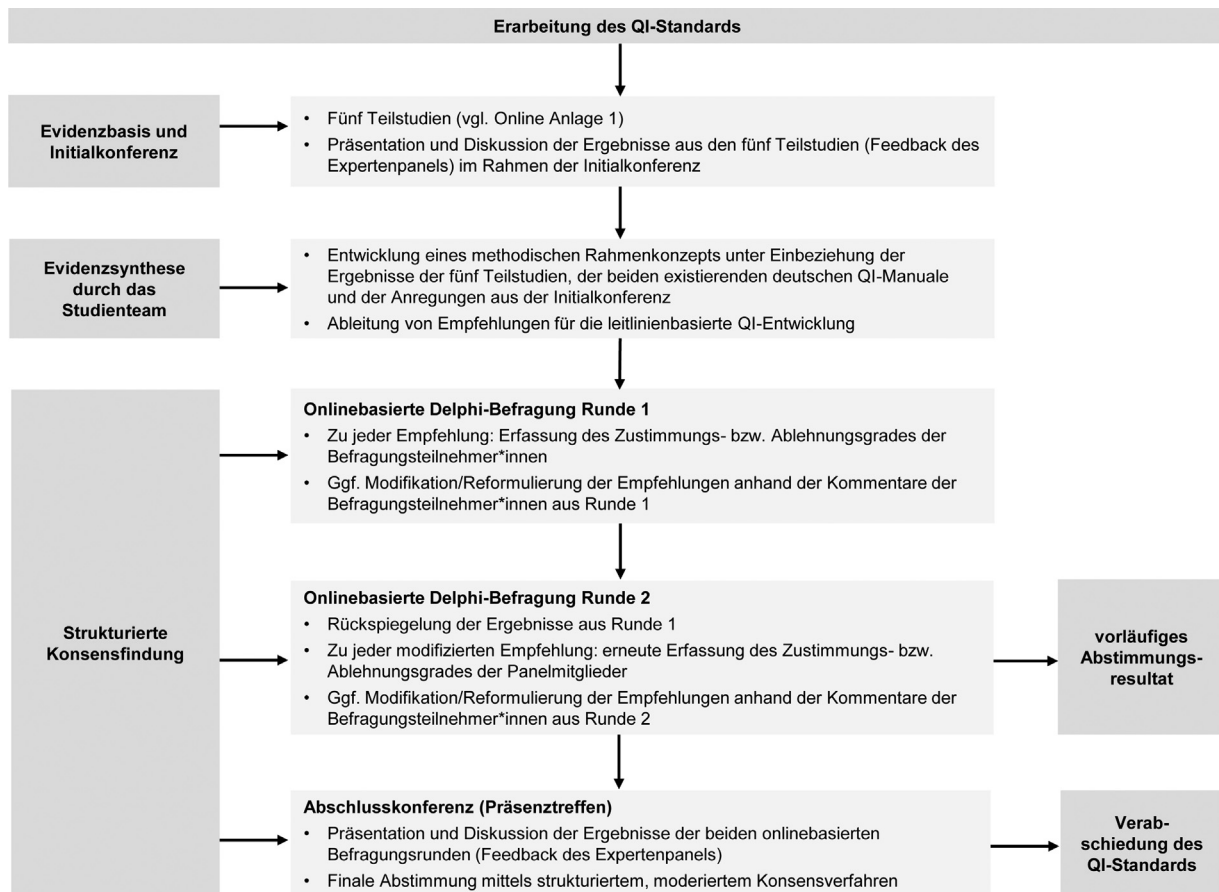


Abbildung 2. Vorgehen bei der Erarbeitung des QI-Standards.

zum Thema QI, insbesondere zur spezifischen Entwicklungsmethodik, welche alle Mitglieder der QI-Entwicklergruppe erhalten sollen, empfohlen (Empfehlung B5)

- Um ein aufeinander abgestimmtes Vorgehen beim Erarbeiten von LL-Empfehlungen und darauf basierenden QI zu erzielen, wurde empfohlen, dass die LL-Gruppe und die QI-Entwicklergruppe zeitlich und personell eng aufeinander abgestimmt zusammenarbeiten soll (Empfehlung B6).

Bereich C: Entwicklung der potenziellen QI

Eine grundlegende Änderung der gegenwärtig angewendeten QI-Entwicklungsprozesse [4,5] beinhaltet das spezifische Vorgehen der QI-Ableitung aus LL-Empfehlungen. Während lt. NVL- und OL-Programm ausschließlich starke Empfehlungen bzw. Empfehlungen des Empfehlungsgrades A zugrunde gelegt werden sollen, wurde für den QI-Standard ein iteratives Vorgehen empfohlen, welches vorsieht, dass

- im Prozess zunächst jede LL-Empfehlung unabhängig vom Empfehlungsgrad in einen oder mehrere potenzielle QI überführt werden soll (Empfehlung C1). Intention dieses Schrittes ist es, den LL-Gruppen zu ermöglichen, die grundsätzliche Abbildbarkeit der formulierten LL-Empfehlungen und damit in der Gesamtschau auch die grundsätzliche Abbildbarkeit der Leitlinienadhärenz kritisch zu reflektieren, da ausschließlich klare, eindeutig formulierte LL-Empfehlungen eine Grundvoraussetzung für die Ableitung von operationalisierbaren QI sind (Empfehlung C1).
- einzelne QI, die aufgrund einer anfangs mangelhaften Operationalisierbarkeit nicht aus LL-Empfehlungen formuliert werden können, nicht per se zum Ausschluss aus der Liste der potenziellen QI führen sollen, sondern eine erneute Überprüfung und ggf.

Neuformulierung der vorgeschlagenen LL-Empfehlungen vorgenommen werden soll. Dieser zu wiederholende iterative Prozess sollte gemeinsam von der LL-Gruppe und QI-Entwicklergruppe vorgenommen werden (Empfehlung C2).

- Für diejenigen Empfehlungen sollen prioritär potenzielle QI weiterentwickelt werden, bei deren Umsetzung - bezogen auf die Versorgungsziele der LL - der größte Gesundheitsgewinn für die Patientenpopulation erwartet wird (Empfehlung C6).
- Ob einzelne QI aus der Liste der potenziellen QI übernommen und weiterentwickelt werden, soll zunächst unabhängig von der aktuellen Datenverfügbarkeit entschieden werden (Empfehlung C7).
- Die Zweckbestimmung(en) der potenziellen QI soll(en) konkret genannt werden. Dies umfasst auch, bestimmte Einsatzzwecke explizit auszuschließen (Empfehlung C8).

Nach Abschluss der Formulierung von aus den LL-Empfehlungen inhaltlich und methodisch gut ableitbaren „potenziellen QI“ erfolgt im nächsten Schritt die Weiterentwicklung der QI hin zu „vorläufigen QI“ durch die QI-Entwicklergruppe.

Bereich D: Bewertung der potenziellen QI

Zunächst wurde über die bisher im NVL- und OL-Programm eingesetzten Bewertungskriterien des QUALIFY Instruments [12] abgestimmt.

Bereich E: Verabschiedung und Publikation der vorläufigen QI

Da vorläufige QI Bestandteil der veröffentlichten S3-LL sein werden und diese in der Breite akzeptiert werden sollen, wird es als wichtig erachtet, dass sowohl Mitglieder der stimmberechtigten

Tabelle 1

Übersicht der beteiligten Experten- und Interessengruppen (in alphabetischer Reihenfolge) innerhalb der Initialkonferenz und einzelnen Konsentierungsphasen.

	Beteiligte Institutionen/Fachgesellschaften ¹	Initialkonferenz (11/2017)	Onlinebasierte Delphi- Befragungen		Abschlusskonferenz (04/2018)	
			Runde 1 (02-03/2018)	Runde 2 (03-04/2018)	Tag 1	Tag 2
1	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF)	●	●	●	●	●
2	Ärztliches Zentrum für Qualität in der Medizin (ÄZQ)	●	●	●	●	●
3	Bundesärztekammer (BÄK)	●	●	●	●	●
4	Bundespsychotherapeutenkammer (BPTK)	●	●	●	●	● ²
5	Bundesministerium für Gesundheit, Referat Qualitätssicherung und Evidenzbasierte Medizin (BMG)	●	●	●	●	●
6	Deutsche Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM)	●	●	●	●	●
7	Deutsche Gesellschaft für Anästhesiologie und Intensivmedizin (DGAI)	●	●	●	●	●
8	Deutsche Gesellschaft für Chirurgie (DGCH)	●	●	●	●	●
9	Deutsche Gesellschaft für Kinder- und Jugendmedizin (DGKJ)	●	●	●	●	●
10	Deutsche Gesellschaft für Pflegewissenschaft (DGP)	●	●	●	●	●
11	Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde (DGPPN)	●	●	●	●	●
12	Deutsche Krebsgesellschaft (DKG)	●	●	●	● ³	●
13	Deutsche Rentenversicherung Bund (DRV)	●	●	●	●	●
14	Deutscher Pflegerat	●	●	●	●	●
15	Deutscher Verband der Ergotherapeuten (DVE)	●	●	●	●	●
16	Deutscher Verband für Physiotherapie (ZVK)	●	●	●	●	●
17	Deutsches Netzwerk Evidenzbasierte Medizin (EbM-Netzwerk)	●	●	●	●	●
18	Deutsches Netzwerk Versorgungsforschung (DNVF)	●	●	●	●	● ⁴
19	Gesellschaft für Qualitätsmanagement im Gesundheitsmanagement (GQMG)	●	●	●	●	●
20	Kassenärztliche Bundesvereinigung (KBV)	●	●	●	●	●
21	Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen (aQua-Institut)	●	●	●	●	●
22	Institut für Qualität & Patientensicherheit (BQS)	●	●	●	●	●
23	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG)	●	●	●	●	●
24	Kompetenz-Centrum Qualitätssicherung und Qualitätsmanagement des MDK (KCQ)	●	●	●	●	●
25	Leitlinienprogramm Onkologie	●	●	●	●	●
26	Institutionsloser Vertreter	●	●	●	●	●
27	Patientenvertretung Gemeinsamer	●	●	●	●	●
28	Bundesausschuss (G-BA)	●	●	●	●	●
29		●	●	●	●	●
30		●	●	●	●	●
	Beteiligungsquote [in %]	77	100	97	70	70

Anmerkungen: ¹Von insgesamt 29 kontaktierten Fachgesellschaften und Institutionen beteiligten sich 27 an dem strukturierten Konsensprozess. Lediglich Vertreter*innen des Gemeinsamen Bundesausschusses (G-BA) sowie der Deutschen Gesellschaft für Innere Medizin (DGIM) nahmen nicht teil. Nach der Initialkonferenz traten weiterhin zwei Patientenvertreter*innen aus persönlichen Gründen von ihrem Mandat zurück (Kontaktstelle für Selbsthilfegruppen und Deutsche Diabetes Hilfe). Mandatsübertragungen an: ²KBV, ³Leitlinienprogramm Onkologie, ⁴BMG

LL-Gruppe als auch der QI-Entwicklergruppe diese verabschieden (Empfehlung E1). Bisher erfolgt dies lt. QI-Manualen ausschließlich durch die QI-Entwicklergruppe.

In S3-LL veröffentlichte „vorläufige QI“ können nach einer erfolgreichen Pilotierung/Testung den Status „finale QI“ erlangen.

Bereich F: Pilotierung/Testung der vorläufigen QI

- Bevor leitlinienbasierte QI regelhaft eingesetzt werden, soll eine Pilotierung/Testung in ausgewählten Versorgungsbereichen erfolgen (Empfehlung F1). Eine Pilotierung/Testung der QI, d.h. datenbasierte Analyse einzelner methodischer Eigenschaften in ausgewählten Settings, ist im OL-Programm derzeit nicht regelhaft vorgesehen.

Strukturierte Konsensfindung

Ziel der strukturierten Konsensfindung war es, anhand eines iterativen Prozesses individuelle Meinungen zu den durch das Studienteam vorgeschlagenen leitlinienbasierten QI-Entwicklungsschritten in eine Gruppenmeinung zu überführen, um damit die Inhalte des QI-Standards zu konsentieren und zu verabschieden. Der Konsensprozess zur Verabschiedung des QI-Standards erfolgte in drei Schritten (vgl. Abb. 2):

Onlinebefragungen (Februar bis April 2018)

Es erfolgten zwei anonymisierte, internetbasierte Befragungsrunden des Expertenpanels auf der Grundlage der „RAND/UCLA

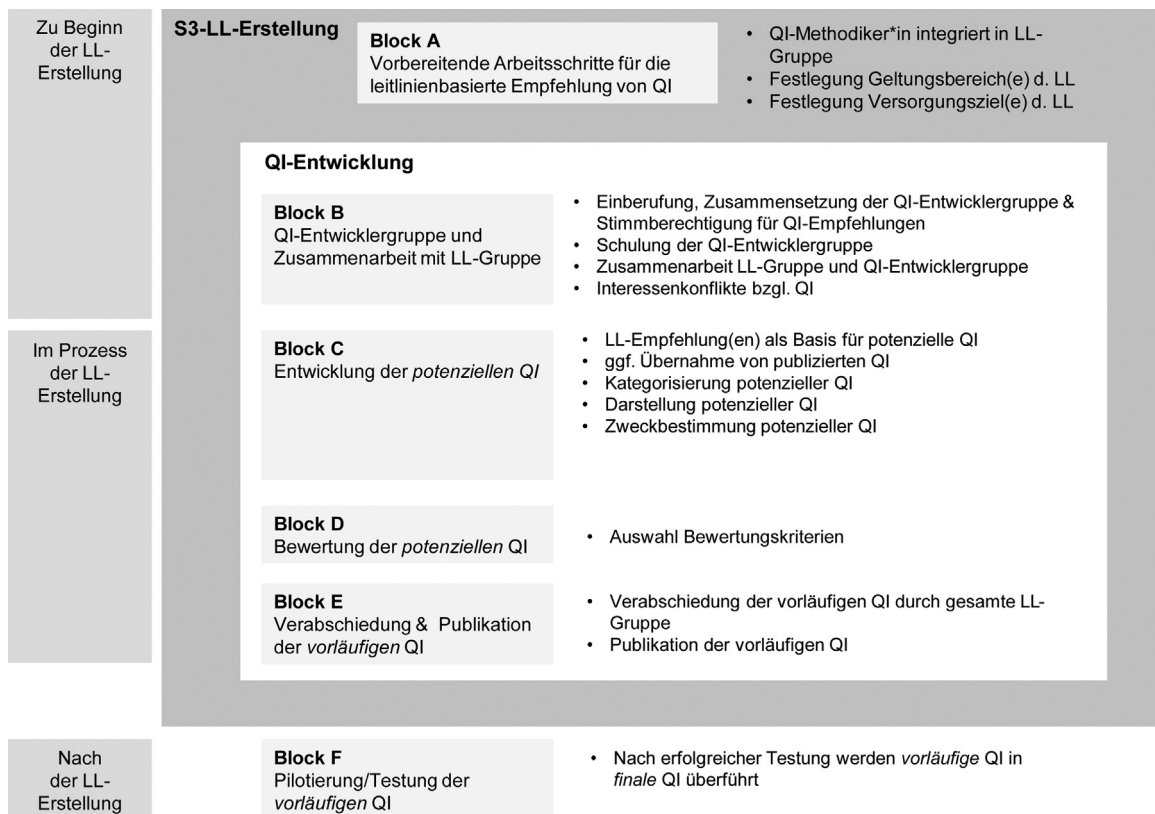


Abbildung 3. Methodisches Rahmenkonzept des QI-Standards.

Anmerkung: Der Bereich A fokussiert zum Teil auf Aspekte, die bereits Gegenstand des AWMF-Regelwerks Leitlinien sind [13]. Auch wenn deren Berücksichtigung im Rahmen der S3-LL-Erstellung vorausgesetzt werden kann, sollten diese Inhalte aus Sicht des Studienteams und Expertenpanels dezidiert im QI-Standard ausgewiesen sein, da sie als zentrale Grundlage für die QI-Entwicklung angesehen werden.

„*Appropriateness Method*“ für Delphi-Verfahren [14]. Zu jeder Empfehlung konnte pro Runde der Grad der Zustimmung auf einer 9-Punkte-Skala (1 bis 3 = „stimme nicht zu“; 4 bis 6 = „stimme teilweise zu“; 7 bis 9 = „stimme zu“) zum Ausdruck gebracht werden. Es bestand außerdem die Möglichkeit der Enthaltung. Bei Enthaltungen oder Gegenstimmen (Skalenbereich 1 bis 3) wurden die Teilnehmer*innen gebeten, dies jeweils zu begründen und Änderungsvorschläge zu geben. In Abweichung von der Originalmethode [14] galt als Konsens-Kriterium eine Zustimmung (Skalenbereich 7 bis 9) von >75% der Befragungsteilnehmer*innen pro Empfehlung, wobei die onlinebasierten Abstimmungen als vorläufig galten. Jede*r Teilnehmer*in erhielt individuelle Zugangsdaten für das Bearbeiten der Onlinebefragungen und ggf. pro Runde zwei schriftliche Erinnerungen.

Onlinebasierte Delphi-Befragung Runde 1

Der Fokus der ersten Delphi-Befragung lag darauf, über die vom Studienteam entworfenen Empfehlungen zur leitlinienbasierten QI-Entwicklung abzustimmen und ggf. weitere bisher nicht genannte QI-Entwicklungsschritte zu ergänzen. Um zu gewährleisten, dass das Expertenpanel eine informierte Abstimmung vornehmen konnte, wurde jeder Empfehlung eine knappe Zusammenfassung der Evidenzbasis aus den jeweiligen Teilprojekten und Inhalten aus bestehenden QI-Manualen [4,5] vorangestellt. Falls dies aufgrund fehlender Daten nicht möglich war, fasste das Studienteam seine Einschätzungen und Begründungen bzgl. der Notwendigkeit und Wichtigkeit der entsprechenden Empfehlungen zusammen. Zur besseren Nachvollziehbarkeit des Vorgehens sind in der [Appendix A, Online Anlage 3](#) Beispiele aufgeführt.

Onlinebasierte Delphi-Befragung Runde 2

Basierend auf Runde 1 erhielt jede*r Befragungsteilnehmer*in in Anlehnung an die etablierte Methodik des Delphi-Verfahrens [14] pro Empfehlung Feedback über das individuelle sowie gruppenbezogene Abstimmungsergebnis in Form von Balkendiagrammen. Zusätzlich wurde je Empfehlung eine vom Studienteam erstellte Zusammenfassung der Kommentare der Befragungsteilnehmer*innen zur Verfügung gestellt ([Appendix A, Online Anlage 3](#)). Danach wurde jedem*r Teilnehmer*in ermöglicht, ihre*seine Bewertungen aus der vorherigen Welle zu bestätigen oder auf Basis der abgebildeten vorläufigen Gruppenmeinung zu revidieren.

In Abhängigkeit der durch die Befragungsteilnehmer*innen aufgeführten Kommentare aus Runde 1 wurden die durch das Studienteam spezifizierten bzw. modifizierten Empfehlungen innerhalb der zweiten Welle erneut abgestimmt ([Appendix A, Online Anlage 2](#)).

Zweitägige Abschlusskonferenz (April 2018)

Ziel der moderierten Abschlusskonferenz war die finale Abstimmung und Verabschiedung des QI-Standards. Grundsätzlich war eine Beteiligung an der Abschlusskonferenz nur für diejenigen Mandatsträger*innen möglich, die bei den vorangegangenen onlinebasierten Abstimmungen teilgenommen hatten. Zusätzlich war eine persönliche Teilnahme Voraussetzung für die finale Abstimmung während der Abschlusskonferenz; eine Mandatsübertragung war jedoch möglich.

Die Ergebnisvorstellung der Onlinebefragungen (vorläufiger Konsens vs. Dissens) sowie themenbezogene Impulsvorträge erfolgten durch das Studienteam. Schrittweise wurden im Anschluss die einzelnen Empfehlungen in moderierten

Tabelle 2
Charakteristik des Expertenpanels (n = 30).

Kategorien	Anzahl (%)
<i>Alter</i>	
30-39 Jahre	3 (10)
40-49 Jahre	5 (17)
50-59 Jahre	15 (50)
60-69 Jahre	5 (17)
70-79 Jahre	2 (6)
<i>Geschlecht</i>	
männlich	16 (53)
weiblich	14 (47)
<i>Rolle innerhalb des Expertenpanels¹</i>	
Ärztliches Personal	4 (13)
Nicht-ärztliches medizinisches Personal	3 (10)
Patientenvertreter*innen	4 (13)
Gesundheitspolitische Vertreter*innen	2 (7)
Methodiker*innen im Bereich Qualitätsmanagement	11 (37)
Vertreter*innen Versorgungs- /Qualitätsforschung	6 (20)
<i>Erfahrungen mit S3-Leitlinien</i>	
Ja	19 (63)
Nein	11 (37)
<i>Erfahrungen mit leitlinienbasierten QI</i>	
Ja	21 (70)
Entwicklung ²	8 (38)
Anwendung	4 (19)
Entwicklung ² und Anwendung	9 (43)
Nein	9 (30)
<i>Erfahrungen mit QI außerhalb von Leitlinienprojekten</i>	
Ja	19 (63)
Entwicklung ²	0 (0)
Anwendung	2 (10)
Entwicklung ² und Anwendung	17 (90)
Nein	11 (37)

Anmerkung:

¹ Zuordnung erfolgte durch die teilnehmenden Expert*innen;

² inklusive Testung, Evaluation und Implementierung; QI - Qualitätsindikatoren

Kleingruppen- und Paneldiskussionen diskutiert und final abgestimmt. Als Konsens-Kriterium galt auch hier jeweils eine Zustimmung von >75% der stimmberechtigten Expert*innen pro Empfehlung.

Ergebnisse

Zusammensetzung des Expertenpanels und Teilnahmequoten

Insgesamt 27 der 29 kontaktierten Fachgesellschaften und Institutionen beteiligten sich mit insgesamt 30 Mandatsträger*innen (vgl. Tab. 1) an der strukturierten Konsensfindung. Neben Vertreter*innen des NVL- und OL-Programms, der AWMF und Mandatsträger*innen wissenschaftlicher medizinischer Fachgesellschaften beteiligten sich Vertreter*innen gesundheitspolitischer sowie im Bereich der Qualitätsförderung auch methodisch tätiger Institutionen sowie Patientenvertreter*innen. Eine Übersicht über die Zusammensetzung des Expertenpanels in den jeweiligen Schritten der strukturierten Konsensfindung findet sich in der [Tabelle 1](#). Die Beteiligungsquote der onlinebasierten Delphi-Befragungen und Abschlusskonferenz variierte zwischen 70 und 100%.

Charakteristische Merkmale des Expertenpanels sind in [Tabelle 2](#) aufgeführt. Die Hälfte des Entscheidungsgremiums war zwischen 50 und 59 Jahre alt. Erfahrungen mit leitlinienbasierten QI wiesen 70% der Expert*innen auf.

Onlinebasierte Delphi-Befragungen

Runde 1

In Runde 1 stimmten die Expert*innen über 30 Empfehlungen zur leitlinienbasierten QI-Entwicklung ab. Dabei variierten die vorläufigen Zustimmungsraten zwischen 20% und 97%. Ein vorläufiger Konsens wurde für 17 Empfehlungen erreicht, für acht weitere wurde der Konsens nur knapp verfehlt (Zustimmung von 73%) (vgl. [Appendix A, Online Anlage 2](#)).

Bei dem überwiegenden Teil der Empfehlungen des Themenschwerpunktes *Entwicklung der potenziellen QI* (Bereich C) wurde kein vorläufiger Konsens erreicht. Sechs der acht Empfehlungen erhielten keine ausreichende Zustimmung. Dies betraf vor allem Empfehlungen, die die QI-Ableitung als iterativen Prozess, der eine erneute Überprüfung der LL-Empfehlung in Bezug auf ihre Operationalisierbarkeit nach anfänglich negativer Einschätzung vorsah, adressierten und sich damit grundlegend vom gegenwärtigen QI-Entwicklungsprozess [4,5] unterschieden. Wesentliche, durch die Expert*innen kommentierte Gründe für die Ablehnung dieser Empfehlungen waren, dass die eingeschränkte Operationalisierbarkeit der QI einerseits auf Defizite in der Formulierung der LL-Empfehlung zurückzuführen sein könne, ggf. seien diese aber auch bei optimaler Formulierung nicht praxistauglich operationalisierbar. Primär gehe es bei LL darum, Empfehlungen für die Versorgung zu formulieren und nicht für die Messung der Versorgungsqualität. Die Möglichkeit der Operationalisierbarkeit der QI könne nicht definieren, was gute Versorgung ist (Empfehlung C2, Zustimmung 40%). QI unabhängig von Datenverfügbarkeit zu erarbeiten sei Verschwendung von Ressourcen und nütze den Patient*innen nichts (Empfehlung C7, Zustimmung 50%). Über den iterativen Ansatz hinausgehend wurde angemerkt, dass durch den Einbezug aller LL-Empfehlungen, nicht nur der starken, der Empfehlungsgrad an Bedeutung verlieren würde und schwache Empfehlungen grundsätzlich nicht für die Beurteilung einer Qualitätsverbesserung geeignet seien (Empfehlung C1, Zustimmung 20%).

Runde 2

Zum einen gab es Empfehlungen, für die bereits eine hohe vorläufige Zustimmung (>75%) ohne relevante Verbesserungsvorschläge bestand, zum anderen Empfehlungen, bei denen sich ein hoher Diskussionsbedarf zeigte, wie bspw. bei der Auswahl der QUALIFY Bewertungskriterien [12]. Neun dieser Empfehlungen wurden nicht erneut abgestimmt, sondern als Themenschwerpunkte für die Abschlusskonferenz vorgesehen (vgl. [Appendix A, Online Anlage 2](#)).

Die Expert*innen stimmten in Runde 2 erneut über 21 Empfehlungen ab, die entweder im Vergleich zur ersten Runde sprachlich präzisiert, inhaltlich zusammengeführt oder anhand der Expertenkommentare aus Runde 1 modifiziert wurden. Die Verteilung der Zustimmung (Skalenbereich 7-9) reichte von 14% bis 96%. Eine vorläufige Zustimmung erzielten 14 Empfehlungen. Basierend auf den aus der ersten Runde modifizierten Aussagen wurde bei acht Empfehlungen nun ein vorläufiger Konsens erreicht. Für die ergänzten Kriterien zur Auswahl von LL-Empfehlungen - Grad der Empfehlung bzw. evidenzbasiert und/oder auch konsensbasiert - bestanden unterschiedliche Präferenzen (Zustimmung 14 bis 46%).

Zum anderen wurde der bereits in Runde 1 beschriebene iterative Ansatz abgelehnt (Zustimmung 57%). Gleiches galt auch für den Aspekt der Datenverfügbarkeit (Zustimmung 71%).

Abschlusskonferenz

Der erarbeitete QI-Entwicklungsprozess wurde in weiten Teilen final verabschiedet nach einigen sprachlichen Präzisierungen sowie der Zusammenlegung von inhaltlich assoziierten

Empfehlungen und einigen redaktionellen Umstellungen von Empfehlungen (Abb. 3).

Der evidenzgeleitete und konsensbasierte QI-Standard setzt sich aus insgesamt 30 Empfehlungen zur leitlinienbasierten QI-Entwicklung zusammen, die sich auf die untenstehenden sechs Bereiche (A-F) verteilen:

- A) Vorbereitende Arbeitsschritte für die leitlinienbasierte Empfehlung von QI (drei Empfehlungen)
- B) QI-Entwicklergruppe und Zusammenarbeit mit der LL-Gruppe (sieben Empfehlungen)
- C) Entwicklung der potenziellen QI (sechs Empfehlungen)
- D) Bewertung der potenziellen QI (acht Empfehlungen)
- E) Verabschiedung sowie Publikation der vorläufigen QI (vier Empfehlungen)
- F) Pilotierung/Testung der vorläufigen QI (zwei Empfehlungen)

Die einzelnen verabschiedeten Empfehlungen pro Bereich sind in der [Tabelle 3](#) aufgelistet.

In Ergänzung zur bisherigen QI-Entwicklungsmethodik des NVL- und OL-Programms wurden 10 Empfehlungen verabschiedet. Diese sind in [Tabelle 3](#) kursiv hervorgehoben.

Der vom Studienteam vorgeschlagene Ansatz, dass zunächst jede LL-Empfehlung in potenzielle QI überführt werden soll (C1, [Appendix A, Online Anlage 2](#)) wurde nicht angenommen. Stattdessen wurde verabschiedet, dass „als Grundlage für potenzielle QI evidenz- und/oder konsensbasierte, nachvollziehbar begründete, starke Empfehlungen (Empfehlungsgrad A/„soll“) heranzuziehen“ sind (Zustimmung 100%). Mit einer Zustimmung von allen abstimmenden Expert*innen wurde die Empfehlung mit Fokus auf „erneute Überprüfung der LL-Empfehlung in Bezug auf ihre Operationalisierbarkeit“ (C2, [Appendix A, Online Anlage 2](#)) gestrichen.

Für zwei Empfehlungen (Prüfen der Möglichkeit der Übernahme von bereits publizierten QI (C3) und Kategorisierung der QI in Struktur-, Prozess- und Ergebnisqualität (C4)) konnte auch abschließend nur eine mehrheitliche Zustimmung (71%) erreicht werden. Sie wurden mit entsprechendem Hinweis in den QI-Standard integriert ([Appendix A, Online Anlage 2](#)).

Eine Zustimmung $\leq 75\%$ wurde weiterhin für folgende während der Abschlusskonferenz diskutierten QI-Bewertungskriterien erzielt: Bewertung der potenziellen QI bzgl. Nutzenpotenzial/erwarteter Nutzen (Zustimmung 55%) und Augenscheinvalidität, d.h. eine auch für Laien unmittelbar einleuchtende Plausibilität eines QI (Zustimmung 40%) ([Appendix A, Online Anlage 2](#)).

Diskussion

Der bisher für die QI-Entwicklungsmethodik im Rahmen von S3-LL fehlende einheitliche „Goldstandard“ und die damit einhergehende geringe bzw. fehlende Evidenz zur Standardisierung der Entwicklung, zu den Anforderungskriterien an die methodische Güte und zur Evaluation von QI, stellt eine wesentliche Barriere für eine adäquate Messung der medizinischen Versorgungsqualität dar und erschwert dadurch die Umsetzung der gesundheitspolitischen Ziele einer qualitativ hochwertigen Versorgung (§135a SGB V).

Ziel der präsentierten Arbeit war es, einen übergreifenden, für alle S3-Leitliniengruppen anwendbaren QI-Standard zu erarbeiten und zu verabschieden, der an etablierte QI-Entwicklungsmethoden anknüpft.

Stärken und Schwächen

Mit dem vorgestellten Forschungsprojekt wurde anhand eines iterativen evidenzgeleiteten Konsensprozesses erstmals ein methodischer Standard (bestehend aus 30 Empfehlungen)

für die Entwicklung von leitlinienbasierten QI für S3-LL mit Einbeziehung wesentlicher Experten- und Interessengruppen formal abgestimmt. Der verabschiedete QI-Standard basiert auf der best-verfügbaren Evidenz zur Entwicklung von QI im Rahmen der LL-Erstellung und der Expertise sowie den Erwartungen von Patient*innen, Leistungserbringer*innen, Leitlinienautor*innen, Methodiker*innen und Repräsentant*innen der Gesundheitspolitik zu bzw. an leitlinienbasierten QI. Eine solche breite und formale, mehrstufige Abstimmung des Vorgehens zu leitlinienbasierten QI ist auch im internationalen Vergleich unseres Wissens nach ein Novum.

Zu Beginn des Konsensprozesses wurde durch das Studienteam eine Auswahl von Fachgesellschaften und Institutionen getroffen, mit dem Ziel, möglichst alle relevanten Experten- und Interessengruppen einzubinden. Mit Ausnahme gesundheitspolitischer Vertreter*innen wurde darauf geachtet, dass ein möglichst ausgewogenes Verhältnis zwischen den o.g. Expertisen erzielt wird. Vor dem Hintergrund, dass ein Großteil des Entscheidungsgremiums mehrere Expertisen aufweist, legten die Mandatsträger*innen während der ersten Delphi-Befragung selbst fest, welche primäre Rolle im Rahmen dieser strukturierten Konsentierung eingenommen wird.

Die einzelnen Teilprojekte trugen zur Ableitung der Empfehlungen für den QI-Standard bei, jedoch war eine direkte Ableitung einer Empfehlung aus den spezifischen Ergebnissen der Teilprojekte nicht immer möglich, da ein systematisches „critical appraisal“ der Evidenzgrundlagen, wie bei der Ableitung von LL-Empfehlungen bspw. aus klinischen Studien u.a. mit Hilfe des GRADE-Prozesses [15], nicht möglich war. Auch zehn Jahre nach der systematischen Aufbereitung der für die Entwicklung von leitlinienbasierten QI angewandten Methodik [6] existieren nach derzeitigem Kenntnisstand keine vergleichenden Studien, die untersuchten, inwiefern aus verschiedenen methodischen QI-Entwicklungsansätzen qualitativ gute oder eben weniger gute QI resultieren.

Vor dem Hintergrund eines bisher national wie international fehlenden „Goldstandards“ zur leitlinienbasierten QI-Entwicklung, war es nicht das Ziel, die in anderen Ländern teils mit differierenden Strukturen (vgl. u.a. The National Institute for Health and Care Excellence (NICE) [16]) angewandte QI-Entwicklungsmethodik vergleichend zu analysieren. Als zielführender wurde erachtet, die bereits in Deutschland etablierte QI-Methodik weiterzuentwickeln, indem durch einen national wie international ausgerichteten Mixed-Methods-Ansatz der Status quo der QI-Entwicklung kritisch reflektiert werden sollte, um Anhaltspunkte für Verbesserungspotenziale zu erhalten.

Ob der QI-Standard im Vergleich zum gegenwärtigen Vorgehen des NVL- und OL-Programms tatsächlich besser geeignet ist, um wissenschaftlich fundierte und für die Versorgung relevante QI zu entwickeln, muss weiterführend für verschiedene Leitlinienthemen geprüft werden. In diesem Zusammenhang kann auch untersucht werden, inwiefern von Leitliniengruppen beschriebene Barrieren (u.a. Möglichkeiten der QI-Messung bei der klinischen Routinedokumentation) während der Ableitung und des Einsatzes von QI minimiert werden können [17]. Die kürzlich geschaffenen Fördermöglichkeiten der Versorgungsforschung zu LL des Innovationsfonds nach §92 des Digitale-Versorgung-Gesetzes (DVG) stellen eine Chance dar, die Entwicklung, Pilotierung und Implementierung leitlinienbasierter QI projektbezogen zu fördern [18]. In diesem Zusammenhang wurde die regelhafte Ableitung von Vorschlägen für QI-Empfehlungen von der AWMF empfohlen [19].

Inhalte des QI-Standards

Durch die vorliegende Arbeit wurde die bisher in Deutschland angewandte leitlinienbasierte QI-Entwicklungsmethodik von

Tabelle 3
Konsentierete Empfehlungen zur leitlinienbasierten QI-Entwicklung.

Bereich A

Vorbereitende Arbeitsschritte für die leitlinienbasierte Empfehlung von QI

- A1 Wir empfehlen, dass zu Beginn der Leitlinienerstellung mindestens eine Person mit methodischer Erfahrung im Bereich Qualitätsbestimmung die gesamte Leitliniengruppe über QI (Konzept, Verwendungszwecke und allgemeine Entwicklungsschritte) informiert.
- A2 Wir empfehlen, in Vorbereitung auf die Erarbeitung von leitlinienbasierten QI zu Beginn der S3-Leitlinienerstellung den Geltungsbereich der LL durch die explizite Nennung der dazugehörigen Professionen, Fachgebiete, Versorgungsbereiche, (Patienten-)Population und Interventionen darzulegen.
- A3 Wir empfehlen, in Vorbereitung auf die Erarbeitung von leitlinienbasierten QI zu Beginn der S3-Leitlinienerstellung Versorgungsziele mit Bezug auf den Geltungsbereich der LL zu formulieren.

Bereich B

Zusammensetzung der QI-Entwicklergruppe und Zusammenarbeit mit der LL-Gruppe

- B1 Wir empfehlen frühzeitig, d.h. zu Beginn der S3-Leitlinienerstellung, festzulegen, welche Mitglieder der Leitliniengruppe für die Entwicklung und Formulierung der QI verantwortlich sind.
- B2 Wir empfehlen, an der Entwicklung und Formulierung von QI Vertreter mindestens der verschiedenen betroffenen Gesundheitsberufe / Fachdisziplinen, Patientenvertreter und Methodiker aus dem Bereich Qualitätsbestimmung zu beteiligen.
- B3 Wir empfehlen, in die QI-Entwicklergruppe Vertreter von Institutionen/Einrichtungen oder Fachgesellschaften hinzuzuziehen, die Expertise für die zukünftige Umsetzung der QI aufweisen, auch wenn diese nicht der LL-Gruppe angehören. Die Beteiligung kann für den gesamten Prozess, für bestimmte Entwicklungsschritte oder als Einholung externer Expertise erfolgen.
- B4 Wir empfehlen, dass die Vertreter der Gesundheitsberufe / Fachdisziplinen sowie Patientenvertreter und Methodiker der QI-Entwicklergruppe im Rahmen der Priorisierung und vorläufigen Bewertung der potenziellen QI stimmberechtigt sind.
- B5 Wir empfehlen, allen Mitgliedern der QI-Entwicklergruppe eine Schulung zum Thema QI (spezifische Entwicklungsmethodik) anzubieten.
- B6 Wir empfehlen, dass die LL-Gruppe und die QI-Entwicklergruppe zeitlich und personell eng aufeinander abgestimmt zusammenarbeiten.
- B7 Wir empfehlen, dass alle Mitglieder der QI-Entwicklergruppe ihre Interessen entsprechend dem Regelwerk Leitlinien der AWMF erklären. Das Verfahren zur Bewertung der Erklärungen und zum Umgang mit Interessenkonflikten soll entsprechend des AWMF-Regelwerks erfolgen.

Bereich C

Entwicklung der potenziellen QI

- C1 Wir empfehlen, als Grundlage für potenzielle QI evidenz- und/oder konsensbasierte, nachvollziehbar begründete, starke Empfehlungen (Empfehlungsgrad A / „soll“) heranzuziehen.
- C2 Wir empfehlen, im Zuge der Entwicklung von leitlinienbasierten QI die Übernahme von bereits veröffentlichten und implementierten QI zu prüfen.
- C3 Wir empfehlen, potenzielle QI auf Versorgungsziele auszurichten. Hierbei ist eine Kategorisierung in Struktur-, Prozess- und Ergebnisqualität hilfreich.
- C4 Wir empfehlen, potenzielle QI als geeignete Messgröße, wie z.B. Anteil, Rate, Quotient (Ratio) oder Sentinel-Event, darzustellen.
- C5 Wir empfehlen, dass die Entwicklung potenzieller QI aus Leitlinien-Empfehlungen prioritär für Versorgungsbereiche mit einem hohen Potenzial für patientenrelevante Qualitätsverbesserung erfolgt.
- C6 Wir empfehlen, die beabsichtigte(n) Zweckbestimmung(en) der potenziellen QI konkret zu benennen. Zweckbestimmungen können bspw. Qualitätsverbesserung, Patientinformation, Monitoring, Planung, Benchmarking, Evaluation der Leitlinienumsetzung, Versorgungs- und Qualitätsforschung sowie Vergütung sein. Wir empfehlen, auch begründet auszuweisen, für welche Einsatzzwecke ein potenzieller QI nicht geeignet erscheint.

Bereich D

Bewertung der potenziellen QI

- D1 Wir empfehlen, die QUALIFY-Kriterien für die Bewertung von potenziellen QI zuzüglich der Kriterien Reflexion kultureller Werte, Akzeptanz und Datenmanagement zugrunde zu legen.
- D2 Wir empfehlen, für die Bewertung der potenziellen QI die Domäne Relevanz mit den folgenden Kriterien (entsprechend des Instruments QUALIFY) explizit anzuwenden:
- Bedeutung für das Versorgungssystem (Der Indikator erfasst wesentliche Aspekte der Lebensqualität, Morbidität oder Mortalität bzw. benennt hierfür wesentliche Versorgungsprozesse oder Versorgungsstrukturen.)
- Berücksichtigung potenzieller Risiken/ Nebenwirkungen (Es sind keine Risiken für einen Anreiz zur Fehlsteuerung bekannt oder die bekannten bzw. vermuteten Risiken durch den Einsatz des Indikators sind beschrieben und erforderlichenfalls berücksichtigt.)
- D3 Wir empfehlen, für die Bewertung der potenziellen QI die Domäne Wissenschaftlichkeit mit dem folgenden Kriterium (entsprechend des Instruments QUALIFY) explizit anzuwenden:
- Klarheit der Definitionen (Der Indikator ist klar und eindeutig definiert.)
- D4 Darüber hinaus empfehlen wir - auch in Vorbereitung einer möglichen Pilotierung/ Testung der QI - Aspekte der Domäne Wissenschaftlichkeit mit dem folgenden Kriterium (entsprechend des Instruments QUALIFY) deskriptiv darzulegen
- Potenzielle Risikoadjustierung
- D5 Darüber hinaus empfehlen wir - auch in Vorbereitung einer möglichen Pilotierung/ Testung der QI - Aspekte der Domäne Wissenschaftlichkeit (nicht in QUALIFY enthalten) mit dem folgenden Kriterium deskriptiv darzulegen
- Reflexion kultureller Werte
- D6 Wir empfehlen, für die Bewertung der potenziellen QI die Domäne Praktikabilität mit dem folgenden Kriterium (entsprechend des Instruments QUALIFY) explizit anzuwenden:
- Beeinflussbarkeit der Indikatorsprägung (Der Qualitätsindikator bezieht sich auf einen Versorgungsaspekt, der von den bewerteten Akteuren beeinflusst werden kann.)
- D7 Wir empfehlen, für die Bewertung der potenziellen QI die Domäne Praktikabilität mit den folgenden Kriterien (entsprechend des Instruments QUALIFY) deskriptiv darzulegen
- Verständlichkeit und Interpretierbarkeit für Patienten und interessierte Öffentlichkeit
- Verständlichkeit für Ärzte und andere Gesundheitsberufe
- Implementierungsbarrieren
- Datenverfügbarkeit
zzgl. empfehlen wir, dass (nicht in QUALIFY enthaltene) Kriterium
- Vermutete Akzeptanz (Inhaltliche Akzeptanz) deskriptiv darzulegen.
- D8 Wir empfehlen, auszuweisen, welche potenziellen QI erhebbar sind und für welche QI Erhebungsmethoden entwickelt werden müssten.

Überführung der potenziellen QI in vorläufige QI

Bereich E

Verabschiedung sowie Publikation der vorläufigen QI

- E1 Wir empfehlen, für die Konsentierung der vorläufigen QI ein formales Konsensverfahren anzuwenden.
- E2 Wir empfehlen, dass an der Verabschiedung der vorläufigen QI alle stimmberechtigten Mitglieder der QI-Entwicklergruppe und LL-Gruppe beteiligt sind.
- E3 Wir empfehlen, für die Annahme eines vorläufigen QI ein Cut-off von $\geq 75\%$ pro explizitem Bewertungskriterium festzulegen. Abweichungen von diesem Wert sollen begründet werden, Sondervoten sollen dokumentiert werden.

Tabelle 3 (Fortsetzung)

E4	Wir empfehlen, dass die Veröffentlichung vorläufiger QI nach deren Verabschiedung durch die Leitliniengruppe und QI-Entwicklergruppe erfolgt. Dabei soll ausdrücklich auf die Vorläufigkeit der QI hingewiesen werden.
Bereich F	
Pilotierung/Testung der vorläufigen QI und in Abhängigkeit der Ergebnisse Überführung in finale QI	
F1	Wir empfehlen, bevor leitlinienbasierte QI regelhaft eingesetzt werden, eine Pilotierung/Testung in ausgewählten Versorgungsbereichen durchzuführen.
F2	Wir empfehlen, dass eine Pilotierung/Testung in engem Austausch mit Vertretern der Leitliniengruppe vorgenommen wird, die bei der QI-Entwicklung beteiligt waren.

Anmerkung: kursiv hervorgehobene Empfehlungen stellen Ergänzungen bzw. Änderungen zu den in Deutschland existierenden QI-Manualen [4,5] dar.

relevanten Experten- und Interessengruppen aus dem Gesundheitswesen bestätigt.

Von den insgesamt 30 verabschiedeten Empfehlungen sind 10 eine Ergänzung bzw. Änderung zu den in den beiden deutschen QI-Manualen beschriebenen Vorgehen, die aus den Ergebnissen der Teilstudien vom Studienteam abgeleitet wurden. Im Prozess der strukturierten Konsensfindung wurden keine Empfehlungen zusätzlich vom Expertenpanel vorgeschlagen bzw. zur Diskussion gestellt.

Aspekte der ergänzenden Empfehlungen beinhalten zum einen personelle bzw. LL-Gruppen betreffende Erweiterungen bzw. Änderungen (4 Empfehlungen bzgl. Involvierung methodischer Expertise, Schulungsangebote, Erweiterung der Stimmberechtigungen, Pilotierung/Testung der QI im Austausch mit der QI-Entwicklergruppe). Neben dem Angebot einer Schulung zum Thema QI, die bspw. durch den Aufbau neuer Kooperationsfelder realisiert werden könnte, wird der Einbezug von Personen mit methodischer Erfahrung im Bereich Qualitätsbestimmung erforderlich sein, die S3-Leitliniengruppen zu Beginn über die QI-Entwicklung informieren. Die Erstellung von S3-LL erfordert von den Autor*innen neben ihrer klinischen auch eine methodische Expertise im Bereich der evidenzbasierten Medizin, v.a. bei der kritischen Studienbewertung (z.B. Anwendung des GRADE Prozesses). Durch die kontinuierliche methodische Weiterentwicklung von bspw. Studiendesigns der klinischen und Versorgungsforschung, Instrumenten zur Bewertung der Studienevidenz [20] oder des Auswahlprozesses von Outcomes anhand von Core Outcome Sets [21] wird LL-Autor*innen zunehmend mehr methodisches Fachwissen, welches fortlaufend weitergebildet werden muss, abverlangt. Durch die im QI-Standard enthaltene Empfehlung, methodische Kompetenzen hinsichtlich QI zu erweitern, sind LL-Gruppen zusätzlich gefordert. Ziel sollte es sein, dass LL-Gruppen von methodisch aus- bzw. weitergebildetem Personal im Prozess der Leitlinienerstellung unterstützt werden. Dies erfordert in der Regel eine finanzielle Unterstützung der Leitliniengruppen, wie sie beispielsweise im Leitlinienprogramm Onkologie oder bei der Ausschreibung von Leitlinienprojekten im Rahmen des Innovationsfonds realisiert ist [22].

Zwei Empfehlungen, die im Prozess der Konsensfindung vorsahen, dass LL-Empfehlungen mit fehlender Operationalisierung durch eine iterative Zusammenarbeit der LL- und QI-Entwicklergruppe überprüft und ggf. umformuliert werden sollen, wurden aufgrund des erhöhten Aufwands bei unklarem Nutzen gestrichen (Zustimmung <75%). Aus der Sicht des Studienteams erscheinen zeitlich und personell recht aufwendig erarbeitete LL-Empfehlungen nur dann zum Monitoring und zur Verbesserung der Versorgungsqualität nützlich, wenn deren praktische Umsetzung evaluiert, d.h. via QI gemessen werden kann. Die unterschiedlichen Einschätzungen zeigen, dass der Spagat zwischen der Intention einer leitlinienbasierten Versorgung und die Umsetzung einer adäquaten QI-Messung auch künftig eine Herausforderung bleiben wird.

Im QI-Standard wurde ergänzt, dass die beabsichtigte(n) Zweckbestimmung(en) der potenziellen QI konkret benannt werden sollen (Bereich C), um deutlich zu machen, dass die empfohlenen

QI nur für solche von der Leitliniengruppe intendierten Zwecke angewendet werden.

Da bisher keine systematische Evaluation der methodischen Güte eines oder mehrerer QI-Bewertungsinstrumente vorliegt, wurde zumindest über die in den Instrumenten enthaltenen Bewertungskriterien, die auch ohne eigene Datenerhebung von der QI-Entwicklergruppe einschätzbar sind, abgestimmt. Die kritische Bewertung der vorläufigen QI wurde um eine Beurteilung der „vermuteten Akzeptanz“ und „Reflexion kultureller Werte“ (Bereich D) erweitert. Für die Bewertung der potenziellen QI bzgl. Nutzenpotenzial/erwarteter Nutzen sowie der Augenscheinvalidität wurde auch abschließend kein Konsens (Zustimmung 40%) erzielt.

Darüber hinaus sieht der QI-Standard vor, dass ergänzend zur Beurteilung der Datenverfügbarkeit (Bewertung durch die Krebsregister und Zertifizierungssysteme der Deutschen Krebsgesellschaft, seit 2017) [5] bzw. Kommentierung durch die LL-Gruppe (NVL)) [4] ausgewiesen werden soll, welche potenziellen QI bereits erhebbar sind und für welche QI Erhebungsmethoden entwickelt werden müssten (Bereich D). Dass ein beachtlicher Anteil an leitlinienbasierten QI derzeit noch nicht für die Überprüfung der leitlinienkonformen Versorgung genutzt werden kann, zeigt eine aktuelle Analyse onkologischer QI, deren Berechnung mit dem Basisdatensatz der Arbeitsgemeinschaft Deutscher Tumorzentren und der Gesellschaft der epidemiologischen Krebsregister in Deutschland untersucht wurde. Basierend auf 17 S3-LL, die 151 QI enthalten, konnten nur 54% aus den vorliegenden Daten berechnet werden [23]. Der Aufwand der QI-Entwicklung für Leitliniengruppen steht deshalb nach Einschätzung vieler Leitlinienautor*innen nicht in Relation zu ihrem Nutzen bzw. dieser ist nach Fertigstellung der Leitlinie aufgrund der unzureichenden Messung und somit ausbleibenden Anwendung der QI nicht garantiert [9].

Die Einführung von QI ist eine komplexe Intervention und bedarf selbst der Hinterlegung mit Evidenz, um sicher zu gehen, dass die Umsetzung der QI zu einer patient*innenrelevanten Verbesserung der Versorgung führt. Aus diesem Grund wird es eine anspruchsvolle Aufgabe sein, für leitlinienbasierte QI eine systematische Pilotierung in ausgewählten Versorgungsbereichen durchzuführen (2 Empfehlungen). Neben einer politischen Debatte zu den Verantwortlichkeiten und der Finanzierung dieser weiterführenden datenbasierten Testung der QI muss auf wissenschaftlicher Ebene die methodische Güte von QI-Bewertungsinstrumenten analysiert sowie die Relevanz von einzelnen Bewertungskriterien möglichst breit abgestimmt werden. Eine in diesem Kontext erstellte Forschungsagenda zeigt weitere zentrale Aufgabenfelder auf [24].

Ausblick

Der evidenzgeleitete und konsensbasierte QI-Standard fokussiert auf die Entwicklung von leitlinienbasierten QI und endet in der Regel mit ihrer Veröffentlichung innerhalb einer S3-Leitlinie. Selbst für das OL-Programm, welches ein umfassendes etabliertes Konzept mit ineinandergreifenden Strukturen für die Überprüfung der definierten QI in der Routineversorgung erarbeitet hat, wurde festgestellt, dass nur für einzelne dieser QI die Berechnung

möglich und somit ein Teil der in S3-Leitliniengruppen erarbeiteten QI nur eingeschränkt nutzbar ist. Die zukünftige Entwicklung von QI im Rahmen von S3-LL wird demnach sehr stark davon abhängen, wie gut es gelingen wird, ein Gesamtkonzept für die Entwicklung, Pilotierung und Implementierung von QI zu erarbeiten.

In einem nächsten Schritt soll der QI-Standard in ausgewählten S3-Leitlinienprojekten erprobt werden. Bei erfolgreicher Pilotierung kann der QI-Standard bei allen künftigen S3-LL zur Implementierung empfohlen werden. Vom Einsatz der Empfehlungen des QI-Standards durch Leitliniengruppen können wissenschaftlich fundierte und für die Versorgung relevante QI erwartet werden. Neben den methodischen Anforderungen an die leitlinienbasierte QI-Entwicklung ist es erforderlich, dass die notwendigen Strukturen und Ressourcen zur Verfügung stehen, um die erarbeiteten Empfehlungen umzusetzen.

Um den politischen Willen des Gesetzgebers in Bezug auf Qualitätssicherung und -verbesserung durch die Leistungserbringer*innen umzusetzen (siehe §135a SGBV), ist es darüber hinaus wesentlich, dass QI-Vorschläge aus Leitliniengruppen von den Institutionen, die für die Entwicklung und Umsetzung von QI vom Gesetzgeber benannt sind, entsprechend aufgegriffen und geprüft werden. Insbesondere ist hier das Institut für Qualität und Transparenz im Gesundheitswesen zu nennen, aber auch der Gemeinsame Bundesausschuss z.B. für QI aus Disease Management Programmen.

Förderung

Die vorliegende Arbeit wurde im Rahmen des Projektes „Systematische Analyse der Translation von Leitlinienempfehlungen in Qualitätsindikatoren und Erarbeitung eines evidenz- und konsensbasierten Standards“ von der Deutschen Forschungsgemeinschaft (DFG) finanziell gefördert (Projektnummer GEPRIS-Datenbank: 289625106).

Danksagung

Wir bedanken uns bei allen Teilnehmer*innen der strukturierten Konsensfindung für ihr großes Engagement und die konstruktiven Diskussionen.

Unser Dank gilt auch Andreas Büscher (Deutscher Pflegerat), Mark Coburn (Deutsche Gesellschaft für Anästhesiologie und Intensivmedizin), Peter Falkai (Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde), Hans-Joachim Meyer (Deutsche Gesellschaft für Chirurgie), Corinna Schäfer (Bundesärztekammer), Erika Sirsch (Deutsche Gesellschaft für Pflegewissenschaft), Christof Veit (Institutionsloser Vertreter) und Wolfgang Wodarg (Patientenvertretung G-BA) für ihre Teilnahme an den onlinebasierten Delphi-Befragungen.

Weiterhin danken wir Margarete Ostholt-Corsten (Deutsche Rentenversicherung Bund) und Herrn Lutz Altenhofen (Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen) für die Teilnahme an den onlinebasierten Delphi-Befragungen und der Abschlusskonferenz.

Die Autoren bedanken sich zudem bei Thomas Petzold (zum Zeitpunkt der Antragstellung ZEGV, Universitätsklinikum Dresden) für die Unterstützung bei der Erarbeitung des Studienkonzeptes und Toni Lange (ZEGV, Universitätsklinikum Dresden) für die technische Erstellung und Auswertung der onlinebasierten Delphi-Befragungen.

Interessenkonflikt

Aus dem beruflichen Kontext heraus wies ein Großteil der Teilnehmer*innen einen potenziellen Interessenkonflikt auf. Die Auswahl der beteiligten Expert*innen bzw. die zu repräsentierende

Einrichtung/ Institution erfolgte seitens des Studienteams absichtsvoll, um sicherzustellen, dass weitgehend alle Experten- und Interessengruppen am Entwicklungsprozess des QI-Standards partizipieren.

Mitglied von im Zusammenhang mit der Leitlinienentwicklung relevanten Fachgesellschaften/Berufsverbänden, Mandatsträger*in im Rahmen der Leitlinienentwicklung: TH, TN, RT, KV

Methodische Beratung für die Erstellung von LL und leitlinienbasierten QI: FD, MF, IK, CM, MN, SS, SW

Autor*in von im vorliegenden Projekt zugrunde gelegten Manuskripten zur QI-Entwicklung oder Instrumenten zur QI-Bewertung: FD, KD, MF, MG, IK, MN, MKS, SW

Kein Interessenkonflikt: Alle anderen

Autorenschaft

Stefanie Deckert: Konzipierung; Auswahl der angewendeten Methoden; Durchführung, Auswertung und Validierung der Ergebnisse; Entwurf und Revision des Manuskripts; Projektkoordination; Fördermittelakquise. Katrin Arnold, Monika Becker, Marie Brombach, Jessica Breuing, Marie Bolster: Durchführung und Auswertung; Revision des Manuskripts. Max Geraedts: Durchführung; Validierung der Ergebnisse; Revision des Manuskripts. Edmund Neugebauer: Konzipierung; Auswahl der angewendeten Methoden; Revision des Manuskripts; Fördermittelakquise. Dawid Pieper: Auswertung; Revision des Manuskripts; Projektleitung. Monika Nothacker: Konzipierung; Auswahl der angewendeten Methoden; Durchführung, Auswertung und Validierung der Ergebnisse; Revision des Manuskripts; Projektleitung; Fördermittelakquise. Jochen Schmitt: Konzipierung; Auswahl der angewendeten Methoden, Auswertung und Validierung der Ergebnisse; Revision des Manuskripts; Projektleitung, Fördermittelakquise. Alle anderen Autor*innen: Durchführung; Revision des Manuskripts.

Appendix A. Zusätzliche Daten

Zusätzliche Daten verbunden mit diesem Artikel finden sich in der Online-Version unter: [doi:10.1016/j.zefq.2020.11.008](https://doi.org/10.1016/j.zefq.2020.11.008).

Literatur

- [1] Nothacker M, Muche-Borowski C, Kopp IB. 20 Jahre ärztliche Leitlinien in Deutschland - was haben sie bewirkt? *Z. Evid. Fortbild. Qual. Gesundh. wesen* 2014;108(10):550–9.
- [2] Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Quality & safety in health care* 2002;11(4):358–64.
- [3] Schmitt J, Petzold T, Eberlein-Gonska M, Neugebauer EAM. Anforderungsprofil an Qualitätsindikatoren Relevanz aktueller Entwicklungen der Outcomes Forschung für das Qualitätsmanagement. *Z. Evid. Fortbild. Qual. Gesundh. wesen* 2013;107(8):516–22.
- [4] Ärztliches Zentrum für Qualität in der Medizin (ÄZQ). Manual Qualitätsindikatoren 2009. Berlin: ÄZQ 2009 Verfügbar: <https://www.aezq.de/mdb/edocs/pdf/schriftenreihe/schriftenreihe36.pdf>. (Zugriff am 3. März 2019).
- [5] Leitlinienprogramm Onkologie (Deutsche Krebsgesellschaft (DKG), Deutsche Krebshilfe (DKH), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF)): Entwicklung von leitlinienbasierten Qualitätsindikatoren. Methodenpapier für das Leitlinienprogramm Onkologie. Version 1.0. 2013. Verfügbar: <https://docplayer.org/2528307-Entwicklung-von-leitlinien-basierten-qualitaetsindikatoren.html>. (Zugriff am 10. Januar 2019).
- [6] Kötter T, Blozik E, Scherer M. Methods for the guideline-based development of quality indicators - a systematic review. *Implementation science* 2012;7:21.
- [7] Nothacker M, Stokes T, Shaw B, Lindsay P, Sipila R, Follmann M, Kopp I. Reporting standards for guideline-based performance measures. *Implementation science* 2016;11:6.
- [8] Deckert S, Steudtner M, Becker M, Brombach M, Neugebauer E, Nothacker M, Schmitt J. (Wie) erfolgt die Ableitung von Qualitätsindikatoren zur Messung und Bewertung der Versorgungsqualität im Rahmen von S3-Leitlinien? Eine Übersichtsarbeit. *Z. Evid. Fortbild. Qual. Gesundh. wesen* 2019;147–148:45–57.
- [9] Arnold A, Breuing J, Becker M, Nothacker M, Neugebauer E, Schmitt J, Deckert S. Entwicklung leitlinienbasierter Qualitätsindikatoren: Ergebnisse

- einer qualitativen Studie zu Barrieren und förderlichen Faktoren aus der Sicht von S3-LeitlinienautorInnen. *Z. Evid. Fortbild. Qual. Gesundh. Wesen* 2019;147–148:34–44.
- [10] Nothacker M, Bolster M, Steudtner M, Arnold A, Deckert S, Becker M, Kopp I, Schmitt J. International experiences in the development and implementation of guideline-based quality indicators - a qualitative study. *BMJ Open* accepted 2020.
- [11] Becker M, Breuing J, Nothacker M, Deckert S, Brombach M, Schmitt J, Neugebauer E, Pieper D. Guideline-based quality indicators-a systematic comparison of German and international clinical practice guidelines. *Implementation science* 2019;14(1):71.
- [12] Reiter A, Fischer B, Kötting J, Geraedts M, Jäckel WH, Döbler K. QUALIFY: Ein Instrument zur Bewertung von Qualitätsindikatoren. *Z. Evid. Fortbild. Qual. Gesundh. Wesen* 2008;101(10):638–88.
- [13] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) - Ständige Kommission Leitlinien. *AWMF-Regelwerk, Leitlinien**. 1. Auflage 2012. Verfügbar: <http://www.awmf.org/leitlinien/awmf-regelwerk.html>. (Zugriff am 22. Januar 2019).
- [14] Fitch K, Bernstein SJ, Aguilar MD, et al. *The RAND/UCLA. Appropriateness Method User's Manual*. Santa Monica: RAND; 2001.
- [15] Conrad S, Kaiser L, Kallenbach M, Meerpohl J, Morche J. [GRADE: Evidence to Decision (EtD) frameworks - a systematic and transparent approach to making well informed healthcare choices. 2 Clinical guidelines]. *Z. Evid. Fortbild. Qual. Gesundh. Wesen* 2019;140:63–73.
- [16] Petzold T, Deckert S, Williamson PR, Schmitt J. Quality Measurement Recommendations Relevant to Clinical Guidelines in Germany and the United Kingdom: (What) Can We Learn From Each Other? *Inquiry* 2018;55, 46958018761495.
- [17] Scherer M, Wollny A, Vollmar HC, Chenot J-F, Baum F, Barzel A, Beyer M, Egidi G, Kaufmann-Kolle P, Mainz A, Popert U, Schmiemann G, Voigt K, Wagner H-O, Szecsenyi J. Entwicklung und Bewertung von Qualitätsindikatoren aus den Leitlinien Halsschmerz Nackenschmerz und Demenz der Deutschen Gesellschaft für Allgemeinmedizin und Familienmedizin (DEGAM). *Z. Evid. Fortbild. Qual. Gesundh. Wesen* 2013;107:74–86.
- [18] Der Bundestag. Gesetz für eine bessere Versorgung durch Digitalisierung und Innovation vom 19.12.2019 (Digitale-Versorgung-Gesetz), Bundesgesetzblatt Teil 1 Nr. 49, S. 2562ff. Bundesanzeiger-Verlag 2019.
- [19] Kopp I. Politische Forderungen und Umsetzungsvorschlag der AWMF: Finanzierung von Leitlinien über den Innovationsfonds des G-BA Vortrag, 30. Leitlinienkonferenz der AWMF 2019. Verfügbar: <https://www.awmf.org/leitlinien/ll-veranstaltungen/30-leitlinien-konferenz.html>. (Zugriff am 4. Juni 2020).
- [20] Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;366:14898.
- [21] Dodd S, Harman N, Taske N, Minchin M, Tan T, Williamson PR. Core outcome sets through the healthcare ecosystem: the case of type 2 diabetes mellitus. *Trials* 2020;21:570.
- [22] Gemeinsamer Bundesausschuss. Innovationsausschuss Förderbekanntmachung Versorgungsforschung – Medizinische Leitlinien (MedLL). Veröffentlicht am 12. Oktober 2020, <https://innovationsfonds.gba.de/foerderbekanntmachungen/foerderbekanntmachung-versorgungsforschung-medizinische-leitlinien-medll.32>. (Zugriff am 9. November 2020).
- [23] Gebauer A, Barlag H, Wesselmann S, Biedenweg L, Weitmann K, Klinkhammer-Schalke M, Hoffmann W. Berechenbarkeit der Qualitätsindikatoren der onkologischen S3-Leitlinien mit dem ADT-/GEKID-Basisdatensatz und seiner Zusatzmodule. *Gesundheitswesen* 2020;82(08/09):710–5.
- [24] Döbler K, Schrappe M, Kuske S, Schmitt J, Sens B, Boywitt D, Misselwitz B, Nothacker M, Geraedts M. Eignung von Qualitätsindikatorensets in der Gesundheitsversorgung für verschiedene Einsatzgebiete – Forschungs- und Handlungsbedarf. *Gesundheitswesen* 2019;81(10):781–7.