

4 Leitlinienerevaluation: Konzepte zur Definition und Messung von Leitlinienumsetzung und -adhärenz

Max Geraedts

C. Günster | J. Klauber | D. Klemperer | M. Nothacker | B.-P. Robra | C. Schmucker (Hrsg.) Versorgungs-Report. Leitlinien – Evidenz für die Praxis.

DOI 10.32745/9783954668007-4, © MWV Medizinisch Wissenschaftliche Verlagsgesellschaft Berlin 2023

Die Leitlinienerevaluation folgt den Konzepten zur Evaluation komplexer Interventionen, indem zum einen formativ die Leitlinienumsetzung geprüft wird, das heißt deren Verbreitung, Bekanntheit und Akzeptanz in der Praxis ermittelt wird. Zum anderen wird die Leitlinienadhärenz analysiert, indem Qualitätsindikatoren formuliert werden, die summativ den Grad der Befolgung der Leitlinienempfehlungen quantifizieren. Als zentrales Evaluationsinstrument müssen für leitlinienbasierte Qualitätsindikatoren u.a. Referenzbereiche und die zur Risikoadjustierung notwendigen Faktoren eruiert sowie potenzielle Datenquellen und die Machbarkeit in der Versorgungspraxis getestet werden. Dabei kommen zur Quantifizierung der Leitlinienadhärenz Primärdaten, zunehmend aber auch Sekundärdaten zum Einsatz, die im Rahmen von prä-post-Designs bis hin zu Cluster-randomisierten Studien den Umsetzungsgrad und die Wirkung einer leitlinienkonformen Behandlung prüfen. Zu beachten ist, dass konfligierende Leitlinienempfehlungen bei multimorbiden Patientinnen und Patienten sowie deren Compliance den Umsetzungsgrad beeinflussen können. Eine aktuelle Analyse der S3-Leitlinien in Deutschland zeigt, dass bisher nur ein Viertel der Leitlinien die für die Evaluation notwendigen Qualitätsindika-

toren aufweisen, wobei diese fast ausschließlich aus der Onkologie stammen. Als Grund dafür nennen Leitlinienautorinnen und -autoren vor allem mangelnde Ressourcen. Daher kann erhofft werden, dass die nun im Rahmen des Innovationsfonds mögliche Finanzierung der Leitlinienentwicklung auch die zur Evaluation essenzielle Entwicklung leitlinienbasierter Qualitätsindikatoren fördert.

Guideline evaluation follows the concepts for the evaluation of complex interventions by, on the one hand, formatively examining guideline implementation, i.e., determining their dissemination, awareness, and acceptance in practice. Second, guideline adherence is analyzed by formulating quality indicators that summatively quantify the degree of adherence to guideline recommendations. As a central evaluation tool, reference ranges and the factors necessary for risk adjustment must be determined for guideline-based quality indicators, and potential data sources and feasibility in health care practice must be tested. To quantify guideline adherence, primary data are used, but increasingly also secondary data, which test the degree of implementation and the effect of guideline-compliant treatment within the framework of pre-post designs

up to cluster-randomized studies. It should be noted that conflicting guideline recommendations in multimorbid patients and their compliance may influence the degree of implementation. A current analysis of S3 guidelines in Germany shows that so far only a quarter of the guidelines have the quality indicators required for evaluation, and these are almost exclusively from the field of oncology. The main reason cited by guideline authors is a lack of resources. Therefore, it can be hoped that the funding of guideline development now possible under the Innovation Fund will also promote the development of guideline-based quality indicators essential for evaluation.

4.1 Implementierung von Leitlinien: eine komplexe Intervention

Von Leitlinien, die nach allen Regeln der Kunst entwickelt wurden, erwartet man gemäß der AWMF-Definition, dass sie die Entscheidungsfindung von Ärztinnen und Ärzten sowie Angehörigen von weiteren Gesundheitsberufen und Patientinnen und Patienten für eine angemessene Versorgung bei spezifischen Gesundheitsproblemen unterstützen. Die angemessene Versorgung – also eine dem aktuellen Stand der Wissenschaft entsprechende, qualitativ hochwertige Behandlung – soll letztlich ein Mehr an Gesundheit und Wohlbefinden ermöglichen: Dies betrifft sowohl die patientenrelevanten Ergebnisse (PRO) als auch Erfahrungen (PRE).

Um zu erfahren, ob erstens angemessen versorgt wird und zweitens die erwarteten Ergebnisse erzielt werden, bedarf es einer Evaluation. Konzeptionell handelt es sich bei der Implementierung von Leitlinienempfehlungen um komplexe Interventionen, bei denen mehrere Teilkomponenten und verschiedene Akteure involviert sind und zudem die jeweiligen Kontextbedingungen zu beachten sind. Die Evaluation sollte sich also an den Konzepten zur Evaluation solcher komplexer Interventionen orientieren (Skivington et al. 2021).

Zentrale Evaluationsinstrumente stellen dabei Messgrößen dar, mit deren Hilfe zum einen der Grad der Umsetzung der Leitlinien im Allgemeinen und zum anderen der Grad der Befolgung einzelner Empfehlungen im Speziellen operationalisiert werden kann.

Der folgende Beitrag konzentriert sich vor allem auf diese zentralen Bausteine der Evaluation, also die Entwicklung und Operationalisierung von leitlinienbasierten Messgrößen in Form von Qualitätsindikatoren. Dabei rekurriert der Beitrag vor allem auf grundlegende Publikationen zu dieser Thematik aus dem deutschen Versorgungskontext. Zu nennen sind die Zusammenstellung methodischer Gütekriterien für klinische Messgrößen (Geraedts et al. 2003) und die darauf aufbauende Entwicklung des Indikatoren-Bewertungsinstruments „QUALIFY“ (Reiter et al. 2007), das Manual zur Entwicklung von Qualitätsindikatoren im Rahmen des Programms für nationale Versorgungs-Leitlinien von BÄK, KBV und AWMF (Altenhofen et al. 2009) und die methodischen Standards für die Entwicklung von Qualitätsindikatoren im Rahmen von S3-Leitlinien (Deckert et al. 2021) und onkologischen Leitlinien (Leitlinienprogramm Onkologie 2021a) sowie die aktuellen „Methodischen Grundlagen Version 2.0“ des Instituts für Qualitätssicherung und Transparenz im Gesundheitswesen, IQTIG (IQTIG 2022). Zudem werden ausgewählte, nicht systematisch recherchierte Studienbeispiele zur Evaluation der Leitlinienadhärenz benannt.

4.2 Definitionen

Leitlinienadhärenz bzw. „guideline adherence“ definiert die National Library of Medicine der USA als Medical Subject Heading (MeSH) wie folgt: „Conformity in fulfilling or following official, recognized, or institutional requirements, guidelines, recommendations, protocols, pathways, or other standards“ (National

Library of Medicine 2022). Wichtig ist dabei zu beachten, dass Leitlinien laut AWMF als „Handlungs- und Entscheidungskorridore“ zu verstehen sind, von denen in begründeten Fällen abgewichen werden kann oder sogar muss (Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften [AWMF] – Ständige Kommission Leitlinien 2020). Dabei wird diskutiert, dass die Gründe für eine Abweichung jeweils dokumentiert werden sollten.

Von der Leitlinienadhärenz oder Leitlinienkonformität – also dem Grad der Befolgung einzelner Empfehlungen aus Leitlinien oder aller Empfehlungen gemeinsam bei der konkreten Versorgung – kann die **Leitlinienumsetzung** abgegrenzt werden. Hierunter kann beispielsweise die Verbreitung von Leitlinien, die Kenntnis bzw. Bekanntheit der Leitlinien durch die jeweils betroffenen Gesundheitsprofessionen und deren Haltung bzw. Akzeptanz gegenüber den Leitlinienempfehlungen aufgefasst werden (Hasenbein u. Wallesch 2007). Umfassende Evaluationen betrachten also auch das sogenannte KAP-GAP-Phänomen, die Lücke (gap) zwischen Kenntnis/Bekanntheit (knowledge), der Haltung/Akzeptanz (attitude) und der tatsächlichen Umsetzung/Durchführung (performance) der Leitlinienempfehlungen.

Bei der Definition des Begriffs „**Qualitätsindikator**“ findet sich zum einen die international gebräuchliche, ins Deutsche übertragene Umschreibung „Ein Indikator ist ein quantitatives Maß, welches zum Monitoring und zur Bewertung der Qualität wichtiger Leitungs-, Management-, klinischer und unterstützender Funktionen genutzt werden kann, die sich auf das Behandlungsergebnis beim Patienten auswirken. Ein Indikator ist kein direktes Maß der Qualität. Es ist mehr ein Werkzeug, das zur Leistungsbewertung benutzt werden kann, das Aufmerksamkeit auf potenzielle Problembereiche lenken kann, die einer intensiven Überprüfung innerhalb einer Organisation bedürfen

könnten“ (Sens et al. 2018). Zum anderen definiert das IQTIG den Begriff „Qualitätsindikator“ in seinen „Methodischen Grundlagen“ (IQTIG 2022) kurz als „Quantitative Größe, die Aussagen über die Erfüllung konkreter Qualitätsanforderungen mittels eines Messverfahrens und eines Bewertungskonzepts ermöglicht. Qualitätsindikatoren umfassen die drei Komponenten Qualitätsziel, Messverfahren und Bewertungskonzept“. Diese Komponenten werden in den „Methodischen Grundlagen“ ausführlich erläutert.

4.3 Entwicklung leitlinienbasierter Qualitätsindikatoren

Für das IQTIG beginnt die Entwicklung von Qualitätsindikatoren mit der Formulierung eines Qualitätsmodells, das anhand von Qualitätsaspekten beschrieben wird. Darunter werden Themen der Versorgungspraxis verstanden, die die Qualität der Gesundheitsversorgung in einem bestimmten Bereich beschreiben (IQTIG 2022). Diese Qualitätsaspekte werden anschließend in Form von Qualitätsmerkmalen konkretisiert, die wiederum zu Qualitätsindikatoren operationalisiert und zu einem Indikatorenset zusammengeführt werden, das einer Machbarkeitsprüfung unterzogen wird (IQTIG 2022).

Zum Verständnis dieses Vorgehens muss das Rahmenkonzept für Qualität des IQTIG erwähnt werden, das sich wiederum auf die Definition von Qualität bezieht: Demnach ist Qualität der Gesundheitsversorgung der Grad, in dem die Versorgung von Einzelpersonen und Populationen Anforderungen erfüllt, die patientenzentriert sind und mit professionellem Wissen übereinstimmen (IQTIG 2022). Die heute international bei der Gesundheitsversorgung berücksichtigten Anforderungen (oft als Qualitätsdimensionen oder -kriterien bezeichnet) werden im Rahmenkonzept wie folgt benannt: Wirksamkeit der Versorgung, Patien-

tensicherheit, Ausrichtung der Versorgungsgestaltung an den Patientinnen und Patienten (international eher Patientenorientierung/responsiveness), Rechtzeitigkeit und Verfügbarkeit, Angemessenheit (betrifft vor allem die Evidenzbasierung und Indikationsstellung) sowie Koordination und Kontinuität der Versorgung.

Je nach Qualitätsaspekt werden einzelne oder alle dieser Qualitätsdimensionen bei der Formulierung von Qualitätsmerkmalen und -indikatoren berücksichtigt.

Soll beispielsweise bei der Brustkrebsfrüherkennung die Qualität der Diagnostik als Qualitätsaspekt betrachtet werden, dann wäre denkbar, auf der Basis der entsprechenden S3-Leitlinie solche Qualitätsdimensionen genauer zu betrachten, zu denen hier Empfehlungen formuliert wurden. Eine Empfehlung mit einem Empfehlungsgrad A (level of evidence 1b) aus dem Bereich der Diagnosemitteilung lautet dabei „Die Art der Vermittlung von Informationen und der Aufklärung der Patientin soll nach folgenden Grundprinzipien einer patientenzentrierten Kommunikation, die eine partizipative Entscheidungsfindung ermöglicht, erfolgen ...“ (diese werden anschließend konkret benannt) (Leitlinienprogramm Onkologie 2021b). Die hier angesprochene Qualitätsdimension wäre also die Patientenorientierung. Ein entsprechendes Qualitätsmerkmal würde konkret als „patientenorientierte Diagnosemitteilung beim Verdacht auf Brustkrebs“ und der darauf beruhende Qualitätsindikator als „Anteil der Patientinnen, die angeben, dass ihre Diagnosemitteilung patientenorientiert erfolgte“ formuliert werden können.

Dieses Beispiel verdeutlicht die wesentlichen Unterschiede zwischen einer leitlinienbasierten Entwicklung von Qualitätsindikatoren und deren de novo Entwicklung ohne vorhandene Leitlinien (Nothacker et al. 2016; Piggott et al. 2022). Die leitlinienbasierte Indikatorentwicklung kann die ersten fünf Prozessschritte überspringen und sofort mit dem sechsten Schritt beginnen (s. folgende Übersicht).

Prozessschritte der Entwicklung von Qualitätsindikatoren

1. Kriterienbasierte Auswahl von Qualitätsaspekten
2. Prozessanalyse: Analyse des Versorgungspfads und der am Versorgungsprozess Beteiligten (Qualitäts-/Wirkmodell)
3. Reflektion bedeutsamer Qualitätsdimensionen
4. Evidenzsuche für Anforderungen an einzelne Prozessschritte/Qualitätsdimensionen
5. Ableitung von Qualitätszielen und Qualitätsmerkmalen
6. Formulierung von Qualitätsindikatoren und Referenzbereichen
7. Reflektion der nicht von Leistungserbringern zu beeinflussenden Einflussgrößen auf die Zielerreichung und Festlegung der zur Risikoadjustierung notwendigen Faktoren
8. Operationalisierung der Datenerhebung und -analyse (Datenquellen, Datenfelder, Berechnungsverfahren)
9. Machbarkeitstestung des Verfahrens
10. Methodische Gütetestung der Indikatoren (QUALIFY)

Dagegen beginnt die nicht-leitlinienbasierte Entwicklung von Qualitätsindikatoren zunächst mit der kriterienbasierten **Auswahl von Qualitätsaspekten**, die bewertet werden sollen. Oftmals werden die von der Joint Commission on Accreditation of Healthcare Organizations, JCAHO, genannten Kriterien zugrunde gelegt (JCAHO 1990): high volume, high risk, problem prone, zum Teil ergänzt um Kriterien wie „bekannte Versorgungsvariabilität mit Verdacht auf Über- oder Unterversorgung“, „kürzliche Versorgungsveränderungen“, „hohe Kosten“ oder „nationales Versorgungsziel“ (Altenhofen et al. 2009).

Daran schließt sich mit der **Prozessanalyse** der wesentliche Schritt der Indikatorentwicklung an: Die einzelnen Schritte der Versorgung und alle Beteiligten werden detailliert beschrieben, wobei auch Kontextfaktoren berücksichtigt werden sollten. Dieser Schritt ähnelt der Formulierung eines Wirkmodells bei Interven-

tionen und bietet im späteren Verlauf – bei einem nicht den Sollwerten entsprechenden Ist-Zustand – also der Abweichung der Indikatorenausprägung von vorab definierten Referenzbereichen im Rahmen der Evaluation der Versorgung – Ansatzpunkte für qualitätsverbessernde Maßnahmen.

Die Prozessanalyse kann auch genutzt werden, um die bei der jeweils betrachteten Versorgung bzw. dem Qualitätsaspekt bedeutsamen Qualitätsdimensionen festzulegen.

Der Frage danach, welche diagnostischen und therapeutischen Maßnahmen und welche Gestaltung der Gesundheitsversorgung mit dem höchsten Nutzen für Patientinnen und Patienten einhergehen, wird auf der Basis einer umfassenden **Evidenzsuche** beantwortet (s. Kap. 2). Die Methodik der Informationsbeschaffung für die Entwicklung von Qualitätsindikatoren wird im neunten Kapitel der Methodischen Grundlagen des IQTIG sehr gut beschrieben und wird daher hier nicht weiter ausgeführt – dabei spielt die Recherche qualitativ hochwertiger Leitlinien bei der Evidenzsuche eine wesentliche Rolle (IQTIG 2022).

Auf der Basis der Rechercheergebnisse können übergeordnete **Qualitätsziele** (z.B. „Wiederherstellung der Teilhabe durch eine Kniegelenksendoprothese“) und einzelne **Qualitätsmerkmale** (z.B. „Gehfähigkeit nach Kniegelenksendoprothesen-Implantation“) abgeleitet werden.

Der Schritt von einem Qualitätsmerkmal hin zur Formulierung eines **Qualitätsindikators**, also einer messbaren Größe, erfordert zunächst Überlegungen dazu, ob eine ratenbasierte Messgröße auf der Basis aggregierter Daten oder ein Einzelereignis-Indikator („Sentinel Event-Indikator“ zur Erfassung schwerwiegender Ereignisse) passend zur Bewertung der Versorgungsqualität bzw. des speziellen Qualitätsmerkmals ist. Im vorgenannten Beispiel könnte ein ratenbasierter Indikator heißen „Anteil der Patientinnen und Patienten nach Kniegelenksendoprothesen-Implantation

mit verbesserter Gehfähigkeit 6 Monate nach der Operation“. Weiterhin muss überlegt werden, welche Populationen von Zähler und Nenner erfasst werden sollen, ob also beispielsweise bestimmte Fälle aus der Nenner- oder Zählerpopulation entfernt werden sollen („exception reporting“; hier evtl. Ausschluss Verstorbener im Sechsmonatszeitraum, das könnte aber kontrovers diskutiert werden). Bei diesem Schritt wird im Allgemeinen auch empfohlen, bereits publizierte Qualitätsindikatoren zu recherchieren und hier u.a. die einschlägigen Indikator-Datenbanken zu nutzen (Altenhofen et al. 2009; Geraedts et al. 2020).

Damit verknüpft werden **Referenzbereiche** definiert, die entweder als absolute Referenzwerte eine minimal zu erreichende Qualität abgrenzen oder aber sich als statistische Referenzbereiche an der Verteilung der Indikatorenausprägungen aller Leistungserbringenden orientieren und ein demnach aktuell erreichbares und zeitabhängig veränderliches Maß darstellen (hier: z.B. „mindestens 30 Meter ohne Stock/Rollator“; oder: „Verlängerung der Strecke, die ohne Hilfsmittel zurückgelegt werden kann, um 100%“).

Damit Qualitätsindikatoren einen fairen Vergleich Leistungserbringender, zum Beispiel im Hinblick auf deren Leitlinienadhärenz, ermöglichen, müssen im Rahmen einer **Risikoadjustierung** diejenigen Faktoren berücksichtigt werden, die die Indikatorenausprägung beeinflussen, jedoch nicht von den Leistungserbringern kontrolliert werden können, hier also z.B. die Berücksichtigung des Schweregrads einer evtl. Herzinsuffizienz. Zudem muss das jeweils zur Risikoadjustierung geeignete methodische Verfahren festgelegt werden.

Im nächsten Schritt der Entwicklung leitlinienbasierter Qualitätsindikatoren werden die einzelnen Daten zur Berechnung der Indikatoren definiert und potenzielle **Datenquellen** gesucht. Dieses Thema wird im nächsten Abschnitt behandelt. Zuletzt geht es bei diesem Schritt noch darum festzulegen, nach welchen

Rechenregeln und wie die Indikatorausprägungen berechnet werden soll, unter anderem, wie statistische Unsicherheiten berücksichtigt werden. Das IQTIG legt dieser Berechnung bayesianische Modelle zugrunde (IQTIG 2022).

Die Entwicklung von Qualitätsindikatoren selbst stellt bereits ein aufwändiges Unterfangen dar. Im Vergleich zu dem Aufwand, der für die Datenerhebung und Berechnung von Indikatorausprägungen getätigt werden muss – gerade, wenn diese zu Zwecken der Qualitätssicherung und nicht nur für eine einmalige Leitlinienerhebung eingesetzt werden – ist der Entwicklungsaufwand jedoch zu vernachlässigen. Deshalb sollte eine **Machbarkeitstestung** in der Routine immer Teil der Indikatorentwicklung sein. Diese Machbarkeitstestung geht einher mit der methodischen **Gütetestung** der Indikatoren, da im Rahmen der Machbarkeitstestung auch Daten für die Gütetestung erhoben werden. Der Fokus der Machbarkeitstestung liegt darin, zu überprüfen, ob mit einem vertretbaren Aufwand eine objektive, reliable und valide Datenerhebung möglich ist und dieser Aufwand von den Betroffenen akzeptiert wird. Diese Aspekte sind auch Teil der methodischen Güteprüfung, die in Deutschland oftmals in Anlehnung an das Bewertungsinstrument QUALIFY durchgeführt wird, das von der Bundesgeschäftsstelle Qualitätssicherung gGmbH (BQS) als Vorläuferorganisation des IQTIG entwickelt wurde (Reiter et al. 2007). Überprüft werden die Kategorien Relevanz, Wissenschaftlichkeit (u.a. Evidenzbasierung) und Praktikabilität. Für einige der Kriterien (z.B. statistische „Unterscheidungsfähigkeit“) sind Daten notwendig, die – wenn nicht als Routinedaten vorhanden – oftmals nicht vorliegen und daher im Rahmen einer Machbarkeitstestung erhoben werden müssen.

Auf der Grundlage der beschriebenen Methodik wurde kürzlich ein breit konsentrierter methodischer Standard für die Entwicklung leitlinienbasierter Qualitätsindikatoren abgestimmt (Deckert et al. 2021). Speziell für die On-

kologie existiert ebenfalls eine ausführliche methodische Handreichung (Leitlinienprogramm Onkologie 2021a).

Wie erwähnt, entfallen bei diesen Standards die ersten Schritte der Auswahl von Versorgungsaspekten bis hin zur Ableitung von Qualitätszielen und -merkmalen, da dies schon im Rahmen der Leitlinienentwicklung bearbeitet wurde. Stattdessen wird zunächst Wert darauf gelegt, dass die Entwickler von Qualitätsindikatoren methodisch informiert und breit repräsentativ für die Betroffenen sind. Zudem wird in den Standards betont, dass die Autorinnen und Autoren von Leitlinien aufgrund der genannten Komplexität einzelner Entwicklungsschritte eine Unterstützung durch in der Methodik Versierte benötigen. Die Indikatorentwicklung selbst besteht dann aus einer kriteriengestützten Auswahl „potenzieller Qualitätsindikatoren (QI)“, die nach Konsentierung als „vorläufige QI“ verabschiedet und in der S3-LL veröffentlicht werden sollen. Nach anschließender erfolgreicher Testung erlangen Indikatoren dann den Status „finale QI“ (Deckert et al. 2021).

Dabei beschränken beide Standards die Indikatorentwicklung aus Effizienzüberlegungen auf solche Leitlinienempfehlungen, die mit einem Empfehlungsgrad A (= soll) konsentiert wurden.

4.4 Datenquellen für eine leitlinienbasierte Evaluation

Die genannten methodischen Standards empfehlen, jeweils zu prüfen, welche Datenquellen für die Berechnung der Indikatoren bereits vorhanden sind und ob neue Erhebungsmethoden zu entwickeln sind.

Grundsätzlich stehen folgende Quellen zur Diskussion:

1. Befragungen der Patientinnen und Patienten zu ihren Versorgungserfahrungen (siehe oben: „patientenorientierte Diagnose-

4 Leitlinienevaluation: Konzepte zur Definition und Messung von Leitlinienumsetzung und -adhärenz

- mitteilung“) oder aber zu den Langzeitergebnissen (siehe oben: „Gehfähigkeit 6 Monate postoperativ“).
2. Eigens für die Evaluation erhobene Daten, die von den Leistungserbringenden dokumentiert werden – entweder nur zum Zweck der Qualitätssicherung oder aber im Rahmen von Registern oder Panels, die zum Monitoring der Versorgung oder zur Forschung mit versorgungsnahen Daten angelegt wurden.
 3. Routinedaten, die zum Zweck der Leistungsabrechnung dokumentiert werden.
 4. In Papierform oder elektronisch dokumentierte Krankenakten-Daten, aus denen die zur Berechnung der Indikatoren notwendigen Daten extrahiert werden.

Alle genannten Datenquellen haben Vor- und Nachteile, die bei der Evaluation der Leitlinienadhärenz zu beachten sind. Grundsätzlich ist zu bemerken, dass nur bei erwartbar gleichförmiger Dokumentation und Kategorisierung über die Leistungserbringer hinweg bzw. gleichförmiger Interpretation von Fragen/Datenfeldern/Items mit diesen Daten verlässliche Aussagen zur Leitlinienadhärenz ableitbar sind.

Kurzgefasst sind folgende Punkte bei den verschiedenen Datenquellen wesentlich:

- Bei Befragungen sind vor allem die Gestaltung des Befragungsinstrumentes, die Durchführungsmodalität (persönlich, telefonisch, online) und die Repräsentativität der Stichprobe bedeutsam.
- Bei den von Leistungserbringern selbst eigens für die Evaluation erhobenen Daten ist vor allem die Akzeptanz des Aufwands und die Problematik möglicher Manipulation zu beachten.
- Routinedaten sind vor dem Hintergrund des primären Erhebungszwecks und der Dokumentationsqualität zu interpretieren. Für Deutschland hat die kürzlich abgeschlosse-

ne, vom Innovationsfonds geförderte, leider noch nicht abschließend publizierte OPTIMISE-Studie diese Problematik eindrucksvoll belegt, indem die zum Teil sehr niedrige Übereinstimmung zwischen den in Krankenakten dokumentierten Sepsisfällen auf der Basis Sepsis-relevanter Befunde und den in den Abrechnungsdaten dokumentierten Sepsisfällen aufgedeckt wurde (Schwarzkopf et al. 2020, 2022).

4.5 Zielparameter und Studientypen der Leitlinienevaluation

Die verschiedenen Datenquellen sind in unterschiedlichem Maß geeignet, die Leitlinienumsetzung, die Ausrichtung der Versorgung an den Leitlinien sowie die Effekte einer leitlinienkonformen Versorgung auf die für Patientinnen und Patienten relevanten Ergebnisse („outcomes“) zu messen.

Bei der **Messung der Leitlinienumsetzung** geht es darum, die in der oben genannten Definition erwähnten Parameter, also die Verbreitung von Leitlinien, die Kenntnis bzw. Bekanntheit der Leitlinien durch die jeweils betroffenen Gesundheitsprofessionen und die Haltung bzw. Akzeptanz gegenüber den Leitlinienempfehlungen zu analysieren. Solche Analysen werden am ehesten mit Methoden der empirischen Sozialforschung durchgeführt. Dokumentenanalysen könnten die genutzten Disseminationsformen zur Implementierung von Leitlinien untersuchen, während zur Bekanntheit und Akzeptanz von Leitlinien Leistungserbringende und im Hinblick auf die zum Teil vorhandenen Patienten-Leitlinien auch Patientinnen und Patienten befragt werden könnten. Letztlich wären hier das Repertoire der Implementationsforschung und die Methoden zur Evaluation komplexer Interventionen zu nutzen, da eine Leitlinienimplementierung typischerweise aus verschiedenen Elementen besteht, die in einem komplexen Kontext ange-

wendet werden (Wirtz et al. 2019; Skivington et al. 2021). Als experimentelles Forschungsdesign, mit dem die Umsetzung einer Leitlinie in Organisationen oder Regionen überprüft werden kann, bieten sich daher auch Cluster-randomisierte Studien an.

Bei der **Untersuchung der Leitlinienadhärenz** stehen zunächst die Versorgungsprozesse oder auch -strukturen im Vordergrund. Die Grundfrage ist, ob die Versorgung den jeweils empfohlenen Prozessen folgt und in den von einer Leitlinie empfohlenen Strukturen stattfindet. Als geeignete Zielparameter werden oftmals Anteile der Patientinnen und Patienten analysiert, die leitlinienkonform behandelt oder in den geeigneten Strukturen behandelt werden. Zur Beantwortung dieser Frage eignen sich verschiedene Studiendesigns. Zunächst kämen Querschnittsstudien auf der Basis aller oben genannten Datenquellen infrage, die jedoch keine Aussage darüber erlauben, ob eine leitlinienkonforme Behandlung das Resultat einer Leitlinieneinführung oder das Resultat anderer Interventionen oder aber ein Zufallsprodukt darstellt. Die datengestützte, sogenannte externe Qualitätssicherung, bei der oftmals aus Leitlinien abgeleitete Qualitätsindikatoren jährlich vom IQTIG berichtet werden, stellt ein prominentes Beispiel für Querschnittsstudien dar (siehe <https://iqtig.org/veroeffentlichungen/qidb/>).

Als Beobachtungsstudien kämen aber auch Fall-Kontroll-Studien und Kohortenstudien infrage. In einer Fall-Kontroll-Studie würde bei zwei Populationen, die leitlinienkonform bzw. nicht leitlinienkonform behandelt wurden, nachträglich eruiert, ob die jeweils Leistungserbringenden eine Leitlinie angewendet haben. Oder es würden prospektiv zwei Populationen, die sich darin unterscheiden, dass in einer eine Leitlinie eingeführt wurde und in der zweiten nicht, im Hinblick auf die Leitlinienkonformität der Behandlung untersucht. Diese Beobachtung kann von Beginn an geplant stattfinden oder aber in Form eines natürlichen Ex-

periments erfolgen, wenn vergleichbare Organisationen oder Populationen nachträglich zu identifizieren sind, die sich nur darin unterscheiden, dass in jeweils einer der Organisationen oder Populationen eine Leitlinie eingeführt wurde, ohne dies evaluativ zu begleiten.

Ein gängiges Studiendesign bei Leitlinienadhärenzprüfungen ist das prä-post-Design, mit dem in einer Organisation/Region vor und nach der Einführung einer Leitlinie der Anteil leitlinienkonform behandelter Patientinnen und Patienten bewertet wird. Auch bei diesem Design eignen sich im Grunde alle vorher genannten Datenquellen. Ein konsequentes frühes Beispiel aus Deutschland wurde von Albert et al. vorgestellt, die als eine der ersten Gruppen aus neu formulierten Leitlinienempfehlungen Qualitätsindikatoren abgeleitet haben und deren Einhaltung in Form eines prä-post-Designs evaluieren wollten (Albert et al. 2004). Zuletzt sind auch bei Untersuchungen der Leitlinienadhärenz Cluster-randomisierte Studien angezeigt, mit denen sich die Anteile einer den Leitlinien entsprechenden Behandlung am ehesten kausal auf die Intervention „Leitlinienimplementierung“ zurückführen lassen (Köpke et al. 2012).

Die Analyse der Leitlinienadhärenz stößt an ihre Grenzen, wenn multimorbide Patientinnen und Patienten betroffen sind, bei denen eine optimale Versorgung die Anwendung verschiedener Leitlinien voraussetzt. Dabei besteht oftmals die **Problematik konfligierender Leitlinienempfehlungen**, auf die u. a. Boyd et al. aufmerksam machten (Boyd et al. 2005). Die meisten Leitlinien berücksichtigen diese Problematik unzureichend (Hughes et al. 2013). Konflikte entstehen vor allem dadurch, dass die Befolgung der verschiedenen, für individuelle Patientinnen und Patienten jeweils zutreffenden Leitlinien eine zu hohe Belastung darstellt oder eine Polypharmazie mit steigender Gefahr problematischer Arzneimittelinteraktionen bedingt (Cohen-Stavi et al. 2022, 2021; Pedersen et al. 2019). Eine singuläre Betrachtung der Leit-

linienadhärenz ohne Berücksichtigung einer eventuell vorliegenden Multimorbidität wird demnach regelmäßig zu Fehlschlüssen führen. Deshalb ist eine Risikoadjustierung bei Leitlinienerhebungen unerlässlich.

Die Notwendigkeit der Risikoadjustierung spielt auch bei der **Evaluation der Leitlinienerfektivität**, also der Analyse potenzieller Leitlinienerfekte auf die an Patientinnen und Patienten „ablesbaren“ Ergebnisse (z.B. Sterblichkeit, Komplikationen, Messung von Morbiditäts-/Funktionalitätsparametern) oder von diesen berichteten Ergebnisse der Versorgung (z.B. Lebensqualität) eine wesentliche Rolle. Ein schwieriges Unterfangen stellt bereits die Verfügbarkeit objektiver, valider und reliabler Daten zu patientenseitigen physischen und psychischen Einflussfaktoren auf die Ergebnisse der Versorgung dar. Noch schwieriger ist die Berücksichtigung der Patientencompliance, wenn nur durch eine aktive Mitarbeit der Patientinnen und Patienten langfristig positive Ergebnisse erzielt werden können. Wenn beispielsweise die oben genannte Teilhabe auf der Basis der Gehfähigkeit 6 Monate nach einer leitlinienkonformen Kniegelenksendoprothesen-Implantation als Zielparame-ter der Leitlinienerfektivität betrachtet wird, dann werden ohne Trainingsfähigkeit, -bereitschaft und tatsächliches Training der Betroffenen kaum positive Effekte bzw. Unterschiede je nach Grad der Leitlinienkonformität der Behandlung messbar sein.

4.6 Qualitätsindikatoren in Leitlinien – Entwicklungsstand

Evaluationen der Leitlinienadhärenz sind also insgesamt aufwändig, aber möglich. Als Voraussetzung gilt, dass zu den einzelnen Leitlinien-Empfehlungen Messgrößen in Form von Indikatoren formuliert wurden. Letztlich sind solche Indikatoren – nach dem Motto „what gets measured gets done“ – ein wesentliches Mittel zur Förderung der Leitlinienimplementierung.

Wie steht es nun um die Umsetzung der in den methodischen Standards sowie im Leitlinienprogramm der Onkologie genannten Forderung, dass Leitlinien immer auch Qualitätsindikatoren enthalten sollten?

Zur Beantwortung dieser Frage wurden für diesen Beitrag alle im November 2022 im Leitlinienregister der AWMF (<https://register.awmf.org/de/start>) als S3-Leitlinie mit Stand 2019–2022 aufgeführten Leitlinien dahingehend untersucht, ob Qualitätsindikatoren aufgeführt sind. Das Ergebnis dieser Analyse zeigt die Tabelle 1.

Insgesamt zählt das Register 113 S3-Leitlinien mit Stand 2019–2022. Darunter befinden sich 26 Leitlinien des Leitlinienprogramms Onkologie der Deutschen Krebsgesellschaft, die übrigen 87 Leitlinien betreffen nicht-onkologische Krankheitsbilder. In 29 Leitlinien (25,7%) sind Qualitätsindikatoren zu finden, die vor allem aus den onkologischen Leitlinien stam-

Tab. 1 Umsetzungsstand Qualitätsindikatoren (QI) in S3-Leitlinien (LL) mit Stand ≥ 2019

| Aktualität/ „Stand“ | Anzahl LL | LL mit QI insgesamt | darunter QI im LL-Programm Onkologie |
|---------------------|-----------|---------------------|--------------------------------------|
| 2019–2022 | 113 | 29 | 23 |
| 2019 | 32 | 7 | 5 |
| 2020 | 25 | 6 | 5 |
| 2021 | 42 | 10 | 8 |
| 2022 | 14 | 6 | 5 |

men: In 88% der onkologischen, aber nur 7% der nicht-onkologischen Leitlinien sind Qualitätsindikatoren zu finden. Ein zeitlicher Trend hin zu einem Mehr an Indikatoren in den aktuellsten Jahren ist nicht ersichtlich. Schaut man sich die im Kapitel „Qualitätsindikatoren“ der Leitlinien aufgeführten Messgrößen genauer an, dann fällt darüber hinaus auf, dass für die im Onkologieprogramm genannten Indikatoren jeweils Zähler, Nenner und Referenzbereich definiert und die Informationsgrundlage angegeben wird, während bei den nicht-onkologischen Indikatoren zum Teil nur Statements zu finden sind, die noch einer exakten Ausformulierung bedürfen (z.B. Leitlinie „Durchführung des Epikutantests mit Kontaktallergenen und Arzneimitteln“).

4.7 Ausblick

Die Analyse des aktuellen Stands der Indikatorentwicklung in Leitlinien verdeutlicht, dass weder das 2009 publizierte Manual für Leitlinienautoren des Ärztlichen Zentrums für Qualität in der Medizin (ÄZQ) (Altenhofen et al. 2009), noch der kürzlich publizierte methodische Standard für die Entwicklung von Qualitätsindikatoren im Rahmen von S3-Leitlinien (Deckert et al. 2021) bisher eine breitere Resonanz bei Autorinnen und Autoren nicht-onkologischer Leitlinien erfahren haben. Arnold et al. führen als Hauptgründe dafür das Fehlen verbindlicher Vorgaben und Ressourcen an (Arnold et al. 2019). Dagegen findet sich in den onkologischen Leitlinien größtenteils eine Befolgung der entsprechenden Empfehlungen des „Methodenpapiers für das Leitlinienprogramm Onkologie“ (Leitlinienprogramm Onkologie 2021a), sodass die Deutsche Krebsgesellschaft die bisher publizierten Qualitätsindikatoren jeweils aktuell zusammenstellen kann (Leitlinienprogramm Onkologie 2022). Hier sind entsprechende Ressourcen und Strukturen vorhanden (s. Kap. 16).

Es scheint also nicht unmöglich, im Rahmen von Leitlinien auch Qualitätsindikatoren zu entwickeln. Ein wahrscheinlich wesentlicher Hinderungsgrund – die mangelhafte Förderung dieser zusätzlichen, aufwändigen Arbeit für Leitlinienautorinnen und -autoren – wurde mit der nun im Rahmen des Innovationsfonds möglichen finanziellen Förderung der Leitlinienentwicklung ausgeräumt. Diese Förderung könnte einen Schub für die Indikatorentwicklung und letztlich auch Leitlinienadhärenz bedeuten. Zumindest weisen die Kurzdarstellungen einiger – jedoch nicht aller – der geförderten Vorhaben darauf hin, sich an das Methodenpapier der Onkologie halten zu wollen (<https://innovationsfonds.g-ba.de/projekte/versorgungsforschung/>).

Für eine durchgängige Entwicklung von Qualitätsindikatoren im Rahmen von Leitlinienentwicklungen wäre bei Förderentscheidungen darauf zu achten, dass immer auch Qualitätsindikatoren unter Berücksichtigung der methodischen Standards entwickelt werden sollen. Der Implementierung von Leitlinien, der Messung der Leitlinienadhärenz und letztlich einer Verbesserung der Versorgung der Patientinnen und Patienten könnte so Vor-schub geleistet werden.

Literatur

- Albert U-S, Koller M, Lorenz W, Doherty J, Schulz K-D, Wagner U, Kopp I (2004) Implementierung und Evaluation von Leitlinien auf nationaler Ebene: Entwicklung eines Konzeptes für die Stufe-3-Leitlinie „Brustkrebs-Früherkennung in Deutschland“. Zeitschrift für ärztliche Fortbildung und Qualitätssicherung 98(5), 347–359
- Altenhofen I, Blumenstock G, Diel F, Döbler K, Geraedts M, Jäckel WH et al. (2009) Programm für nationale VersorgungsLeitlinien von BÄK, KBV und AWMF. Qualitätsindikatoren – Manual für Autoren. Neukirchen: Verl. Make a Book (ÄZQ-Schriftenreihe, Bd. 36). URL: <https://www.aezq.de/mdb/edocs/pdf/schriftenreihe/schriftenreihe36.pdf> (abgerufen am 03.08.2020)
- Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) – Ständige Kommission Leitlinien (2020) AWMF-Regelwerk „Leitlinien“. Berlin. URL: <https://>

4 Leitlinienevaluation: Konzepte zur Definition und Messung von Leitlinienumsetzung und -adhärenz

www.awmf.org/leitlinien/awmf-regelwerk.html (abgerufen am 26.11.2022)

Arnold K, Breuing J, Becker M, Nothacker M, Neugebauer E, Schmitt J, Deckert S (2019) Entwicklung leitlinienbasierter Qualitätsindikatoren: eine qualitative Studie zu Barrieren und förderlichen Faktoren aus der Sicht von S3-LeitlinienautorInnen. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 147–148, 34–44. DOI: 10.1016/j.zefq.2019.09.001

Boyd CM, Darer J, Boulton C, Fried LP, Boulton L, Wu AW (2005) Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *JAMA* 294(6), 716–724 DOI: 10.1001/jama.294.6.716

Cohen-Stavi CJ, Givon S, Key C, Molcho T, Balicer R, Shadmi E (2021) Guideline deviation and its association with specific chronic diseases among patients with multimorbidity: a cross-sectional cohort study in a care management setting. *BMJ Open* 11(1), e040961. DOI: 10.1136/bmjopen-2020-040961

Cohen-Stavi CJ, Key C, Molcho T, Yacobi M, Balicer RD, Shadmi E (2022) Mixed Methods Evaluation of Reasons Why Care Deviates From Clinical Guidelines Among Patients With Multimorbidity. *Medical Care Research and Review* 79(1), 102–113. DOI: 10.1177/1077558720975543

Deckert S, Arnold K, Becker M, Geraedts M, Brombach M, Breuing J et al. (2021) Methodischer Standard für die Entwicklung von Qualitätsindikatoren im Rahmen von S3-Leitlinien – Ergebnisse einer strukturierten Konsensfindung. *Z Evid Fortbild Qual Gesundhwes* DOI: 10.1016/j.zefq.2020.11.008

Geraedts M, Mehl C, Schmitz J, Siegel A, Graf E, Stelzer D et al. (2020) Entwicklung eines Indikatorensets zur Evaluation der Integrierten Versorgung Gesundes Kinzigtal. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 150–152, 54–64. DOI: 10.1016/j.zefq.2020.04.001

Geraedts M, Selbmann H-K, Ollenschläger G (2003) Critical appraisal of clinical performance measures in Germany. *Int J Qual Health Care* 15(1), 79–85. DOI: 10.1093/intqhc/15.1.79

Hasenbein U, Wallech C-W (2007) Was ist Leitlinienkonformität? Theoretische und methodische Überlegungen zu einem aktuellem Konzept der Versorgungsforschung und Qualitätssicherung. *Gesundheitswesen* 69(8–9), 427–437. DOI: 10.1055/s-2007-985388

Hughes LD, McMurdo MET, Guthrie B (2013) Guidelines for people not for diseases: the challenges of applying UK clinical guidelines to people with multimorbidity. *Age Ageing* 42(1), 62–69. DOI: 10.1093/ageing/afs100

IQTIG (2022) Methodische Grundlagen. Version 2.0. Unter Mitarbeit von Dr. D. Boywitt et al. Hg. v. IQTIG – Institut für Qualitätssicherung und Transparenz im Gesundheitswesen. Berlin. URL: https://iqtig.org/downloads/berichte-2/meg/IQTIG_Methodische-Grundlagen_Version-2.0_2022-04-27_barrierefrei.pdf (abgerufen am 13.02.23)

Joint Commission on Accreditation of Healthcare Organizations (1990) *Primer on indicator development and application. Measuring quality in health care.* Oakbrook Terrace, Ill.: JCAHO

Köpke S, Mühlhauser I, Gerlach A, Haut A, Haastert B, Möhler R, Meyer G (2012) Effect of a guideline-based multicomponent intervention on use of physical restraints in nursing homes: a randomized controlled trial. *JAMA* 307(20), 2177–2184. DOI: 10.1001/jama.2012.4517

Leitlinienprogramm Onkologie (2021a) Entwicklung von leitlinienbasierten Qualitätsindikatoren. Methodenpapier für das Leitlinienprogramm Onkologie. Hg. v. Deutsche Krebsgesellschaft e.V., Stiftung Deutsche Krebshilfe, Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) e.V. URL: <http://www.leitlinienprogramm-onkologie.de/methodik/informationen-zur-methodik/> (abgerufen am 30.11.2022)

Leitlinienprogramm Onkologie (2021b) Interdisziplinäre S3-Leitlinie für die Früherkennung, Diagnostik, Therapie und Nachsorge des Mammakarzinoms. Langversion 4.4 – Juni 2021, AWMF-Registernummer: 032-045OL. URL: https://register.awmf.org/assets/guidelines/032-045OL_S3_Mammakarzinom_2021-07.pdf (abgerufen am 28.11.2022)

Leitlinienprogramm Onkologie (2022) Onkologische Qualitätsindikatoren. Leitlinienbasierte Qualitätsindikatoren im Leitlinienprogramm Onkologie (OL). Hg. v. Deutsche Krebsgesellschaft e.V., Stiftung Deutsche Krebshilfe, Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) e.V. URL: https://www.leitlinienprogramm-onkologie.de/fileadmin/user_upload/Onkologische_Qualitätsindikatoren_6.0.pdf (abgerufen am 30.11.2022)

National Library of Medicine (2022) MeSH. Guideline Adherence. URL: <https://www.ncbi.nlm.nih.gov/mesh/68019983> (abgerufen am 26.11.2022)

Nothacker M, Stokes T, Shaw B, Lindsay P, Sipilä R, Follmann M, Kopp I (2016) Reporting standards for guideline-based performance measures. *Implementation Sci* 11, 6. DOI: 10.1186/s13012-015-0369-z

Pedersen RA, Petursson H, Hetlevik I (2019) Stroke follow-up in primary care: a Norwegian modelling study on the implications of multimorbidity for guideline adherence. *BMC family practice* 20(1), 138. DOI: 10.1186/s12875-019-1021-9

Piggott T, Langendam MW, Parmelli E, Adolffson J, Akl EA, Armstrong D et al. (2022) Integrating Quality Assurance and Quality Improvement With Guidelines: Systematic Stakeholder-Driven Development of an Extension of the Guidelines International Network-McMaster Guideline Development Checklist. *Annals of internal medicine* 175(5), 735–739. DOI: 10.7326/M21-3977

Reiter A, Fischer B, Köttling J, Geraedts M, Jäckel WH, Döbler K (2007) QUALIFY: Ein Instrument zur Bewertung von Qualitätsindikatoren. *Z Arztl Fortbild Qualitätssich* 101(10), 683–688. DOI: 10.1016/j.zgesun.2007.11.003

Schwarzkopf D, Fleischmann-Struzek C, Schlattmann P, Dorow H, Quart D, Edel A et al. (2020) Validation study of German inpatient administrative health data for epidemiological surveillance and measurement of quality of care for sepsis: the OPTIMISE study protocol. *BMJ Open* 10(10), e035763. DOI: 10.1136/bmjopen-2019-035763

I Grundlagen

- Schwarzkopf D, Rüdell H, Brinkmann A, Fleischmann-Struzek C, Friedrich ME, Glas M et al. (2022) The German Quality Network Sepsis: Evaluation of a Quality Collaborative on Decreasing Sepsis-Related Mortality in a Controlled Interrupted Time Series Analysis. *Frontiers in medicine* 9, 882340. DOI: 10.3389/fmed.2022.882340
- Sens B, Pietsch B, Fischer B, Hart D, Kahla-Witzsch HA, von Friedrichs V et al. (2018) Begriffe und Konzepte des Qualitätsmanagements – 4. Auflage. *GMS Medizinische Informatik, Biometrie und Epidemiologie* 14(1):Doc04. DOI: 10.3205/MIBE000182
- Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM et al. (2021) A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ* 374, n2061. DOI: 10.1136/bmj.n2061
- Wirtz MA, Bitzer EM, Albert U-S, Ansmann L, Bögel M, Ernstmann N et al. (2019) DNVF-Memorandum III – Methoden für die Versorgungsforschung, Teil 4 – Konzept und Methoden der organisationsbezogenen Versorgungsforschung. Kapitel 3 – Methodische Ansätze zur Evaluation und Implementierung komplexer Interventionen in Versorgungsorganisationen. *Gesundheitswesen* 81(3), e82-e91. DOI: 10.1055/a-0862-0588



Prof. Dr. med. Max Geraedts, M.Sc.

Max Geraedts, Arzt und Gesundheitswissenschaftler, ist Leiter des Instituts für Versorgungsforschung und Klinische Epidemiologie am Fachbereich Medizin der Universität Marburg. Arbeitsschwerpunkte: qualitätsbezogene Gesundheitsversorgungsforschung, Evaluation gesundheitspolitischer Maßnahmen.