ZUKUNFT DURCH QUALITÄT

Sektorenübergreifende Qualitätssicherung im Gesundheitswesen nach §137a SGB V

Allgemeine Methoden

im Rahmen der sektorenübergreifenden Qualitätssicherung
im Gesundheitswesen nach §137a SGB V

Version 4.0

Stand: 17. Februar 2015

AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH

Impressum

Herausgeber:

AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH

Thema:

Allgemeine Methoden im Rahmen der sektorenübergreifenden Qualitätssicherung im Gesundheitswesen nach §137a SGB V. Version 4.0

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum der Abgabe:

13. November 2014, redaktionelle Überarbeitung vom 17. Februar 2015

Signatur:

14-SQG-029

Hinweis:
Aus Gründen der leichteren Lesbarkeit wird im Folgenden auf eine geschlechtsspezifische Differenzierung verzichtet. Entsprechende Begriffe gelten im Sinne der Gleichbehandlung für beide Geschlechter.

Anschrift des Herausgebers:

AQUA – Institut für angewandte Qualitätsförderung und
Forschung im Gesundheitswesen GmbH
Maschmühlenweg 8-10 · 37073 Göttingen

Telefon: (+49) 0551 - 789 52 -0
Telefax: (+49) 0551 - 789 52-10

office@aqua-institut.de
www.aqua-institut.de

© 2015 AQUA-Institut GmbH

Inhaltsverzeichnis

Tabellenverzeichnis ... 8

Abbildungsverzeichnis ... 9

Abkürzungsverzeichnis ... 11

Präambel ... 15

1 Qualität, Qualitätssicherung und Qualitätsförderung im Gesundheitswesen ... 17

1.1 Gesetzliche Grundlage ... 17

1.2 Qualität und Qualitätsdimensionen ... 17

1.3 Messung der Versorgungsqualität mit Indikatoren ... 19

1.4 Anforderungen und Ziele ... 21

1.5 Sektorenübergreifende Qualitätssicherung ... 21

1.6 Förderung der Versorgungsqualität ... 22

1.7 Aktuelle Grenzen und Perspektiven der Qualitätssicherung ... 26

2 Aufgaben im Rahmen der einrichtungsübergreifenden Qualitätssicherung ... 28

2.1 Beteiligung an der Durchführung der einrichtungsübergreifenden Qualitätssicherung ... 28

2.1.1 Datenannahme für QS-Dokumentationen der Leistungserbringer ... 28

2.1.2 Datenannahme für Sozialdaten bei den Krankenkassen ... 29

2.1.3 Datenauswertung ... 29

2.1.4 Datenvalidierung ... 29

2.1.5 Bewertung von Auffälligkeiten und Durchführung von Qualitätssicherungsmaßnahmen ... 30

2.2 Weiterentwicklung der bestehenden einrichtungsübergreifenden Qualitätssicherung ... 32

2.2.1 Systempflege ... 33

2.2.2 Sektorspezifische Weiterentwicklung, Neuausrichtung und Neuentwicklung ... 34

2.2.3 Weiterentwicklung oder Neuausrichtung zu einem sektorenübergreifenden Verfahren ... 35

2.3 Neuentwicklung sektorenübergreifender QS-Verfahren ... 36

2.3.1 Allgemeine Ziele ... 36

2.3.2 Entwicklung von Indikatoren, Instrumenten und der notwendigen Dokumentation ... 36

2.3.2.1 Projektierung ... 36

2.3.2.2 Themenerschließung ... 37

2.3.2.3 Indikatorenrecherche ... 37

2.3.2.4 Auswahl und Bewertung von Indikatoren ... 38

2.3.2.5 Entwicklung der Instrumente und der Dokumentation ... 38

2.3.2.6 Empfehlung zur Umsetzung ... 38

2.3.3 Themenspezifische Bestimmungen ... 39

2.3.4 Erprobung von Indikatoren, Instrumenten und der notwendigen Dokumentation ... 39

© 2015 AQUA-Institut GmbH

2.4 Themenvorschläge und Themenauswahl ... 39
3 Auswahl und Beteiligung von Experten ... 41
3.1 Offenlegung von Interessenkonflikten ... 41
3.2 Wissenschaftlicher Beirat ... 41
3.3 Scoping-Workshop ... 41
3.4 RAM-Panel ... 42
3.4.1 Aufgaben und Ziele ... 42
3.4.2 Ausschreibung ... 42
3.4.3 Auswahlkriterien ... 43
3.4.4 Zeitbedarf und Honorierung der Experten ... 44
3.5 Bundesfachgruppen ... 44
3.6 Informelle Experten und ad-hoc-Arbeitsgruppen ... 45
3.7 Sonstige Formen der Einbeziehung von Expertise ... 45
4 Themenerschließungen ... 46
4.1 Ziele ... 46
4.1.1 Einschätzung der Versorgungsrelevanz ... 46
4.1.2 Einschätzung der Umsetzbarkeit ... 47
4.1.3 Einschätzung der Potenziale für Qualitätssicherung und -verbesserung ... 48
4.2 Inhalte ... 48
4.2.1 Definition der Zielpopulation ... 48
4.2.2 Medizinisch-pflegerische Versorgungsaspekte ... 49
4.2.3 Allgemeine Qualitätsziele ... 49
4.2.4 Versorgungspolitischer Rahmen und Rechtslage ... 49
4.2.5 Qualitätsprojekte und -initiativen ... 50
4.2.6 Versorgungspfad ... 50
4.2.7 Qualitätspotenziale ... 52
4.3 Methoden ... 52
4.3.1 Recherche nach aggregierter Evidenz ... 52
4.3.2 Einbeziehung von externer Expertise ... 54
4.3.3 Prüfung möglicher Datenquellen ... 54
5 Indikatorenrecherche ... 56
5.1 Indikatorenrecherche bei Institutionen und Agenturen des Gesundheitswesens ... 56
5.2 Systematische Literaturrecherche ... 57
5.2.1 Suchstrategie und Auswahl der Datenbanken ... 57
5.2.2 Suchmodell und datenbankspezifische Suchbegriffe ... 58
5.2.3 Auswahl der datenbankspezifischen Suchbegriffe ... 59
5.2.4 Testset ... 59

5.2.5 Auswahl der relevanten Literatur ... 60
5.2.6 Recherchebericht ... 60
5.2.7 Datenextraktion ... 60
6 Erstellung des Indikatorenregisters ... 61
6.1 Filterung des vorläufigen Indikatorenregisters ... 61
6.2 Ergänzungen des Registers ... 61
6.2.1 Ableitung von Indikatoren auf Basis von Leitlinien ... 62
6.2.2 Neuentwicklung von Indikatoren ... 62
6.2.3 Indikatoren zur Patientenperspektive ... 63
6.3 Aufarbeitung der Indikatoren des Registers ... 63
7 Entwicklung der Erhebungsinstrumente ... 67
7.1 Datenerhebungen bei/über Einrichtungen ... 67
7.1.1 Auslösung einer Dokumentation (QS-Auslösung) ... 68
7.1.2 Vollzähligkeit und Vollständigkeit ... 68
7.2 Nutzung von Routinedaten ... 69
7.2.1 Hintergrund ... 69
7.2.2 Routinedatennutzung in der einrichtungsübergreifenden Qualitätssicherung ... 71
7.2.3 Methodische Aspekte ... 73
7.3 Datenerhebungen bei/über Patienten ... 75
7.3.1 Ziele und Anforderungen von Patientenbefragungen ... 75
7.3.2 Entwicklung der Befragungsinstrumente ... 75
7.3.2.1 Itemgenerierung ... 76
7.3.2.2 Recherche nach Befragungsinstrumenten ... 77
7.3.2.3 Durchführung von Fokusgruppen ... 78
7.3.2.4 Einbindung von Fachexperten und Patientenvertretern ... 79
7.3.2.5 Kognitive Pretests ... 79
7.3.2.6 Durchführung eines Panelprozesses ... 80
7.3.3 Festlegung des Befragungszeitpunktes ... 80
7.3.3.1 Methoden der Veränderungsmessung für die Ergebnisqualität ... 80
7.3.3.2 Methodik zur Ermittlung des Befragungszeitpunkts (bzw. Befragungszeitraums) ... 81
8 Auswahl und Bewertung von Indikatoren ... 83
8.1 Bewertungskategorien und Gütekriterien ... 83
8.1.1 Relevanz ... 84
8.1.2 Klarheit/Verständlichkeit ... 85
8.1.3 Praktikabilität (Umsetzbarkeit) ... 85
8.1.4 Erhebungsaufwand ... 86
8.1.5 Risikoadjustierung ... 87

8.1.6 Messeigenschaften ... 88
8.1.7 Eignung für die öffentliche Berichterstattung ... 88
8.1.8 Nicht berücksichtigte Gütekriterien ... 89
8.2 RAND Appropriateness Method (RAM) ... 89
8.2.1 Grundlagen ... 89
8.2.2 Ablauf des RAM-Prozesses ... 90
8.2.3 Auswertung ... 93
8.3 Würdigung des Indikatorensets ... 94
8.4 „Verkürzte“ Auswahl- und Bewertungsprozesse ... 95
9 Übergreifende Aspekte von Entwicklungsleistungen ... 96
9.1 Stichproben ... 96
9.1.1 Bestandsaufnahme und grundsätzliche Kriterien zur Nutzung von Stichproben ... 96
9.1.2 Methodische Grundlagen zur Bildung von Zufallsstichproben ... 97
9.2 Datenvalidität und Dokumentationsqualität ... 99
9.2.1 Allgemeine Prüfkriterien der Dokumentationsqualität ... 99
9.2.2 Bewertung der Prüfergebnisse ... 101
9.2.3 Datenvalidierung ... 101
10 Vorbereitung der Umsetzung von Entwicklungsleistungen ... 103
10.1 Informationstechnische Aufbereitung ... 103
10.1.1 Spezifikation für die QS-Dokumentation ... 104
10.1.2 Spezifikation für Sozialdaten bei den Krankenkassen ... 109
10.2 Erprobung von Erhebungsinstrumenten ... 115
10.2.1 Erprobung der QS-Dokumentation ... 115
10.2.2 Erprobung von Patientenbefragungen ... 117
10.2.2.1 Deskriptive Analysen ... 117
10.2.2.2 Analysen zur Aussagekraft der Stichprobe (Non-Responder-Analyse) ... 117
10.2.2.3 Faktorenanalyse ... 117
10.2.2.4 Prüfung der Diskriminationsfähigkeit ... 117
10.2.2.5 Prüfung der Reliabilität ... 117
10.2.2.6 Empirische Überprüfung der Risikofaktoren ... 117
11 Auswertungen ... 118
11.1 Rechenregeln ... 118
11.2 Risikoadjustierung ... 118
11.2.1 Adjustierungsvariablen ... 118
11.2.2 Adjustierungsverfahren ... 119
11.2.3 Stand der Risikoadjustierung in der Qualitätssicherung ... 122
11.2.4 Weiterentwicklung der Risikoadjustierung ... 123

11.2.5 Anwendungsfälle ... 126
11.3 Area-Indikatoren ... 126
11.4 Qualitätsindizes ... 128
11.4.1 Fallzahl-Prävalenz-Problem ... 128
11.4.2 Lösungsansätze ... 130
11.4.3 Vor- und Nachteile der Verwendung von Indizes ... 131
11.5 Referenzbereiche ... 132
12 Berichterstattung ... 134
12.1 Hintergrund ... 134
12.2 Adressaten der Berichterstattung ... 134
12.2.1 Laien ... 135
12.2.2 Interessierte Öffentlichkeit ... 135
12.2.3 Leistungserbringer und Fachpublikum der Qualitätssicherung ... 135
12.2.4 Wissenschaftliche Experten ... 136
12.3 Berichtsarten ... 137
12.3.1 Berichte im Rahmen der gegenwärtigen einrichtungsübergreifenden Qualitätssicherung ... 137
12.3.2 Auswertungskonzepte für sektorenübergreifende QS-Verfahren ... 139
12.4 Formen der Ergebnisdarstellung ... 140
12.4.1 Aktuell verwendete Darstellungen ... 140
12.4.2 Weiterentwicklung der Qualitätsanalysen und -darstellungen ... 145
12.4.3 Allgemeinverständliche Aufbereitung von Qualitätsinformationen im Internet ... 148
Glossar ... 151
Literatur ... 165

Tabellenverzeichnis

Tabelle 1: Nomenklatur und Zuordnungsmerkmale der Weiterentwicklung der bestehenden einrichtungsübergreifenden QS-Verfahren 32

Tabelle 2: Beispiele für die kontinuierliche Anpassung von QS-Verfahren (Systempflege) 34

Tabelle 3: Entwicklungsarten der sektorspezifischen Weiterentwicklung 35

Tabelle 4: Matrix zur Einschätzung der (potenziellen) Verfügbarkeit von Qualitätsinformationen 54

Tabelle 5: Basisangaben des Indikatordatenblattes 63

Tabelle 6: Evidenzklassen und Studientypen 65

Tabelle 7: Für die QS nutzbare Datenbestände 72

Tabelle 8: Methoden kognitiver Interviews 79

Tabelle 9: Unterschiedliche Studienbewertungen des quasi-indirekten Vergleichs als Methode der Veränderungsmessung 81

Tabelle 10: Kategorien und Phasen der Auswahl und Bewertung von Indikatoren 83

Tabelle 11: Unterscheidung gebräuchlicher Verfahren zur Einbeziehung von Expertenmeinungen 89

Tabelle 12: Datenkategorien in der Datenlieferung 113

Tabelle 13: Teildateien der Datenlieferungen 115

Tabelle 14: Qualitätsindikator „Sterblichkeit“ für den Leistungsbereich Hüftgelenknahe Femurfraktur (Ergebnis 2009) 120

Tabelle 15: Systematik der Referenzbereiche in der externen stationären Qualitätssicherung im Erfassungsjahr 2013 133

Abbildungsverzeichnis

Abbildung 1: OECD-Qualitätsmodell (Arah et al. 2006) ... 18

Abbildung 2: Das Verfahren der Datenvalidierung am Beispiel der stationären Qualitätssicherung (Schubert et al. 2014) ... 29

Abbildung 3: Ablauf des Strukturierten Dialogs in der stationären QS (Köppen et al. 2014) ... 32

Abbildung 4: Meilensteine einer beauftragten Neuentwicklung (Indikatoren, Instrumente und Dokumentationen) ... 39

Abbildung 5: Erhebung qualitätsrelevanter Informationen ... 48

Abbildung 6: Beispiel für die Darstellung eines Versorgungspfades (hier: Nierenersatztherapie) ... 51

Abbildung 7: Beispielhaftes Modell für eine Suchstrategie ... 59

Abbildung 8: Datenquellen und Erhebungsinstrumente der Qualitätssicherung ... 67

Abbildung 9: Ablauf des RAM-Prozesses mit Zuordnung zu den einzelnen Bewertungsrunden ... 91

Abbildung 10: Bewertungsfeld für die Bewertungsrunde 1 ... 92

Abbildung 11: Bewertungsfeld für die Bewertungsrunde 2 ... 93

Abbildung 12: Serielles Datenflussmodell für die sektorenübergreifende Qualitätssicherung ... 105

Abbildung 13: Bereiche der Spezifikation der sektorenübergreifenden Qualitätssicherung ... 106

Abbildung 14: Beispielansicht eines XML-Schemas ... 108

Abbildung 15: GPacker, interaktive Oberfläche zur Verschlüsselung der XML-Daten ... 109

Abbildung 16: Spezifikationsdatenbank und die Ableitung der Anwendersichten ... 110

Abbildung 17: Datenfluss bei den Krankenkassen im Rahmen der Qesü-RL bzw. QSKH-RL ... 110

Abbildung 18: Serielles Datenflussmodell Sozialdaten bei den Krankenkassen (Qesü-RL, Anlage zu Teil 1, Abb. 2) ... 111

Abbildung 19: Prozesse bei den Krankenkassen ... 112

Abbildung 20: Risikoadjustierungsmethode der Qualitätsindikatoren (König et al. 2014) ... 123

Abbildung 21: Verteilung krankenhausspezifischer Sterblichkeitsraten ... 128

Abbildung 22: Unterscheidungsfähigkeit der Qualitätsindikatoren der externen stationären Qualitätssicherung ... 129

Abbildung 23: Krankenhausspezifischer Prävalenz des Qualitätsindex ... 131

Abbildung 24: Beispiel für ein Fehlerbalkendiagramm (Stichprobe) ... 141

Abbildung 25: Beispiel für ein Liniendiagramm (Bundesergebnisse 2013: N = 276 Fälle und 2012: N = 295 Fälle) ... 142

Abbildung 26: Beispiele für ein Fehlerbalkendiagramm (Benchmarkreport) ... 142

Abbildung 27: Benchmarkdiagramm (Beispiel für Krankenhäuser mit mindestens 20 Fällen) ... 143

Abbildung 28: Beispiel für einen Box-and-Whisker-Plot (Kastengrafik) ... 143

Abbildung 29: Beispiel für ein Säulendiagramm (Sentinel-Events) ... 144

Abbildung 30: Beispiel für ein Säulendiagramm (Differenz O – E) ... 144

Abbildung 31: Beispiel Geodarstellung ... 145

© 2015 AQUA-Institut GmbH

Abbildung 32: Beispiel eines Trichterdiagramms (Funnel Plot) 145

Abbildung 33: Beispiele eines Summenplots 146

Abbildung 34: Indikatorverlauf bei vierteljährlicher Auswertung und Verwendung eines gleitenden Durchschnitts von 4 Quartalen (Eckardt 2008) 147

Abbildung 35: Überlebenszeitanalyse nach Implantation von Hüftendoprothesen bei Patienten mit und ohne Diabetes 147

Abbildung 36: Beispiel für eine mögliche Darstellung der Patientenbefragungsergebnisse 148

Abbildung 37: Patientengerechte Darstellung der Krankenhaussuche 149

Abbildung 38: Darstellung der Fachinformationen 150

Abkürzungsverzeichnis

Abkürzung	Beschreibung
AES	Advanced Encryption Standard
AHRQ	Agency for Healthcare Research and Quality (USA)
AIRE	Appraisal of Indicators through Research and Evaluation
AOK	Allgemeine Ortskrankenkasse
AQUA	Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen
AQUIK	Ambulante Qualitätsindikatoren und Kennzahlen
ASA	American Society of Anesthesiologists
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
ÄZQ	Ärztliches Zentrum für Qualität in der Medizin
BAS	Bundesauswertungsstelle
BDSG	Bundesdatenschutzgesetz
BQS	Bundesgeschäftsstelle Qualitätssicherung. Institut für Qualität & Patientensicherheit GmbH
CINAHL	Cumulative Index to Nursing & Allied Health Literature
CSV	Comma Separated Values (Dateityp)
CUSUM	Cumulative sum; kumulative Summe
DAHTA	Deutsche Agentur für Health Technology Assessment
DAS	Datenannahmestelle
DAS-KK	Datenannahmestelle für die Krankenkassen
DIMDI	Deutsches Institut für medizinische Dokumentation und Information
DKG	Deutsche Krankenhausgesellschaft
DMP	Disease-Management-Programm
DRG	Diagnosis Related Groups (Diagnosebezogene Fallgruppen)
EbHC	Evidence-based Health Care
EBM	Einheitlicher Bewertungsmaßstab
Embase	Excerpta Medica Database
EPA	Europäisches Praxisassessment
esQS	Externe stationäre Qualitätssicherung
G-BA	Gemeinsamer Bundesausschuss
gematik	Gesellschaft für Telematikanwendungen der Gesundheitskarte mbH
G-I-N	Guideline International Network
GKV	Gesetzliche Krankenversicherung
GKV-SV	GKV-Spitzenverband
HTA	Health Technology Assessment

© 2015 AQUA-Institut GmbH

Abkürzung	Beschreibung
ICD	International Statistical Classification of Diseases and Related Health Problems
ID	Identifikationsnummer
IK	Institutionskennzeichen
IMIBE	Institut für Medizinische Informatik, Biometrie und Epidemiologie
INAHTA	International Network of Agencies of Health Technology Assessment
InEK	Institut für das Entgeltsystem im Krankenhaus
IQ Healthcare	Scientific Institute for Quality of Healthcare (Nijmegen, Niederlande)
IQM	Initiative Qualitätsmedizin e.V.
IQWiG	Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen
KBV	Kassenärztliche Bundesvereinigung
KHEntgG	Krankenhausentgeltgesetz
KID	Krankenkassenidentifizierende Daten
KKG	Kuratorium für Fragen der Klassifikation im Gesundheitswesen
KKS	Krankenkassenkommunikationssystem
KV	Kassenärztliche Vereinigung
KZBV	Kassenzahnärztliche Bundesvereinigung
LAG	Landesarbeitsgemeinschaft
LE	Leistungserbringer
LID	Leistungserbringeridentifizierende Daten
LKG	Landeskrankenhausgesellschaft(en)
LQS	Landesgeschäftsstelle Qualitätssicherung
MCMC	Markov Chain Monte Carlo
MEDLINE	Bibliografische Datenbank der National Library of Medicine
MeSH	Medical Subject Headings
Morbi-RSA	Morbiditätsorientierter Risikostrukturausgleich (§303 a–e SGB V)
MVZ	Medizinisches Versorgungszentrum
NHS	National Health Service
NICU	Neonatal Intensive Care Unit
O/E	Observed to Expected Ratio
OECD	Organisation for Economic Co-operation and Development
OPS	Operationen- und Prozedurenschlüssel
OR	Odds Ratio
PatBeteiligungsV	Patientenbeteiligungsverordnung
PCI	Perkutane Koronarintervention
PDCA	Plan-Do-Check-Act
PDF	Portable Document Format (Dateiformat)

© 2015 AQUA-Institut GmbH

Abkürzung	Beschreibung
PID	Patientenidentifizierende Daten
PKCS	Public Key Cryptography Standards
PKI	Public-Key-Infrastruktur
PKV	Private Krankenversicherung
PROM	Patient Reported Outcome Measurement
PROQOLID	Patient-Reported Outcome and Quality of Life Instruments Database
QEP	Qualität und Entwicklung in Praxen
Qesü-RL	Richtlinie (des G-BA) gemäß §92 Abs. 1 Satz 2 Nr. 13 i.V.m. §137 Abs. 1 Nr. 1 SGB V über die einrichtungs- und sektorenübergreifenden Maßnahmen der Qualitätssicherung
QiSA	Qualitätsindikatorensystem für die ambulante Versorgung
QNeu-RL	Vereinbarung über Maßnahmen zur Qualitätssicherung der Versorgung von Früh- und Neugeborenen
QOF	Quality and Outcomes Framework (Großbritannien)
QPot	Qualitätspotenzial
QS	Qualitätssicherung
QSD-RL	Richtlinie (des G-BA) zur Sicherung der Qualität von Dialyse-Behandlungen nach den §§136 und 137 Abs. 1 Nr. 1 SGB V
QSKH-RL	Richtlinie (des G-BA) gemäß §137 Abs. 1 SGB V i.V.m. §135a SGB V über Maßnahmen der Qualitätssicherung für nach §108 SGB V zugelassene Krankenhäuser
QSR	Qualitätssicherung mit Routinedaten
QUALIFY	Methode zur strukturierten Evaluation von Qualitätsindikatoren im Gesundheitswesen (BQS)
RAM	RAND Appropriateness Method
RAND	Research and Development (RAND Health, USA)
RCT	Randomized Controlled Trial
RL	Richtlinie
ROC	Receiver Operating Characteristic
RSA	Verfahren zur Datenverschlüsselung, entwickelt von R. Rivest, A. Shamir und L. Adleman
SGB	Sozialgesetzbuch
SMR	Standardisierte Mortalitätsrate
SQG	Sektorenübergreifende Qualitätssicherung im Gesundheitswesen
UCLA	University of California, Los Angeles, USA
VPN	Virtual Private Network
VST	Vertrauensstelle
VStG	Versorgungsstrukturgesetz
W3C	World Wide Web Consortium
WHO	World Health Organisation

© 2015 AQUA-Institut GmbH

Abkürzung	Beschreibung
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language Transformations

Präambel

Das AQUA-Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH ist ein interessenunabhängiges und neutrales Dienstleistungsunternehmen, das sich auf Qualitätsförderungsprojekte spezialisiert hat. Im Rahmen einer Kooperation von Wissenschaftlern der Universitäten Göttingen und Hannover wurde im Jahr 1993 die „Arbeitsgemeinschaft Qualitätssicherung in der ambulanten Versorgung" gegründet. Aus dieser Arbeitsgemeinschaft entstand 1995 das AQUA-Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH mit Sitz in Göttingen. Verankert in der wissenschaftlichen Qualitätsforschung entwickelt das Institut konkrete Konzepte und Strategien zur angewandten Qualitätsförderung sowie zur Qualitätssicherung und übernimmt die Umsetzung komplexer Großprojekte.

Im Anschluss an ein europaweit ausgeschriebenes Vergabeverfahren erhielt das AQUA-Institut vom Gemeinsamen Bundesausschuss (G-BA) am 1. September 2009 den Auftrag, ihn als „fachlich unabhängige Institution" im Sinne des §137a SGB V[1] bei der Durchführung seiner Qualitätssicherungsaufgaben zu unterstützen.

Zur wissenschaftlichen Fundierung der Entwicklung von Indikatoren, Instrumenten und der geforderten Dokumentation sind folgende feste Kooperationspartner als Nachunternehmer seit Ausschreibungsbeginn in den Auftrag eingebunden:

- Scientific Institute for Quality of Healthcare (IQ Healthcare), St. Radboud Universiteit Nijmegen, Niederlande (Prof. Michel Wensing). Das Institut unterstützt Recherchearbeiten zu Qualitätsindikatoren und Instrumenten sowie die Verfahrens- und Methodenentwicklung einschließlich Konzepten zur Evaluation und zu Patientenbefragungen.
- Abteilung Allgemeinmedizin und Versorgungsforschung, Universitätsklinikum Heidelberg (Prof. Joachim Szecsenyi). Das Institut unterstützt Recherchearbeiten zu Qualitätsindikatoren und Instrumenten (z.B. Patientenbefragungen) sowie die Entwicklung der Dokumentation.
- Institut für Medizinische Biometrie und Informatik, Universitätsklinikum Heidelberg (Prof. Meinhard Kieser). Das Institut unterstützt die Weiterentwicklung biometrischer Modelle, insbesondere zur Risikoadjustierung.

Details und aktuelle Informationen zum Aufbau der Qualitätssicherungsaufgaben des AQUA-Instituts sind auf der Webseite www.sqg.de abrufbar.

Die „Allgemeinen Methoden im Rahmen der sektorenübergreifenden Qualitätssicherung im Gesundheitswesen nach §137a SGB V" (kurz: Methodenpapier) beschreiben die gesetzlichen und wissenschaftlichen Grundlagen der Beteiligung des Instituts an der gesetzlichen Qualitätssicherung sowie den Prozess der geforderten Entwicklungsleistungen nach §137a Abs. 2 Satz 1 Nr. 1 und Nr. 2 SGB V.

Das Methodenpapier legt dar, welche Aufgaben das AQUA-Institut im Rahmen der Durchführung der einrichtungsübergreifenden Qualitätssicherung hat und auf welche Weise sektorenübergreifende Verfahren der Qualitätssicherung entwickelt werden. Es dient als verbindliche Arbeitsgrundlage für das Institut, seine Mitarbeiter, die Nachunternehmer und die externen Kooperationspartner. Weiterhin dient es dem Gemeinsamen Bundesausschuss als eine Grundlage zur Überprüfung der Entwicklungsleistungen des Instituts. Nicht zuletzt ermöglicht die transparente Darstellung grundlegender Arbeitsweisen einen öffentlichen Diskurs über die strukturierte Weiterentwicklung der externen Qualitätssicherung. In diesem Sinne ist das Methodenpapier keine statische Arbeitsgrundlage. Vielmehr wird es im Zuge gewonnener Erkenntnisse bei der Verfahrensentwicklung und neuer Forschungsergebnisse angepasst. Spätestens alle zwei Jahre, bei Bedarf auch früher, wird deshalb eine aktuelle, weiterentwickelte Fassung des Methodenpapiers erstellt. Die bis zur offiziellen Abnahme des Methodenpapiers 4.0 gültige Version (3.0) des Methodenpapiers wurde am 15. Februar 2013 veröffentlicht.

Für den vorliegenden Entwurf der Version 4.0 gilt die in der nicht mehr rechtskräftigen Fassung des §137a SGB V mit Stand vom 1. Januar 2012 festgelegte Regelung, nach der die „Kassenärztlichen Bundesvereinigun-

[1] Hier und im Folgenden wird im Regelfall Bezug genommen auf §137a SGB V in der Fassung vom 1. Januar 2012, da dieser die Grundlage des Vertrags zwischen dem G-BA und dem AQUA-Institut ist.

© 2015 AQUA-Institut GmbH

gen, die Deutsche Krankenhausgesellschaft, der Spitzenverband Bund der Krankenkassen, der Verband der privaten Krankenversicherung, die Bundesärztekammer, die Bundeszahnärztekammer, die Bundespsychotherapeutenkammer, die Berufsorganisationen der Krankenpflegeberufe, die wissenschaftlichen medizinischen Fachgesellschaften, die für die Wahrnehmung der Interessen der Patientinnen und Patienten und der Selbsthilfe chronisch kranker und behinderter Menschen maßgeblichen Organisationen auf Bundesebene sowie der oder die Beauftragte der Bundesregierung für die Belange der Patientinnen und Patienten“ bei der Entwicklung der Inhalte zu beteiligen sind.

Nach Zusendung des Entwurfs des Methodenpapiers erhalten die genannten Institutionen acht Wochen Zeit für schriftliche Stellungnahmen.

Die in diesem Methodenpapier erläuterten wissenschaftlichen und inhaltlichen Grundlagen werden in einer möglichst allgemeinverständlichen Form beschrieben. An einigen Stellen ist jedoch ein gewisses Maß an Fachwissen vorausgesetzt. Fragen zu den Ausführungen, aber auch Anregungen und Hinweise können jederzeit per E-Mail an das AQUA-Institut gerichtet werden (methoden@aqua-institut.de).

© 2015 AQUA-Institut GmbH

1 Qualität, Qualitätssicherung und Qualitätsförderung im Gesundheitswesen

Mit den Begriffen dieser Überschrift ist ein normativer Rahmen der Qualitätssicherung im Gesundheitswesen abgesteckt, in dem sich das AQUA-Institut als beauftragte Institution gemäß §137a SGB V bewegt. Das allerdings, was Qualität im engeren Sinne ist, welche Ziele der Qualitätssicherung inhärent sind und wo möglicherweise eine eher statisch vorgestellte „Sicherung" in die praktische Qualitätsförderung übergeht, ist weder gesetzlich geregelt noch gibt es dazu verbindliche Aussagen. Auch darüber, wie und wo Qualität mit welchen Erhebungsinstrumenten gemessen werden sollte oder überhaupt gemessen werden kann, gehen die Meinungen oft auseinander.

1.1 Gesetzliche Grundlage

Die gesetzlichen Grundlagen zur „Sicherung der Qualität der Leistungserbringung" (Qualitätssicherung) finden sich im Neunten Abschnitt des SGB V (§§135–139c). Nach §135a SGB V sind alle dem SGB V unterliegenden Leistungserbringer in Deutschland zur Qualitätssicherung bzw. zur Teilnahme an Qualitätssicherungsmaßnahmen verpflichtet. Soweit es den vertragsärztlichen, vertragszahnärztlichen und den Bereich der zugelassenen Krankenhäuser betrifft, fällt dem G-BA die Aufgabe zu, die konkreten Maßnahmen der Qualitätssicherung zu beschließen bzw. in Form von Richtlinien festzulegen.

Nach §137a Abs. 1 SGB V bestimmt der G-BA für die vertragsärztliche Versorgung und für zugelassene Krankenhäuser grundsätzlich einheitlich für alle Patienten durch Richtlinien nach §92 Abs. 1 Satz 2 Nr. 13 insbesondere die verpflichtenden Maßnahmen der Qualitätssicherung nach den §§135a Abs. 2, 115b Abs. 1 Satz 3 und 116b Abs. 3 Satz 3.

Seit dem 1. September 2009 wird er dabei durch das AQUA-Institut, als fachlich unabhängige Institution gemäß §137a SGB V, unterstützt. Demnach umfasst der Auftrag insbesondere folgende Aufgaben:

- Messung und Darstellung der Versorgungsqualität bei bestehenden Verfahren und möglichst sektorenübergreifende Entwicklung abgestimmter Indikatoren und Instrumente zur Vorbereitung neuer Verfahren
- Entwicklung der notwendigen Dokumentation für die einrichtungsübergreifende Qualitätssicherung unter Berücksichtigung des Gebots der Datensparsamkeit
- Beteiligung an der Durchführung der einrichtungsübergreifenden Qualitätssicherung und, soweit erforderlich, Einbeziehung entsprechender weiterer Einrichtungen
- Veröffentlichung der Ergebnisse der Qualitätssicherungsmaßnahmen in geeigneter Weise und in einer für die Allgemeinheit verständlichen Form

Das AQUA-Institut hat im Rahmen der gesetzlichen Qualitätssicherung keinerlei normsetzende Kompetenzen. Als den G-BA bei seiner Arbeit unterstützende Institution stellt das AQUA-Institut überwiegend (versorgungs-) wissenschaftliche Grundlagen für (versorgungs-)politische Entscheidungen bereit. Mithin fokussieren die Konzepte und Methoden des Instituts nicht primär auf die medizinische Gesundheitsversorgung, sondern auf die Gewinnung, die Synthese und die Bewertung gesundheitsrelevanter Informationen (Pfaff et al. 2011).

1.2 Qualität und Qualitätsdimensionen

Auch wenn der Ruf nach Qualität in der Medizin bzw. in der Gesundheitsversorgung allenthalben zu vernehmen ist und überdies Einigkeit darüber besteht, dass in jedem Fall eine „hohe" Qualität der Versorgung geboten ist und dass Qualität stetig „verbessert" werden soll, so entbehrt der Qualitätsbegriff in der Gesundheitsversorgung einer einheitlichen Definition (einen diesbezüglich noch immer aktuellen Überblick über die Diskussion bietet Gerlach (2001)). Lediglich auf einer abstrakten Ebene, wie sie beispielsweise in der DIN EN ISO 2005 formuliert ist, findet sich eine mehr oder weniger allgemeingültige Qualitätsdefinition. Demnach ist Qualität der Grad, in dem ein Satz inhärenter Merkmale (festgelegte) Anforderungen erfüllt (DIN EN ISO 2005).

© 2015 AQUA-Institut GmbH

Diese abstrakte Definition macht deutlich, dass Qualität ein Begriff ist, der nur vor dem Hintergrund konkret formulierter Anforderungen oder Zielsetzungen bestimmbar ist. Auch die weithin anerkannte Qualitätsdefinition des Institute of Medicine – „[...] quality of care is the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge“ (IOM et al. 1990) – lässt offen, was genau „erwünschte Ergebnisse” sind und wer diese unter welchen Bedingungen festlegt. In einer deutschen Version dieser Definition wird diesbezüglich gefordert, dass es sich hier um „vom Patienten erwünschte“ Ergebnisse handeln sollte (Klemperer 1996).

Damit ist die eigentliche Problematik des Qualitätsbegriffs im Gesundheitswesen angedeutet, denn unterschiedliche Akteure in unterschiedlichen Zusammenhängen haben oft sehr unterschiedliche, nicht selten sogar diametrale Ansichten über „erwünschte Ergebnisse“, also die „richtigen“ Ziele der Gesundheitsversorgung – ganz zu schweigen von einer Einigkeit über die „richtigen“ Wege zu den Zielen. Unter den gegebenen Bedingungen begrenzter Ressourcen gleicht eine Konsensfindung über das, was Versorgungsqualität beinhalten sollte, schon zwischen Leistungserbringern, Leistungsempfängern und Kostenträgern der viel zitierten Quadratur des Kreises.

Die praktisch vorherrschenden und teilweise grundsätzlich unterschiedlichen Vorstellungen darüber, worin sich die Qualität der Versorgung manifestiert, haben sich in der Qualitäts- bzw. Versorgungsforschung in mehrdimensionalen Modellen niedergeschlagen.

Die OECD beispielsweise konkretisiert den Qualitätsbegriff vor dem Hintergrund der Frage, welche Faktoren ausschlaggebend für die Leistungsfähigkeit eines Gesundheitssystems sind (Arah et al. 2006). Kerndimensionen der Qualität sind aus dieser Perspektive:

- Effektivität
- Sicherheit
- Patientenorientierung

Darüber hinaus sehen die Autoren den Zugang zur und die Koordination der Versorgung sowie auch den Aufwand (Kosten) als weitere Dimensionen der Leistungsfähigkeit des Gesundheitssystems. Übergeordnet sind allen Dimensionen die Kategorien Gerechtigkeit (z.B. gerechte Verteilung der Gesundheitsversorgung über die gesamte Population) und Effizienz (sowohl auf mikroökonomischer als auch auf makroökonomischer Ebene).

Leistungsfähigkeit des Gesundheitssystems						
	Dimension der Leistungsfähigkeit					
Versorgungsbedarf	Qualität			Zugang	Kosten/Abgaben	↑ Gerechtigkeit ↓
↓	Effektivität	Sicherheit	Patientenorientierung	Zugang zur Versorgung		
gesund bleiben						
gesund werden						
mit einer Krankheit oder Behinderung leben						
das Lebensende bewältigen						
← Effizienz →						

Abbildung 1: OECD-Qualitätsmodell (Arah et al. 2006)

Andere Autoren schlagen weitere Qualitätsdimensionen wie Angemessenheit, Rechtzeitigkeit (Commonwealth Fund 2004) und Mitarbeiterorientierung (Groene 2006) vor. Campbell et al. (2000) sehen den Zugang zur Versorgung als Teil der Qualität. Sie weisen außerdem darauf hin, dass die Relevanz verschiedener Qualitätsdimensionen auch davon abhängt, ob man auf der Ebene der Versorgung individueller Patienten oder gesamter Populationen argumentiert. Folglich sind sogar Konflikte zwischen diesen Ebenen denkbar:

> *Care for whole populations may potentially conflict with care for individual patients, because of prioritisation of resources towards particular health needs.* (Campbell et al. 2000: 1617)

Neben der grundsätzlichen Betrachtungsebene spielen auch persönliche und/oder professionelle Hintergründe der Betrachter eine Rolle bei der Frage danach, welche Aspekte der Versorgung eher als wichtig oder unwichtig angesehen werden. Je nach Perspektive kann dabei z.B. der ökonomischen Effizienz (Ökonom), der Patientenorientierung (Patient) oder der therapeutischen Effektivität (Arzt, Wissenschaftler) eine vorrangige Bedeutung zufallen. Auch hier sind also Konflikte denkbar, beispielsweise wenn eine als effektiv erachtete (evidenzbasierte) Therapie vom Patienten nicht gewünscht wird. Die Beispiele zeigen, wie vielschichtig und nicht selten durchaus gegensätzlich der Inhalt von Versorgungsqualität bestimmt wird.

1.3 Messung der Versorgungsqualität mit Indikatoren

In den letzten Jahrzehnten sind Qualitätsindikatoren, auf internationaler wie nationaler Ebene, zu den wichtigsten Mess- und Bewertungsinstrumenten der Gesundheitsversorgung avanciert. In der einfachsten Form können Qualitätsindikatoren als spezifische und messbare Elemente der Versorgung definiert werden, die zur Bewertung der Qualität verwendet werden können (Schneider et al. 2003; Marshall et al. 2002). Allerdings zeigt ein Indikator (lat. „indicare“ – zeigen) die Qualität zunächst nur an – d.h. er ist selbst kein direktes Maß für Qualität (Sens et al. 2007). Die Interpretation und Bewertung ist ein Prozess, in den die Perspektiven der verschiedenen Akteure, Experten und Nutzer des Gesundheitswesens einfließen. Genauer gesagt messen Indikatoren den Erfüllungsgrad von relevanten Qualitätsanforderungen, beispielsweise in Bezug auf die Effektivität der Versorgung, die Patientensicherheit, die Patientenorientierung und den Zugang zur Versorgung bzw. deren Koordination. Sie werden beispielsweise auf Basis von Patientenakten oder Befragungen ermittelt. Sie unterscheiden sich dadurch z.B. von (evidenzbasierten) Leitlinien, die Handlungskorridore für eine gute medizinische Vorgehensweise aufzeigen und für eine prospektive Steuerung der Versorgung verwendet werden können. Qualitätsindikatoren unterscheiden sich damit auch deutlich von Standards, die ihrerseits durchaus in Leitlinien oder klinischen Behandlungspfaden enthalten sein können und die ebenfalls prospektiv auf das Handeln einwirken sollen.

Die durch den Einsatz von Indikatoren in Aussicht gestellte Möglichkeit, auch sehr unterschiedliche Qualitätsaspekte wie Strukturen, Prozesse oder Ergebnisse zu messen und damit Qualität objektivieren zu können, hat zu einer nahezu flächendeckenden Anwendung im Gesundheitswesen geführt. Qualitätsindikatoren dienen z.B. der Bewertung und dem Vergleich von medizinischen Leistungen oder werden zur Gesundheitsberichterstattung und zur Qualitätsverbesserung eingesetzt. Sie spielen nicht nur eine wichtige Rolle in der Forschung, sondern werden auch bei Wahlentscheidungen für Patienten (öffentliche Berichterstattung), bei der Honorierung von Leistungen (Pay for Performance/P4P), im Vertragsmanagement sowie bei Akkreditierungs- und Zertifizierungsvorgängen berücksichtigt.

Trotz dieser steilen Karriere darf nicht vergessen werden, dass die Qualitätsmessung mit Indikatoren erst auf eine sehr kurze Geschichte zurückblickt. Es gibt weder ein „Kochbuch“ zur Erarbeitung und Anwendung von Indikatoren und Indikatorensystemen noch können mit ihnen alle Probleme der Gesundheitsversorgung gelöst werden. Die grundsätzliche Stärke von Indikatoren, nämlich die Möglichkeit, mit ihrer Hilfe die Komplexität der Realität auf einzelne und wiederholt messbare Größen zu reduzieren, ist zugleich auch ihre Schwäche.

Einerseits können Qualitätsindikatoren:

- Grundlage oder Voraussetzung für Steuerungsprozesse und Entscheidungsfindungen sein
- Vergleiche zwischen Leistungsanbietern ermöglichen – im zeitlichen Verlauf oder gegenüber einem Goldstandard (Benchmarking)
- Schwächen der Versorgung entdecken und ein konkretes Verbesserungspotenzial aufzeigen
- eine Diskussion über die Qualität der Versorgung und den effektiven Einsatz von Ressourcen anregen
- effektivere Instrumente zur Qualitätsbewertung sein als andere Methoden
- über erhöhte Transparenz Vertrauen fördern
- den Erfolg von Umstrukturierung oder Ablaufänderungen messen und belegen

© 2015 AQUA-Institut GmbH

Andererseits können Qualitätsindikatoren aber auch:

- eine fragmentierte und/oder einseitige Betrachtung der Medizin begünstigen (engsichtig)
- nur leicht messbare Aspekte der Versorgung einbeziehen und die eher subjektiven und schwerer messbaren Aspekte außer Acht lassen
- schwer interpretierbar sein – z.B. könnten scheinbare Versorgungsunterschiede mit zufälligen Schwankungen oder willkürlichen Fallverteilungen („case-mix") zusammenhängen, sodass sie keine wirklichen Unterschiede in der Qualität der Versorgung aufzeigen (Interpretationsspielraum)
- kostspielig und zeitaufwendig in der Entwicklung und Erhebung sein (Frage nach der Effizienz und der ökonomischen Rechtfertigung)
- Schuldzuweisungen fördern und dadurch die Motivation in den medizinischen Professionen senken
- dazu verleiten, sich auf gemessene Aspekte der Versorgung zu stützen und deswegen eher kurzfristige Ziele zu verfolgen, als eine Langzeitstrategie zu entwickeln

Ob oder inwieweit Qualitätsindikatoren zur Feststellung und ggf. Lösung von Problemen bzw. zur Erreichung von Zielen geeignet sind, hängt von einer realistischen Beurteilung nicht nur der Chancen, sondern auch der Grenzen und Risiken des Einsatzes von Indikatoren ab.

So können beispielsweise auch „gute" Qualitätsindikatoren in der praktischen Konsequenz Effekte bewirken, die den eigentlichen Intentionen entgegengesetzt sind. Hierzu gibt es in der Literatur zahlreiche Hinweise (Lester et al. 2011; Bratzler et al. 2007; Rodriguez et al. 2007; MacLean et al. 2006). Ein sehr anschauliches Beispiel gibt Schiavo-Campo (1999: 26):

> *If hospital subsidies are based on the length of patient waiting lists, hospital managers and doctors will keep non-critical cases waiting as long as possible and will spend inordinate amounts of time on other cases (higher quality for some, none for others); if performance instead is assessed by a number of patients treated, the quality of care may suffer.*

Leider sind nicht alle unerwünschten Auswirkungen von Indikatoren so leicht zu erkennen. Es ist allerdings generell zu erwarten, dass Anreize für ein unerwünschtes Verhalten mit dem Ausmaß der drohenden negativen Sanktionen für „schlechte" Indikatorergebnisse steigen.

Jenseits der Bewirkung von möglicherweise unerwünschten Effekten unterliegen auch Qualitätsindikatoren selbst Qualitätskriterien. Nach der sogenannten RUMBA-Regel (Wilson 1987: 75) sollten Indikatoren die folgenden Bedingungen erfüllen:

- **R**elevant for the selected problem
- **U**nderstandable for providers and patients
- **M**easurable with high reliability and validity
- **B**ehaviourable i.e. changeable by behavior
- **A**chievable and feasible

Ausgehend von diesem allgemein verständlichen und leicht nachvollziehbaren Kriterienkatalog stellt sich indes die weit kompliziertere Frage nach den konkreten Methoden der Bewertung und Prüfung von Indikatoren auf Relevanz, Verständlichkeit, Messbarkeit, Beeinflussbarkeit und Erreichbarkeit.

Auch unter evidenzbasierten Gesichtspunkten konkurrieren hier mehrere Methoden mit zum Teil ganz unterschiedlichen Schwerpunkten und Zielsetzungen. Soweit es um die Bewertung von „neuen" Indikatoren bzw. die Entwicklung von Indikatorensystemen im Rahmen der evidenzbasierten Gesundheitsversorgung geht, orientiert sich das AQUA-Institut an der RAND/UCLA Appropriateness Method (RAM), die sich gut bewährt hat und deswegen zum internationalen Goldstandard geworden ist (siehe Abschnitt 8.2).

Angesichts des medizinischen Fortschritts sowie zahlreicher sich ändernder Bedingungen, unter denen Indikatoren zum Einsatz kommen, müssen Indikatoren, die bereits angewendet werden oder sich in der Erprobung befinden, prinzipiell unter dem Vorbehalt empirisch-statistischer Prüfungen (siehe Abschnitt 8.1.6) stehen.

1.4 Anforderungen und Ziele

Auch wenn es grundsätzlich sinnvoll ist, unterschiedliche Perspektiven und Dimensionen für eine Konkretisierung der Qualitätsanforderungen zu berücksichtigen (Ose et al. 2008; Rubin et al. 2001) und eine möglichst breitflächige Messung der Qualität eine der Voraussetzungen zur Erschließung von Verbesserungspotenzialen ist (SVR-G 2007; SVR-E 2005; McGlynn 2003; SVR-KAG 2003), gibt es letztlich kein System, das alle genannten Qualitätsdimensionen misst. Das landesweit in Großbritannien implementierte Quality and Outcomes Framework (QOF) (Campbell et al. 2009; Campbell et al. 2007; McDonald et al. 2007) fokussiert überwiegend auf medizinische Indikatoren („clinical indicators"), die etwa 70 % aller Indikatoren ausmachen. Ergänzt werden die medizinischen Indikatoren durch Strukturindikatoren (etwa 20 %) und Indikatoren zur Patientenperspektive (etwa 10 %). Besonders die stärkere Einbeziehung der Patientensicherheit und der Patientenperspektive insgesamt werden in der jüngeren internationalen Diskussion gefordert (Grol et al. 2008).

Die Berücksichtigung von Qualitätsdimensionen allein reicht nicht aus, um den ebenso komplexen wie abstrakten Qualitätsbegriff greifbar zu machen. Bei Berücksichtigung sehr vieler unterschiedlicher Qualitätsanforderungen besteht möglicherweise sogar die Gefahr, wesentliche Ziele zu vernachlässigen oder widersprüchliche Qualitätsforderungen gleichrangig in den Raum zu stellen. Qualität braucht daher klare Ziele, um bestimmbar zu werden, Qualitätssicherung braucht Ziele, um Maßnahmen ergreifen zu können. Dabei stellt sich die Frage, auf welche Weise sinnvolle und gute Ziele für die Gesundheitsversorgung formuliert werden können.

Für das AQUA-Institut hat sich aus den Erfahrungen der letzten Jahre immer deutlicher herauskristallisiert, dass sich die „guten", d.h. versorgungsrelevanten Zielbestimmungen weniger aus theoretisch-abstrakten Dimensionen ableiten lassen, als dass sie auf der Grundlage konkret benannter Qualitätsziele sowie ggf. bekannter -defizite für definierte Themen gewonnen werden müssen (Broge et al. 2014; Willms et al. 2013). Die Perspektive der Qualitätsdimensionen dagegen ist eher eine sekundäre Perspektive, die auf „blinde" Flecken oder prinzipiell problematische Bereiche der Gesundheitsversorgung hinweisen kann.

Dieser Tatsache wird mittlerweile unter dem Begriff der Qualitätspotenziale Rechnung getragen, die auf einem Versorgungspfad abgebildet werden, der Patientenwege und idealtypische Versorgungsprozesse nachzeichnet.

1.5 Sektorenübergreifende Qualitätssicherung

Die Qualitätssicherungsaufgaben des AQUA-Instituts finden im Kontext der evidenzbasierten Gesundheitsversorgung (Evidence-based Health Care – EbHC) statt. Die EbHC trägt dem Sachverhalt Rechnung, dass die medizinische Gesundheitsversorgung der Gegenwart nicht mehr allein eine Frage der Effektivität medizinischer Technologien ist, sondern untrennbar mit den demografischen Entwicklungen sowie den organisatorischen und politischen Rahmenbedingungen des Gesundheitssystems verknüpft ist (Schrappe 2015).

Die sozialen Determinanten der Gesellschaft ändern sich. So nimmt das Durchschnittsalter der Patienten seit Jahren stetig zu. Dadurch steigt auch die Bedeutung von chronischen Krankheiten und Multimorbiditäten. Der medizinisch-technische Fortschritt eröffnet neue oder alternative Behandlungsmethoden, die es u.a. ermöglichen, viele Operationen mittlerweile ambulant durchzuführen. So hat sich die Anzahl der ambulanten Operationen im Krankenhaus seit Anfang der 1990er Jahre in etwa vervierfacht, während sich im gleichen Zeitraum die durchschnittliche Verweildauer fast halbiert hat (Bölt et al.). Gleichzeitig steigt die Bedeutung von neuen Versorgungsmodellen. Disease Management Programme (DMP) und Modelle zur integrierten Versorgung reagieren auf den gestiegenen Koordinationsbedarf der Versorgung, und „zwischen" den Sektoren entstehen Medizinische Versorgungszentren (MVZ) und spezialfachärztliche Versorgungen.

Der Gesetzgeber hat diesen Wandlungen durch §137a SGB V Rechnung getragen, in dem u.a. bestimmt ist, dass die zu entwickelnden Verfahren der Qualitätssicherung „möglichst sektorenübergreifend anzulegen sind". Damit verändert sich auch die bisherige, überwiegend „klinische" Perspektive der Qualitätssicherung, was Auswirkungen auf die „Messung" der Versorgung und damit auch die Konzeption von Indikatoren bzw. Indikatorensystemen hat. Durch die sektorenübergreifende Ausrichtung und die damit einhergehende Möglichkeit von Längsschnittbetrachtungen steht auch die Qualitätssicherung vor zahlreichen neuen Herausforderungen.

© 2015 AQUA-Institut GmbH

Dies betrifft vor allem die fall- bzw. patientenbezogene Zusammenführung von Daten, die an unterschiedlichen Orten in unterschiedlichen Sektoren und zu unterschiedlichen Zeitpunkten erhoben werden. Dabei sind die rechtlichen Anforderungen an den Datenschutz und die Veränderung von Datenflüssen nur der technisch-formale Teil der Herausforderung. In inhaltlicher Hinsicht stellen Längsschnittbetrachtungen oftmals ganz andere Anforderungen an Indikatoren, an Erhebungsinstrumente und Auswertungen. Der Wunsch nach möglichst sektorenübergreifender Messung trifft auf ein Gesundheitssystem, das, wie beispielsweise die Ausgestaltung des Vergütungssystems oder die unterschiedliche Verwendung von Kodiersystemen zeigen, sektorspezifisch ausgerichtet ist. Dies erschwert die Vergleichbarkeit von Leistungen, führt zu Problemen bei der QS-Auslösung bzw. den Erhebungsinstrumenten und erhöht den Aufwand für eine Datenvalidierung, während gleichzeitig schlanke, datensparsame und aufwandsarme QS-Verfahren gefordert werden.

Um diesen Herausforderungen zu begegnen, müssen innovative Konzepte entwickelt und erprobt werden, weil die bisherigen Erfahrungen zeigen, dass eine einfache Erweiterung der Methoden der externen stationären Qualitätssicherung in Richtung der sektorenübergreifenden QS nicht funktioniert. Das AQUA-Institut hat in den letzen Jahren mehrere innovative Konzepte für die sektorenübergreifende QS vorgestellt. Genannt seien hier Konzepte zur QS-Auslösung sowie zur Nutzung der elektronischen Gesundheitskarte; ein Konzept zur Datenvalidierung sowie Konzepte zur Nutzung von Routinedaten und von Patientenbefragungen.

Qualitätssicherung ist mithin die Summe aller Maßnahmen, die geplant und systematisch eingesetzt werden, um definierte Qualitätsanforderungen zu überprüfen, sicherzustellen bzw. sie zu erreichen. Bezogen auf die Gesundheitsversorgung einer ganzen Gesellschaft weist diese Definition darauf hin, dass Qualitätssicherung eine Aufgabe ist, die die größtmögliche Zusammenarbeit aller Beteiligten erfordert: Der Gesetzgeber muss die rechtlichen Rahmenbedingungen für die Qualitätssicherung schaffen und durchführende Institutionen benennen, medizinische Leistungserbringer müssen die Qualität ihrer Leistungen transparent machen, Kostenträger setzen Prioritäten, und Patienten müssen nicht nur ihre Perspektiven einbringen, sondern sich auch im Rahmen von Befragungen beteiligen. Und zuletzt müssen auch das Wissen und die Erfahrung von zahlreichen anderen Akteuren des Gesundheitswesens (Qualitätsinitiativen, medizinische Forschung, Versorgungsforschung usw.) eingebunden werden.

Da in diesem komplexen Geflecht verständlicherweise nicht immer alle Akteure an einem Strang bzw. zumindest nicht immer exakt in dieselbe Richtung ziehen, entspricht die oben genannte „Summe aller Maßnahmen“ zwangsläufig nicht immer dem Optimum. Gleichwohl ist das Bemühen um eine möglichst große Summe alternativlos und damit eine zentrale Herausforderung der Entwicklung und Weiterentwicklung der bundesweiten Qualitätssicherung.

Um die bundesweite Qualitätssicherung zu optimieren, macht das AQUA-Institut seine Methoden transparent, erstattet über seine Ergebnisse öffentlich Bericht, begründet seine Empfehlungen wissenschaftlich fundiert und bemüht sich darum, alle relevanten Akteure mit Expertise in die Arbeit der Qualitätssicherung einzubinden.

1.6 Förderung der Versorgungsqualität

Qualitätssicherung ist kein Selbstzweck, sondern soll so gestaltet sein, dass sie erwünschte Effekte im Gesundheitswesen erzielen kann und geeignet ist, die Versorgungsqualität positiv zu beeinflussen. Grundsätzliche Ziele sind:

- Kontinuierliche Verbesserung der Versorgungsqualität, insbesondere wenn Defizite bekannt sind
- Unterstützung bei der Einführung neuer Behandlungsgrundsätze, z.B. als flankierende Maßnahme zur Veröffentlichung von Leitlinien
- Vermeidung von Überversorgung (Ein Sonderfall möglicher Überversorgung liegt in einer unbegründeten Ausweitung der Indikationsstellung.)
- Aufdecken bzw. Kompensieren von Fehlanreizen und Sicherstellung, dass die Versorgung trotz Fehlanreizen (z.B. durch fallpauschalierte Vergütung) ausreichend ist
- Erkennen von Unterversorgung
- Erkennen von und Gegensteuerung bei Problemen mit der Patientensicherheit

© 2015 AQUA-Institut GmbH

An diesen Anforderungen muss sich die Ausgestaltung der Qualitätssicherung ausrichten. Voraussetzung ist zunächst die Erfassung bzw. Entwicklung entsprechender Indikatoren. Die alleinige Erfassung der Daten, die Ermittlung der Ergebnisse und deren Veröffentlichung erfüllen lediglich die Ansprüche eines Kontrollsystems und sind für sich allein genommen wenig geeignet, die Versorgungsqualität zu verbessern.
Folgende Elemente sind daher in einem Fördersystem mehr oder weniger relevant:

- Benchmarking (Wo steht die eigene Einrichtung im Vergleich bzw. wie schaffen es die Besten, ihre Ergebnisse zu erreichen und was kann ich davon lernen?)
- Feedback an den Leistungserbringer
- Eignung der Feedbackergebnisse für die Integration in ein internes Qualitätsmanagement der Leistungserbringer
- Benennung realistischer Ziele, um auf Ergebnisse zu reagieren und diese beeinflussen zu können
- Nachvollziehbarkeit beim Leistungserbringer, dass er für die Ergebnisse verantwortlich ist (Attribuierbarkeit)
- Möglichkeit der Sanktionierung, wenn trotz fortgesetzt vorhandener Qualitätsprobleme keine Veränderungen erfolgen oder die Mitarbeit verweigert wird

Maßnahmen zur Umsetzung der Qualitätssicherung beziehen sich generell auf das Grundmodell der Qualitätsverbesserung, den sogenannten Plan-Do-Check-Act (PDCA-Zyklus), in anderen Spielarten auch „Deming Wheel“ oder „Kreislauf der Qualitätsförderung“ genannt (Gerlach 2001). Demnach sind Qualitätsverbesserungen das Ergebnis eines geplanten Prozesses, bei dem – ausgehend von einem zielorientierten Planungsprozess und einer systematischen Analyse des Erreichungsgrades – Veränderungen schrittweise umgesetzt werden und dann festgestellt wird, ob tatsächlich eine Verbesserung eingetreten ist. In der Realität stößt man jedoch immer wieder auf zahlreiche Barrieren, die einer Veränderung entgegenstehen (Grol et al. 2003).

Viele Studien zeigen, dass bloßes Feedback oder die Herstellung von Transparenz über Versorgungsdaten geringe oder allenfalls moderate Verbesserungseffekte haben (Jamtvedt et al. 2006) und die alleinige frontale Fortbildung von Ärzten gar keine oder nur sehr wenig nachhaltige Veränderungen bewirken (Bloom 2005). Von daher kommen inzwischen zunehmend mehrperspektivische Ansätze zum Tragen (Grol et al. 2005b; Grol et al. 2003), in denen Informationen über organisatorische und/oder klinisch-medizinische Indikatoren z.B. mit Informationen aus Patientenbefragungen und/oder der Perspektive von Versorgungsteams verknüpft werden. Wichtig erscheinen auch das Konzept des sozialen Einflusses, die Reflexion über eigenes Handeln, Austausch und Unterstützung in Qualitätszirkeln (Wensing et al. 2009), ein persönliches und strukturiertes Gespräch mit Fachkollegen (Peer-Review, Kolloquium) oder persönliche Besuche vor Ort (O’Brian et al. 2007).

Welch deutliches Potenzial zur Qualitätsförderung wirksam genutzt werden kann, wenn Elemente wie ein mehrperspektivisches, indikatorengestütztes Feedback und Reflexionen im Team mit Unterstützung durch einen externen Visitor zusammengebracht werden, zeigt eine aktuelle Publikation aus dem ambulanten Bereich (Szecsenyi et al. 2011). Auch aus der externen stationären Qualitätssicherung sind einschlägige Beispiele bekannt. Bei der Versorgung von Patienten mit ambulant erworbener Pneumonie konnte nach der Einführung einer entsprechenden AWMF-S3-Leitlinie gezeigt werden, dass die Sterblichkeit im Krankenhaus bundesweit zwischen den Jahren 2006 und 2013 kontinuierlich, von 14,4 % (201.715 Patienten) auf 7,9 % (230.951 Patienten), gesunken ist (AQUA 2014e; BQS 2007). Dieses erfreuliche Ergebnis kann auch an der besseren Umsetzung von Kernempfehlungen der S3-Leitlinie nachvollzogen werden, die durch entsprechende Qualitätsindikatoren abgebildet sind.

Grundlage der bisher etablierten Qualitätssicherungsmaßnahmen sind einrichtungsindividuelle Auswertungen zu den Ergebnissen der Qualitätsindikatoren und ergänzende Informationen. Diese richten sich zunächst an alle Einrichtungen und sollen dort, bspw. im Rahmen des einrichtungsinternen Qualitätsmanagements, verwendet werden. Weitergehende Maßnahmen zielen auf Einrichtungen, die auffällige Ergebnisse bei den Qualitätsindikatoren aufweisen. Die konkrete Ausgestaltung der Maßnahmen ist zwar je nach Richtlinie unterschiedlich, aber allgemein geht es darum, einen Prozess zu initiieren, der zur Klärung der Auffälligkeiten führt, um ggf. geeignete Maßnahmen zur Qualitätsförderung einzuleiten.

Der Strukturierte Dialog, der als Grundlage zur Umsetzung qualitätsfördernder Prozesse in der stationären Qualitätssicherung dient, wirft hinsichtlich seiner Ergebnisse allerdings auch Fragen auf. Auch nach der Einführung eines eindeutigeren Bewertungssystems stellt sich ein beachtlicher Teil der statistisch auffälligen Ergebnisse nach der qualitativen Prüfung als unauffällig oder aufgrund fehlerhafter Dokumentation als nicht beurteilbar heraus. Von den 17.686 rechnerischen Auffälligkeiten im Erfassungsjahr 2012 erwiesen sich nur 10,2 % als qualitativ auffällig, 10,4 % der Auffälligkeiten waren nicht bewertbar(AQUA 2014f; Köppen et al. 2014).

Eine mögliche Ursache der gegenwärtigen geringen Spezifizität statistisch auffälliger Ergebnisse ist die geringe Fallzahl pro Einrichtung und Indikator. Diese Problematik ist international bekannt und aufgrund der geringen Prävalenz von Qualitätsindikatoren (Mor 2005: 342; Credé et al. 1990:199; Luft et al. 1986) zumeist unvermeidbar. Perspektivisch kann dieses Problem in Zukunft mithilfe von Indexbildungen, zum Teil auch durch längere Analyseintervalle und spezielle statistische Verfahren entschärft werden, da Qualitätsindizes eine höhere Prävalenz aufweisen (vgl. König et al. 2014; Dimick et al. 2012; Mukerji et al. 2012; Staiger et al. 2009) und somit geringere Fallzahlen für statistisch sichere Ergebnisse erfordern (vgl. Abschnitt 11.4).

Ein weiterer Ansatz zur Verbesserung der Effektivität der qualitätssichernden Maßnahmen könnte eine stärker edukativ ausgelegte Gestaltung des Strukturierten Dialogs sein. Insbesondere der Ansatz des Peer Review im Rahmen von Einrichtungsbesuchen gewinnt derzeit in Deutschland an Aufmerksamkeit durch Einführung in verschiedenen Qualitätsinitiativen[2] und wird unterstützt durch ein Curriculum der Bundesärztekammer (BÄK 2013).

Besuche von Einrichtungen werden in vielen Ländern, beispielsweise in Australien, Großbritannien, Kanada und Neuseeland, seit Längerem erfolgreich als Instrument im Rahmen der Qualitätsförderung und im Rahmen von Qualitätsmanagementverfahren genutzt. Dabei wird das Konzept in unterschiedlichen Gestaltungsformen in ambulanten hausärztlichen (Contencin et al. 2006; Grol et al. 2005a; van den Hombergh et al. 1998) und in fachärztlich-spezialisierten Praxen (Lombarts et al. 2003), in Praxen anderer Gesundheitsberufe (Hofhuis et al. 2006) sowie in Akutkrankenhäusern (Loer 2009) eingesetzt. In Deutschland gibt es bereits umfangreiche Erfahrungen im Bereich der Qualitätssicherung der stationären Rehabilitation (Baumgarten 2007; Meixner 2006).[3] Überdies liegen auch Erfahrungen aus mehr als 250 Praxen und medizinischen Versorgungszentren (MVZ) aus dem Zertifizierungsprogramm zu „Qualität und Entwicklung in Praxen“ (QEP) (KBV 2014)und aus mehr als 2.000 Praxen und MVZ im Rahmen des Europäischen Praxisassessments (EPA) vor (Broge 2008). Dieses Konzept findet eine hohe Akzeptanz bei den beteiligten Einrichtungen und hat nachweislich zu deutlichen Qualitätsverbesserungen geführt (Szecsenyi et al. 2011; Szecsenyi et al. 2009).

In einem aktuellen Review der Cochrane Collaboration (Ivers et al. 2012) wird vermerkt, dass die Kombination von Besuchen (audit) und Feedback auf professionelles Verhalten und patientenbezogene Outcomes in verschiedenen Studien als unterschiedlich effektiv eingeschätzt wird. Die Autoren kommen zu der Einschätzung:

> *Audit and feedback may be most effective when:*
>
> 1. *the health professionals are not performing well to start out with;*
> 2. *the person responsible for the audit and feedback is a supervisor or colleague;*
> 3. *it is provided more than once;*
> 4. *it is given both verbally and in writing;*
> 5. *it includes clear targets and an action plan.*

Die Diskussion um eine Effektivierung der Maßnahmen zur Qualitätssicherung wurde vom G-BA insbesondere mit der Beauftragung zur Entwicklung eines Qualitätssicherungsverfahrens „Arthroskopie am Kniegelenk“ aufgegriffen.[4] Neben der Entwicklung von Qualitätsindikatoren und Erhebungsinstrumenten wurden im Abschlussbericht zu diesem Projekt auch praktikable Ansätze zu einer sogenannten „Externen Begutachtung“ aufgezeigt,

[2] Vgl. die Empfehlungen zur Implementierung fachgruppenspezifischer und speziell auf die Rahmenbedingungen der vertragsärztlichen Versorgung ausgerichteter Peer-Review-Verfahren. http://www.kbv.de/html/10848.php (Zugriff am 16.02.2015).

[3] Allerdings werden hier, im Gegensatz zu den vorher genannten Verfahren, die Mitarbeiter des Leistungsträgers als Visitoren eingesetzt und nicht einrichtungsfremdes Personal.

[4] Beschluss des Gemeinsamen Bundesausschusses über eine Beauftragung der Institution nach §137a SGB V: Entwicklungen für ein Qualitätssicherungsverfahren zur Arthroskopie am Kniegelenk vom 21. März 2013. https://www.g-ba.de/downloads/39-261-1689/2013-03-21_AQUA-Beauftragung_QS-Verfahren-ASK.pdf (Zugriff am 14.10.2014).

die wesentliche der oben genannten Elemente für eine effektive Förderung und Verbesserung der Qualität enthält (AQUA 2014h).

Die Entwicklungsschritte für das Instrument der Externen Begutachtung erfolgten in Anlehnung an bereits bestehende Qualitätssicherungsinstrumente (Stichprobenprüfung der KV gem. §136 Abs. 2 SGB V, Peer-Review-Verfahren der Initiative Qualitätsmedizin e.V. (IQM) sowie Strukturierter Dialog der externen stationären Qualitätssicherung (§§11 und 12 QSKH-RL)). Hierbei wurden positive und bereits bewährte Teilaspekte synergistisch für das neu zu entwickelnde Instrument „Externe Begutachtung" genutzt.

Im vorgeschlagenen Verfahren der „Externen Begutachtung" wird das rechnerisch auffällige Ergebnis nicht aus einem einzelnen Indikator, sondern aus einem Qualitätsindex abgeleitet (vgl. Abschnitt 11.4). Eine wie bisher ausschließlich fallbezogene Prüfung dahingehend, ob die rechnerische Auffälligkeit auch tatsächlich mit einer qualitativen gleichzusetzen ist, steht nicht im Fokus. Es wird vielmehr davon ausgegangen, dass bei den durch den Qualitätsindex rechnerisch auffälligen Leistungserbringern (z.B. 95. Perzentil) Verbesserungspotenziale bestehen. Etwaige Defizite sollen analysiert und behoben werden. Die Erkenntnisse der „Externen Begutachtung" (identifizierte Verbesserungspotenziale und vereinbarte Verbesserungsmaßnahmen) können allen Leistungserbringern anonymisiert zurück gespiegelt werden und so gewinnbringend von allen Beteiligten genutzt werden.

Während der Entwicklung wurden deutlich, dass zwei Rahmenbedingungen die Ausgestaltung des Instruments stark beeinträchtigen:

- In einer Prüfung der rechtlichen Rahmenbedingungen kam die Rechtsabteilung des G-BA im August 2013 zu der Einschätzung, dass alle im Rahmen einer „Externen Begutachtung" angeforderten Daten (hier: Patientenakten) ausschließlich pseudonymisiert zur Verfügung gestellt und eingesehen werden dürfen. Diese Vorgabe erschwert den angestrebten Prüfprozess der ärztlichen Primärdokumentation deutlich.
- In den Sozialdaten bei den Krankenkassen ist derzeit sowohl für den stationären als auch für den ambulanten Bereich keine Fallnummer verfügbar. Daher können nicht gezielt jene Fälle ausgewählt werden, die zur rechnerischen Auffälligkeit geführt haben.

Unter diesen Bedingungen wurde das vorgeschlagene Instrument „Externe Begutachtung" so ausgestaltet, dass es sich aus vier Elementen zusammensetzt. Das erste Element sind die Ergebnisse der Einrichtung. Auf Basis dieser Ergebnisse (Rückmeldebericht) können sowohl der Leistungserbringer selbst als auch die externen Begutachter erste Qualitätsanalysen bzgl. der erbrachten Leistung vornehmen. Das zweite Element ist eine Selbsteinschätzung des Leistungserbringers. Alle Leistungserbringer mit einem auffälligen Indexergebnis werden um eine Selbsteinschätzung gebeten. In dieser sollen sie darlegen, wie sie die aufgezeigten Auffälligkeiten interpretieren (Selbst-/Fehleranalyse) und Vorschläge einbringen, wie sie ihre Qualität mithilfe gezielter Qualitätsmaßnahmen und -projekte verbessern wollen. Das dritte Element ist die Prüfung der Indikations-, Dokumentations- und Prozessqualität anhand von Auszügen aus pseudonymisierten Patientenakten. Das vierte Element ist ein kollegiales Gespräch zwischen Vertretern der Einrichtung und den Kollegen, die die „Externe Begutachtung" durchführen. Basis dieses Gesprächs sind die Rückmeldeberichte zu den Indikatoren aus den Sozialdaten bei den Krankenkassen und der Patientenbefragung, die Selbsteinschätzung des Leistungserbringers und die Ergebnisse der Fallanalysen. In diesem Gespräch werden die festgestellten Verbesserungspotenziale, aber auch die positiven Aspekte gemeinsam erörtert. Am Ende des Gespräches werden in einer Zielvereinbarung konsentierte Maßnahmen zur Qualitätsverbesserung und -förderung festgeschrieben (AQUA 2014h).

Das vorgeschlagene Instrument ist unter Berücksichtigung der vorhandenen Evidenz ein vielversprechender Weg zur Effektivierung der Maßnahmen zur Qualitätssicherung. Allerdings sind Erprobungs- und Evaluierungsphasen notwendig, um zu analysieren, ob diese Erwartung tatsächlich eintritt.

Insgesamt kann im Bezug auf sinnvolle Maßnahmen zur Qualitätsverbesserung festgehalten werden, dass zwar methodische Grundlagen auf Basis wissenschaftlicher Erkenntnisse bereitgestellt werden, aber bisher noch keine spezifischen Modelle und Kriterien zur Auswahl der am besten geeigneten Maßnahmen in einem konkreten Kontext bestehen. Faktisch lehnen sich Empfehlungen im Rahmen von Entwicklungsprojekten daher auch an verfügbare Strukturen und bereits etablierte Qualitätsinitiativen in dem jeweils betrachteten Bereich an.

1.7 Aktuelle Grenzen und Perspektiven der Qualitätssicherung

Die vorliegenden allgemeinen Methoden der Qualitätssicherung zielen darauf ab, unter den gegebenen Rahmenbedingungen der Qualitätssicherung und auf dem Stand der wissenschaftlichen Erkenntnisse umsetzbare und nützliche Verfahren mit möglichst geringem Aufwand zu etablieren. Das Aufwand-Nutzen-Verhältnis der Qualitätssicherung hängt allerdings auch von Rahmenbedingungen ab, die nicht von den Methoden selbst beeinflussbar sind. Nachfolgend sollen wesentliche dieser Aspekte skizziert werden, um Perspektiven für eine Weiterentwicklung der Qualitätssicherung aufzuzeigen.

Evidenzgrundlagen der Indikatorenentwicklung

Recherchen bilden eine wesentliche Grundlage der Entwicklung von Qualitätsindikatoren (siehe Abschnitt 4.3 und Kapitel 5). Sie sind die grundlegende Methode, um vorhandene Qualitätsdefizite zum Zeitpunkt der Entwicklung belegen zu können, und somit frühzeitig Aussagen über den Nutzen möglicher Qualitätssicherungsverfahren treffen zu können. Dabei ist die alleinige Information über Qualitätsdefizite im Kontext der Qualitätssicherung nicht ausreichend, sondern sollte ergänzt werden durch Informationen zu Qualitätsunterschieden zwischen Einrichtungen des Gesundheitswesens, zur Beeinflussbarkeit der Ergebnisse und zu geeigneten Maßnahmen der Qualitätsverbesserung.

Allerdings zeigen die Erfahrungen der bisherigen Entwicklungsprojekte, dass die Fragestellungen der Qualitätssicherung auf Basis vorhandener Studien und Leitlinien nur teilweise zufriedenstellend beantwortet werden können. Insofern ist es aus Sicht der Qualitätssicherung dringend erforderlich, die Leitlinienentwicklung zukünftig weiter zu intensivieren und Forschungsprojekte, insbesondere solche zur Versorgungsforschung, auch auf ihren möglichen Nutzen für die Qualitätssicherung hin auszurichten.

Weiterentwicklung der Erhebungsinstrumente

Auch wenn gegenwärtig intensiv an der Etablierung neuer Erhebungsinstrumente (Routinedaten, Patientenbefragungen) als Datengrundlagen für die Qualitätssicherung gearbeitet wird, sind Datenerhebungen in den Einrichtungen (QS-Dokumentation) auch zukünftig eine wichtige Grundlage zur Abbildung von Qualitätsindikatoren. Zumindest Teile der qualitätsrelevanten, klinischen Informationen lassen sich absehbar nicht über Routinedaten oder Patientenbefragungen abbilden.

Im sektorenübergreifenden Kontext hat sich die QS-Dokumentation allerdings bisher als nur eingeschränkt nutzbar erwiesen. Dies gilt insbesondere im ambulanten Bereich und hier sowohl im kollektiv- als auch im selektivvertraglichen Kontext. Gegenwärtig sind ambulante Abrechnungsdiagnosen häufig nicht geeignet, um die Zielpopulation eines Qualitätssicherungsverfahrens ausreichend genau zu identifizieren. Dies liegt einerseits daran, dass Diagnosen zeitlich nicht genau zuzuordnen sind und andererseits generell an bundesweit unterschiedlichen Abrechnungsvorschriften (vor allem im selektivvertraglichen Rahmen). Weil die Auslösung einer Dokumentationspflicht letztlich aber an abgerechnete Leistungen geknüpft ist, ist damit die Möglichkeit zur Etablierung eines längsschnittlichen QS-Verfahrens stark einschränkt. Derzeit werden sektorenübergreifende Ansätze daher vor allem in den Bereichen des ambulanten Operierens verfolgt. Dort gelingt auf Grundlage der verpflichtend zu dokumentierenden Operationen- und Prozedurenschlüssel (OPS) zumeist eine exaktere Abbildung des Leistungsgeschehens.

Die wichtigste Forderung, um die QS-Dokumentation hinsichtlich der genannten Einschränkungen zu verbessern, ist, dass qualitätsrelevante Parameter in der Abrechnung besser abgebildet sein müssen als es bislang der Fall ist. Entsprechende Vorschläge vonseiten der Qualitätssicherung waren jedoch noch nicht erfolgreich – weder im Zusammenhang mit einer besseren Abbildung der Nachsorge für Kataraktoperationen noch bei der vorgeschlagenen Einführung eines „Present On Admission“-Kennzeichens im Leistungsbereich *Pflege: Dekubitusprophylaxe* wurde den Erfordernissen der Qualitätssicherung bisher Rechnung getragen.

Die Abbildung zentraler qualitätsrelevanter Inhalte in Routinedaten ist sowohl für die Nutzung der QS-Dokumentation (z.B. Auslösung von Dokumentationen) als auch für die zukünftige Nutzung von Routinedaten von zentraler Bedeutung. Gegenwärtig finden Weiterentwicklungen der Abrechnungssysteme und der Qualitätssicherung jedoch weitgehend unabhängig voneinander statt. Dabei könnten unterschiedliche Sichtweisen auf das

gleiche Ereignis grundsätzlich gut vereinbar sein. Beispielsweise besteht eine Tendenz im Rahmen von Abrechnungssystemen, bestimmte Komplikationen aufgrund der Gefahr einer systematischen Höherstufung der Fallschwere (Case-Mix) mit dem Ziel der Erlössteigerung nicht zu honorieren. Bei gleichzeitiger Verbindung mit einer Qualitätsmessung könnte diese Gefahr verringert werden. Zwar führt die Kodierung einer Komplikation dann zu höheren Erlösen, umgekehrt erscheint sie aber auch als Qualitätsproblem.

Um die Abrechnungssysteme mit der Qualitätssicherung besser zu verbinden, könnte eine stärkere institutionelle Verankerung und Abstimmung der Expertise in beiden Systemen erfolgen. Eher kontraproduktiv wäre es dagegen, wenn jedes System für sich versucht, Expertise der anderen Seite in hohem Umfang aufzubauen. Dies würde voraussichtlich darin enden, dass Parallelstrukturen entstehen, in denen Abrechnungsexperten versuchen, das bessere Qualitätssystem aufzubauen und umgekehrt.

Eine zweite, ggf. komplementäre Möglichkeit, längsschnittliche Erhebungen zu ermöglichen, wäre es, auf der elektronischen Gesundheitskarte (eGK) zu vermerken, dass eine qualitätssicherungsrelevante, d.h. dokumentationspflichtige Leistung erbracht wurde. Ambulante und stationäre Nachbehandlungen könnten dann auch ohne spezifische Abrechnungsziffern erkannt und dokumentiert werden.

Solange die skizzierten Möglichkeiten zur Weiterentwicklung der QS-Dokumentation nicht umgesetzt werden, können qualitätsrelevante Informationen im sektorenübergreifenden Kontext nur lückenhaft abgebildet werden.

Regionale Versorgungskonzepte

In der wissenschaftlichen Diskussion um die Weiterentwicklung der Qualität im Gesundheitswesen spielen regionale Versorgungskonzepte und Kooperationen eine herausragende Rolle (Schrappe 2015; SVR-G 2007). Auch für eine Qualitätsmessung bieten solche Ansätze grundsätzlich Vorteile: Während es in einem arbeitsteiligen Gesundheitssystem oft schwierig ist, Verantwortung für Ergebnisse einer einzelnen Einrichtung zuzuordnen, kann dies in regionalen Versorgungsverbünden besser gelingen. Obwohl das hierbei zugrunde liegende Konzept der Area-Indikatoren grundsätzlich etabliert ist (vgl. Abschnitt 11.3), spielen diese Indikatoren in der Praxis bisher kaum eine Rolle. Entweder sind die notwendigen regionalen Versorgungsstrukturen bisher nicht vorhanden oder sie sind außerhalb der gesetzlichen Qualitätssicherung etabliert. Weil es deshalb de facto in der gesetzlichen Qualitätssicherung keine Ansprechpartner gibt, die man mit regionalen Versorgungsergebnissen konfrontieren könnte, muss bisher auf entsprechende Qualitätsindikatoren in der Regel verzichtet werden.

Dabei ist die gesetzliche Qualitätssicherung grundsätzlich dezentral aufgebaut. Zwar werden die Qualitätsindikatoren und andere Rahmenbedingungen einheitlich auf Bundesebene entwickelt. Die Umsetzung qualitätssichernder Maßnahmen erfolgt jedoch auf Länderebene. Damit sind grundlegende Strukturen für eine eher regionale Betrachtungsweise eigentlich vorhanden (vgl. Abschnitt 11.3. Für die Länderebene stehen bereits spezifische Auswertungen (Geodarstellungen, vgl. Abschnitt 12.4.1) zur Verfügung, die für einen Prozess des „Lernens von Besten“ genutzt werden könnten. Eine darüber hinausgehende systematische Einbeziehung neuer, regionaler Versorgungskonzepte in diese Strukturen wird gegenwärtig allerdings nicht diskutiert. Daher ist nicht absehbar, wann die Qualitätssicherung einen Beitrag zur Weiterentwicklung dieser Versorgungskonzepte liefern kann.

© 2015 AQUA-Institut GmbH

2 Aufgaben im Rahmen der einrichtungsübergreifenden Qualitätssicherung[5]

Gemäß §137a Abs. 1 SGB V ist das AQUA-Institut beauftragt, „Verfahren zur Messung und Darstellung der Versorgungsqualität für die Durchführung der einrichtungsübergreifenden Qualitätssicherung nach §115b Abs. 1, §116b Abs. 3 Satz 3, §137 Abs. 1 und §137f Abs. 2 Nr. 2 zu entwickeln, die möglichst sektorenübergreifend anzulegen sind.“ Nach §137 Abs. 2 Satz 1 Nr. 3 SGB V, ist das AQUA-Institut darüber hinaus an der Durchführung von Maßnahmen zur Qualitätssicherung beteiligt.

Neben den in der entsprechenden Richtlinie nach §92 Abs. 1 Nr. 13 i.V.m. §137a Abs. 2 Nr. 1 SGB V (Qesü-RL) formulierten und in themenspezifischen Bestimmungen vom G-BA zu konkretisierenden Aufgaben für die sektorenübergreifende Qualitätssicherung umfasst die Beteiligung an der Durchführung wesentlich die Weiterführung der bestehenden sektorspezifischen Qualitätssicherungsmaßnahmen in Krankenhäusern. Maßgeblich ist hier die Richtlinie des G-BA gemäß §137 Abs. 1 SGB V i.V.m. §135a SGB V über Maßnahmen der Qualitätssicherung für nach §108 SGB V zugelassene Krankenhäuser (QSKH-RL). Je nach Themenfeld und Aufgabenstellung kommen auch weitere Richtlinien des G-BA zum Tragen (bspw. QNeu-RL, QSD-RL). Umfang und Inhalt der Aufgaben sind vertraglich zwischen dem G-BA und dem AQUA-Institut geregelt.

Die folgenden Abschnitte beschreiben die aus diesem gesetzlich-vertraglichen Rahmen hervorgehenden konkreten Aufgaben des AQUA-Instituts und geben einen Überblick über die technischen und organisatorischen Prozesse, die in diesem Rahmen stattfinden. Die wissenschaftlichen und methodischen Grundlagen der Entwicklungs- und Weiterentwicklungsaufgaben der Qualitätssicherung sind nicht Gegenstand dieses Kapitels. Sie werden differenziert in den Kapiteln 4ff. dargelegt.

2.1 Beteiligung an der Durchführung der einrichtungsübergreifenden Qualitätssicherung

Das AQUA-Institut nimmt die zur Qualitätssicherung erforderlichen Daten bzw. Datensätze der Leistungserbringer von den zuständigen Landesstellen zur Auswertung entgegen und hält die dafür notwendige technische Struktur vor. Soweit dies Daten einrichtungs- und sektorenübergreifender Verfahren betrifft, werden die jeweiligen Modalitäten der Datenannahme und -auswertung über die Qesü-RL hinaus durch themenspezifische Bestimmungen des G-BA geregelt. Für die bestehenden sektorspezifischen Verfahren ist die Vorgehensweise einheitlich in der QSKH-RL geregelt.

Bei allen Qualitätssicherungsverfahren ist zwischen länderbezogenen (indirekten) und bundesbezogenen (direkten) Verfahren zu unterscheiden. Die konkreten Datenflüsse der sektorenübergreifenden QS-Verfahren sind in der Anlage 1 zur Qesü-RL geregelt, deren allgemeine Vorgaben in den themenspezifischen Bestimmungen eines Verfahrens konkretisiert, ergänzt oder modifiziert werden. Die Datenflüsse der sektorspezifischen QS-Verfahren sind in den §§6 und 7 der QSKH-RL geregelt.

2.1.1 Datenannahme für QS-Dokumentationen der Leistungserbringer

Die Datenannahme für Dokumentationen der Leistungserbringer in länderbezogenen Verfahren erfolgt durch die jeweiligen Datenannahmestellen nach §9 Abs. 1 Qesü-RL sowie den §§6 und 7 und Anlagen 2 und 3 der QSKH-RL. Diese Stellen müssen die auf Plausibilität, Vollzähligkeit und Vollständigkeit geprüften Daten zur weiteren Auswertung für die Bundesauswertungsstelle vorhalten bzw. über eine Vertrauensstelle weiterleiten. Für die bundesbezogenen (direkten) Verfahren der externen stationären Qualitätssicherung im Rahmen der QSKH-RL ist das AQUA-Institut die Datenannahmestelle. Bei den neuen, bundesbezogenen sektorenübergrei-

[5] Angesichts der unterschiedlichen Begrifflichkeiten bzw. deren unterschiedlicher Verwendung im Rahmen der Aufgabenbeschreibung in der Anlage 1.1 des Vertrags zwischen dem G-BA und dem AQUA-Institut, dem §137a SGB V mit Stand vom 1. Januar 2012 (rechtliche Grundlage des Vertrags zwischen dem G-BA und dem AQUA-Institut) und dem §137a SGBV in der aktuellen Fassung, lassen sich begriffliche Unschärfen im Kapitel 2 nicht vermeiden.

© 2015 AQUA-Institut GmbH

fenden Verfahren kann auch die Landesebene (oder eine andere beauftragte Stelle) als Datenannahmestelle fungieren.

2.1.2 Datenannahme für Sozialdaten bei den Krankenkassen

Die Datenannahme für die von den Krankenkassen für ein QS-Verfahren im Rahmen der Qesü-RL zu liefernden Sozialdaten (DAS-KK) erfolgt bundeseinheitlich durch eine vom G-BA beauftragte Annahmestelle nach §9 Abs. 1 Satz 5 Qesü-RL. Sie muss die auf Plausibilität, Vollzähligkeit und Vollständigkeit geprüften Daten zur weiteren Auswertung für die Bundesauswertungsstelle vorhalten bzw. über die Vertrauensstelle an sie weiterleiten, nachdem sie die leistungserbringeridentifizierenden Daten pseudonymisiert bzw. anonymisiert hat.[6]

2.1.3 Datenauswertung

Derzeit fungiert das AQUA-Institut als Datenannahmestelle gem. §7 QSKH-RL bzw. als auswertende Stelle gem. §8 QSKH-RL. Es übernimmt dementsprechend die statistische Auswertung der Daten nach bundeseinheitlichen Rechenvorschriften bzw. -regeln (vgl. Abschnitt 11.1), die das AQUA-Institut jährlich neu veröffentlicht. Für zukünftige sektorenübergreifende Verfahren findet sich die analoge Regelung in den §§7 und 8 Qesü-RL bzw. in den jeweils zu beschließenden themenspezifischen Bestimmungen der einzelnen Verfahren.

2.1.4 Datenvalidierung

Das AQUA-Institut prüft die Validität der von den Leistungserbringern übermittelten Daten der Qualitätssicherung. Für die bestehenden sektorspezifischen Verfahren ist die Datenvalidierung in §9 QSKH-RL geregelt (siehe Abbildung 2). Regelungen für sektorenübergreifende Verfahren finden sich in §15 Abs. 2 Qesü-RL bzw. sie werden zukünftig in entsprechenden themenspezifische Bestimmungen festgelegt.[7]

Im Rahmen der externen stationären Qualitätssicherung übernimmt das AQUA-Institut die Durchführung

- einer statistischen Basisprüfung (Vollzähligkeits-, Vollständigkeits- und Plausibilitätsprüfung) und
- eines Stichprobenverfahrens mit Datenabgleich, in dem die Übereinstimmung (Konkordanz) mit der Dokumentation des Leistungserbringers (Patientenakte) überprüft wird.

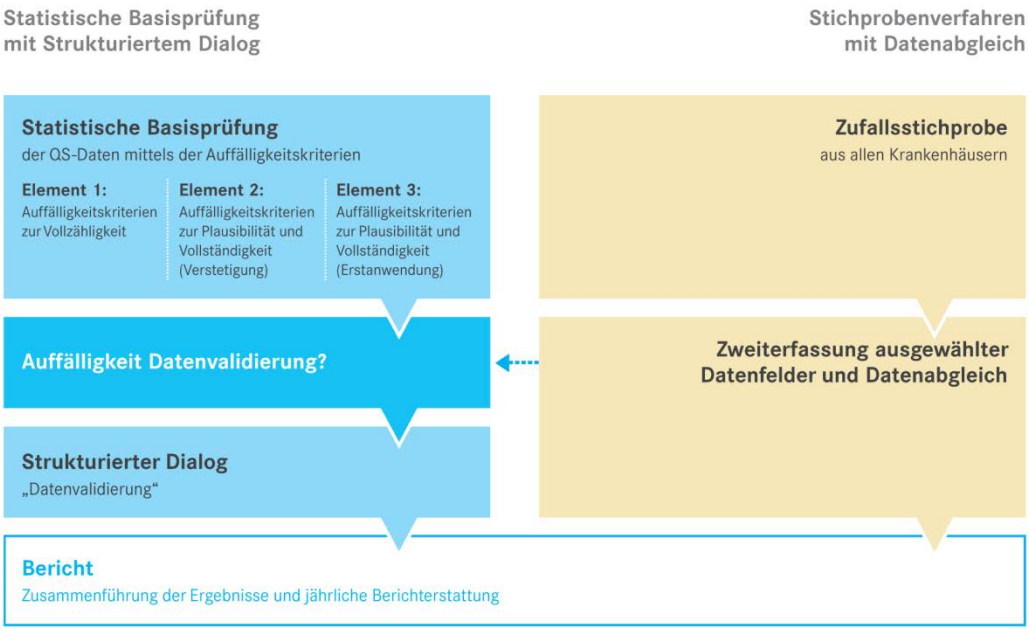

Abbildung 2: Das Verfahren der Datenvalidierung am Beispiel der stationären Qualitätssicherung (Schubert et al. 2014)

[6] Eine entsprechende Ergänzung der QSKH-RL um die erforderlichen Regelungen zur Nutzung von Sozialdaten befindet sich noch in Vorbereitung.
[7] Nachfolgende Ausführungen beziehen sich vornehmlich auf den stationären Sektor, da bisher noch kein sektorenübergreifendes Datenvalidierungsverfahren etabliert wurde.

Die Datenvalidierung der externen stationären Qualitätssicherung wird in der Regel für drei Leistungsbereiche pro Jahr durchgeführt, die gemäß §9 Abs. 2 QSKH-RL vom Unterausschuss Qualitätssicherung des G-BA (UA QS) festzulegen sind. Das AQUA-Institut und die Projektgruppe Datenvalidierung unterbreiten dem UA QS auf Grundlage der vom G-BA beschlossenen Kriterien entsprechende Vorschläge.

Dem Stichprobenverfahren mit Datenabgleich in den indirekten Verfahren sind pro Bundesland und pro Leistungsbereich jeweils 5 % aller Krankenhausstandorte zu unterwerfen. Standorte mit weniger als 4 Fällen in dem betroffenen Leistungsbereich sind davon ausgenommen.

Dem Stichprobenverfahren mit Datenabgleich in den direkten Verfahren sind pro Leistungsbereich mindestens 5 % der Standorte in den ausgewählten Leistungsbereichen zu unterwerfen. Hierbei sind je Leistungsbereich mindestens 4 Standorte und insgesamt mindestens 40 Fälle, bezogen auf die von allen Krankenhäusern in dem ausgewählten Leistungsbereich dokumentierten Fälle, in das Stichprobenverfahren einzubeziehen. Für die sektorenübergreifende Qualitätssicherung sollen analog bundesweit einheitliche Regelungen gelten.

Der Statistischen Basisprüfung sind alle in den festgelegten Leistungsbereichen verpflichteten Krankenhäuser und Arztpraxen zu unterwerfen. Ein Stichprobenverfahren erfolgt in der Regel im zweiten Jahr der Datenerhebung. Nach einer Bewertung der Ergebnisse der Datenvalidierung wird ggf. eine Empfehlung des AQUA-Instituts an den Unterausschuss Qualitätssicherung ggf. für eine erneute Datenvalidierung ausgesprochen (z.B. zwei oder drei Jahre später).

Das AQUA-Institut ermittelt bis Mitte April des dem Erfassungsjahr nachfolgenden Jahres mittels eines zu dokumentierenden Zufallsverfahrens für jeden einzelnen ausgewählten Leistungsbereich die Krankenhäuser und Arztpraxen, die dem Datenabgleich unterworfen werden. Zum Abgleich der gemeldeten Qualitätssicherungsdaten führen die auf Landesebene beauftragten Stellen mittels Einsicht in die Patientenakte eine Zweiterhebung von Qualitätssicherungsdaten anhand von 20 zufällig ausgewählten Patientenakten je Krankenhaus und Arztpraxis durch und vergleichen die Ergebnisse mit den zuvor übermittelten Daten. Nicht verfügbare Patientenakten sind zu dokumentieren und vom Leistungserbringer zu begründen (vgl. §9 Abs. 4 QSKH-RL).[8]

Gemäß §15 Abs. 1 QSKH-RL haben „die Institution nach §137a SGB V und die auf Landesebene beauftragten Stellen [...] über den Strukturierten Dialog und das Datenvalidierungsverfahren bis zum 15. März des auf den Beginn des Strukturierten Dialogs folgenden Jahres zu berichten. Dieser Bericht der auf Landesebene beauftragten Stellen wird nach bundeseinheitlichen Vorgaben als Datenbank erstellt und ihr Inhalt elektronisch, in maschinenlesbarer Form an die Institution nach §137a SGB V übersendet. Die Spezifikation für die Übersendung nach Satz 2 wird auf der Homepage der Institution nach §137a SGB V bis zum 30. Juni bekanntgegeben."

Gemäß §15 Abs. 2 fasst die Institution nach §137a SGB V „die Berichte nach Absatz 1 zusammen und erstellt bis zum 15. Mai des auf den Beginn des Strukturierten Dialogs folgenden Jahres einen Abschlussbericht über die Ergebnisse des Strukturierten Dialogs sowie des Datenvalidierungsverfahrens und veröffentlicht diese nach Freigabe durch den Unterausschuss Qualitätssicherung auf ihrer Homepage."

2.1.5 Bewertung von Auffälligkeiten und Durchführung von Qualitätssicherungsmaßnahmen

Liegt das berechnete Ergebnis von Qualitätsindikatoren außerhalb eines vorgegebenen Bereichs, dann wird dies als rechnerische Auffälligkeit bezeichnet. Die hierfür notwendigen Referenzbereiche (vgl. Abschnitt 11.5) werden gemäß §10 Abs. 1 Satz 1 QSKH-RL vom AQUA-Institut festgelegt. Analoge Festlegungen für sektorenübergreifende Verfahren erfolgen gem. §3 Satz 2 Nr. 9 Qesü-RL im Rahmen der themenspezifischen Bestimmungen. Einrichtungen können sowohl durch die Analyse der Qualitätssicherungsdaten als auch im Rahmen der Datenvalidierung rechnerisch auffällig werden. Ergänzende Regelungen werden im Rahmen der Qesü-RL zudem in den themenspezifischen Bestimmungen festgelegt.

[8] Nur für Krankenhäuser gilt: Soweit eine direkte Einsicht in Patientenakten aus Gründen des im jeweiligen Bundesland geltenden Datenschutzrechts nicht zulässig ist, ist das Krankenhaus aufgefordert, eine qualifizierte Mitarbeiterin oder einen qualifizierten Mitarbeiter zu bestimmen, der die Informationen aus der Patientenakte unter Wahrung der Anonymität der Patientin oder des Patienten auf Fragen der die auf Landesebene beauftragte Stelle vertretenden Person in deren Anwesenheit unmittelbar weitergibt. Für das Betreten des Krankenhauses durch Personen, die die auf Landesebene beauftragte Stelle vertreten, ist das vorherige Einverständnis des Krankenhauses einzuholen.

Referenzbereiche können entweder durch einen festen Wert definiert (fixer Referenzbereich) oder durch die Verteilung der Krankenhausergebnisse festgelegt sein (Perzentil-Referenzbereich). Da eine rechnerische Auffälligkeit nicht zwingend auf tatsächliche Qualitätsmängel zurückzuführen ist, werden die rechnerischen Auffälligkeiten von den zuständigen Stellen bewertet bzw. im Dialog mit den Einrichtungen geklärt.

Eine Sonderrolle spielen die sog. Sentinel-Event-Indikatoren, die sehr seltene und schwerwiegende Ereignisse erfassen. Diese Ereignisse werden grundsätzlich als auffällig eingestuft, d.h. jeder aufgetretene Fall muss bewertet werden.

Die Bewertung von rechnerischen Auffälligkeiten und die Durchführung von Qualitätssicherungsmaßnahmen soll bei allen länderbezogenen Verfahren im Rahmen der QSKH-RL den beauftragten Stellen auf Landesebene (§14 Abs. 1 QSKH-RL) obliegen bzw. im Rahmen der Qesü-RL den Landesarbeitsgemeinschaften (§17 Abs. 1 Qesü-RL) Die Ergebnisse der durchgeführten QS-Maßnahmen sollen der Bundesstelle übermittelt und von dieser zusammen mit den Ergebnissen der direkten Verfahren veröffentlicht werden.

Für die direkten Verfahren ermittelt die Auswertungsstelle auf Bundesebene die rechnerischen Auffälligkeiten und leitet die entsprechenden Maßnahmen nach den §§11 und 12 QSKH-RL (für die bestehenden sektorspezifischen Verfahren) bzw. nach §17 Abs. 2–6 Qesü-RL ein (siehe Abbildung 3).

Vor der Einleitung von qualitätssichernden bzw. qualitätsfördernden Maßnahmen – und unabhängig von Unterschieden in der Terminologie der beiden zur Anwendung kommenden Richtlinien – wird den betreffenden Leistungserbringern prinzipiell die Gelegenheit zur schriftlichen Stellungnahme eingeräumt, um die Auffälligkeiten zu erklären oder auf bereits unternommene Schritte zur Beseitigung erkannter Probleme hinzuweisen. Gegebenenfalls kann das schriftliche Stellungnahmeverfahren durch ein persönliches Gespräch zwischen der Einrichtung und der für die QS-Maßnahme zuständigen Stelle ergänzt werden. Sind bereits im Vorfeld Defizite bekannt, können gezielte Maßnahmen empfohlen werden, die über den Strukturierten Dialog hinausgehen (Teilnahme an Qualitätszirkeln, Peer Reviews usw.; vgl. §17 Abs. 3 Qesü-RL).

Wenn die Auffälligkeiten im genannten Rahmen nicht geklärt werden können, entscheiden die zuständigen Stellen in Absprache mit den in §14 QSKH-RL bzw. §17 Abs. 1 Qesü-RL benannten Stellen bzw. den entsprechend benannten Expertengruppen über die zu veranlassenden Maßnahmen nach §11 QSKH-RL oder die Qualitätssicherungsmaßnahmen der Maßnahmenstufe 1 nach §17 Abs. 3 Qesü-RL .

Sollten sich Hinweise auf schwerwiegende einzelne Missstände ergeben und/oder kann die betroffene Einrichtung etwaige Zielvereinbarungen der Maßnahmenstufe 1 nicht erfüllen oder verweigert sie die Zusammenarbeit, können die zuständigen Stellen über die Anwendung von Maßnahmen nach den §§12 und 13 QSKH-RL oder Maßnahmen der Stufe 2 nach §17 Abs. 4 Qesü-RL entscheiden.

Zum Abschluss des Strukturierten Dialogs nach QSKH-RL bzw. der qualitätssichernden Maßnahmen nach Qesü-RL müssen die Ergebnisse bewertet werden. Dafür stehen unterschiedliche Einstufungen zur Verfügung, die am Beispiel der bereits durchgeführten Maßnahmen nach QSKH-RL dargestellt werden sollen: Für die abschließende Bewertung sind nicht nur das Indikatorergebnis und die dazugehörige Stellungnahme entscheidend. Es werden auch die Ergebnisse der anderen Indikatoren des Leistungsbereichs berücksichtigt. Außerdem werden soweit möglich immer auch die Ergebnisse und die Bewertungen der Vorjahre betrachtet, um ein Gesamtbild des jeweiligen Leistungserbringers zu erhalten.

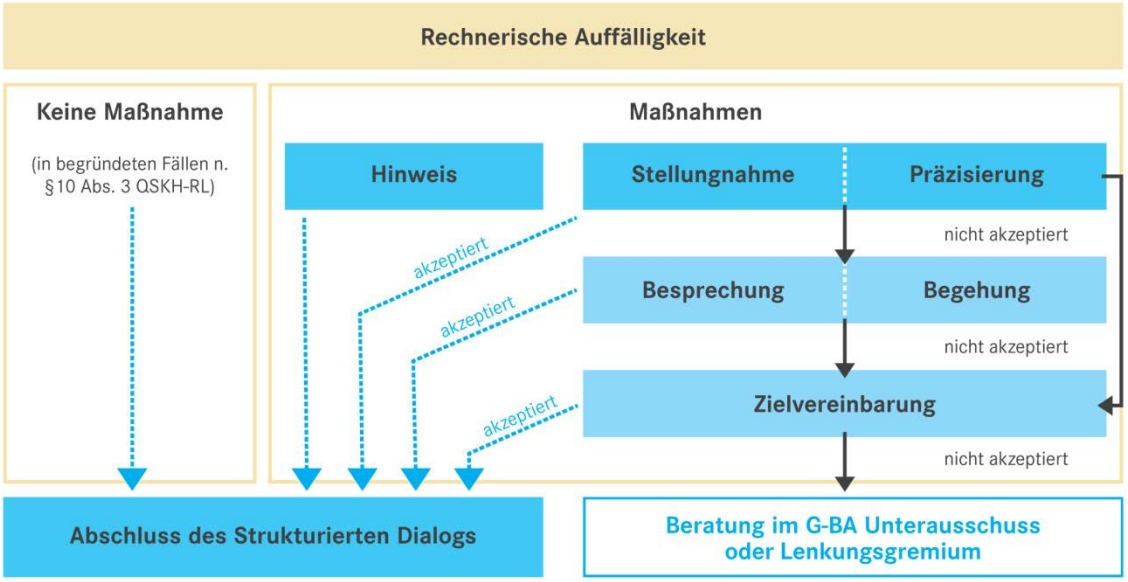

Abbildung 3: Ablauf des Strukturierten Dialogs in der stationären QS (Köppen et al. 2014)

2.2 Weiterentwicklung der bestehenden einrichtungsübergreifenden Qualitätssicherung

Im Zuge des Auftrags zur Beteiligung nicht nur an der Durchführung, sondern auch an der Weiterentwicklung der bestehenden einrichtungsübergreifenden Qualitätssicherung ergeben sich vielfältige Aufgaben. Im Sinne des Verständnisses von Qualitätssicherung als einem kontinuierlichen Prozess ergibt sich stetig Änderungs- bzw. Entwicklungsbedarf von unterschiedlicher Reichweite. Das Spektrum reicht hier von routinemäßig durchgeführten Anpassungen der QS-Verfahren in den derzeit bestehenden stationären Leistungsbereichen im Rahmen der Systempflege über die Transformation von sektorspezifischen in sektorenübergreifende QS-Verfahren bis zur vollständigen Neuentwicklung und -einführung von QS-Verfahren.

Tabelle 1 vermittelt einen Überblick über die Arten der Pflege, Entwicklung und Weiterentwicklung laufender QS-Verfahren, die damit verbundenen Inhalte bzw. charakteristischen Merkmale sowie den anhängigen Regelungs- bzw. Richtlinienbereich.

Tabelle 1: Nomenklatur und Zuordnungsmerkmale der Weiterentwicklung der bestehenden einrichtungsübergreifenden QS-Verfahren

	Entwicklungsart	Inhalt	Richtlinie
I	Systempflege	Überprüfung der bestehenden Leistungsbereiche und Feststellung des Weiterentwicklungs-, Änderungs- und Verbesserungsbedarfs sowie Umsetzung von Änderungen bei QS-Filter und -Dokumentation	QSKH- oder Qesü-RL
II	Sektorspezifische Weiterentwicklung	Erweiterung oder wesentliche Änderung eines bestehenden sektorspezifischen Verfahrens	QSKH- oder QSD-RL
	Sektorspezifische Neuausrichtung	Änderung der Widmung des Verfahrens (Bsp.: Umstellung von Prozeduren auf Diagnosen)	QSKH- oder QSD-RL
	Sektorspezifische Neuentwicklung	Entwicklung eines erstmaligen Verfahrens für einen bestimmten Leistungsbereich	Neue RL oder neuer Leistungsbereich der QSKH-RL

	Entwicklungsart	Inhalt	Richtlinie
III	Weiterentwicklung zum sektorenübergreifenden Verfahren	Erweiterung eines sektorspezifischen Verfahrens um einen weiteren Sektor, bspw. als sektorgleiches Verfahren oder durch ein sektorenübergreifendes Follow-up; dadurch gekennzeichnet, dass das bisherige sektorspezifische Verfahren auch im anderen Sektor anwendbar ist, bzw. dass es nur geringere Anpassungen gibt (Qualitätsindikatoren für beide Sektoren ohne wesentliche Änderungen nutzbar).	Qesü-RL
	Neuausrichtung zum sektorenübergreifenden Verfahren	Neuentwicklung eines bestehenden sektorspezifischen Verfahrens und Erweiterung dieses Verfahrens um einen weiteren Sektor (z.B. als sektorgleiches Verfahren oder sektorenübergreifendes Follow-up); dadurch gekennzeichnet, dass nicht nur für den weiteren Sektor etwas Neues entwickelt wird, sondern auch für den bisherigen Sektor eine wesentliche Änderung erfolgt.	Qesü-RL

2.2.1 Systempflege

Die Systempflege von laufenden QS-Verfahren besteht aus einer routinemäßigen und kontinuierlichen Evaluation der Qualitätsindikatoren. Im Wesentlichen wird überprüft, ob sie weiterhin sach- und zeitgerecht sind und der störungsfreie Betrieb der Leistungsbereiche gewährleistet ist (Steen et al. 2014).

Im Einzelnen überprüft das AQUA-Institut die bestehenden Leistungsbereiche und aktuellen Qualitätsindikatoren auf folgende mögliche Weiterentwicklungs-, Änderungs- und Verbesserungsbedarfe:

- Bewertung und ggf. Berücksichtigung eingereichter Änderungsvorschläge: Zur systematischen Erfassung und Bewertung eingereichter Vorschläge stellt das AQUA-Institut ein Formular zu Verfügung. Über dieses Formular haben alle an der externen Qualitätssicherung Beteiligten die Möglichkeit, sich hinsichtlich der Weiterentwicklung einzubringen.
- Überprüfung von Veränderungen relevanter Klassifikationskataloge: Beispielsweise werden die Internationale Klassifikation der Krankheiten (ICD) und die Operationen- und Prozedurenschlüssel (OPS) daraufhin überprüft, ob sie für die Leistungsbereiche relevant sind. Aus diesem Grund nehmen Vertreter des Instituts auch an Sitzungen der entsprechenden Arbeitsgruppen des Kuratoriums für Fragen der Klassifikation im Gesundheitswesen (KKG) beim Deutschen Institut für Medizinische Dokumentation und Information (DIMDI) teil.
- Möglichkeiten zur Harmonisierung mit anderen Datenerhebungen im Gesundheitswesen (z.B. Krebsregister, DMP usw.).
- Möglichkeiten zur Reduzierung des Dokumentationsaufwands.
- Bedarf und ggf. Möglichkeiten zur Anpassung an neue/überarbeitete Leitlinien im Rahmen der Systempflege
- Bedarf und ggf. Möglichkeiten zur Weiterentwicklung der Risikoadjustierung im Rahmen der Systempflege
- Bewertung und ggf. Berücksichtigung sonstiger systematischer Analysen (z.B. Befragungen der Landesebene zu speziellen Aspekten eines Leistungsbereichs)
- Sofern ein Datenvalidierungsverfahren stattgefunden hat: Berücksichtigung von Ergebnissen der Datenvalidierung.
- Messeigenschaften der Qualitätsindikatoren

Alle gesammelten Informationen werden mit Fachgruppen – im Rahmen der externen stationären Qualitätssicherung sind dies die Bundesfachgruppen unter Leitung des AQUA-Instituts – erörtert. Über die Beratungsergebnisse bzw. über die unterschiedlichen Empfehlungen zu Anpassungen und/oder grundlegenden Änderungen wird der G-BA in einem jährlichen Bericht informiert. Der G-BA entscheidet über den Handlungsbedarf bzw. über die Umsetzung der Empfehlungen. Die positiv beschiedenen Änderungen setzt das AQUA-Institut in der nächstmöglichen Spezifikation um. Das entsprechende Vorgehen für Verfahren der Qesü-RL wird derzeit (Feb-

ruar 2015) im Rahmen der Formulierung der themenspezifischen Bestimmungen für das QS-Verfahren *Perkutane Koronarintervention (PCI) und Koronarangiographie* (AQUA 2014c; AQUA 2014d) entwickelt und abgestimmt.

Tabelle 2: Beispiele für die kontinuierliche Anpassung von QS-Verfahren (Systempflege)

Sachverhalt	Konsequenz
Neue wissenschaftliche Erkenntnisse bezüglich zu stellender Anforderungen	Anpassung des Referenzbereichs
Über Änderungsvorschläge berichtete Dokumentationsprobleme	Änderung der Ausfüllhinweise bzw. der Indikatorenbeschreibungen
Einführung bzw. Weiterentwicklung der Risikoadjustierung	Anpassung der Berechnung
Auffälligkeiten aus der Datenvalidierung	Anpassung der Datenerfassung/Anpassung der Rechenregeln
Veränderungen der Kataloge für ICD und OPS	Anpassung von Softwarespezifikation (QS-Filter) und Auswertungsregeln

In praktischer Hinsicht spiegeln sich die Ergebnisse der Systempflege im Wesentlichen in der jährlich neu zu beschließenden und zu veröffentlichenden Spezifikation und in den veröffentlichten Rechenregeln wider. Sofern als Ergebnis der Prüfungen umfassende Weiterentwicklungen notwendig sind, wird deren Form in Abstimmung mit dem G-BA konsentiert.

Sofern im Rahmen der Bewertung einzelner Weiterentwicklungs-, Änderungs- und Verbesserungsbedarfe deutlich wird, dass eine sachgerechte Aufarbeitung im Rahmen der Systempflege nicht möglich ist, wird dies ebenfalls offengelegt, und es werden in einem jährlichen Bericht Empfehlungen des AQUA-Instituts in Abstimmung mit den Bundesfachgruppen unterbreitet, wie diese notwendigen Änderungen umgesetzt werden könnten. Nachfolgend werden konkrete Beispiele für solche Entwicklungsarten dargestellt.

2.2.2 Sektorspezifische Weiterentwicklung, Neuausrichtung und Neuentwicklung

Unter diese Entwicklungsarten sind Änderungs- und Verbesserungsbedarfe zu einem bestehenden Leistungsbereich subsummiert, die den Rahmen der Systempflege deutlich überschreiten würden und die sich allein auf einen Sektor beziehen (vgl. Punkt II in Tabelle 1). Die beiden letzten der unter Punkt II genannten Weiterentwicklungsarten – sektorspezifische Neuausrichtung und sektorspezifische Neuentwicklung – sind als Neuentwicklungen von QS-Verfahren zu betrachten,[9] weshalb allgemein die ab Kapitel 4 geschilderten Methoden zur Anwendung kommen.

In Bezug auf die sektorspezifische Weiterentwicklung sind verschiedene Fallkonstellationen denkbar. Als wesentliche sektorspezifische Weiterentwicklungen[10] gelten derzeit die Entwicklung von Follow-up-Indikatoren, die Neuentwicklung von Indikatorensets zu einem bestehenden Leistungsbereich sowie die Zusammenführung bisher getrennter Leistungsbereiche (vgl. Tabelle 3). Für diese Weiterentwicklungsarten der bestehenden Leistungsbereiche der stationären QS besteht die Herausforderung darin, die vorhandenen Methoden der Neuentwicklung von QS-Verfahren (Themenerschließung, Indikatorenrecherche, RAM-Prozess, Routinedatenanalyse, Entwicklung einer Risikoadjustierung usw.) so anzupassen, dass sie innerhalb der existierenden Strukturen anwendbar sind. Zusätzlich müssen Wege gefunden werden, um Routinedaten systematisch berücksichtigen zu können. Die folgende Tabelle gibt einen Überblick über die genannten Fälle und erläutert sie ggf. methodisch.

[9] Das wesentliche zusätzliche Element bei einer Neuentwicklung gegenüber einer Weiterentwicklung ist die systematische Themenerschließung und die Erarbeitung von Qualitätspotenzialen. Infolgedessen sollte die Entscheidung darüber, ob für ein bestehendes QS-Verfahren eine Neuentwicklung oder eine Weiterentwicklung angemessen ist, mit der Frage verknüpft werden, welcher zusätzliche Nutzen von einer Themenerschließung erwartet wird.

[10] Die nachfolgend geschilderten Weiterentwicklungsprozesse gelten zunächst nur für die bestehenden Verfahren der externen stationären Qualitätssicherung.

Tabelle 3: Entwicklungsarten der sektorspezifischen Weiterentwicklung

Sektorspezifische Weiterentwicklung ...	Methodische Aspekte	Richtlinie
a) Follow-up-Indikatoren	Die Entwicklung von Follow-up-Indikatoren entspricht methodisch der Entwicklung neuer Qualitätsindikatoren. Im Rahmen bestehender QS-Verfahren kann hierfür auf einen verkürzten Auswahl- und Bewertungsprozess zurückgegriffen werden (vgl. Abschnitt 8.4).	QSKH- oder QSD-RL
b) Patientenbefragung	Sofern ein bestehendes QS-Verfahren um Qualitätsindikatoren ergänzt werden soll, die auf einer Patientenbefragung basieren, erfolgt diese Entwicklung gemäß der dargelegten Methodik (siehe Abschnitte 7.3 und 10.2.2).	QSKH- oder QSD-RL
c) durch wesentliche Neuerung oder eine Zusammenführung von Leistungsbereichen	Die Art der Methodik ist abhängig von der Art der Neuerung.	QSKH- oder QSD-RL

2.2.3 Weiterentwicklung oder Neuausrichtung zu einem sektorenübergreifenden Verfahren

Wenn ein bestehendes sektorspezifisches QS-Verfahren um Leistungen aus einem anderen Sektor erweitert wird, wird aus diesem ein sektorenübergreifendes Verfahren. Hier kann man grundsätzlich zwischen folgenden Entwicklungsarten unterscheiden:

Weiterentwicklung zu einem sektorenübergreifenden Verfahren
Die Weiterentwicklung fokussiert auf Entwicklungsleistungen für den neuen Sektor, beispielsweise durch die Entwicklung eines sektorgleichen QS-Verfahrens oder durch Hinzufügung eines sektorenübergreifenden Follow-up. Entscheidend ist, dass das bestehende QS-Verfahren nicht grundsätzlich geändert wird.

Diesen Weiterentwicklungen liegt die Vermutung zugrunde, dass es sektorspezifische QS-Verfahren geben könnte, die mit einem geringen Aufwand auf den anderen Sektor ausgeweitet werden können.

Die Weiterentwicklung zu einem sektorenübergreifenden Verfahren ist in methodischer Hinsicht erläuterungsbedürftig, weil sie teilweise auf angepassten oder gänzlich neuen Methoden basieren muss. Die Entscheidung darüber, wann welche Methoden in welchem Umfang zum Einsatz kommen, hängt von der Klärung folgender Fragen ab:

- Gelten für den neu einzubeziehenden Sektor die gleichen Voraussetzungen hinsichtlich der Qualitätspotenziale?
- Gelten für den neu einzubeziehenden Sektor die gleichen Voraussetzungen hinsichtlich der Datenlage?
- Gelten für den neu einzubeziehenden Sektor die gleichen Voraussetzungen hinsichtlich der Umsetzung qualitätssichernder Maßnahmen?

Methodisch gesehen kann eine solche Prüfung in einem Prozess stattfinden, der dem einer erweiterten Themenerschließung (Konzeptskizze) entspricht (vgl. Kapitel 4).

Neuausrichtung zu einem sektorenübergreifenden Verfahren
Die Neuausrichtung umfasst neben den Entwicklungsleistungen für den neuen Sektor auch wesentliche Änderungen am bestehenden Verfahren. Die Neuausrichtung zu einem sektorenübergreifenden Verfahren ist analog der Neuentwicklung eines QS-Verfahren zu betrachten, weshalb allgemein die ab Kapitel 4 geschilderten Methoden zur Anwendung kommen.

2.3 Neuentwicklung sektorenübergreifender QS-Verfahren

Nach §137a Abs. 2 Satz 1 Nr. 1 und 2 SGB V entwickelt das AQUA-Institut sowohl Indikatoren und Instrumente, mit denen die Versorgungsqualität gemessen und dargestellt werden kann als auch, unter Berücksichtigung des Gebotes der Datensparsamkeit, die notwendige Dokumentation für die einrichtungsübergreifende Qualitätssicherung. Nach Maßgabe des Gesetzes sollen die beauftragten Entwicklungsleistungen möglichst sektorenübergreifend angelegt sein.

2.3.1 Allgemeine Ziele

Orientiert am Patientennutzen verfolgt die Qualitätssicherung folgende grundlegende Ziele:

- systematische Identifizierung von Qualitätspotenzialen (Qualitätsziele und –defizite), d.h. von Ansätzen zur Qualitätssicherung und -verbesserung medizinischer Leistungsbereiche
- Unterstützung einer systematischen, kontinuierlichen und berufsgruppenübergreifenden einrichtungsinternen Qualitätssicherung (internes Qualitätsmanagement),
- insbesondere durch die Entwicklung von (ggf. risikoadjustierten) Indikatoren faire Leistungserbringervergleiche zu ermöglichen
- durch signifikante, valide und vergleichbare Erkenntnisse – insbesondere zur Indikationsstellung, zur Angemessenheit der Leistung und zur Ergebnisqualität – die Methoden der Qualitätssicherung zu verbessern

2.3.2 Entwicklung von Indikatoren, Instrumenten und der notwendigen Dokumentation

Voraussetzung für eine sach- und zeitgerechte Entwicklung von Indikatoren, Instrumenten und der notwendigen Dokumentation ist ein hinreichend konkretisierter Auftrag. Wenn ein Auftrag vonseiten des G-BA nicht hinreichend konkretisiert ist bzw. eine Konkretisierung durch das AQUA-Institut vorgenommen werden soll, kann zu den beauftragten Entwicklungsleistungen eine Konzeptskizze z.B. zur Prüfung der Versorgungsrelevanz oder zur Nutzbarkeit von Routinedaten erstellt werden. Da es sich hier um eine mögliche Zielvorgabe für den eigentlichen Auftrag handelt, ist hierfür ein angemessener und ggf. zusätzlich zu bemessender Zeitrahmen einzukalkulieren.[11] Da eine solche Auftragskonkretisierung nicht Gegenstand des allgemeinen Vertrags zwischen dem G-BA und dem AQUA-Institut ist, bedarf eine entsprechende Auftragskonkretisierung der Beauftragung durch den G-BA.

Entwicklungsleistungen beginnen mit der Auftragserteilung und enden mit der Abnahme des Abschlussberichts jeweils durch den G-BA. Im Anschluss daran kann der G-BA beschließen, die entwickelten Indikatoren, Instrumente und Dokumentationen vor einer Überführung in den Regelbetrieb zu testen. Dazu kann er das AQUA-Institut mit einer Erprobung (Abschnitt 10.2) gesondert beauftragen.

Ausgangspunkt aller Entwicklungsleistungen ist das vorliegende, mit dem G-BA abgestimmte Methodenpapier sowie ein ebenfalls abgestimmter Bearbeitungsrahmen für die beauftragten Entwicklungsleistungen.

2.3.2.1 Projektierung

Für jedes beauftragte Thema bestimmt das AQUA-Institut eine qualifizierte Projektleitung. Die Projektleitung strukturiert ein beauftragtes Thema inhaltlich und chronologisch und ordnet ihm die spezifische Expertise aus den unterschiedlichen Abteilungen des AQUA-Instituts zu. Zum Projektteam im engeren Sinne gehören typischerweise Mediziner und Gesundheitswissenschaftler, die die medizinischen Inhalte entwickeln sowie Sozialwissenschaftler, Epidemiologen, Biometriker oder Statistiker, die mit der Erstellung von Auswertungskonzepten und der Berichterstattung befasst sind. Projektbegleitend werden Mitarbeiter mit Expertise für Methodenanwendung und Recherchen hinzugezogen sowie Mitarbeiter, die mit der Entwicklung von Erhebungsinstrumenten betraut sind. Die Projektleitung ist für die Planung und Durchführung aller wesentlichen im Folgenden genannten Entwicklungsschritte verantwortlich.

[11] Ein solcher Zeitrahmen wurde dem AQUA-Institut erstmals mit dem Auftrag zur Indikatorenentwicklung für die Versorgung bei psychischen Erkrankungen (2012) eingeräumt.

© 2015 AQUA-Institut GmbH

2.3.2.2 Themenerschließung

Ziel der Themenerschließung (vgl. Kapitel 4) ist die Beurteilung der Versorgungsrelevanz und der potenziellen Umsetzbarkeit eines Themas (s.u.). In diesem Sinne eröffnet die Themenerschließung einen orientierenden Handlungsrahmen für die weitere Entwicklungsarbeit. Veranschaulicht werden diese Ergebnisse anhand dargelegter Qualitätspotenziale und anhand eines Versorgungspfades.

Die Themenerschließung besteht zunächst aus einer Recherche nach aggregierter Evidenz in dafür geeigneten Informationsquellen (vgl. Abschnitt 4.3.1). Inhaltlich fokussiert die Themenerschließung einerseits auf die zentralen (medizinischen) Versorgungsaspekte – also auf die möglichen Zielpopulationen, die Versorgungsleistungen sowie die therapie- und patientenrelevanten Endpunkte – andererseits auf die Qualitätspotenziale (vgl. Abschnitt 4.2). Die Themenerschließung wird – im Regelfall unter Einbezug externer Experten – von einer Kontextrecherche bzw. einer Sammlung von Kontextinformationen, die die gesundheitspolitischen und strukturellen Rahmenbedingungen der Versorgung betreffen (vgl. Abschnitt 4.2.4), begleitet. Darunter fällt auch eine erste Prüfung der Umsetzbarkeit, d.h. eine Prüfung der Verfügbarkeit von Datenquellen und Erhebungsinstrumenten. An dieser Stelle können die Versorgungsrelevanz und die potenzielle Umsetzbarkeit eines Themas kritisch reflektiert werden (vgl. Abschnitt 4.3.3).

Der auf Basis der Versorgungsrelevanz aufgestellte Versorgungspfad – bzw. in alternativer Perspektive die Qualitätspotenziale – soll (idealerweise) auf einem Scoping-Workshop diskutiert bzw. bestätigt werden. Die an die Experten des Workshops zu richtenden Fragen beziehen sich insbesondere darauf, die Kongruenz des konkret beauftragten QS-Verfahrens mit den allgemeinen Zielstellungen der Qualitätssicherung zu prüfen:

- Ist das Thema tatsächlich hinreichend versorgungsrelevant (epidemiologischer Nutzen)?
- Ist die versorgungsrelevante Zielpopulation eindeutig definiert?
- Sind der durch den Pfad abgebildete Versorgungsweg und die konkreten Versorgungsleistungen hinreichend evidenzbasiert bzw. bildet der Pfad eine bestmögliche reale Versorgung ab?
- Berücksichtigt der Pfad die patientenrelevanten Endpunkte (Patientennutzen)?
- Berücksichtigt der Versorgungspfad alle relevanten Qualitätsziele, -potenziale und -defizite?
- Ist der Versorgungspfad auf das deutsche Versorgungsgeschehen anwendbar?
- Sind die erforderlichen QS-Daten, insbesondere an den Schnittstellen/Sektorenübergängen, grundsätzlich verfügbar/erhebbar?

Zu diesem Workshop werden alle gesetzlich zur Stellungnahme berechtigten Organisationen (§137a Abs. 7 SGB V) und die für das Thema relevanten Fachgesellschaften, bekannte Fachexperten sowie auch Vertreter einschlägiger Qualitätsprojekte eingeladen. Auf der Grundlage des erarbeiteten und auf dem Workshop zur Diskussion gestellten Versorgungspfads und seiner etwaigen Besonderheiten wird die Indikatorenrecherche unternommen.

2.3.2.3 Indikatorenrecherche

Für die Indikatorenentwicklung recherchiert das AQUA-Institut nach national und international vorhandenen Indikatoren (vgl. Abschnitt 5.1). Neben den üblichen Indikatoren zur Struktur-, Prozess- und Ergebnisqualität werden auch Indikatoren zur Indikationsqualität und zur Berücksichtigung der Patientenperspektive systematisch einbezogen. Ziel der Recherche ist die Erstellung eines themenspezifischen Indikatorenregisters, das als Arbeitsgrundlage für den weiteren Auswahl- und Bewertungsprozess der Indikatoren (im Rahmen des sog. RAM-Prozesses, vgl. Abschnitt 8.2) dient. Eine gezielte Suche nach Indikatoren erfolgt sowohl indirekt über Datenbanken als auch direkt bei entsprechenden Institutionen und Agenturen aus dem Gesundheitswesen.

Vervollständigt wird die Indikatorenrecherche durch eine systematische Literaturrecherche (vgl. Abschnitt 5.2). Das Suchmodell wird in Anlehnung an den konkreten Versorgungspfad und die Qualitätspotenziale formuliert, der Maßstab für die Vorauswahl sowie ggf. für eine Ergänzung der Indikatoren ist.

Sollten mit Blick auf die im Versorgungspfad genannten Ansatzpunkte für eine Qualitätssicherung keine Indikatoren gefunden werden, wird das Projektteam ein Indikatorenregister aus selbstentwickelten Indikatoren zusammenstellen. Dies kann beispielsweise durch die Ableitung von Indikatoren aus Leitlinien (vgl. Abschnitt

6.2.1), auf der Basis von Literatursynthesen oder auf Grundlage von Routinedatenanalysen (vgl. Abschnitt 6.2.2) geschehen.

2.3.2.4 Auswahl und Bewertung von Indikatoren

Zur Auswahl und Bewertung der recherchierten Indikatoren bildet das AQUA-Institut eine Sachverständigengruppe (Panel) aus externen Experten (vgl. Abschnitt 3.4). Diese bewerten die Indikatoren nach einer standardisierten Konsensmethode. Dabei handelt es sich um eine adaptierte Version der RAND Appropriateness Method (RAM), dem gegenwärtigen Goldstandard der Indikatorenvalidierung (vgl. Abschnitt 8.2.1).

Das RAM-Panel bewertet die Indikatoren in einer ersten Bewertungsrunde nach den Kriterien *Relevanz* sowie *Klarheit/Verständlichkeit*. In einer zweiten Bewertungsrunde wird vom Panel die *Praktikabilität* bewertet. Alle Bewertungen erfolgen schriftlich und werden vom Projektteam ausgewertet.

Die Patientenperspektive wird über einen verfahrensspezifisch zu entwickelnden Fragebogen erhoben (siehe Abschnitt 7.3). Dieser wird in einem parallelen Prozess von einem durch externe Expertise verstärkten Projektteam des AQUA-Instituts und unter Einbeziehung der Ergebnisse der Themenerschließung und der systematischen Literaturrecherche entwickelt. Wenn die Entwicklung einer Patientenbefragung im Rahmen einer Verfahrensneuentwicklung mit beauftragt ist, werden auch die Indikatoren der Patientenperspektive im RAM-Prozess bewertet. Das im RAM-Panel konsentierte Set relevanter und praktikabler Indikatoren wird ausführlich und abschließend vom AQUA-Institut bzw. dem Projektteam und den Panelteilnehmern dahingehend geprüft, ob alle Qualitätspotenziale durch das abgestimmte Indikatorenset angemessen abgebildet sind (vgl. Abschnitt 8.3). Am Ende des Recherche-, Auswahl- und Bewertungsprozesses steht dann ein Indikatorenset, das vom AQUA-Institut zur Umsetzung bzw. zur Erprobung empfohlen wird.

2.3.2.5 Entwicklung der Instrumente und der Dokumentation

Parallel zum Auswahl- und Bewertungsprozess der Indikatoren definiert das Projektteam die Inhalte der Erhebungsinstrumente (vgl. Kapitel 7) und arbeitet die Rahmenbedingungen für die Implementierung der Indikatoren aus. Dazu gehören:

- die exakte Beschreibung der Erhebungsinstrumente für die Indikatoren
- die Differenzierung der beteiligten Sektoren, Einrichtungen und Leistungserbringer
- die Festlegung der Kriterien für die QS-Auslösung

Anschließend werden die für die Qualitätssicherung erforderlichen Dokumentationsparameter konkretisiert bzw. die Datenfelder und Items für spezifische Erhebungsinstrumente werden formuliert und beschrieben.

Zur Entwicklung der Dokumentation gehören:

- die Entwicklung der durch die Leistungserbringer zu dokumentierenden Datenfelder (QS-Dokumentation)
- die Erarbeitung von Ausfüllhilfen für diese Datenfelder
- die Entwicklung von Kriterien zur Plausibilitätsprüfung (Vollzähligkeit und Vollständigkeit)
- die Spezifizierung der zu übermittelnden Datenfelder (Routinedatennutzung)
- die Entwicklung von Items (Patientenbefragung)

2.3.2.6 Empfehlung zur Umsetzung

Sämtliche entwickelte Indikatoren, Instrumente und Dokumentationen haben grundsätzlich den Charakter von begründeten Vorschlägen. Sollte die Würdigung des Indikatorensets am Ende ergeben, dass wesentliche Qualitätsziele unter den gegebenen Rahmenbedingungen (noch) nicht erreicht werden können oder dass ein entsprechendes QS-Verfahren unter dem Gesichtspunkt von Aufwand-Nutzen problematisch ist, dann werden dem G-BA ggf. andere Optionen zu einem weiteren Vorgehen empfohlen.

Für die entwickelten Indikatorensets werden grundsätzliche Angaben zur weiteren Umsetzung (vgl. Kapitel 10) sowie – soweit zu dieser frühen Zeit möglich und sinnvoll – zu Einzelheiten der Auswertung (vgl. Kapitel 11) und zur Berichterstattung (in Form eines vorläufigen Auswertungskonzepts, vgl. Abschnitt 12.3.2) gegeben, um dem G-BA die Formulierung etwaig notwendiger themenspezifischer Bestimmungen zu ermöglichen (s.u. Abschnitt 2.3.3)

Die Ergebnisse werden zunächst in Form eines Vorberichts den in §137a Abs. 3 SGB V genannten und zur Stellungnahme berechtigten Institutionen vorgelegt. Diese haben in einem Zeitraum von acht Wochen Gelegenheit, zu den Ergebnissen Stellung zu nehmen (vgl. Abbildung 4).

Das AQUA-Institut prüft die Stellungnahmen und passt die eigenen Vorschläge und Empfehlungen an, wenn die Stellungnahmen sachgerechte und begründete Einwände oder Korrekturen enthalten. Der ggf. überarbeitete Vorbericht wird um die Dokumentation des Stellungnahmeverfahrens ergänzt und schließlich als Abschlussbericht dem G-BA zur Abnahme übergeben.

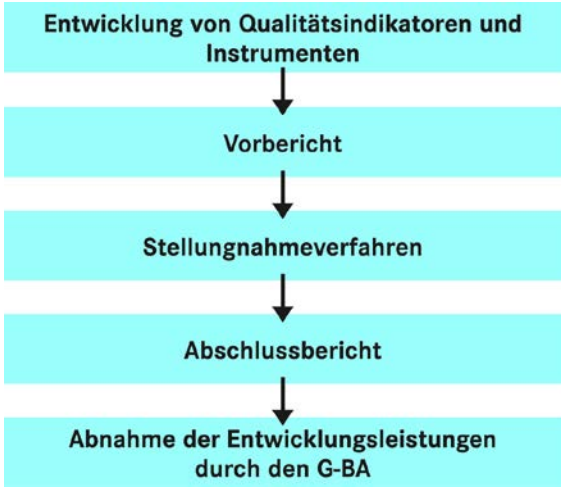

Abbildung 4: Meilensteine einer beauftragten Neuentwicklung (Indikatoren, Instrumente und Dokumentationen)

2.3.3 Themenspezifische Bestimmungen

Zusammen mit den Empfehlungen des Abschlussberichts bzw. den Indikatordatenblättern enthalten die Ergebnisse eines Entwicklungsberichts im Zusammenhang mit der informationstechnischen Aufarbeitung (Abschnitt 10.1) und etwaig durchgeführten Erprobungen (Abschnitt 10.2) die konkreten Informationen zum Beschluss themenspezifischer Bestimmungen (durch den G-BA) gemäß §3 Qesü-RL. Sie ergänzen die Richtlinien und Beschlüsse des G-BA sowie die bundeseinheitlich vom AQUA-Institut vorzugebenden Regelungen und bilden die Grundlage der Spezifikationen.

2.3.4 Erprobung von Indikatoren, Instrumenten und der notwendigen Dokumentation

Sofern eine Erhebung der Daten über Einrichtungen (QS-Dokumentation) oder Patienten (Patientenbefragungen) geplant ist, besteht darüber hinaus vor Einführung eines Regelbetriebs zumeist Klärungsbedarf dahingehend, ob die Erhebungsinstrumente in der gewünschten Art und Weise funktionieren und valide Aussagen liefern. Bei einem solchen Bedarf kann der G-BA das AQUA-Institut mit entsprechenden Erprobungsprojekten beauftragen (siehe Abschnitt 10.2).

2.4 Themenvorschläge und Themenauswahl

Das AQUA-Institut kann bezüglich der Themenauswahl für neue Qualitätssicherungsverfahren Vorschläge einbringen oder den G-BA bei der wissenschaftlichen Aufbereitung von Themen unterstützen.

Eine wesentliche Grundlage der Themenfindung und der Priorisierung von Themen sind sogenannte Kriterienkataloge, anhand derer potenzielle Themen für die Qualitätssicherung bewertet werden. Diese Grundlage hat sich in der Vergangenheit allerdings oft als nicht ausreichend erwiesen, sodass eher Qualitätsindikatoren mit einem schlechten Aufwand-/Nutzenverhältnis entwickelt wurden. Dies liegt u.a. darin begründet, dass aus den in den Kriterienkatalogen dargelegten Qualitätspotenzialen bzw. Qualitätsdefiziten nicht ersichtlich wird, welche Versorgungsrelevanz die Themen tatsächlich haben, wie hoch voraussichtlich der Aufwand zur Messung von Quali-

tätsdefiziten sein wird und ob eine Zuordnung der Ergebnisverantwortung in ausreichend hohem Maße möglich sein wird, um diese für Maßnahmen der Qualitätssicherung nutzen zu können.

Die erfolgreiche Implementierung von Qualitätssicherungsverfahren ist von zahlreichen Bedingungen abhängig, die nicht erst im Verlauf einer konkreten Verfahrensentwicklung oder während der Erprobung des Verfahrens thematisiert werden sollten. Im Zuge dieser sukzessiv gewonnen Erkenntnis wurden immer mehr Elemente der ursprünglichen Methode der Verfahrensentwicklung immer früher eingesetzt. Dies hat in der Konsequenz dazu geführt, dass die frühere „Vorrecherche“ sich in Richtung einer immer umfassenderen „Themenerschließung“ entwickelt hat, in der nun dezidiert Qualitätspotenziale recherchiert und Beurteilungen von teilweise sehr komplexen Rahmenbedingungen einer Verfahrensentwicklung vorgenommen werden. In den letzten Entwicklungen wurde auch die Einschätzung der Abbildbarkeit von verfahrensspezifischen Qualitätspotenzialen mittels verschiedener Erhebungsinstrumente sowie eine darauf bezogene allgemeine Einschätzung von Aufwand und Nutzen in die Themenerschließung verlagert. Angesichts der Tatsache, dass eine sektorenübergreifende QS-Dokumentation (Datenerhebung beim Leistungserbringer) unter den gegenwärtigen Bedingungen kaum umsetzbar scheint, gewinnt auch die Möglichkeit der Nutzung von Routinedaten an Gewicht. Auch dies ist ein Aspekt, der idealerweise einer eigentlichen (Weiter-) Entwicklung von QS-Verfahren vorgeschaltet sein sollte.

Mit der Berücksichtigung der genannten neuen Aspekte kann die Themenerschließung auch als eine mehr oder weniger eigenständige und in sich abgeschlossene Methode betrachtet werden, die den Charakter einer Vorstudie aufweist und formal als Konzeptskizze beauftragt werden kann. Aufgrund des hohen Aufwands kann eine solche Konzeptskizze die Kriterienkataloge nicht ersetzen, aber sie kann im Vorfeld von Beauftragungen als wichtige Entscheidungsgrundlage für eine Auftragskonkretisierung, aber auch als Grund für einen Entwicklungsstopp, fungieren.

Die zentralen Aspekte einer im obigen Sinne ausgestalteten Konzeptskizze sind:

- Prüfung und ggf. Ergänzung der im Kriterienkatalog/Auftrag genannten Qualitätspotenziale (Ziele und Defizite) und der Versorgungsrelevanz durch Recherchen
- Darlegung der vorhandenen Evidenz zu den jeweiligen Qualitätspotenzialen als einer wichtigen Komponente des potenziellen Nutzens
- Einschätzung der Umsetzbarkeit und der Validität einer Messung der Qualitäts- und Verbesserungspotenziale mit den vorhandenen Erhebungsinstrumenten als einer wesentlichen Komponente zur Abschätzung des zu erwartenden Aufwand-/Nutzenverhältnisses
- Einschätzung der Zurechenbarkeit von Verantwortung für Prozesse und Ergebnisse auf spezifische Einrichtungen als einer weiteren wesentlichen Komponente des potenziellen Nutzens

3 Auswahl und Beteiligung von Experten

Gemäß der eigenen Qualitätsphilosophie und mit Blick auf die vielfältigen Aufgaben des AQUA-Instituts ist die Einbindung externer Expertise nicht nur unerlässlich, sondern auch gewollt. Die folgenden Abschnitte geben Auskunft über die Grundsätze der Einbeziehung externer Expertise, die Zusammensetzung von Expertengruppen bzw. die Auswahl von Experten sowie Art und Umfang der Beteiligung.

3.1 Offenlegung von Interessenkonflikten

Als freies, fachlich unabhängiges und interessenneutrales Unternehmen gehören Transparenz und Nachvollziehbarkeit der relevanten Abläufe zu den grundsätzlichen Arbeitsprinzipien des AQUA-Instituts. Während die Interessenneutralität für die Mitarbeiter und Nachunternehmer des Instituts vertraglich geregelt ist, kann dies nicht in gleicher Weise für andere, an der Entwicklung von Indikatoren, der externen Qualitätssicherung und weiteren entscheidenden Prozessen Beteiligte, d.h. für externe Experten, vorausgesetzt werden.

Auch wenn das AQUA-Institut nicht im besonders sensiblen Bereich der klinischen Forschung tätig ist (Ganser 2011), müssen externe Experten, die aktiv in die Prozesse der Indikatorenentwicklung und/oder -bewertung eingebunden sind, ihre möglichen Interessenkonflikte untereinander sowie gegenüber dem AQUA-Institut und dem G-BA offenlegen (vgl. auch §137a Abs. 5 SGB V). Mit dem Vorliegen eines Interessenkonflikts sind nicht zwingend Maßnahmen verbunden, denn ein Interessenkonflikt bezeichnet zunächst nur das Risiko, dass das notwendige und erwünschte professionelle Urteilsvermögen unangemessen beeinflusst wird (DNEbM o.J. (2011): 9; Lo et al. 2009: 46). Insofern dient die Offenlegung insbesondere der Schaffung von Vertrauen durch Transparenz (Lieb et al. 2011).

Das jeweils aktuelle Formular zur Offenlegung von Interessenkonflikten ist den Unterlagen zur Interessenbekundung bezüglich der Teilnahme an einer Verfahrensentwicklung angefügt und kann jederzeit auf den Webseiten des Instituts eingesehen werden.

3.2 Wissenschaftlicher Beirat

Für seine Aufgaben im Bereich der sektorenübergreifenden Qualitätssicherung hat das AQUA-Institut einen Wissenschaftlichen Beirat mit namhaften Experten aus den Bereichen Qualitätsmanagement, Gesundheitswissenschaft, Versorgungsforschung, Patientensicherheit, Pflegewissenschaft, Medizin und Zahnmedizin als beratendes Gremium eingerichtet. Die zentrale Aufgabe des Beirats ist die unabhängige Beratung des AQUA-Instituts zu wissenschaftlich-methodischen Fragestellungen im Rahmen der Aufgaben nach §137a SGB V. Die Beratungsthemen werden von der Geschäftsleitung des Instituts oder vom Beirat vorgeschlagen. Das AQUA-Institut sieht im Wissenschaftlichen Beirat ein Gremium, in dem externe Expertise in einen offenen wissenschaftlichen Diskurs zu methodischen Fragestellungen eingeht. Anregungen und Empfehlungen aus dem Beirat fließen in die Meinungsbildung im Institut ein, haben jedoch keinen verbindlichen Charakter. Darüber hinaus soll durch den Beirat die Kommunikation mit den wissenschaftlichen Fachgesellschaften gefördert werden.

Die Mitglieder des Beirats werden vom AQUA-Institut für einen Zeitraum von jeweils drei Jahren berufen. Die konstituierende Sitzung des Beirats fand am 10. Mai 2010 statt. Auf der Webseite des Instituts sind die Namen der Mitglieder (www.sqg.de/hintergrund/wissenschaftlicher_beirat) einsehbar. Der Wissenschaftliche Beirat wird mindestens einmal pro Jahr zusammengerufen. Über die Beiratssitzungen wird ein Protokoll angefertigt, das den Mitgliedern anschließend zur Verfügung gestellt wird.

3.3 Scoping-Workshop

Der sogenannte Scoping-Workshop ist eine Veranstaltung, die im Rahmen beauftragter Entwicklungsleistungen für neue Qualitätssicherungsverfahren zumindest dann stattfindet, wenn keine eigenständige Themenerschließung im Sinne einer Konzeptskizze angefertigt wurde (vgl. Kapitel 4). Die als Expertenforum konzipierte Veranstaltung mit Vorträgen und Diskussionen findet in der Endphase der Auftragskonkretisierung statt und verfolgt

© 2015 AQUA-Institut GmbH

den Zweck, die Ziele der anstehenden Entwicklungsleistungen mit einer möglichst breiten externen Expertise zu untermauern. Idealerweise wird auf dem Workshop ein entwicklungsleitender Versorgungspfad im Sinne der „face validity“ (Anastasi et al. 1997) konsentiert.[12]

Primäre Adressaten der Scoping-Workshops sind Fachexperten aus dem Bereich Qualitätssicherung sowie medizinisch-therapeutische und pflegerische Experten und Patientenvertreter. Der eintägige Scoping-Workshop ist eine offene Veranstaltung, die auf der Webseite des Instituts angekündigt wird. Die an den Entwicklungsverfahren zu beteiligenden Organisationen nach §137a Abs. 3 SGB V und ausgewiesene Fachexperten werden gesondert auf den Scoping-Workshop hingewiesen. Die Teilnahme erfordert eine Anmeldung. Externe Referenten auf diesem Workshop müssen ihre Interessenkonflikte gegenüber dem AQUA-Institut und den Teilnehmern des Workshops offenlegen.

Beschränkungen des Teilnehmerkreises sind nur durch die organisatorischen Rahmenbedingungen bzw. durch die Nichteinhaltung der Anmeldefrist begründet.

3.4 RAM-Panel

3.4.1 Aufgaben und Ziele

Ein wichtiger Bestandteil der Indikatorenentwicklung im Rahmen beauftragter Entwicklungsleistungen (Neu- und Weiterentwicklungen) ist das Auswahl- und Bewertungsverfahren der recherchierten Indikatoren durch unabhängige Experten. Die Aufgabe der Experten im RAM-Panel ist es, wissenschaftlich fundierte und belastbare Indikatoren für spezifische Versorgungsprozesse auszuwählen und ggf. dem Versorgungspfad anzupassen (z.B. mittels Ein- und Ausschlusskriterien, Verbesserung der Verständlichkeit etc.).

Die Experten des Panels sind einerseits Fachexperten aus verschiedenen Bereichen der Versorgung und der beteiligten Sektoren, aber auch Personen mit übergreifender Expertise (z.B. aus der Versorgungsforschung, Epidemiologie). Darüber hinaus gibt es Experten aus den Patientenvertretungen. Fachexperten und Patientenvertreter agieren als gleichberechtigte Teilnehmer des Panels. Eine berufs- oder fachgruppenspezifische Einschränkung von Stimmrechten ist mit Blick auf die Stringenz des Verfahrens ebenso wenig vorgesehen, wie eine Gewichtung besonderer Expertise. Im Gegenteil soll gerade eine obligatorische Einführungsveranstaltung bewirken, dass fach- und berufsgruppenspezifische Einzelinteressen (die prinzipiell nie ausgeschlossen werden können) hinter das gemeinsame Interesse zurücktreten, wissenschaftlich fundierte und belastbare Indikatoren für den gesamten Versorgungsprozess auszuwählen. Ziel der Einführungsveranstaltung ist es, ein gemeinsames Bewusstsein dafür zu erzeugen, dass es sich bei den Qualitätssicherungsverfahren um mehrperspektivische, schnittstellen- und sektorenübergreifende Gesamtbetrachtungen von Versorgungsabläufen handelt, die mit einem umfassenden Set von Indikatoren beschrieben werden sollen.

3.4.2 Ausschreibung

Die Besetzung eines RAM-Panels für ein neues Entwicklungsverfahren wird auf der SQG-Webseite des AQUA-Instituts angekündigt, wo neben dem Ausschreibungstext auch nähere Informationen zum Verfahren und die Bewerbungsunterlagen zu finden sind. Gleichzeitig wird die Ausschreibung dem G-BA und seinen beteiligten Organisationen bekannt gegeben.

Zur Besetzung des Panels mit Fachexperten werden alle für das jeweilige Verfahren relevanten Gruppen (Fachgesellschaften, Leistungserbringer auf Bundes- und Landesebene, zur Stellungnahme berechtigte Organisationen usw.) angeschrieben und auf die Möglichkeit individueller Bewerbungen hingewiesen. Dabei werden auch die am Verfahren beteiligten Heil- und Hilfsberufe sowie Organisationen und Gesellschaften aus dem Bereich der Gesundheitswissenschaft berücksichtigt (Biometrie, Epidemiologie, Public Health, Gesundheitsökonomie). Darüber hinaus werden etwaige, dem AQUA-Institut bekannte Fachexperten (z.B. durch Vorschläge der Fachgesellschaften, Experten bestehender Fachgruppen oder aus der Literaturrecherche bekannte und infrage kommende Fachexperten) persönlich angesprochen und zu einer Bewerbung aufgefordert. Auch die Teilnehmer des

[12] Angesichts langfristiger Zeitplanungen und nicht immer flexibler oder frei wählbarer Termine ist der geschilderte Ablauf wünschenswert, aber nicht immer in dieser idealen Form realisierbar. Vor allem besteht die Möglichkeit, dass die Themenerschließung auf dem Scoping-Workshop noch nicht vollständig abgeschlossen ist, sodass kein „fertiger“ Versorgungspfad präsentiert werden kann.

© 2015 AQUA-Institut GmbH

vorher abgehaltenen Scoping-Workshops werden im Rahmen der Themenkonkretisierung gebeten, sich zu bewerben.

3.4.3 Auswahlkriterien

Die Teilnahme am Panel erfordert eine individuelle und ausführliche Bewerbung der Interessenten, inklusive der Offenlegung ihrer Interessen. Grundsätzliche Teilnahmebedingung ist überdies, dass sich Bewerber zur Wahrnehmung der vorgegebenen künftigen Termine bereit erklären, damit der Zeitplan für die Verfahrensentwicklung eingehalten werden kann.

Ziel der Ausschreibung eines Expertenpanels ist es, Teilnehmerinnen und Teilnehmer zu gewinnen, die das adressierte Thema in einer möglichst großen Breite abdecken können. Je nach Komplexität des Verfahrens sollen, entsprechend den Vorgaben der RAND-Methodik und den bisherigen Erfahrungen sowie im Interesse der Arbeitsfähigkeit der Gruppe, mindestens 9 aber höchstens 15 Experten in ein Panel eingebunden werden. Das AQUA-Institut geht von einer durchschnittlichen Panelgröße von 12 Experten aus. In Ausnahmefällen, etwa bei sehr komplexen Themen oder Verfahren, kann von der Höchstzahl nach oben abgewichen werden, wenn anders nicht gewährleistet werden kann, dass alle entscheidend beteiligten Professionen und Disziplinen berücksichtigt werden. Abgesehen davon, dass immer zwei Patientenvertreter dabei sein sollen, existiert keine Regelung zur Mindest- oder Höchstanzahl von Teilnehmern eines bestimmten Bereichs.

Aus Erfahrung mit den bisherigen Verfahren übersteigt die Zahl der Bewerbungen die zu besetzenden Plätze regelmäßig, weshalb das AQUA-Institut aus einer Vielzahl geeigneter Bewerber auswählen muss.

Die Auswahl der Experten aus dem Bewerberpool erfolgt durch das jeweilige mit dem Verfahren betraute Projektteam des AQUA-Instituts anhand der eingereichten Unterlagen und der Anwendung der nachstehend genannten allgemeinen Kriterien. Vorrangiges Ziel ist es, Personen auszuwählen, welche die Auswahl von Indikatoren mit der bestmöglichen Evidenz unterstützen können. Ein wesentliches Kriterium hierfür ist die Fach- und Methodenkompetenz. Diese kann basieren auf:

- praktischer klinischer Erfahrung,
- wissenschaftlich-thematischer Fachexpertise (nachgewiesen bspw. durch Publikationen, Teilnahme an Leitlinienentwicklungen) oder
- übergeordneter Expertise mit Bezug zum Thema (z.B. Biometriker/Epidemiologen, Gesundheitsökonomen, Sozialmediziner, Pharmazeuten usw.).

Da die Panelteilnehmer einen möglichst breiten Querschnitt der Versorgungslandschaft repräsentieren sollen, werden neben der fach- und themenspezifischen Expertise auch unterschiedliche Versorgungssettings (ambulant/stationär; Grund-, Regel- und Maximalversorgung) sowie regionale und Genderaspekte berücksichtigt.

Für den unwahrscheinlichen und bisher nicht eingetretenen Fall, dass Bewerber über die exakt gleichen Qualifikationen verfügen und auch das Projektteam hier keine begründbaren Prioritäten zu setzen vermag, wird eine Zufallsauswahl getroffen. Sollte ein Bewerber unerwartet ausscheiden, wird entsprechend der Bewertung durch das Projektteam eine Nachrückerliste generiert. Beim Ausscheiden eines ausgewählten Teilnehmers wird bevorzugt ein Nachrücker von der Liste ausgewählt, dessen Qualifikationsprofil dem des ausgefallenen Bewerbers so weit wie möglich entspricht.

Das Ergebnis des Auswahlverfahrens bleibt für den Zeitraum der Auswahl und Bewertung der Indikatoren vertraulich. Erst im Vorbericht werden die Namen der Experten inkl. ihrer Angaben zu den Interessenkonflikten veröffentlicht. Zusätzlich wird in einem Raster dargestellt, wie sich die o.g. Kriterien im Panel widerspiegeln.

Die Auswahl der Patientenvertreter unterliegt nicht den genannten Kriterien des AQUA-Instituts. Sie werden von den für die Wahrnehmung der Interessen der Patientinnen und Patienten und der Selbsthilfe chronisch kranker und behinderter Menschen maßgeblichen Organisationen auf Bundesebene einvernehmlich benannt

(vgl. §§2 bis 4 PatBeteiligungsV[13]). Gleichwohl müssen auch Patientenvertreter ihre Interessenkonflikte darlegen.

3.4.4 Zeitbedarf und Honorierung der Experten

Für die Teilnahme am RAM-Panel besteht ein Zeitbedarf von etwa vier bis fünf Tagen je Experte und Verfahren. Dieser teilt sich auf folgende Meilensteine auf:

- Auftaktveranstaltung
- Bewertungsrunde 1 (postalisch/online)
- Bewertungsrunde 1 (Paneltreffen)
- Bewertungsrunde 2 (postalisch-online)
- Bewertungsrunde 2 (Paneltreffen)
- Gesamtwürdigung [optionales Treffen, bei Bedarf]

Für die Arbeit im Rahmen der Indikatorenentwicklung zahlt das AQUA-Institut eine Aufwandspauschale. Die Erstattung der Reisekosten ist in einer entsprechenden Reisekostenregelung festgelegt. Genaue Angaben zum voraussichtlichen Zeitbedarf und zur Honorierung werden in den Ausschreibungsunterlagen für das jeweilige Verfahren dargelegt.

3.5 Bundesfachgruppen

Die Bundesfachgruppen des AQUA-Instituts sind beratende Expertengruppen für die gegenwärtig bestehenden Verfahren der externen stationären Qualitätssicherung. Die derzeit (Stand 2014) 14 Bundesfachgruppen werden vom AQUA-Institut geleitet und koordiniert. Sie bestehen aus Fachexperten, die durch die Träger des G-BA, die maßgeblichen Organisationen der Patientenvertretung, die LQS/LKG (für indirekte QS-Verfahren) und das AQUA-Institut benannt werden. Auch für Bundesfachgruppenmitglieder gilt die Pflicht zur Offenlegung von Interessenkonflikten.

Die personelle Besetzung der Bundesfachgruppen ist auf der Webseite des AQUA-Instituts (www.sqg.de/hintergrund/bundesfachgruppen) einsehbar.

Die Bundesfachgruppen treffen sich derzeit regulär zwei- bis dreimal im Jahr mit jeweils konkreten Arbeitsprogrammen. In den sogenannten Frühjahrssitzungen werden

- die Ergebnisse der Bundesauswertung beraten,
- die Qualitätsindikatoren mit unterschiedlichen Handlungsbedarfen kommentiert,
- die Inhalte des anstehenden Qualitätsreports beraten,
- die Ergebnisse der Datenvalidierung und des Strukturierten Dialogs besprochen,
- die Effekte vorgenommener Änderungen an der Spezifikation/den Rechenregeln überprüft,
- Vorschläge für Änderungen in zukünftigen Spezifikationen diskutiert,
- ggf. Sonderauswertungen zu einzelnen Leistungsbereichen/Indikatoren beantragt.

Zwischen August und Oktober finden weitere Treffen zur Durchführung des Strukturierten Dialogs für die direkten Verfahren statt. Hier werden die Ergebnisse der Datenvalidierung und der Sonderauswertungen beraten. Ggf. werden Expertengespräche geführt. In den sogenannten Herbsttreffen stehen besonders die Basisarbeiten der Systempflege auf der Tagesordnung. Diskutiert und beraten werden:

- die Ergebnisse der Sonderauswertungen (insbesondere für die indirekten Verfahren),
- die Weiterentwicklungsmöglichkeiten der Leistungsbereiche,
- notwendige Anpassungen der Spezifikation (für die übernächste Veröffentlichung),

[13] Patientenbeteiligungsverordnung vom 19. Dezember 2003 (BGBl. I S. 2753), die durch Artikel 3 des Gesetzes vom 20. Februar 2013 (BGBl. I S. 277) geändert worden ist.

- Änderung/Anpassung von Qualitätsindikatoren,
- Anpassung der Referenzbereiche und der Rechenregeln (für das Folgejahr),
- Änderungsbedarfe bzgl. der Dokumentationen (Datenfelder/Ausfüllhinweise),
- ggf. der Bedarf einer Datenvalidierung.

3.6 Informelle Experten und ad-hoc-Arbeitsgruppen

Angesichts einer Fülle unterschiedlicher Projekte und oftmals eng gesteckter Zeitpläne, ist das AQUA-Institut immer wieder auch mit Aufgaben konfrontiert, die zwar der Einbeziehung externer Expertise bedürfen, die aber nicht im Rahmen bestehender oder formal zu berufender Expertengremien bearbeitet werden können.

Aus diesem Grund besteht zuweilen die Notwendigkeit, weitere Expertengespräche zu führen oder ad-hoc-Arbeitsgruppen zu bilden. Die Auswahl von Experten und die Zusammensetzung dieser Gruppen erfolgt themen- und fachspezifisch, oft aber auf informellem Wege. Soweit Experten oder ad-hoc-Arbeitsgruppen Ergebnisse und/oder Empfehlungen vorlegen, werden die betreffenden Personen in den entsprechenden Berichten namentlich genannt (mit Ausnahme einbezogener Patienten). Zu dieser namentlichen Nennung und dem Hinweis darauf, dass die Expertise in anonymer und aggregierter Form in Berichte einfließt, wird eine Einverständniserklärung eingeholt. Weiterhin unterschreiben alle informell einbezogenen Experten eine Vertraulichkeitserklärung.

3.7 Sonstige Formen der Einbeziehung von Expertise

Bei der Beteiligung an der Durchführung der Qualitätssicherung besteht ein hoher Abstimmungsbedarf mit der Länderebene (bei indirekten, länderbezogenen QS-Verfahren). In der stationären Qualitätssicherung wurde zu diesem Zweck eine Arbeitsgruppe zur Abstimmung mit den LQS gebildet. Diese Arbeitsgruppe verfolgt verschiedene Ziele:

- Einbindung der Expertise der Länder aus der Umsetzung der Qualitätssicherung für Prozesse der Systempflege und der Weiterentwicklung
- Abstimmung der Maßnahmen im Rahmen Strukturierter Dialoge und der dazugehörigen Berichterstattung
- Interpretation der Ergebnisse der Qualitätssicherung

Zur Vorbereitung der Datenvalidierung ist ebenfalls eine Arbeitsgruppe gegründet worden, in der neben Vertretern der Bänke des G-BA und Patientenvertretern auch Ländervertreter einbezogen sind. Ländervertreter bzw. Vertreter aus Fachgruppen der Länder engagieren sich auch in den Bundesfachgruppen.

Die enge Abstimmung mit der Länderebene soll helfen, Anregungen für Verbesserungen der QS-Verfahren möglichst zügig umzusetzen. Umgekehrt soll verhindert werden, dass auf Bundesebene Maßnahmen beschlossen werden, die sich in der Praxis nicht umsetzen lassen. Weil die Ländervertreter auch direkt an der Systempflege und Weiterentwicklung von Qualitätsindikatoren beteiligt sind, besteht auch hier die oben dargelegte Notwendigkeit einer Offenlegung von Interessenkonflikten.

Weitere externe Expertise wird durch das AQUA-Institut bei der Vorbereitung der Spezifikation für die Datenerfassungen einbezogen. Beispielsweise werden mit betroffenen Institutionen (z.B. Vertrauensstelle, Softwareanbieter) und Vertretern des G-BA technische Methoden zur Umsetzung der Dokumentation bei den Leistungserbringern abgestimmt.

Bei der Neu- und Weiterentwicklung von QS-Verfahren bezieht das AQUA-Institut im Regelfall externe Kodierexperten ein, insbesondere wenn Routinedaten genutzt werden sollen.

© 2015 AQUA-Institut GmbH

4 Themenerschließungen

Eine Themenerschließung ist der eigentlichen Entwicklung von Indikatoren und Instrumenten für ein QS-Verfahren vorgeschaltet; entweder explizit im Sinne eines gesonderten Auftrags zur Erstellung einer sogenannten Konzeptskizze oder implizit im Rahmen einer Verfahrensneuentwicklung. Sie dient dazu, Aspekte des Aufwands und Nutzens möglicher QS-Maßnahmen frühzeitig zu ermitteln und auf dieser Basis Schwerpunkte des weiteren Entwicklungsprozesses zu begründen.

Der Leitgedanke einer effektiven, auf wissenschaftlichen Grundlagen basierenden und sektorenübergreifend angelegten Qualitätsförderung ist die Betrachtung der Versorgung als ganzheitliches, einzelne Sektorengrenzen überschreitendes und am Patienten orientiertes Geschehen; daher sind Themenerschließungen entsprechend breit angelegt. Als mögliche Informationsquellen neben den Recherchen nach Evidenz und der Einbeziehung von Experten und Patienten kommen systematisch gesammelte Daten infrage, beispielsweise des Statistischen Bundesamtes, des Zentralinstituts für die Kassenärztliche Versorgung, der Kassenärztlichen Bundesvereinigung und Daten der einrichtungsübergreifenden Qualitätssicherung. Sofern die Beauftragung dies umfasst, können zusätzlich anzufordernde Abrechnungsdaten des InEK und anonymisierte Routinedaten von Krankenkassen in die Analysen einbezogen werden (vgl. Abschnitt 7.2).

In Abhängigkeit von der Komplexität des Themas, der konkreten Aufgabenstellung und den zur Verfügung stehenden zeitlichen Ressourcen werden die einzelnen Bausteine der Informationsgewinnung vom AQUA-Institut sachgerecht eingesetzt und flexibel genutzt. Nur im Rahmen der Entwicklung von Indikatorensets für neue QS-Verfahren werden nach Möglichkeit alle folgend geschilderten Bausteine verwendet. Im Rahmen der Weiterentwicklung bestehender QS-Verfahren (z.B. Ergänzung eines bestehenden Leistungsbereichs um ein Follow-up) können Themenerschließungen teilweise deutlich enger gefasst werden.

Bei Beauftragungen zu Neuentwicklungen von Indikatoren und Instrumenten für ein QS-Verfahren mündet die Themenerschließung im Versorgungspfad und der Darlegung von Qualitätspotenzialen, die als Arbeitsgrundlage der Indikatorenentwicklung dienen. Wenn solchen Beauftragungen keine Konzeptskizze zugrunde liegt, geht das AQUA-Institut zunächst von einer anderweitig – z.B. vom G-BA oder seinen Gremien – festgestellten Versorgungsrelevanz aus. Allerdings ist die Betrachtung der Versorgungsrelevanz auch im weiteren Verlauf von Entwicklungsarbeiten wichtig. Wenn beispielsweise Indikatoren im RAM-Prozess von den Panel-Experten (vgl. Kapitel 8) als nicht relevant beurteilt werden oder die Basis zur Erhebung benötigter Qualitätsdateninformationen ungenügend ist, reduziert sich möglicherweise die Versorgungsrelevanz eines QS-Verfahrens. Spätestens bei der abschließenden Würdigung des vom RAM-Panel zur Umsetzung empfohlenen Indikatorensets ist die Frage der Versorgungsrelevanz erneut zu stellen.

4.1 Ziele

Aus der Qualitätssicherungsperspektive liegen die Ziele einer Themenerschließung nicht darin, eine bestmögliche oder leitliniengerechte medizinisch-pflegerische Versorgung darzustellen, sondern in der Beurteilung der in dieser Versorgung liegenden Potenziale für Qualitätssicherung und -verbesserung.

Wenn eine Themenerschließung mit der Zielstellung einer möglichen Auftrags- oder Themenkonkretisierung bzw. der Unterbreitung von Entwicklungsvorschlägen an den G-BA verbunden ist („Konzeptskizze“), impliziert dies die Einschätzung der Versorgungsrelevanz des Themas und der potenziellen Umsetzbarkeit sowie die Darlegung von Qualitätspotenzialen (Qualitätsziele und -defizite sowie Potenziale zur Qualitätsverbesserung).

4.1.1 Einschätzung der Versorgungsrelevanz

Ein der Themenerschließung inhärentes Ergebnis ist die Möglichkeit, Aussagen zum Nutzen bzw. zur Versorgungsrelevanz eines Themas und damit ggf. auch zu den Zielsetzungen eines Auftrages zu treffen. Die Bewertungskategorie der Versorgungsrelevanz umfasst prospektive Gütekriterien in Bezug auf die spezifischen Zielsetzungen und Qualitätspotenziale eines zu entwickelnden QS-Verfahrens. Im Vordergrund stehen dabei drei Nutzenkategorien:

- Epidemiologischer Nutzen:
  Die Zielsetzungen des QS-Verfahrens betreffen die Versorgung von Populationen, die in ausreichend hoher Zahl oder in besonderer Schwere von einer Erkrankung/Behandlung betroffen sind.
- Patientennutzen:
  Die Qualitätsziele bzw. deren Erreichen stehen in unmittelbarem Zusammenhang mit den primären Kategorien des Patientennutzens (Verringerung von Mortalität und Morbidität, Erhöhung der Lebensqualität).
- Qualitätspotenziale:
  Die Qualitätsziele beziehen sich auf bekannte oder vermutete Qualitätsdefizite und bieten ein tatsächliches Potenzial zur Qualitätsverbesserung bzw. Ansätze für diesbezügliche Maßnahmen.

In einer erweiterten Sichtweise gehören zur Einschätzung der Versorgungsrelevanz auch empirische Informationen über die Datenlage und die Möglichkeit, unter gegebenen versorgungspolitischen Rahmenbedingungen überhaupt Ansätze zur Qualitätssicherung zu implementieren. Typische übergreifende Fragen in diesem Zusammenhang sind beispielsweise:

- Welche und wie viele Leistungserbringer erbringen welche Leistungen zu welchen Zeitpunkten und in welchem Umfang und in welchen Mengen?
- Wie differenziert werden diese Leistungen im Abrechnungsgeschehen abgebildet? Lassen sich ggf. eindeutige Ansatzpunkte zur Etablierung differenzierter Dokumentationen, also sogenannte Auslöser von QS-Dokumentationen, identifizieren?
- (Wie) Können die relevanten Endpunkte abgebildet werden?
- Welche Sektorenübergänge und -schnittstellen sind unter Qualitätsgesichtspunkten relevant und welche Informationen hiervon sind abbildbar (Überweisungen, Arztbriefe usw.)?

4.1.2 Einschätzung der Umsetzbarkeit

Jede datenbasierte Form der Qualitätssicherung muss mit der Einschränkung leben, dass qualitätsrelevante Informationen zum Teil nicht oder nur mit hohem Aufwand erhoben werden können. Abbildung 5 veranschaulicht, dass immer nur Teilmengen von qualitätsrelevanten Informationen verfügbar sind sowie dass mit den (theoretisch)[14] verfügbaren Datenquellen einerseits unterschiedliche Informationen abgebildet werden können, es andererseits aber auch Überlappungen in dem Sinne gibt, dass manche Informationen in mehr als einer Datenquelle vorhanden sind.

Gerade im Zusammenhang mit der sektorenübergreifenden Abbildung von Langzeitverläufen der Patientenversorgung tritt das Problem nicht verfügbarer oder nur mit unvertretbarem Aufwand erhebbarer Qualitätsinformationen auf. Dies hat in der Vergangenheit dazu geführt, dass entwickelte Qualitätsindikatoren mit Verweis auf den entstehenden Aufwand nicht zur Etablierung eines QS-Verfahrens geführt haben.[15] Für zukünftige Entwicklungsleistungen soll dies vermieden werden. Einerseits dadurch, dass weitere Datenquellen für die Qualitätssicherung erschlossen und neue Erhebungsinstrumente entwickelt werden, andererseits dadurch, dass mögliche Probleme in der Umsetzbarkeit von QS-Verfahren frühzeitig erkannt und für die Gestaltung des weiteren Entwicklungsprozesses berücksichtigt werden.

[14] Je nach Auftragskontext können die im Rahmen einer Themenerschließung zu berücksichtigenden Datenquellen eingeschränkt werden. Es ist aber auch möglich, dass der G-BA neue, noch nicht zur Verfügung stehende Datenquellen in die Beauftragung einschließt.
[15] Vgl. dazu das Eckpunktepapier zum „gemeinsamen Verständnis und Handlungsempfehlungen zur sektorenübergreifenden Qualitätssicherung". https://www.g-ba.de/downloads/17-98-3536/Eckpunkte-Handlungsempfehlungen-sQS.pdf (Zugriff am 13.11.2014)

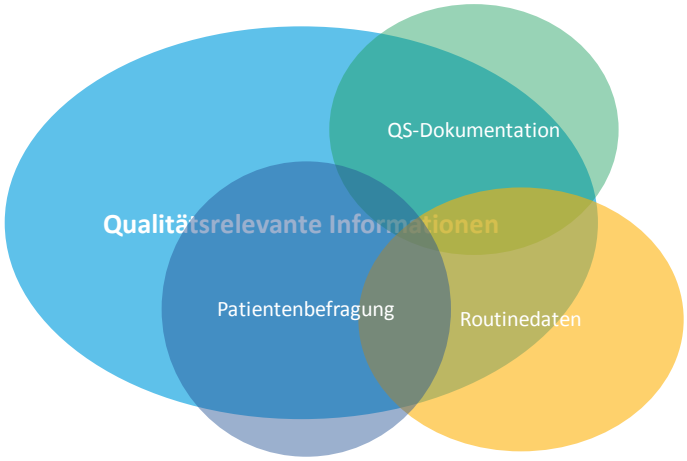

Abbildung 5: Erhebung qualitätsrelevanter Informationen

Wesentliches Ziel der Themenerschließung ist es, unter Abwägung von Aufwand und Nutzen einer Datenerhebung, die potenziellen Erhebungsinstrumente für den weiteren Entwicklungsprozess zu definieren. Darüber hinaus soll auch deutlich gemacht werden, an welchen Stellen noch Lücken in der Erhebbarkeit qualitätsrelevanter Informationen bestehen.

4.1.3 Einschätzung der Potenziale für Qualitätssicherung und -verbesserung

Ein weiterer Inhalt der Themenerschließung bezieht sich auf Qualitätspotenziale, d.h. auf Qualitätsziele und -defizite, die Ansätze zur Qualitätssicherung und -verbesserung bieten. Grundsätzlich erwartet das AQUA-Institut, dass im Themenfindungs- und Priorisierungsverfahren des G-BA Verbesserungspotenziale und Qualitätsdefizite festgestellt werden, weil beides wichtige Gründe für Beauftragungen sein sollten. Dementsprechend sollten Aufträge Hinweise auf etwaige Über- und Unterversorgung sowie ggf. auf mögliche Fehlanreize bei der Indikationsstellung beinhalten (Maynard et al. 1997; Andersen et al. 1990; Wennberg 1990).

Sollten solche Qualitätspotenziale nicht hinreichend aus dem Auftrag entnehmbar sein, kann das AQUA-Institut im Rahmen der Themenerschließung prüfen, ob sich aus der verfügbaren Literatur oder durch Expertenhinweise Hinweise auf Lücken in der Versorgung oder Indizien für Über-, Unter- oder Fehlversorgung feststellen lassen bzw. ob angemessene Versorgungsleistungen nicht oder nicht in der gebotenen Qualität erbracht werden. Gegebenenfalls ist hierfür ein gesondert zu bemessender Zeitraum anzusetzen (Badura et al. 2003).

4.2 Inhalte

Die Inhalte der Themenerschließung beziehen sich zunächst auf die Ermittlung der unmittelbar versorgungs- und damit potenziell qualitätsrelevanten medizinischen Strukturen, Prozesse und Ergebnisse. Dazu gehören die Definition von Zielpopulationen und die Darlegung der diesbezüglich relevanten Versorgungsaspekte mit Blick auf die Behandlungs- und Versorgungsziele. Jede Versorgung findet jedoch unter spezifischen rechtlichen und versorgungspolitischen Rahmenbedingungen statt, die nicht nur erhebliche Auswirkungen auf die praktische Versorgung, sondern auch auf mögliche Qualitätssicherungsverfahren haben. Die von den Rahmenbedingungen ausgehenden Probleme oder auch der Zugang zu Datenquellen bzw. die Verfügbarkeit von Erhebungsinstrumenten wird aber nur selten in der wissenschaftlichen Literatur thematisiert.

Es geht hier also um versorgungspolitische Fragestellungen, die im weiten Rahmen der evidenzbasierten Gesundheitsversorgung (Evidence-based Healthcare) angesiedelt sind, sodass eine Recherche unter dem alleinigen Gesichtspunkt der medizinischen Evidenz zu kurz greift.

4.2.1 Definition der Zielpopulation

Unter der Fragestellung „Wer wird versorgt?“ wird eine Zielpopulation definiert, die in einem bestehenden oder zukünftigen QS-Verfahren qualitätssicherungspflichtig ist oder sein soll. Dies meint die lückenlose Aufzählung aller betroffenen Subpopulationen sowie die Abgrenzung zu Subpopulationen, die aus zu nennenden Gründen nicht Gegenstand des QS-Verfahrens sind oder sein sollen. Zielpopulationen können Menschen mit bestimmten

© 2015 AQUA-Institut GmbH

Diagnosen sein, aber auch Menschen, die von bestimmten medizinischen Eingriffen (zum Beispiel der Anlage von Gefäßkathetern) oder sonstigen Maßnahmen (z.B. Rehabilitationsmaßnahmen) betroffen sind (Cochrane Collaboration; Gail et al. 2000; ICH 1999).

4.2.2 Medizinisch-pflegerische Versorgungsaspekte

Unter der Fragestellung „Wie wird versorgt?“ wird eine Versorgungsleistung bzw. eine in Bezug auf die Zielpopulation zur Anwendung kommende Versorgung ermittelt, die idealerweise evidenzbasiert ist. Dabei handelt es sich einerseits um konkrete medizinisch-therapeutische, um pflegerische oder rehabilitative Leistungen und Prozesse sowie um die Akteure, die für diese Leistung verantwortlich sind. Andererseits sind damit aber auch einrichtungsbezogene oder allgemeine Strukturmerkmale gemeint, die einen Einfluss auf das Erreichen von Versorgungszielen haben (Schwartz et al. 2003).

Insbesondere bei sektorenübergreifenden Versorgungsprozessen kommt der Analyse der Versorgungsstruktur eine herausragende Bedeutung zu. Ermittelt werden die beteiligten Akteure und Sektoren, etwaige Besonderheiten des Versorgungspfades sowie interdisziplinäre und transdisziplinäre Schnittstellen der Versorgung (Beske et al. 2005).

4.2.3 Allgemeine Qualitätsziele

Unter der Fragestellung „Was sind die entscheidenden (Qualitäts-)Parameter einer guten Versorgung“ werden die zur Beurteilung dieser Frage relevanten Versorgungsaspekte ermittelt. Dies ist im Regelfall der schwierigste Konkretisierungsschritt, da es hier um differenziert darzulegende Aspekte der tatsächlichen Versorgung geht, die erfüllt werden müssen, um die angestrebten Behandlungsziele zu erreichen. Es besteht Einigkeit darüber, dass sich eine qualitativ hochwertige Versorgung dadurch auszeichnet, dass sie gesetzte Ziele erreicht – also dass sie (medizinisch) effektiv, sicher und patientenorientiert ist sowie dass alle Zugang zu einer angemessenen Versorgung haben und diese gut koordiniert ist (Badura et al. 2003). Mit Blick auf den zu erstellenden Versorgungspfad und die Qualitätspotenziale sowie die Indikatorenrecherche und das Suchmodell für die systematische Literaturrecherche sind die allgemeinen Qualitätsdimensionen dagegen zu unspezifisch.

Dementsprechend muss Qualität hier in Bezug auf einzelne Aspekte einer angemessenen Versorgung definiert werden. Nach dem Grundsatz, dass sich eine gute Versorgungsqualität im Erreichen patientenrelevanter Versorgungsziele manifestiert und dass Effektivität und Sicherheit im gemeinsamen Interesse von Behandelnden und Behandelten liegt, muss für jedes Thema im Einzelfall konkretisiert werden, wodurch sich die Versorgungsqualität als solche ausweist und wie sie sich abbildet. In Bezug auf Effektivität und Sicherheit können zum Beispiel die richtige Diagnose, eine adäquate Indikationsstellung bzw. die Wahl der angemessenen Therapie, die fachliche Qualifikation des pflegerischen Personals, die Erfahrung eines Operateurs mit bestimmten Eingriffen von besonderer Wichtigkeit sein oder auch die technische Ausstattung des Leistungserbringers, ein effektives Überweisungs- oder Entlassungsmanagement.

Von entscheidender Bedeutung ist die Qualitätsdimension der Patientenorientierung, die in jedem Fall eine Beachtung der Perspektive des Patienten auf „seine“ Versorgung und die gewünschten Endpunkte erfordert, gerade weil patientenrelevante Endpunkte nicht immer kongruent zu einem medizinisch gesetzten Behandlungserfolg sind, z.B. Gelenkbeweglichkeit vs. Schmerzfreiheit; Heilungschance vs. Lebensqualität (Mühlhauser et al. 2009; Badura 1999). Deswegen sind die Aufklärung des Patienten und ggf. seiner Angehörigen, die Beteiligung an den Entscheidungen (Shared-Decision-Making) sowie die Ergebnisbeurteilung aus seiner Sicht überwiegend prinzipielle Qualitätsparameter einer patientenorientierten Gesundheitsversorgung (Mühlhauser et al. 2009; Badura 1999).

4.2.4 Versorgungspolitischer Rahmen und Rechtslage

Bei jeder Themenerschließung werden die versorgungspolitischen Rahmenbedingungen und die Rechtslage eruiert. Insbesondere sind für jedes Thema bzw. für die Zielsetzungen eines QS-Verfahrens zu berücksichtigen:

- die gesetzlichen Rahmenbedingungen des SGB V, in dem die gesetzliche Qualitätssicherung im Gesundheitswesen verankert ist
- die Strukturbedingungen des deutschen Gesundheitssystems

- die vom G-BA verabschiedeten allgemeinen Richtlinien zur Qualitätssicherung
- die versorgungspolitischen Zielsetzungen, wie sie konkret in den Aufträgen des G-BA formuliert sind

4.2.5 Qualitätsprojekte und -initiativen

Neue QS-Verfahren sollen einerseits keine parallelen Qualitätssicherungsstrukturen aufbauen und andererseits sollen auch etwaige Synergieeffekte genutzt werden. Da die Prüfung bestehender Qualitätsprojekte im Wesentlichen eine Prüfung der dort jeweils vorhandenen Datenbasis meint, werden ausschließlich nationale Qualitätsprojekte (z.B. Register) betrachtet. Weil darüber hinaus für die bundesweite externe Qualitätssicherung gewisse Mindestanforderungen an die Datenmenge und -qualität gestellt sind, beschränkt sich die Prüfung auf Qualitätsprojekte, die entweder von bundesweiter Bedeutung sind (z.B. Krebsregister) oder über repräsentative Datenpools (z.B. Disease-Management-Programme) verfügen.

Auch hierbei kann es sich um eine aufwendige Prüfung handeln, insofern als in diesen Fällen die Datenbasis teilweise sehr tief in ihrer Struktur geprüft werden muss.

4.2.6 Versorgungspfad

Der Versorgungspfad bündelt und visualisiert die Ergebnisse der relevanten Versorgungsaspekte entlang der Versorgungskette und über die Sektorengrenzen hinweg, wobei ihm vor allem deskriptive Aufgaben zufallen.

Im Einzelnen dient er

- der strukturierten Darstellung des Themas,
- der differenzierten Darstellung der beteiligten Sektoren,
- der Zuordnung der Verantwortlichkeiten,
- der Verortung von Qualitätspotenzialen im Versorgungsprozess,
- der Zuordnung von Ansatzpunkten der Qualitätssicherung und
- der Darstellung von Erhebungsepisoden mit korrespondierenden Datenquellen und möglichen Erhebungsinstrumenten.

In dieser Visualisierung eines Patientenwegs (ggf. auch mehrerer möglicher Patientenwege) soll der Versorgungspfad auf dem Scoping-Workshop (wenn Bestandteil einer Verfahrensentwicklung) zur Diskussion gestellt, geprüft und ggf. ergänzt werden. Die grafische Darstellung des Pfads fördert die Verständlichkeit und ergänzt die rein textuelle Beschreibung. Insofern bietet er auch für das RAM-Panel eine gute Übersicht über alle relevanten Aspekte der Indikatorenentwicklung.

Je nach Thema kann ein Versorgungspfad unterschiedlich gestaltet werden. Das Grundgerüst wird in einem Koordinatensystem dargestellt. Dabei differenziert die waagerechte Koordinate die Phasen des Krankheitsverlaufs (z.B. Vorgeschichte/prästationäre Phase; stationäre Phase, poststationäre Behandlung/Nachbehandlung; Follow-up). Die senkrechte Koordinate differenziert nach Versorgungseinrichtungen/Leistungserbringern (z.B. Hausarzt, anderer Facharzt, ambulante Therapieeinrichtung, Krankenhaus, Rehabilitationseinrichtung usw.). Gefüllt wird das Koordinatensystem mit dem Versorgungspfad, d.h. mit den spezifischen Leistungen, den Endpunkten bzw. den qualitätsrelevanten Zielen und den ermittelten Qualitätspotenzialen (vgl. dazu die folgende Abbildung).

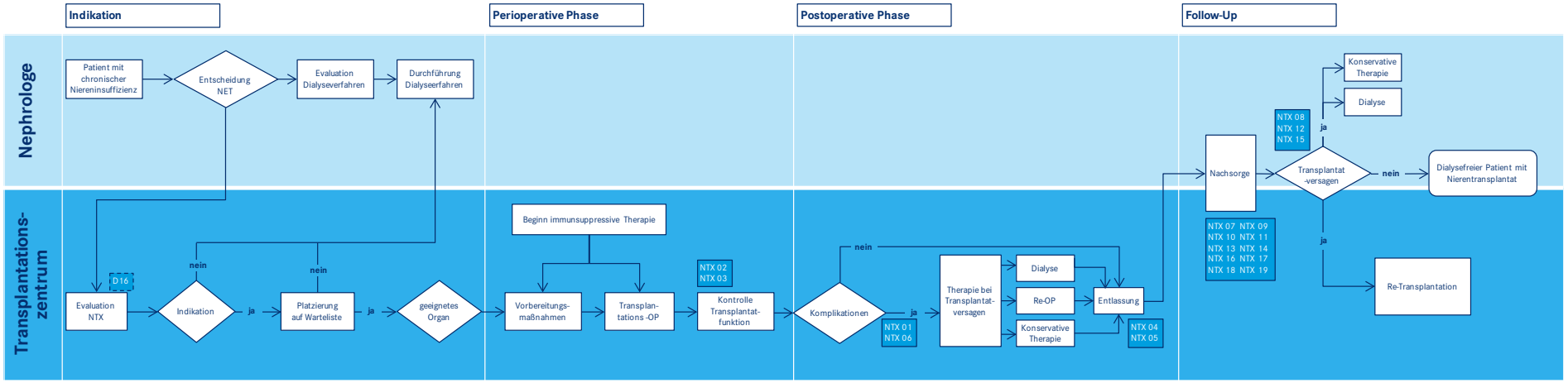

Abbildung 6: Beispiel für die Darstellung eines Versorgungspfades (hier: Nierenersatztherapie)

In späteren Entwicklungsphasen werden dem Versorgungspfad bzw. den darauf abgebildeten Qualitätspotenzialen die ermittelten Indikatoren hinzugefügt, sodass er noch zwei weitere Funktionen erfüllen kann: als Filterinstrument (vgl. Abschnitt 6.1) für das dem RAM-Panel zur Verfügung zu stellende Indikatorenregister sowie als Hilfestellung bei der abschließenden Würdigung von Entwicklungsleistungen durch das Expertenpanel (vgl. Abschnitt 8.3).

4.2.7 Qualitätspotenziale

Mit der Einführung der Qualitätspotenziale als eines die Ziele und Defizite der Versorgung fokussierenden Begriffs wurde der Entwicklungsprozess möglicher QS-Verfahren nochmals geschärft.

Mit der Ausrichtung der Qualitätssicherung an (erreichbaren) Qualitätszielen bzw. darauf bezogenen (vermuteten oder tatsächlichen) Qualitätsdefiziten rückt das Aufwand-Nutzen-Verhältnis der Qualitätsmessung noch stärker in den Mittelpunkt der Entwicklung. Die versorgungstheoretische Ermittlung der Qualitätspotenziale zieht die Frage ihrer Erhebbarkeit nach sich, d.h. der Möglichkeit, das Erreichen oder Verfehlen von Zielen systematisch und einrichtungsübergreifend zu messen. Nicht für jedes ermittelte Qualitätspotenzial können die qualitätsrelevanten Informationen erhoben werden, sei es, weil die Information in den Datenquellen schlecht oder gar nicht abgebildet wird oder (noch) keine geeigneten Datenquellen zur Verfügung stehen.

4.3 Methoden

Die Recherche nach medizinisch-pflegerischen Inhalten im weiteren bzw. nach Versorgungsaspekten im engeren Sinn fokussiert bei Themenerschließungen zunächst auf Quellen mit aggregierter Evidenz. Bei Bedarf werden ergänzend verschiedene Formen der Expertise (Experteninterviews, Fokusgruppen) einbezogen.

4.3.1 Recherche nach aggregierter Evidenz

Die Recherche nach aggregierter Evidenz umfasst die Suche nach systematischen Übersichtsarbeiten, Health Technology Assessments (HTA) und Leitlinien.

Systematische Übersichtsarbeiten

Zur Abschätzung des evidenzbasierten Wissensstandes zu einem medizinischen Thema greift das AQUA-Institut zunächst auf Sekundärliteratur in Form von systematischen Übersichtsarbeiten und Metaanalysen zurück. Dementsprechend ist die Suche nach solchen Arbeiten in der Cochrane Library (Cochrane Collaboration) obligatorisch – und hier insbesondere nach Reviews, die nach den Gütekriterien für Cochrane Reviews erstellt worden sind. Mit Blick auf die bestverfügbare Evidenz zu komplexen Versorgungsthemen sind Cochrane Reviews allerdings nicht immer ausreichend, da sie überwiegend auf die Effektivität medizinischer Prozeduren fokussieren. Dies spiegelt sich in den Evidenzbewertungen wider, die zumeist auf spezifische klinische Studien und hier im Besonderen auf Randomized Controlled Trials (RCTs) oder analoge Studientypen beschränkt sind (Fitch et al. 2001: 10) – also auf Studien, die oft nur einen bestimmten Ausschnitt der Versorgung beleuchten und überdies keineswegs zu jedem Thema verfügbar sind (Diener et al. 2009).

HTA-Recherche

Um zu einer umfassenderen Einschätzung von Entwicklungsstand und Etablierungsgrad, aber auch der Effektivität von Interventionen und angewandten (Versorgungs-)Technologien zu gelangen bzw. um die aggregierte Evidenz auch in Bezug auf soziale und ethische Voraussetzungen und gesundheitspolitische Konsequenzen zu erfassen, recherchiert das AQUA-Institut systematisch nach verfügbaren HTAs.

Die HTA-Recherche des AQUA-Instituts bezieht nationale und internationale HTAs ein. Die Suche nach nationalen HTAs erfolgt über die Datenbank der Deutschen Agentur für Health Technology Assessment (DAHTA) über das Portal des DIMDI. Internationale HTAs werden in der Datenbank des Centre for Reviews and Dissemination (CRD) des National Institute for Health Research (NIHR) recherchiert.

Das AQUA-Institut behält es sich vor, ausschließlich HTAs mit gewissen Mindestanforderungen bezüglich der ihnen zugrunde liegenden Studien bzw. der Studienqualität zu berücksichtigen (Dreier et al. 2010). Gradmesser

der Bewertung vor allem internationaler HTAs sind die Kriterien der DAHTA, mit denen es möglich ist, HTAs unterschiedlichen Evidenzklassen (vgl. Abschnitt 6.3) zuzuordnen.

Leitlinienrecherche

Zur Erfassung der Ausgangslage bezüglich der praktischen Versorgung von Patienten ist auch die Recherche nach aktuellen themenspezifischen Leitlinien obligatorisch. Unabhängig von der ggf. notwendigen Suche nach fachspezifischen und wenig verbreiteten Leitlinien wird die Recherche nach deutschsprachigen Leitlinien prinzipiell über das Internetportal „leitlinien.de“ – der Metadatenbank des Leitlinien-Informations- und Recherchedienstes des Ärztlichen Zentrums für Qualität in der Medizin (ÄZQ) – sowie über das Leitlinienportal der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF)[16] durchgeführt (Ollenschläger et al. 1999; BÄK/KBV 1997). Vervollständigt wird die Leitlinienrecherche durch die obligatorische Suche in der International Guideline Library des Guideline International Network (G-I-N), sowie den in speziellen Suchmaschinen der AHRQ (National Guideline Clearinghouse) und des NHS (NHS-Evidence Guidelines Finder).

Vorrangig wird nach Leitlinien mit hoher methodischer Güte (S3 oder vergleichbare internationale Evidenzklassen) gesucht (Behrens et al. 2006; Ollenschläger et al. 2000). Die Gütebestimmung internationaler Leitlinien erfolgt nach folgenden, allgemeingültigen Kriterien[17]:

- Aktualität: Die Leitlinie ist aktuell bzw. wird aktualisiert; es gibt Angaben zur Ablauffrist.
- Zielgruppen: Anwenderzielgruppen und/oder Anwendungsbereich sind klar definiert; Hintergrund und Ziele sind beschrieben; Patientenpräferenzen wurden berücksichtigt.
- Methodik/Methodische Standards: Entwicklungsprozess wird transparent geschildert; systematische Recherchen wurden durchgeführt; Leitlinie wurde von Experten begutachtet; beteiligte Autoren und Fachgesellschaften sind namentlich genannt; es wurden Angaben zu Interessenkonflikten gemacht.
- Empfehlungen: Die Leitlinie enthält konkrete Empfehlungen; Empfehlungsgrade mit Bezug zur Evidenz sind vorhanden.

Grenzen der Recherche

Publikationen mit aggregierter Evidenz können diverse systematische Fehler (Bias) enthalten, sei es, weil die ihnen zugrunde liegenden Studienergebnisse selbst in vielfacher Hinsicht verzerrt sein können oder weil es allgemein einen Publikationsbias zugunsten positiver Studienergebnisse gibt (Dreier et al. 2010; Thornton et al. 2000). Übersichtsarbeiten wie HTAs und Cochrane Reviews sind zudem oft hinsichtlich der Wahrnehmung von RCTs als alleinigem Evidenzkriterium überbewertet (Fitch et al. 2001: 10). HTAs sind zudem trotz ihres breit angelegten theoretischen Horizonts aufgrund ihres zumeist praktischen Auftragscharakters nicht selten auf spezifische Nutzenfragen von Kostenträgern und Leistungserbringern beschränkt (Bührlen 2010), d.h. gerade patientenrelevante soziale und ethische Aspekte werden oft nicht hinreichend gewürdigt (Droste 2008).

Expertenmeinungen sind ebenso wie Patientenmeinungen zwangsläufig subjektiv – nicht ohne Grund werden sie in der niedrigsten Evidenzklasse verortet) – und oft auch von bestimmten und nicht immer offensichtlichen Interessen beeinflusst. Insofern besteht die Gefahr, durch Expertenmeinungen bestimmte Gewichtungen vorzunehmen, die mit einer evidenzbasierten Vorgehensweise nicht vereinbar sind. Um das erschlossene Thema tatsächlich so gut wie möglich zu validieren, bedarf es daher im Regelfall einer systematischen Literaturrecherche (siehe Abschnitt 5.2). Diese kann die genannten Probleme zwar nicht lösen, doch indem sie darauf zielt, alle verfügbaren relevanten Publikationen (d.h. indikatorrelevante Primärliteratur) zu ermitteln, relativiert sie die Gefahr einseitiger Gewichtungen und beugt einer unkritischen Übernahme ggf. verzerrter aggregierter Evidenzen vor (Dreier et al. 2010). Insofern ist nicht auszuschließen, dass sich im Verlauf der im Anschluss an die Themenerschließung folgenden systematischen Recherchen neue Erkenntnisse ergeben, die auch zu Korrekturen der Qualitätspotenziale führen können.

[16] http://www.awmf.org/leitlinien.html (letzter Zugriff am 12.02.2015)
[17] Eine methodisch tiefergehende Leitlinienbewertung erfolgt nicht, da sie für die Zielsetzungen einer Indikatorenentwicklung nicht zielführend und überdies zu aufwendig ist. Vgl. dazu auch die entsprechenden Anmerkungen bei Semlitsch et al. (2014)

4.3.2 Einbeziehung von externer Expertise

Es gibt zahlreiche Themen, zu denen bisher keine aggregierte Evidenz publiziert ist. Zudem ist aggregierte Evidenz notwendigerweise immer retrospektiv und damit nicht immer aktuell. Darüber hinaus sind vor allem Informationen über Qualitätsdefizite sowie auch über die Realität der Versorgung und die praktischen Auswirkungen von versorgungspolitischen Rahmenbedingungen nur selten der Literatur zu entnehmen.

Zur diesbezüglichen Informationsgewinnung wird auch im Rahmen von Themenerschließungen auf externe Expertise zurückgegriffen. Diese kann je nach Thema und Bedarf unterschiedliche Beteiligungsformen annehmen (z.B. telefonische Leitfadeninterviews) und unterschiedliche Akteure adressieren (Mediziner, Patienten, Kodierexperten) (Sackett et al. 1997). Die möglichen Formen sowie Art und Einbeziehung von externen Experten sind detailliert in Kapitel 3 beschrieben.

Um spezielle Versorgungsaspekte und ggf. -defizite explorativ identifizieren zu können, die nur unter Einbeziehung der Patientenperspektive erhoben und bewertet werden können (z.B. Lebensqualität, Kommunikation und Information, menschlicher Umgang im Versorgungsgeschehen), können verfahrensspezifisch gebildete Fokusgruppen mit Patienten durchgeführt werden. Die Patienten sollen in diesem Rahmen von ihren Erfahrungen mit der medizinischen Versorgung berichten, insbesondere hinsichtlich ihrer Meinung zu besonders wichtigen Versorgungsaspekten oder auch der von ihnen erlebten Defizite (Morgan 2009; Fayers et al. 2000; Capitulo 1998). Die Ergebnisse der Fokusgruppenarbeit gehen in die Entwicklung verfahrensspezifischer Patientenfragebögen ein.

4.3.3 Prüfung möglicher Datenquellen

Die im Rahmen der Erschließung der Themen und Inhalte eines möglichen QS-Verfahrens identifizierten Qualitätsziele und -defizite können einer ersten Einschätzung dahingehend unterworfen werden, ob sie mittels (verfügbarer) Datenquellen abbildbar sind. Zur Beurteilung der potenziellen Abbildbarkeit der verfahrensrelevanten Strukturen, Prozesse und Ergebnisse/Endpunkte (z.B. medizinisch relevante Prozesse, patientenrelevante Endpunkte, Potenziale für Qualitätssicherung und -verbesserung) werden im Rahmen der Themenerschließung und damit bereits im Vorfeld einer Indikatorenrecherche die Fragen gestellt:

- Welche Datenquellen (z.B. Leistungserbringer, Patienten, Routinedaten) stehen zur Abbildung der verfahrensrelevanten Ziele prinzipiell (oder gemäß Auftrag) zur Verfügung?
- Sind bereits Erhebungsinstrumente vorhanden/etabliert bzw. lassen sich unter den gegebenen Bedingungen Erhebungsinstrumente entwickeln?
- Welche Aufwände sind mit der Erhebung der qualitätsrelevanten Informationen verbunden (vgl. Abschnitt 8.1.4)?

Je nach Datenquelle werden zur Einschätzung der Abbildbarkeit von Qualitätspotenzialen Teile der in Kapitel 7 beschriebenen Methoden zur Entwicklung der Erhebungsinstrumente angewendet. Das Ergebnis der Einschätzung wird in einer Matrix festgehalten, in der die Qualitätspotenziale ihrer vermuteten Abbildbarkeit gegenübergestellt werden.

Tabelle 4: Matrix zur Einschätzung der (potenziellen) Verfügbarkeit von Qualitätsinformationen

Qualitätspotenziale	Datenquelle/Erhebungsinstrument			
	Sozialdaten	QS-Dokumentation		Patienten-befragung
		fallbezogen	nicht fallbezogen	
QPot 1	-	-	+	-
QPot 2	(+)	(+)	+	(+)
QPot 3	-	-	-	-
QPot ...	(+)	-	+	+
QPot n	-	+	+	-

Legende:

+ = Die zur Abbildung des Qualitätspotenzials benötigte Qualitätsinformation kann ohne erhebliche Einschränkung aus der Datenquelle mit einem entsprechend spezifizierten Instrument erhoben werden. Dabei kann es aber z.B. sein, dass eine Datenquelle allein nicht ausreicht, sondern die gewünschten Informationen erst im Zusammenspiel mit anderen Datenquellen erhoben werden können.

(+) = Die zur Abbildung des Qualitätspotenzials benötigte Qualitätsinformation kann nur mit erheblichen Einschränkungen aus der Datenquelle erhoben werden. Darunter ist zu verstehen, dass eine Erhebung der gewünschten Informationen entweder nur eingeschränkt möglich ist (z.B. durch Ausschluss von Teilpopulationen) oder aber mit einem unverhältnismäßigen Erhebungsaufwand (z.B. sehr aufwendige Spezifizierung eines Instruments) verbunden ist.

– = Die zur Abbildung des Qualitätspotenzials benötigte Qualitätsinformation ist in der Datenquelle nicht verfügbar bzw. sie kann auch mit einem entsprechend spezifizierten Instrument nicht oder in nicht ausreichender Güte erhoben werden.

Ziele, die im Rahmen der gegebenen versorgungspolitischen Bedingungen nicht oder nur mit nicht vertretbar hohem Aufwand erreichbar sind, werden vom jeweiligen Projektteam des AQUA-Instituts als Ausschlusskriterium bei der Auswahl von Indikatoren für das Indikatorenregister verwendet, um die Umsetzbarkeit der auszuwählenden Indikatoren zu gewährleisten. Allerdings sollen diese Qualitätsaspekte im Rahmen der späteren Weiterentwicklung berücksichtigt werden, soweit geeignete Erhebungsinstrumente vorhanden sind.

5 Indikatorenrecherche

Im Rahmen von Beauftragungen zur Entwicklung von Indikatoren, Instrumenten und der notwendigen Dokumentation nimmt die systematische Suche nach Qualitätsindikatoren einen breiten Raum ein. Das Ziel der Indikatorenrecherche ist die Erstellung eines themenspezifischen Indikatorenregisters mit dem Zweck, dieses für den weiteren Bewertungs- und Auswahlprozess im Rahmen des RAM-Panels zur Verfügung zu stellen.

Jeder Indikatorenrecherche ist eine Themenerschließung vorgeschaltet (vgl. Kapitel 4). Nur auf Basis umfassender Kenntnisse eines Themas und seiner Besonderheiten, können letztlich jene konkreten Fragestellungen und Suchstrategien formuliert werden, die für jede zielgerichtete und systematische Recherche unabdingbar sind.

Die Indikatorenrecherche im Rahmen von beauftragten Entwicklungsleistungen des AQUA-Instituts schließt konzeptuell an die identifizierten Qualitätspotenziale an und besteht aus den folgenden Teilrecherchen:

- Recherche nach vorhandenen Indikatoren bei Institutionen und Agenturen des Gesundheitswesens
- Systematische Literaturrecherche

5.1 Indikatorenrecherche bei Institutionen und Agenturen des Gesundheitswesens

Weltweit wird die Qualitätssicherung der medizinisch-pflegerischen Versorgung mithilfe von Indikatoren betrieben. Diese Indikatoren und Indikatorensysteme werden von zahlreichen, teilweise privatwirtschaftlich tätigen Agenturen und Forschungseinrichtungen, aber vielfach auch von staatlich eingesetzten oder finanzierten sowie auch von unmittelbar an der Versorgung beteiligten Akteuren entwickelt (Kötter et al. 2011). Auch diese Indikatoren werden zunehmend wissenschaftlich, d.h. mit Blick auf publizierte Studienergebnisse entwickelt oder im Zuge ihrer praktischen Anwendung evaluiert und weiterentwickelt. Insofern ist die Recherche nach vorhandenen und ggf. etablierten Indikatoren die Basisarbeit und die Suche in den Datenbeständen von Agenturen mit Erfahrung in der Entwicklung und/oder Anwendung von Indikatoren zwingend.

Die Suchstrategie nach konkreten Indikatoren ist am Versorgungspfad orientiert, der in dieser Weise als Suchmodell verstanden werden muss. Dementsprechend werden Indikatorensets und einzelne Indikatoren im Zusammenhang mit eher wenigen, aber zentralen Begrifflichkeiten entlang dieses Pfads gesucht.

Leider existieren schon auf Ebene der einzelnen Institutionen und Agenturen nicht immer echte Datenbanken mit entsprechend technisch-systematischen oder gar einheitlichen Suchmöglichkeiten. Viele Indikatoren sind lediglich im Rahmen von zusammenhängenden Indikatorensets veröffentlicht, die ihrerseits nur als Dokument und nicht als Datenbank vorliegen. Folglich ist eine „systematische“ Suche nach Indikatoren hier oft nicht mehr als das Durchsuchen der Datenbestände und Publikationen der infrage kommenden Institutionen.

Deutschland betreffend, ist die Indikatorensuche durch die seit März 2011 öffentlich zugängliche Datenbank der GKV Quinth (Qualitätsindikatorenthesaurus) erheblich erleichtert worden. Ähnlich umfangreiche Suchmöglichkeiten bieten auf internationaler Ebene das Indicator and Measurement Registry der WHO sowie die Datenbank des National Quality Measures Clearinghouse (NQMC) des US Department of Health & Human Services. Zu den weiteren Standards im Rahmen der Recherche gehören die Suchen bei der Agency of Healthcare Research and Quality (AHRQ) sowie der Indikatordatenbank des Health and Social Care Information Centre (HSCIC) des NHS. Die jeweilige Liste der im Rahmen der Indikatorenrecherche durchsuchten Institutionen und Agenturen findet sich in den Anhängen zu den Entwicklungsberichten.

Die Auswahl der Indikatoren erfolgt anhand ihrer Versorgungsrelevanz sowie der ermittelten Qualitätspotenziale. Die Kriterien zur Auswahl der Indikatoren werden von zwei unabhängig voneinander bewertenden Wissenschaftlern auf jeden ermittelten Indikator angewendet. Unterschiedliche Bewertungen werden diskutiert und

konsentiert. Da es hierbei prinzipiell um die Bewertung der Relevanz von Indikatoren geht, findet keine parallele Gütebewertung der Quelle des Indikators statt.

5.2 Systematische Literaturrecherche

Im Rahmen der evidenzbasierten Medizin ist die systematische Literaturrecherche ein unverzichtbares Vorgehen (Kunz et al. 2009). Sie zielt darauf ab, den weltweit aktuellen Wissensstand zu einer konkreten Fragestellung möglichst umfassend, präzise und unverzerrt aufzuarbeiten, um diesen als Grundlage für weitere medizinische oder gesundheitspolitische Bewertungs- und Entscheidungsprozesse bereitzustellen (Vespermann et al. 2009).

Die systematische Literaturrecherche folgt im Kern der Methode, die auch von anderen gesundheitswissenschaftlich tätigen Institutionen mit evidenzbasiertem Hintergrund eingesetzt wird (IQWiG 2013: 107; Dreier et al. 2010; Vespermann et al. 2009). Sie beinhaltet:

- die Konkretisierung der Fragestellung(en)
- die Festlegung der Suchstrategie, bestehend aus
  - der Auswahl geeigneter (bibliografischer) Recherchequellen
  - der Konsentierung geeigneter (datenbankspezifischer) Suchbegriffe
  - der Übertragung der Fragestellung in konkrete Suchmodelle
  - der praktischen Durchführung der Suche(n)
  - der Auswahl potenziell relevanter Quellen
  - der Dokumentation der Ergebnisse

Im größeren Kontext der Indikatorenrecherche ist die systematische Literaturrecherche nicht allein als zusätzliche Quelle der Indikatorsuche zu verstehen. Sie dient überdies dem Zweck, Primärstudien und andere Publikationen zu finden, die die Evidenz von bereits entwickelten Indikatoren belegt bzw. infrage stellt.

Betrachtet man die systematische Literaturrecherche in einem weiten Begriffsverständnis der evidenzbasierten Gesundheitsversorgung als umfassend angelegte Suche nach bestmöglicher Evidenz, dann ist sie nicht nur auf medizinische Behandlungsformen im engeren Sinne („Technologien“) gerichtet, sondern auch auf die diese Behandlungsformen umgebenden patientenrelevanten sozialen und ethischen Aspekte.

Mit Blick auf die Vielschichtigkeit vor allem sektorenübergreifend angelegter Versorgungsmodelle lässt sich feststellen, dass Fragestellungen im Kontext der Versorgungsforschung bzw. der Entwicklung von anwendbaren Indikatorensets für die Qualitätssicherung vergleichsweise komplex sind. Allerdings ist auch die systematische Literaturrecherche aufgrund der schier unüberschaubaren Menge weltweit veröffentlichter Literatur sowie einer damit korrespondierenden Entwicklung technischer Recherchemöglichkeiten mittlerweile zu einer Aufgabe für Spezialisten geworden. Das Zusammenwirken von fachspezifisch-medizinischer und bibliothekarisch-bibliografischer Expertise ist mit Blick auf eine konsistente Suchstrategie unverzichtbar (Fitch et al. 2001: 13).

5.2.1 Suchstrategie und Auswahl der Datenbanken

Die Formulierung einer recherchierbaren Forschungsfrage ist die Voraussetzung für eine systematische Literaturrecherche. Wenn eine systematische Literaturrecherche auf enge und sehr konkrete Fragestellungen gerichtet ist – etwa die Überprüfung von einzelnen Indikatoren auf bestehende Risikoadjustierungen – kann auf eine umfängliche Themenerschließung verzichtet werden.

Im Zuge der Entwicklung von Indikatorensets für neue Qualitätssicherungsverfahren basiert die Suchstrategie auf der Themenerschließung und hier insbesondere auf dem Versorgungspfad, den identifizierten Qualitätspotenzialen und den Informationen aus der Recherche nach aggregierter Evidenz.

Am Beginn steht grundsätzlich die Auswahl der geeigneten bibliografischen Datenbanken. Eine systematische bibliografische Recherche wird prinzipiell in der Datenbank der National Library of Medicine (MEDLINE) und der

Excerpta Medica Database (Embase) durchgeführt. Letztere ermöglicht seit 2010 auch die Abfrage der MEDLINE-Datenbank. Obwohl es aufgrund des etwas anders strukturierten Thesaurus teilweise zu leicht abweichenden Ergebnissen kommen kann, wird die MEDLINE-Suche seit 2012 ausschließlich über das Embase-Portal durchgeführt.

Je nach Thema und Fragestellung sind ergänzende Suchen in weiteren bibliografischen Datenbanken – z.B. CINAHL, PSYNDEX usw. – erforderlich, für die ggf. auch die Begrifflichkeiten der Suche angepasst werden müssen. Die entsprechenden Entscheidungen über die zusätzlich auszuwählenden Datenbanken werden themen- und projektspezifisch gefällt. Wenn notwendig, werden die bibliografischen Datenbankrecherchen auch noch um eine Handsuche in ausgewählten Fachzeitschriften ergänzt, sofern diese unzureichend in den ausgewählten Datenbanken indexiert sind. Auch dies muss für den konkreten Einzelfall entschieden werden. Die über diese Handsuche gefundenen Quellen sowie mögliche weitere über Expertenkontakte und Hinweise ermittelte Literatur fließt, entsprechend gekennzeichnet, in das weitere Auswahl- und Bewertungsverfahren ein.

5.2.2 Suchmodell und datenbankspezifische Suchbegriffe

Unabhängig von der Komplexität von Versorgungspfaden und -modellen muss eine recherchierbare Forschungsfrage präzise Auskünfte zu in der Regel mindestens drei Suchkategorien einfordern:

- zur Zielgruppe (Was ist das Subjekt der Fragestellung? Welche Populationen betrifft die Fragestellung?)
- zur Versorgungsleistung (Welchen Maßnahmen oder Eingriffen wird das Subjekt der Fragestellung ausgesetzt/unterzogen?)
- zum Evaluationsinteresse (Erkenntnisinteresse; Unter welchen Gesichtspunkten wird die Versorgungsleistung bewertet?)

Je nach Fragestellung enthalten diese Suchkategorien nicht nur unterschiedliche, sondern vor allem auch mehrere Elemente in unterschiedlich aggregierten Bedeutungszusammenhängen. Soweit sich das Evaluationsinteresse nicht auf eine einzelne Prozedur und ein konkret messbares Ergebnis bezieht, sondern auf einen ganzen Versorgungspfad mit unterschiedlichen Leistungen, und man für diese Versorgungsleistung eine ihrerseits ebenso vielschichtige Qualität messen will, sind die Anforderungen an die Suchstrategie, die Formulierung der Suchbegriffe und die praktische Durchführung der systematischen Literaturrecherche deutlich erhöht.

Wenn nicht für ein einzelnes Thema anders dargelegt, erfolgt die systematische Literaturrecherche im Rahmen einer Indikatorenentwicklung in der Regel über die drei oben genannten Suchkategorien, unten exemplarisch als „Säulen“ dargestellt .

- Die erste Säule enthält die Suchwörter und -begriffe zur definierten Zielgruppe des Themas, beispielsweise Personen mit einer bestimmten Erkrankung oder einer Diagnose, die eine spezielle Versorgungsleistung erfordert.
- Die zweite Säule summiert Begriffe zu dieser definierten Versorgungsleistung bzw. zu einem Versorgungsprozess entlang eines Versorgungspfads, also der in Bezug auf die Zielgruppe zum Einsatz kommenden Behandlungsform(en) und medizinisch-pflegerischen „Technologien“.
- Die dritte Säule schließlich enthält die relevanten Evaluationsfelder, was im Zuge der Entwicklung von Qualitätsindikatoren typischerweise qualitätsrelevante Aspekte meint, ggf. aber auch andere Parameter enthält, sofern es die Fragestellung verlangt.

Zu den qualitätsrelevanten Aspekten gehört mit Blick auf das OECD-Qualitätsmodell (Arah et al. 2006) grundsätzlich auch die Patientenorientierung; d.h. die Berücksichtigung der Patientenperspektive ist ein spezifischer eigenständiger Qualitätsparameter innerhalb der Versorgungsqualität.

Innerhalb der einzelnen Suchsäulen werden sämtliche Begriffe logisch mit dem Boole’schen Operator „OR“ (oder) verknüpft, was einer Addition aller bibliografischen Quellen entspricht, in denen einer der relevanten Begriffe und Termini vorkommt. In einem zweiten Schritt werden die drei Suchsäulen untereinander mit dem Boole’schen Operator „AND“ (und) verknüpft. Auf diese Weise reduziert man die potenziell relevanten Publikationen auf jene, in denen die definierte Zielgruppe, die definierte Versorgungsleistung und die definierten Qualitätsaspekte miteinander verknüpft sind.

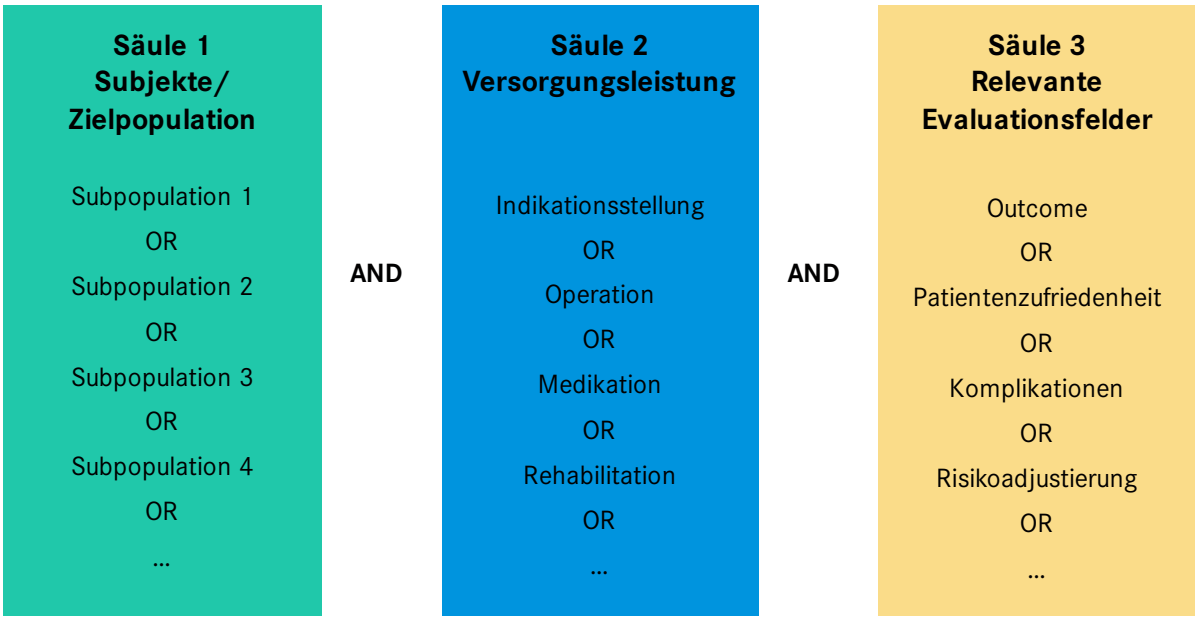

Abbildung 7: Beispielhaftes Modell für eine Suchstrategie

Je nach Thema und Fragestellung kann von diesem Modell aus sachlichen Gründen (teils erheblich) abgewichen werden.

5.2.3 Auswahl der datenbankspezifischen Suchbegriffe

Die Auswahl der datenbankspezifischen Suchbegriffe erfolgt auf Grundlage der im Rahmen der Themenerschließung erarbeiteten themenrelevanten Aspekte des Versorgungspfads.

Soweit Datenbanken die Möglichkeiten dazu bieten, wird auf hierarchisch strukturierte Thesauri wie die „MeSH-Terms“ (MEDLINE) oder den „Emtree“ (Embase) zurückgegriffen, die, wenn sachgerecht oder notwendig, mit spezifischen Schlagwörtern für eine ergänzende Freitextsuche verknüpft werden. Obwohl zentrale Suchbegriffe eigentlich durch den MeSH-Term oder im Emtree abgebildet sind, kommt der ergänzenden Freitextsuche dennoch eine Bedeutung zu, weil die entsprechende Indexierung der Publikationen durch die Datenbankbetreiber keineswegs immer fehlerfrei und vollständig ist (Droste et al. 2010). Selbst dort, wo eine hierarchische Suche zunächst unproblematisch zu sein scheint, weil man annehmen sollte, dass zumindest die Klassifizierung von Publikationen nach einem bestimmten Studientyp bei einer gelisteten Zeitschrift eindeutig ist und es insofern zu keinen Unterschieden zwischen dem „Goldstandard“ einer Handsuche und einer datenbankgestützten Suche kommt, wurden solche Fehler nachgewiesen (Cathey et al. 2006; Nasser et al. 2006).

5.2.4 Testset

Dem Anspruch an die Sensitivität einer Recherche, also der Anforderung, möglichst alle themenrelevanten Publikationen zu ermitteln, steht die Anforderung nach Präzision gegenüber, die sich in einem möglichst hohen (prozentualen) Anteil tatsächlich relevanter Publikationen an allen systematisch ermittelten Publikationen ausdrückt. Die ideale Treffermenge einer systematischen Literaturrecherche kann weder in absoluten noch in relativen Zahlen vorhergesagt werden (Droste et al. 2010) – Sensitivität und Präzision einer Recherche lassen sich erst im Nachhinein bei einer Qualitätskontrolle bestimmen.

Im Regelfall wird vor jeder systematischen Literaturrecherche ein Testset festgelegt. Dabei handelt es sich um eine Auswahl von im Zuge der Themenerschließung bereits gefundenen, relevanten Publikationen, von denen zudem bekannt ist, dass sie in den zu durchsuchenden bibliografischen Datenbanken indexiert sind. Werden diese bei der systematischen Recherche nicht annähernd vollzählig ermittelt, weist dies auf eine mangelnde Sensitivität der Recherche hin und ggf. müssen die Ausschlusskriterien überprüft und angepasst werden.

© 2015 AQUA-Institut GmbH

Im Zuge eines Testlaufs lassen sich auch erste Indizien für eine mangelnde Präzision gewinnen. So sollte eine einfache Zufallsstichprobe ausgewählter Treffer eine Mindestmenge tatsächlich relevanter Publikationen enthalten. Ist dies nicht der Fall, werden die Ausschlusskriterien ggf. verschärft.

Die Frage, wie viele der insgesamt relevanten Publikationen sich überhaupt auf dem Weg der systematischen Literaturrecherche finden lassen, hängt zum einen von der der sicheren Beherrschung der jeweiligen Abfragesprachen der zu durchsuchenden Datenbanken ab (Rauprich et al. 2010). Zum anderen spielt auch der Gegenstand der Suche eine entscheidende Rolle. Sucht man nach weitverbreiteten und etablierten medizinischen Technologien – etwa orthopädischen Standardprozeduren – generiert man mit einer „einfachen" systematischen Literaturrecherche relevante Treffermengen nahe 100 %, nicht zuletzt, weil auch die wesentlichen Fachzeitschriften in den Datenbanken gelistet und die entsprechenden Beiträge zuverlässig indexiert sind (Slobogean et al. 2009). Weniger etablierte und neue Verfahren bzw. Themen, die schwerer zu indexieren sind und ggf. in Zeitschriften erscheinen, die in den Metadatenbanken nicht gelistet sind, werden dementsprechend geringere Treffermengen generieren. Hier erfordert die Formulierung der Suchstrategie die Berücksichtigung von mehr Freitextanteilen. Auch eine ergänzende Handsuche in den entsprechenden Fachzeitschriften ist unverzichtbar (Blümle et al. 2009; Hopewell et al. 2007).

5.2.5 Auswahl der relevanten Literatur

Alle im Rahmen der systematischen und händischen Literaturrecherche ermittelten Quellen werden in ein Literaturverwaltungsprogramm importiert und nach Beseitigung der Duplikate gespeichert.

Aufgrund der notwendigerweise eher sensitiv als präzise angelegten Recherche in den großen bibliografischen Datenbanken ist nicht jede ermittelte Quelle tatsächlich relevant. Das sich anschließende zweistufige, datenbankgestützte Auswahlverfahren wird durch zwei Wissenschaftler vorgenommen, die unabhängig voneinander die Relevanz jeder Quelle beurteilen.

Im ersten Schritt werden anhand der Prüfung der Titel und der Abstracts offensichtlich irrelevante Texte ausgeschlossen. Anschließend werden die verbliebenen, potenziell relevanten Texte beschafft und anhand des Volltextes erneut bewertet. Grundlage der Auswahlmethode sind die im Vorfeld, d.h. in Bezug auf die konkrete Fragestellung des Themas, bestimmten Ausschlusskriterien. Nur wenn es bezüglich der Relevanz von Texten zu unterschiedlichen Bewertungen durch die beiden Gutachter kommt, werden die entsprechenden Fälle konsentiert – im Bedarfsfall unter Hinzuziehung weiterer Expertise.

5.2.6 Rechercherbericht

Wenn und soweit ein Expertengremium (RAM-Panel oder Bundesfachgruppe) für ein Auswahl- und Bewertungsverfahren ausgewählt ist, erhält dieses vor Beginn der tatsächlichen Indikatorbewertungen einen Rechercherbericht. Damit wird einerseits der Zweck verfolgt, frühzeitig Transparenz über die Recherche und die Recherchebedingungen herzustellen. Andererseits soll damit die (seit jeher bestehende) Möglichkeit gestärkt werden, dass Experten frühzeitig auf weitere oder fehlende Studien hinweisen können. Der Rechercherbericht enthält den Auftrag bzw. die Zielsetzungen des zu entwickelnden Verfahrens in Kurzform, das daraufhin entwickelte Suchmodell sowie die Ergebnisse der für eine Volltextanalyse ausgewählten Literatur in Form eines Literaturverzeichnisses.

5.2.7 Datenextraktion

Im letzten Schritt werden aus der tatsächlich relevanten Literatur die Indikatoren und die ihnen direkt zugehörigen bewertungs- und auswertungsrelevanten Angaben (Zähler, Nenner, Zielstellungen, Begründung usw.) extrahiert. Jeder aus der Literatur ermittelte Indikator wird für das Indikatorenregister (siehe Kapitel 6) aufbereitet, d.h. die Angaben zu jedem Indikator werden auf ein standardisiertes Datenblatt übertragen, das alle in der Quelle verfügbaren Informationen zum Indikator zusammenfasst. Des Weiteren werden relevante Primärstudien und andere Publikationen dazu genutzt, die Evidenz von bereits entwickelten bzw. vorhandenen Indikatoren zu unterstützen.

6 Erstellung des Indikatorenregisters

Ziel der Indikatorenrecherche ist die Erstellung eines Indikatorenregisters, das zunächst die Ergebnisse der Recherche abbildet und die Arbeitsgrundlage für das weitere Auswahl- und Bewertungsverfahren im Rahmen des RAM-Prozesses bildet. Der weitere Auswahl- und Bewertungsprozess, dessen Kernelement der RAM-Prozess ist, zielt darauf ab, ein Set von relevanten und umsetzbaren Indikatoren zu entwickeln, das die Qualität zu einem komplexen medizinisch-pflegerischen Themenbereich abbildet.

6.1 Filterung des vorläufigen Indikatorenregisters

Die Erfahrungen aus den bisherigen Entwicklungsprozessen haben gezeigt, dass das Ergebnis der Indikatorenrecherche in dem Sinne unbefriedigend sein kann, dass es auch Indikatoren enthält, die zwar in einem weiteren Begriffsverständnis „relevant“ sind, aber mit den konkreten Qualitätspotenzialen nur wenig oder gar nicht in Verbindung stehen. Bei der Auswahl und Bewertung der Indikatoren durch das RAM-Panel ist es sinnvoll, zumindest die im Sinne der definierten Qualitätspotenziale ungeeigneten Indikatoren nicht in das Indikatorenregister aufzunehmen.

Die zentralen Kriterien, um recherchierte Indikatoren nicht ins Register zu überführen, beziehen sich direkt auf den Versorgungspfad bzw. auf die in ihm abgebildeten Qualitätspotenziale. Um nicht sachgerechten Entscheidungen bezüglich eines Ausschlusses von Indikatoren vorzubeugen, werden alle fraglichen Indikatoren von den Projektmitarbeitern auf die folgenden fünf Punkte hin überprüft:

- Bezieht sich der Indikator auf die im Versorgungspfad genannte Zielpopulation oder eine der als relevant erachteten Subpopulationen?
- Bezieht sich der Indikator auf einen der im Versorgungspfad genannten Bereiche mit vermuteten oder recherchierten Defiziten?
- Bezieht sich der Indikator auf eine Bewertung der im Versorgungspfad genannten und als relevant erachteten Leistungserbringer (was gleichzeitig die Regelungsbereiche einschließt)?
- Bezieht sich der Indikator auf im Versorgungspfad genannte Qualitätspotenziale bzw. auf als thematisch relevant erachtete Strukturen, Prozesse oder Ergebnisse?
- Lässt sich der Indikator mit den vorgesehenen Erhebungsinstrumenten messen bzw. lässt er sich so anpassen, dass er mit ihnen gemessen werden kann?

Wenn die die bewertenden Personen unabhängig voneinander und übereinstimmend zu dem Schluss kommen, dass ein bestimmtes Kriterium von einem betreffenden Indikator nicht erfüllt ist und sich auch keine Anpassungsmöglichkeiten ergeben, kann von einer Übernahme in das Indikatorenregister abgesehen werdenEine möglicherweise schlechte Operationalisierbarkeit von Indikatoren ist an dieser Stelle kein Ausschlussgrund.

6.2 Ergänzungen des Registers

Die bisherigen Erfahrungen des AQUA-Instituts zeigen, dass die in der Indikatorenrecherche gefundenen Indikatoren oft nicht alle Aspekte des Versorgungspfads oder der Qualitätspotenziale abdecken. In diesem Fall ist es notwendig, das Indikatorenregister zu ergänzen. Hierfür kommen zwei Möglichkeiten in Betracht:

- Die Entwicklung oder Ableitung von Indikatoren aus Leitlinien, was aufgrund des relativ hohen Verwandtschaftsgrads der ihnen zugrunde liegenden Entwicklungsprozesse eine vergleichsweise ressourcenschonende Möglichkeit ist (Kötter et al. 2011).
- Die Neuentwicklung auf Grundlage von evidenzbasierter Literatur (bzw. adaptierbaren Indikatoren aus anderen Zusammenhängen) und/oder Routinedatenanalysen.

Beide Ansätze müssen jeweils um die Expertensicht ergänzt werden, weshalb auch abgeleitete oder neu entwickelte Indikatoren zwingend in den RAM-Prozess einmünden.

© 2015 AQUA-Institut GmbH

6.2.1 Ableitung von Indikatoren auf Basis von Leitlinien

Systematisch entwickelte Leitlinien, insbesondere wenn sie von verschiedenen Fachgesellschaften gemeinsam getragen werden, sind aufgrund der Methodik bei der Erstellung und den aus der Evidenz abgeleiteten Empfehlungen eine wichtige Quelle für die Entwicklung von Qualitätsindikatoren. Aus diesem Grunde soll der Leitlinienbezug bei der Entwicklung von Verfahren zur Qualitätssicherung noch weiter gestärkt werden.

Bisher gibt es keine allgemein akzeptierte Vorgehensweise zur Auswahl von Leitlinien oder Einzelempfehlungen aus Leitlinien, die für die Generierung von Qualitätsindikatoren geeignet sind. Ein Review zu diesem Thema weist in einigen Arbeiten angewandte Kriterien aus. Dazu gehören „Methodische Qualität“, „Aktualität“ und „Anwendbarkeit auf den ausgewählten Themenbereich“ (Kötter et al. 2012).

Unabhängig von den übergreifenden Gütekriterien einer Leitlinie, kann nur überprüft werden, ob auf Basis von Einzelempfehlungen der Leitlinie von den Autoren entsprechende Qualitätsindikatoren vorgeschlagen werden und ob für diese eine Entwicklungsmethodik beschrieben ist. Die aus Leitlinien entnommenen Qualitätsindikatoren werden mit den Ergebnissen der Indikatorenrecherche abgeglichen und gehen gesondert gekennzeichnet in das RAM-Panel ein.

Da bisher nur für wenige AWMF-Leitlinien Qualitätsindikatoren seitens der Leitlinienautoren entwickelt wurden, ist der häufigere Fall, dass Qualitätsindikatoren aus den Leitlinienempfehlungen abgeleitet werden müssen. Für diesen Entwicklungsschritt gibt es verschiedene, mehr oder weniger umfassende Ansätze und Empfehlungen (Kötter et al. 2011; ÄZQ 2009; Wensing et al. 2008; Mourad et al. 2007; Ouwens et al. 2007; Wollersheim et al. 2007; Hermens et al. 2006; Eccles et al. 1996), die aber im Kern in das vom Institut gewählte Verfahren zur Bewertung von Indikatoren integrierbar sind. Je präziser die Empfehlungen im Hinblick auf die angegebenen Evidenzlevel, die einzuschließenden Patientengruppen usw. in der Leitlinie beschrieben sind, desto einfacher ist deren Übersetzung in Qualitätsindikatoren.

Bezüglich der Frage, nach welchen Kriterien Empfehlungen aus Leitlinien für die Erstellung von Qualitätsindikatoren ausgewählt werden sollten, konnten Kötter et al. (2012) nur recht allgemeine Empfehlungen wie „Einfluss auf Krankheitslast der Patienten“, „Value for money“, „Einfluss auf Qualität der Versorgung“ und „Umsetzbarkeit“ identifizieren. Das AQUA-Institut ist der Ansicht, dass hier die in der Themenerschließung ermittelten Qualitätspotenziale die ersten Orientierungen bieten sollten.

Leitlinienempfehlungen werden von der Projektgruppe extrahiert, den Qualitätspotenzialen zugeordnet und auf folgende Fragen hin geprüft:

- Entstammt die Empfehlung einer Leitlinie, die allgemeine, zentrale Gütekriterien erfüllt (siehe Abschnitt 4.3.1)?
- Entspricht die Zielgruppe der Einzelempfehlung der Zielgruppe des QS-Verfahrens bzw. des Indikators?
- Ist die Empfehlung aktuell?
- Welchen Stärkegrad hat die Empfehlung?
- Welche Evidenz ist für die Empfehlung ausgewiesen?
- Ist beschrieben, welche Risiken für Patienten mit der Umsetzung der Empfehlung verbunden sein können?

Die Leitlinienempfehlungen werden anschließend mit Angabe des Zählers und Nenners in Indikatorvorschläge „übersetzt“. Ergeben sich hierbei Schwierigkeiten oder Verständnisprobleme, wird mit den Leitlinienautoren Rücksprache gehalten. Diese Vorschläge fließen dann in das Indikatorenregister ein, wobei die jeweils zugrunde liegende konkrete Leitlinienempfehlung mit Angabe des Empfehlungsgrades im Originaltext mit auf dem Indikatordatenblatt ausgewiesen wird. Falls möglich wird dann, wenn eine AWMF-S2- oder -S3-Leitlinie oder eine Nationale Versorgungsleitlinie für das entsprechende Thema vorliegt, ein Leitlinienautor in das RAM-Panel einbezogen.

6.2.2 Neuentwicklung von Indikatoren

Es gibt viele Bereiche in der medizinisch-pflegerischen Versorgung, für die es weder Leitlinienempfehlungen noch sonstige aus der Literaturrecherche ableitbare und somit gut begründete Qualitätsindikatoren gibt. Dabei

können, ausgehend von einer Analyse mit Routine- bzw. Sozialdaten, Hinweise entstehen, dass relevante Qualitätsdefizite in der Versorgung existieren. Ein typisches Beispiel wäre, dass in der Betrachtung von Follow-up Ereignissen eines QS-Verfahrens Diagnosen bzw. Prozeduren identifiziert werden, die als qualitätsrelevante Ereignisse vorangegangener Prozeduren angesehen werden können (z.B. Komplikationen, stationäre Wiederaufnahmen). Das methodische Vorgehen ist in Abschnitt 7.2 beschrieben.

Daneben können auf der Grundlage von Expertenempfehlungen (z.B. beim Scoping-Workshop) oder einer zusätzlichen Literaturrecherche nach Einzelstudien Indikatoren vom Projektteam selbst entwickelt und vorgeschlagen werden. Auch diese Indikatoren fließen – gesondert begründet und entsprechend gekennzeichnet – in das Indikatorenregister und damit in das RAM-Panel ein.

6.2.3 Indikatoren zur Patientenperspektive

Sofern Qualitätspotenziale und Ziele identifiziert wurden, die auf Basis von Patientenaussagen erhoben werden sollen und die Beauftragung die Entwicklung einer Patientenbefragung umfasst, werden hierfür Qualitätsindikatoren in einem eigenständigen Prozess entwickelt und im Indikatorenregister ergänzt (siehe Abschnitt 7.3).

6.3 Aufarbeitung der Indikatoren des Registers

Alle im Zuge der Recherche ermittelten und potenziell für das zu entwickelnde QS-Verfahren geeigneten Indikatoren werden zusammen mit den ggf. aus Leitlinien abgeleiteten oder neu entwickelten Indikatoren in einem Register zusammengestellt. Dieses bildet in entsprechender Aufarbeitung die Arbeitsgrundlage des weiteren Auswahl- und Bewertungsvorgangs im Rahmen des RAM-Prozesses.

Das Indikatorenregister besteht aus einzelnen standardisierten Formularen, auf denen die jeweils für einen Indikator relevanten Informationen zusammengefasst sind. Die einzelnen Datenblätter zu jedem Indikator sind die Informationsgrundlage für die Teilnehmer des RAM-Panels. Im Auswahl- und Bewertungsprozess werden die Indikatordatenblätter mitgeführt und im Sinne eines stets zu aktualisierenden Verlaufs- und Ergebnisprotokolls kontinuierlich ergänzt und modifiziert. Das Ziel ist ein Register mit vollständig operationalisierten Indikatoren, die vom AQUA-Institut zur Umsetzung in einem Set empfohlen werden. Das Indikatorenregister bildet eine wesentliche Informationsgrundlage beispielsweise für die Spezifizierung von Erhebungsinstrumenten (vgl. Abschnitt 10.1) und die spätere Berichterstellung im Regelbetrieb (vgl. Abschnitt 12.3.1)

Tabelle 5: Basisangaben des Indikatordatenblattes

Indikator-ID:	Name des Indikators
Beschreibung	Möglichst allgemeinverständliche Beschreibung des Indikators.
Zähler	Zu messendes Qualitätsmerkmal bzw. definierte Auffälligkeit.
Nenner	Grundgesamtheit bzw. Zielpopulation, auf die sich der Indikator bezieht.
Ausschlusskriterien	Nennung von Kriterien zur Definition von Subpopulationen, die nicht in die Berechnung des Indikators einfließen sollen, z.B. Altersgrenzen oder Komorbiditäten. Bei Indikatoren, die sich auf Einrichtungen beziehen, können Ausschlusskriterien für ganze Einrichtungen definiert werden.
Qualitätsziel	Welches Ziel wird erreicht, wenn der Indikator erfüllt wird? Welche Ausprägung des Indikators stellt eine gute Versorgung dar?
Begründung (Rationale)	Begründung der Relevanz des Indikators und der diesbezüglich hinterlegten Evidenz in der Literatur. Einzelne Aussagen werden hinsichtlich ihrer Evidenz (vgl. Abschnitt XXX) klassifiziert und durch entsprechende Studien belegt. Wenn und soweit es zum Indikator Handlungsempfehlungen (in Form von Leitlinien, Standards oder analogen Quellen) gibt, werden diese im Feld: „Stärke der Empfehlung“ (s.u.) vermerkt.
Prozess-/Ergebnisverantwortung	Wer erbringt die indikatorrelevante Leistung?

Indikator-ID:	Name des Indikators
Dokumentations-verantwortung	Wer dokumentiert die indikatorrelevante Leistung?
Originalformulierung	Formulierung/Zitat des Indikators in der Originalfassung + Quellenangabe.
Indikatorvarianten	Verweise auf ähnliche Indikatoren (ggf. leicht abweichende Formulierungen, Referenzbereiche, Zielpopulationen usw.).
Anmerkungen	Anmerkungen zu relevanten Besonderheiten oder speziellen Eigenschaften des Indikators (Definitionszusätze, Risikoadjustierungsbedingungen, Zielpopulationen usw.) sowie zu redaktionellen Anpassungen, die bei der Erstellung des Indikatorblattes durch das AQUA-Institut vorgenommen wurden.
Indikatortyp	Typisierung des Indikators (Struktur, Prozess, Ergebnis).
Qualitätsdimension	Zuordnung des Indikators zu einer Qualitätsdimension (Effektivität, Patientensicherheit, Patientenorientierung, Zugang zu und Koordination der Versorgung).
Stärke der Empfehlung (gem. Leitlinie oder Standard)	Wenn Empfehlung (Leitlinie, Standard) hierzu vorliegt: Angabe der Literaturquelle und der Einzelempfehlung inkl. ihres Empfehlungsgrades.
Änderungsprotokoll	Hier wird dokumentiert, welche inhaltlichen Veränderungen im Indikatordatenblatt im Verlauf des Entwicklungsprozesses vorgenommen wurden.
Spezifizierung zur Berechnung	
Besonderheiten der Berechnung	Zum Beispiel Angabe zur Gewichtung/Priorisierung bei einer Indexbildung; Zusammenführung von Daten, die zu unterschiedlichen Zeitpunkten erhoben werden usw.
Referenzbereich	Angabe des Referenzbereichs (fixe oder perzentilbasierte Vorgaben) bzw. Angabe darüber, ob ein Referenzbereich erst nach Verfügbarkeit empirischer Daten festgelegt werden soll.
Risikoadjustierung	Ist eine Risikoadjustierung vorgesehen. Welche möglichen Risikofaktoren liegen vor.
Datenquellen	Datenquellen und Datenarten, z.B. Sozialdaten der Krankenkassen

Jeder Indikator verfügt über eine (im Rahmen eines zu entwickelnden Verfahrens) eindeutige *Identifikationsnummer (ID).* Diese wird während des gesamten Entwicklungsprozesses bis zum Abschluss einer etwaigen Erprobung im Kern beibehalten und gibt ggf. über Modifizierungen/Versionen des Indikators Auskunft.[18] Jeder Indikator hat darüber hinaus einen möglichst kurzen, prägnanten und verständlichen *Namen,* der im ersten Feld des Datenblattes einzutragen ist. Die folgende aussagekräftige und möglichst allgemeinverständliche *Beschreibung* des Indikators ist wichtig, weil es das Ziel ist, mittelfristig alle Indikatoren zu veröffentlichen.

In den beiden nächsten Feldern folgen die Basisangaben zur Operationalisierung von Indikatoren, namentlich die Definition eines *Zählers* und eines *Nenners* (Mainz 2003: 524).[19] Der Nenner zählt die Grundgesamtheit der Fälle (z.B. Patienten mit einer bestimmten Krankheit/Diagnose), für die ein Qualitätsmerkmal bzw. eine Auffälligkeit erhoben werden soll. Der Zähler repräsentiert die interessierenden Qualitätsmerkmale, respektive die Auffälligkeiten (JCAHO 1989).[20] Ein Spezialfall der Operationalisierung ist der sogenannte Sentinel-Event, der unabhängig von der Anzahl der interessierenden Ereignisse im Zähler eine eigene Aussagekraft entfaltet, indem

[18] Mit der Übernahme in den Regelbetrieb wird eine neue Indikator-ID vergeben. Hierbei handelt es sich um einen automatisch generierten ganzzahligen Wert. Dieser Wert ist/bleibt über sämtliche Leistungsbereiche und Erfassungsjahre eindeutig.

[19] Die hier zum Tragen kommende Definition von Qualitätsindikatoren impliziert eine weitgehend synonyme Verwendung des Begriffs „Qualitätsindikator“ zum Begriff der „Kennzahl“, nämlich als einer Messgröße, die aus Zähler und Nenner besteht (mit dem Sonderfall des Sentinel-Event). Wie viele Konzepte in der Gesundheitsforschung ist jedoch auch das Konzept der Qualitätsindikatoren nicht eindeutig definiert (Donaldson et al. 1999:18ff; Sens et al. 2007:30ff). Im Qualitätsindikatorensystem der bestehenden stationären Qualitätssicherung (BQS 2009; AQUA 2010) beispielsweise stehen Qualitätsindikator und Kennzahl in einer hierarchischen Beziehung zueinander: Während der Qualitätsindikator das zu messende theoretische Konstrukt definiert, beschreibt die Kennzahl die gewählte Art der Messung(en). Sofern nicht zwingend notwendig, soll diese Unterscheidung im Kontext der sektorenübergreifenden Qualitätssicherung nicht mehr vorgenommen werden.

[20] Im Kontext der Qualitätssicherung wird oft von anteils- bzw. ratenbasierten Indikatoren gesprochen (Mainz 2003: 524). Der Begriff „Rate“ ist im epidemiologischen Sinn erst dann gerechtfertigt, wenn im Nenner eine Personenzeit (also eine Beobachtungszeit) steht (Rothman et al. 2008). Allerdings kann z.B. eine durchschnittliche Verweildauer bei einer gegebenen Nennerpopulation über eine Multiplikation der beiden Elemente als eine Personenzeit und der resultierende Indikator als eine kumulative Inzidenzrate interpretiert werden.

bereits ein einziges Ereignis unmittelbare Konsequenzen (z.B. die Einleitung qualitätssichernder Maßnahmen) nach sich zieht, weil das Eintreten eines solchen Ereignisses als besonders gravierend und qualitätsrelevant angesehen wird (JCAHO 1989).

Es folgt das Feld mit den *Ausschlusskriterien*, in dem Parameter zur Definition von Subpopulationen (oder auch Einrichtungen oder Fallkonstellationen) genannt werden, die aus bestimmten Gründen nicht in die Berechnung des Indikators einbezogen werden sollen (z.B. Altersgrenzen, Komorbiditäten usw.). Soweit in der Originalquelle vorhanden, werden anschließend das *Qualitätsziel* (mit einer „Richtungsangabe") und die in der Studienliteratur genannte *Begründung* für dieses Ziel genannt. Wenn möglich wird die Evidenz der begründenden Aussage(n) entsprechend des folgenden Schemas klassifiziert.

Tabelle 6: Evidenzklassen und Studientypen[21]

Evidenz-klasse	Anforderung
Ia	mindestens ein systematischer Review auf der Basis methodisch hochwertiger kontrollierter, randomisierter Studien (RCTs) [z.B. Metaanalysen, Übersichtsarbeiten]
Ib	mindestens ein ausreichend großer, methodisch hochwertiger RCT
IIa	mindestens eine hochwertige Studie (ohne Randomisierung) [Kohorten-Studie, Kontrollierte Klinische Studie]
IIb	mindestens eine hochwertige Studie eines anderen Typs quasi-experimenteller Studien [Fall-Kontroll-Studie]
III	mehr als eine methodisch hochwertige nichtexperimentelle Studie [Querschnittsstudie, Ökologische Studie, Fallbericht, Vorher-Nachher-Studie, Registerstudie]
IV	Meinungen von angesehenen Autoritäten (aus klinischer Erfahrung), Experten- und Fachkommissionen

In den folgenden Feldern werden die Verantwortlichkeiten der Durchführung (*Prozess-/Ergebnisverantwortung* und der Dokumentation (*Dokumentationsverantwortung*) angegeben. Diese Angaben sind gleichzeitig entscheidende Aspekte der Auswertung bzw. entsprechender Rückmeldeberichte – insbesondere wenn bei Follow-up-Indikatoren die Verantwortlichkeiten nicht bei einem Leistungserbringer allein liegen.

Wenn der Indikator aus einer nicht deutschsprachigen Quelle stammt, wird auch die *Originalformulierung* des Indikators im Indikatordatenblatt aufgeführt. Falls während der Recherche ähnliche Indikatoren bzw. gleiche Indikatoren in mehreren Versionen (z.B. unterschiedliche Referenzbereiche, leicht abweichende Zielpopulationen) entdeckt werden, dokumentiert das Projektteam diese *Indikatorvarianten* und ergänzt diese Informationen ggf. noch um zusätzliche *Anmerkungen*.

Um den Messkontext des Indikators zu beschreiben, wird er nach den international gebräuchlichen Kriterien als Struktur-, Prozess- oder Ergebnisindikator *typisiert* (Hussey 2007; Donabedian 1966) und einer *Qualitätsdimension* im Sinne des OECD-Modells (Arah et al. 2006) zugeordnet. Wenn der Indikator aus einer Leitlinie abgeleitet wurde bzw. einer Leitlinie entnommen wurde oder zu diesem Indikator eine Leitlinienempfehlung existiert, wird die entsprechende *Empfehlungsstärke* der Leitlinienempfehlung nach Vorgabe der Originalquelle aufgeführt. Das Feld *Änderungsprotokoll* hält die im Verlauf des RAM-Prozess vorgenommenen Änderungen am Indikator fest.

In den vier folgenden Feldern werden die Basisangaben, wenn erforderlich, um spezifische Angaben zur Berechnung ergänzt: Zunächst um etwaige *Besonderheiten zur Berechnung* des Indikators (Besonderheiten, die

[21] Die Einteilung bzw. Zuordnung basiert auf einer frühen Leitlinie der Agency for Health Care Policy and Research (AHCPR 1992, Apendix B) und wird in dieser bzw. in leicht abgewandelten Formen u.a. vom Deutschen Netzwerk für evidenzbasierte Medizin (http://www.ebm-netzwerk.de/was-ist-ebm/images/evidenzklassen.jpg/view, Zugriff am 21.10.2014) und der Cochrane Collaboration (http://www.cochrane.de/de/evidenz-empfehlung, Zugriff am 21.10.2014) empfohlen.

außerhalb der üblichen Berechnung anzusiedeln sind) sowie anschließend um *Angaben zum Referenzbereich* (sollte sich kein Referenzwert aus der Literatur oder praktischen Anwendung ermitteln lassen, legt das AQUA-Institut einen perzentilbasierten Referenzbereich fest) und zur *Risikoadjustierung*. Das letzte Feld in diesem Zusammenhang nennt die (möglichen) Datenquellen, über die der Indikator erhoben werden soll.

Der letzte Block des Indikatordatenblattes (oben nicht abgebildet) enthält die Dokumentationsvorgaben bzw. die erforderlichen *Datenfelder zur Berechnung*. Dieser Block ist je nach Erhebungsinstrument (QS-Dokumentation, Patientenbefragung, Spezifikation für Routinedaten) anders gestaltet (vgl. dazu auch die folgenden Abschnitte 7.1 bis 7.3).

Jedes Indikatordatenblatt schließt mit den *Literaturangaben* zum Indikator.

Je nach Prozessschritt, Informationszweck und Adressaten werden die Indikatordatenblätter in unterschiedlichen Aufarbeitungen dargestellt. Für das RAM-Panel werden den Datenblättern für beide Bewertungsrunden zeitweise spezielle Bewertungsfelder hinzugefügt. Die Datenfelder zur Dokumentation bzw. zur Erhebung werden nur für die im RAM-Panel als „relevant“ konsentierten Indikatoren erstellt und damit erst zur zweiten Bewertungsrunde hinzugefügt. In den Berichten über die Entwicklungsleistungen findet sich im Textteil eine inhaltsbezogene Darstellung, die auf die Darstellung der zu erhebenden Datenfelder verzichtet, aber ein zusätzliches Feld enthält, das den Diskussionsverlauf im RAM-Panel und ggf. abweichende Voten zum jeweiligen Indikator dokumentiert.

Die vollständigen Indikatordatenblätter mit den operationalisierten Indikatoren, inklusive der zur Berechnung notwendigen Angaben, bilden das Indikatorenset. Dieses wird als eigenständiges Dokument von den Abschlussberichten abgekoppelt und im Verlauf der weiteren Entwicklung bzw. der Erprobungen mitgeführt und nach jedem Entwicklungsschritt mit fortlaufenden Versionsnummern und einer Änderungsübersicht versehen.

© 2015 AQUA-Institut GmbH

7 Entwicklung der Erhebungsinstrumente

Gegenwärtig stehen im Rahmen der Qualitätssicherung drei unterschiedliche Datenquellen/Erhebungsinstrumente für die Operationalisierung von Qualitätsindikatoren zur Verfügung: Erhebungen bei Leistungserbringern/Einrichtungen (QS-Dokumentation), Nutzung von Routinedaten (z.B. von Krankenkassen) und Erhebungen bei Patienten (siehe Abbildung 8).[22]

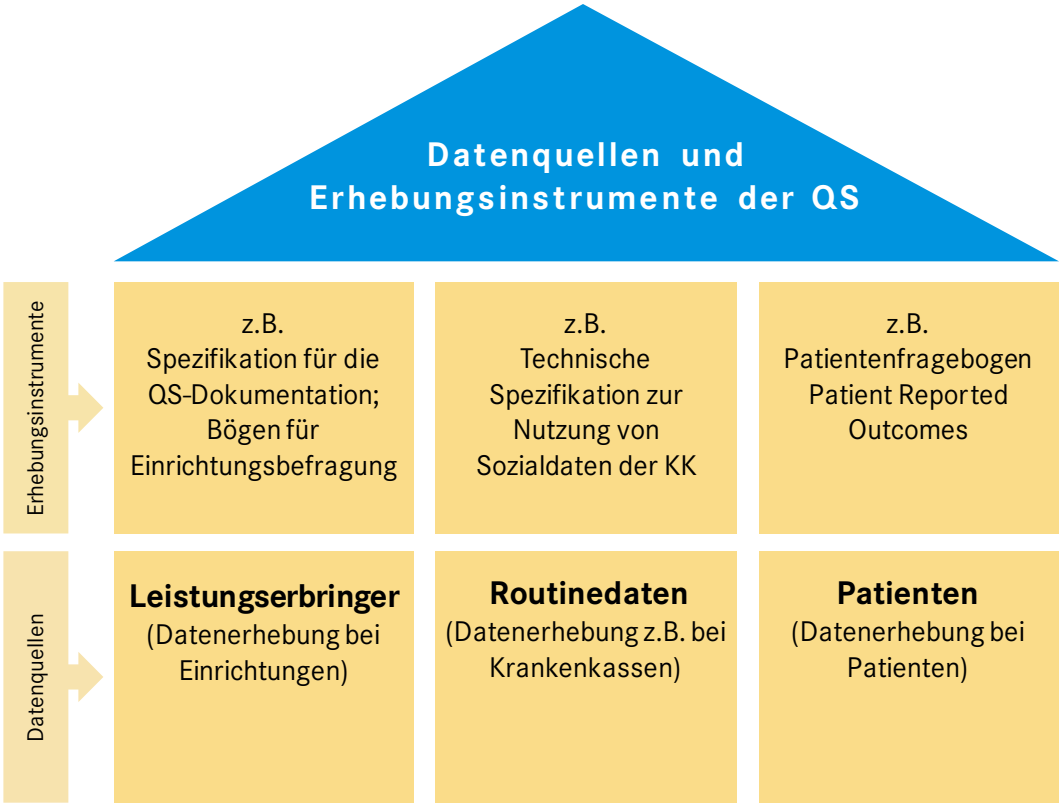

Abbildung 8: Datenquellen und Erhebungsinstrumente der Qualitätssicherung

Für alle im Rahmen der Qualitätssicherung erhobenen bzw. zu nutzenden Daten gelten die gesetzlichen Bestimmungen des Datenschutzes. Insofern ist bei der Entwicklung von Erhebungsinstrumenten stets zu berücksichtigen, dass für jedes zu erhebende Datum ein korrelierender Zweck anzugeben ist.

Für jede der grundsätzlich verfügbaren Datenquellen existiert eine unterschiedliche methodische Vorgehensweise zur Entwicklung der Erhebungsinstrumente. Im Folgenden werden die Grundlagen hierzu beschrieben.

7.1 Datenerhebungen bei/über Einrichtungen

Datenerhebungen bei/über Einrichtungen bilden derzeit die am häufigsten verwendete Datenquelle in der datenbasierten Qualitätssicherung. Es lassen sich bezüglich der Daten folgende Unterscheidungen treffen:

- Angaben der Einrichtungen, die sich auf einen Zeitraum beziehen (z.B. Anzahl behandelter Patienten in einem Leistungsbereich)
- Angaben der Einrichtungen, die nicht fallbezogen erhoben werden können (z.B. Angaben zum Desinfektionsmittelverbrauch)

[22] Im Rahmen der Beauftragung durch den G-BA kann eine Einschränkung der zu verwendenden Datenquellen erfolgen. Darüber hinaus kann eine Operationalisierung für die Datenquellen Routinedaten und Patientenbefragungen nur dann im Zuge des Entwicklungsprozesses umgesetzt werden, wenn eine empirische Prüfung mit Routinedaten bzw. die Entwicklung einer Patientenbefragung explizit Bestandteil der Beauftragung ist. Anderenfalls muss dieser Teil der Entwicklung nachgeholt werden. Ein solches Vorgehen wurde beispielsweise für die Entwicklung des Qualitätssicherungsverfahrens *Perkutane Koronarintervention (PCI) und Koronarangiographie* (AQUA 2014c; AQUA 2014d) gewählt.

- Angaben, die sich auf Patienten oder Behandlungsfälle beziehen:
  - auf Grundlage vorhandener Routinedaten der Einrichtungen (z.B. Abrechnungsdaten, Diagnosen)
  - auf Grundlage der medizinischen Dokumentation bzw. spezifischer Dokumentationsanforderungen der Qualitätssicherung

Für die Operationalisierung entwickelter Qualitätsindikatoren wird dabei versucht, so wenig zusätzlichen Dokumentationsaufwand zu verursachen, wie möglich. Die Erhebung notwendiger Datenfelder für Indikatoren orientiert sich daher an folgender hierarchischer Reihenfolge:

- Ist eine aggregierte, d.h. auf einen Zeitraum bezogene Information zur Berechnung des Qualitätsindikators ausreichend?
- Ist zu erwarten, dass ein notwendiges Datenfeld bereits im Rahmen der Abrechnung dokumentiert wurde?
- Ist zu erwarten, dass ein notwendiges Datenfeld, z.B. aufgrund von Leitlinienempfehlungen oder etablierten freiwilligen Qualitätsprojekten, bereits in der Dokumentation vorhanden ist?
- Sofern nicht zu erwarten ist, dass ein notwendiges Datenfeld bereits erhoben wird: Wie lässt sich dieses möglichst aufwandsarm erheben?

Auch die automatisierte Erhebung vorhandener Daten in Einrichtungen birgt noch ein erhebliches Potenzial zur Vereinfachung der Datenerhebungen. Dies zeigt das Beispiel des QS-Verfahrens *Pflege: Dekubitusprophylaxe*. Nach erfolgreichen Probeberechnungen anhand von stationären Abrechnungsdaten (AQUA 2011b) und der Umstellung der Datenerhebung in diesem Verfahren, konnte der manuelle Dokumentationsaufwand erheblich gesenkt werden (Heller et al. 2014). So werden hier mittlerweile fast alle erforderlichen und über die Abrechnungsdaten hinausgehenden Daten automatisiert aus dem Krankenhausinformationssystem übertragen.

7.1.1 Auslösung einer Dokumentation (QS-Auslösung)

Die sektorenübergreifende Qualitätssicherung zeichnet sich unter anderem dadurch aus, dass Behandlungsverläufe einrichtungs- und sektorenübergreifend abgebildet werden sollen. Die Verantwortung für die Dokumentation obliegt dabei zumeist der Einrichtung, in welcher eine Leistung (z.B. Operation, Nachsorge) erbracht wird. Aus diesem Grund sind für sektorenübergreifende Verfahren der Qualitätssicherung in der Regel mehrere Dokumentationsbögen notwendig.

Analog zum gegenwärtigen Vorgehen in der stationären Qualitätssicherung bilden Abrechnungsdaten zu medizinisch-pflegerischen Leistungen den Startpunkt für eine Dokumentation (sogenannte QS-Auslösung). Einschreibeprozeduren wurden bisher, aufgrund des damit verbundenen administrativen Aufwands, noch nicht als Ausgangspunkt zur Definition einer Dokumentationspflicht in der stationären Qualitätssicherung erwogen.

7.1.2 Vollzähligkeit und Vollständigkeit

Die Messung der Versorgungsqualität in einer sektorenübergreifenden Betrachtungsweise zeichnet sich vor allem durch die Beteiligung mehrerer Leistungserbringer am Versorgungsprozess und die Betrachtung unterschiedlicher Messzeitpunkte aus. Um dokumentationsrelevante Leistungen zu bestimmen, sind vor allem folgende Fragen zu klären:

- Welche Leistungserbringer waren am Versorgungsprozess beteiligt?
- Welche der erbrachten Leistungen stehen im Zusammenhang mit dem betrachteten Versorgungsproblem? Ergänzend: Handelt es sich um einen für die Berechnung des Qualitätsindikators relevanten Behandlungszeitraum?
- Welche relevanten Daten wurden noch nicht dokumentiert und müssen daher erfasst werden bzw. welche relevanten Daten wurden bereits dokumentiert und müssen daher nicht erneut erfasst werden?
- Wer muss ggf. die noch nicht dokumentierten relevanten Daten erheben?

Mit Bezug zu allgemeinen Kriterien der Dokumentationsqualität lassen sich diese Fragen unter die Aspekte Vollzähligkeit und Vollständigkeit der Dokumentation subsummieren (zu weiteren Kriterien der Dokumentationsqualität, siehe Abschnitt 9.2). Vollzähligkeit umfasst dabei die ersten beiden Fragenkomplexe (Welche Leis-

tungserbringer müssen welche Fälle dokumentieren?), Vollständigkeit die beiden letzten (Welche Inhalte müssen von wem dokumentiert werden?).

Die Prüfung der Vollzähligkeit hängt eng mit der QS-Auslösung zusammen: Für alle medizinisch-pflegerischen Leistungen, die zu einer Auslösung der QS-Dokumentation geführt haben, sollte eine Dokumentation vorliegen. Durch Abgleich aller laut Abrechnungsdaten erwarteten Dokumentationen (Soll) mit den tatsächlich vorhandenen Dokumentationen (Ist) kann die Vollzähligkeit bestimmt werden.

Die Prüfung der Vollständigkeit erfolgt über Plausibilitätskontrollen. Inhaltliche Kriterien hierfür sind im Indikatordatenblatt festzuhalten.

7.2 Nutzung von Routinedaten

Angesichts des vergleichsweise hohen Aufwands der QS-Dokumentation für die Leistungserbringer werden derzeit verstärkt Möglichkeiten zur Nutzung von Routinedaten geprüft.

Im deutschen Gesundheitssystem sind Routinedaten zumeist Daten, die zu Abrechnungszwecken erhoben werden. Im stationären Bereich sind dies u.a. Daten, die auf Grundlage des §301 SGB V an die Kostenträger übermittelt werden bzw. die sog. §21-Daten (KHEntgG). Im ambulanten Bereich zählen u.a. Abrechnungsdaten auf Grundlage des §295 SGB V bzw. §300 SGB V dazu. Daneben gibt es aber eine Reihe weiterer Daten, die in Praxen und Krankenhäusern bzw. von weiteren Leistungserbringern erfasst und im Rahmen der Rechnungsstellung weitergeleitet werden (Swart et al. 2014; Heller 2009; Swart et al. 2007).

7.2.1 Hintergrund

Ein Blick auf die internationale Literatur zeigt, dass Routinedaten seit Langem und intensiv für Zwecke der Qualitätssicherung und der Versorgungsforschung genutzt werden. Seit den 1990er Jahren finden sich zunehmend Arbeiten zu diesem Thema, wobei zumeist diagnose- bzw. prozedurenbezogene Indikatoren analysiert werden (Dr. Foster 2013; Bottle et al. 2009; Aylin et al. 2007; Pine et al. 2007; Jha et al. 2005; Williams et al. 2005; DesHarnais et al. 2000; Ghali et al. 2000; Black et al. 1998; Romano et al. 1995). In einigen Ländern werden darüber hinaus globale mortalitätsbasierte Qualitätsindikatoren für Einrichtungen auf der Basis von Routinedaten berechnet und ausgewiesen (Dr. Foster 2013; CIHI 2010; Jarman et al. 2010). Daneben sind verschiedene Qualitätsindikatorensets der Agency for Healthcare Research and Quality (AHRQ 2010), insbesondere deren Indikatoren zur Patientensicherheit, weit verbreitet. Dies ist allerdings dem Umstand geschuldet, dass die zugrunde liegenden Definitionen und Analyseprogramme frei verfügbar sind und ständig weiterentwickelt werden (AHRQ 2010; Mihrshahi et al. 2010; Bottle et al. 2009; Romano et al. 2009; Drösler et al. 2007).

Hinsichtlich der Eignung von Routinedaten für Zwecke der Qualitätssicherung werden an mancher Stelle grundlegende Bedenken angeführt, die nicht nur das Fehlen qualitätsrelevanter Details bemängeln, sondern die Validität dieser Daten bezweifeln. Van der Lei beschreibt dies folgendermaßen:

> *Data as such do not exist in nature; data are man-made artefacts. Real-world entities are observed and described by humans. When usage of the medical record changes, the recorded data change accordingly. Data are collected with a purpose in mind; that purpose has a direct influence on what data are recorded.* (1991: 70)

Abgesehen davon, dass diese Tatsache für alle Daten gilt, spiegelt dies die existierende Kontroverse über die Möglichkeiten einer sinnvollen Verwendung von (stationären) Routinedaten wider (Shahian et al. 2007; Black et al. 1998; Iezzoni 1997). Dabei konnte international allerdings für einige Qualitätsanalysen gezeigt werden, dass eine alleinige Verwendung von Routinedaten zu sehr ähnlichen Ergebnissen führt, wie auch Analysen auf Basis von klinischen Daten oder Registern (Aylin et al. 2007; Pine et al. 2007; Ghali et al. 2000).

Einige Untersuchungen in Deutschland bestätigen diese Kritik. Beispielsweise konnte gezeigt werden, dass ambulante Abrechnungsdiagnosen bislang nur eingeschränkt für andere Zwecke verwendbar sind (Erler 2007; Giersiepen et al. 2007; Gerste et al. 2006). Dem steht allerdings eine Vielzahl von Projekten gegenüber, in denen Qualitätsindikatoren auf der Basis von Routinedaten erfolgreich verwendet werden. Auch in einem wis-

senschaftlichen Gutachten wurden Methoden dargelegt, wie sich ambulante Abrechnungsdaten beispielsweise im Rahmen des Morbi-RSA nutzen lassen (Busse et al. 2007). Ein Beispiel ist das Projekt Qualitätssicherung mit Routinedaten (QSR) der AOK (WIdO et al. 2007). Es zeigt, dass gerade Indikatoren zur Ergebnisqualität (z.B. Sterblichkeit, langfristige Revisionsraten) gut auf dieser Basis abgebildet werden können, insbesondere deswegen, weil Betrachtungen über einen längeren Zeitraum und über verschiedene Leistungserbringer hinweg möglich sind. Dies macht Routinedaten sowohl für die sektorenübergreifende Qualitätssicherung als auch für sektorgleiche Verfahren (z.B. 30-Tage-Sterblichkeit nach Koronarintervention) interessant, weil sie eine Nachbeobachtung ermöglichen, die in der bisherigen externen stationären Qualitätssicherung nicht realisiert werden konnte, da die Datengrundlage dafür fehlte.

Grund dafür sind weniger die o. g. Routinedaten selbst, sondern vielmehr die Möglichkeit der Kostenträger (Kassen), unterschiedliche Behandlungen/Leistungen einer individuellen Person zuzuordnen und so Krankheitsverläufe abzubilden. Formal geschieht dies über eine Verknüpfung der Routinedaten mit Versichertenbestandsdaten (Stammdaten). Dabei sind die Fragen, ob ein Versicherter zu einem gegebenen Zeitpunkt bei einer Kasse versichert ist (Versichertenstatus) und sein Überlebensstatus (einer der möglichen Gründe, warum die Person nicht mehr versichert ist) der obligate und erste Teil jeder Rechnungsprüfung. Beide Informationen liegen daher bei den Kassen gut gepflegt und nahezu vollständig vor. Der Informationsverlust durch Kassenwechsler ist dabei für die meisten Qualitätsanalysen vernachlässigbar, insbesondere dann, wenn es sich um ältere, chronisch kranke Menschen handelt, die erwartungsgemäß eher seltener die Krankenkasse wechseln. So konnte am Beispiel der stationären Versorgung bei Myokardinfarkt anhand von AOK-Daten der Jahre 2004/2005 gezeigt werden, dass weniger als 0,2 % der Patienten 30 Tage nach Aufnahme weder (sicher) verstorben noch bei der AOK versichert waren (Heller et al. 2008), während Bitzer et al. (2008) – auf der Basis von Daten der Gmünder Ersatzkasse – 1-Jahres-Überlebensraten nach koronarer Revaskularisation mit weniger als 0,4 % „Lost to Follow-up“ berichten. Auch wenn man nicht bei allen Kostenträgern eine gleich gute Datenlage unterstellen kann und auch wenn es durch verstärkten Wettbewerb zu mehr Kassenwechslern kommt, so werden Routinedaten weiterhin als ein bisher zu wenig genutztes Potenzial in der Qualitätssicherung angesehen.

Abgesehen vom Überlebensstatus können weitere medizinische Behandlungsanlässe zur Einschätzung der Ergebnisqualität genutzt werden. Prominente Beispiele sind stationäre Wiederaufnahmen wegen typischer Komplikationen oder auch ambulante Behandlungen nach einem Indexereignis (z.B. einem chirurgischen Eingriff), die auf mögliche Komplikationen hinweisen.

Das AQUA-Institut verwendet erfolgreich Routinedaten aus dem ambulanten Bereich im Projektrahmen der *Qualitätszirkel zur Pharmakotherapie* und verschiedener anderer Evaluationen (Joos et al. 2008; Wensing et al. 2004). Auch die Indikatoren des *Qualitätsindikatorensystems für die ambulante Versorgung* (QiSA) (Szecsenyi et al. 2009) können zu etwa 30 % durch Routinedaten abgebildet werden. Allerdings schwankt dieser Wert je nach betrachtetem Thema: Im Bereich Pharmakotherapie sind es etwa 90 % der Indikatoren, im Bereich Prävention dagegen lässt sich keiner der Indikatoren aus Routinedaten ableiten.

Auch bei eingeschränkter Eignung (wobei immer der Kontext und das Ziel zu berücksichtigen sind) sprechen somit einige Argumente für die Nutzung von Routinedaten im Rahmen der Qualitätssicherung.
Zu nennen wären

- der geringe Dokumentationsaufwand,
- die Vollzähligkeit der Datenbasis,
- der extern überprüfte Datenbestand und
- die vielfach erprobte Anwendung von Routinedaten in großen Indikatorenregistern (z.B. AHRQ).

Wesentliche Vorteile der Nutzung von Routinedaten sind demnach, dass beim Leistungserbringer kein zusätzlicher Dokumentationsaufwand entsteht und dass im Fall der Daten nach §299 SGB V Längsschnittbetrachtungen auch über mehrere Leistungserbringer hinweg möglich sind. Da die Abrechnungsdaten jedoch nicht originär für die Zwecke der QS-Dokumentation entwickelt wurden, resultieren bei ihrer Nutzung ggf. spezifische Probleme:

- Vom stationären Leistungserbringer können nur die Leistungen oder Diagnosen kodiert werden, die einen Aufwand in der Behandlung begründen. Zudem besteht für den Leistungserbringer keine Pflicht, Kodes zu Diagnosen oder Prozeduren, die nicht erlösrelevant sind, zu dokumentieren. Aus diesen Gründen findet keine vollständige Erfassung aller Kodes statt.
- Im vertragsärztlichen Bereich gibt es keine definierten Minimalanforderungen für die Kodierung, sodass Abrechnungsdaten hier nur unter Berücksichtigung der Kodierpraxis für die Qualitätssicherung verwendet werden können. Prozeduren werden darüber hinaus außerhalb des ambulanten und belegärztlichen Operierens sowie des fachärztlichen Bereichs in der Regel nicht erfasst, da deren Abrechnung über Ziffern des Einheitlichen Bewertungsmaßstabs (EBM) erfolgt.
- Auch bezüglich der Kodierung von Medikamenten gibt es deutliche Unterschiede: Im stationären Bereich werden Medikamente in der Regel nicht zusätzlich kodiert, da diese in den meisten Fällen im Rahmen der Fallpauschalen abgerechnet werden und in der Kalkulationsmatrix des InEK eingerechnet sind.
- Weiterhin haben kodierte Diagnosen im niedergelassenen und im stationären Bereich keinen Datumsbezug. Im stationären Bereich können sie daher nur einem Krankenhausaufenthalt und im vertragsärztlichen Bereich nur quartalsweise zugeordnet werden. Dies macht es in vielen Fällen unmöglich zu bestimmen, ob es sich bei einer Diagnose um eine Komplikation einer bestimmten Prozedur oder eine Vorerkrankung handelt.
- Es gibt grundsätzliche Kodierunschärfen, da die Strukturen von OPS und ICD historisch gewachsen sind und daher keine durchgehende Systematik aufweisen.
- Bei der Nutzung von Routinedaten im Rahmen des SGB V sind nur Daten von GKV-Patienten einbezogen, d.h. ca. 10 bis 15 % der Patienten bleiben bei Nutzung dieser Datenquelle unberücksichtigt.
- Routinedaten haben außerdem den allgemeinen Nachteil, dass sie für die Zwecke der externen Qualitätssicherung erst nach einer erheblichen Verzögerung (Lieferzeiten von bis zu zehn Monaten und länger) zur Verfügung stehen: Die Daten können quartalsweise, jedoch mit einer Lieferzeit von zehn Monaten und mehr zur Verfügung gestellt werden, die InEK-Daten für den Zeitraum eines Erfassungsjahres vom 1. Januar bis zum 31. Dezember und mit einem halben Jahr Verzug. Dies wirkt sich sowohl auf die Identifizierung der einzuschließenden Fälle als auch auf die Erstellung von Rückmeldeberichten für die Leistungserbringer aus.

Bei der Nutzung von Routinedaten in der Qualitätssicherung gilt es, diese Einschränkungen sachgerecht zu berücksichtigen, indem bezogen auf konkrete Indikatoren hinterfragt wird, inwiefern diese Einschränkungen relevant sind und ob sie eine ausreichend valide Abbildung der Qualität und einen fairen Einrichtungsvergleich verhindern.

7.2.2 Routinedatennutzung in der einrichtungsübergreifenden Qualitätssicherung

Im Kontext der einrichtungsübergreifenden Qualitätssicherung und in Abgrenzung zu anderen Wortverwendungen wird der Begriff Routinedaten nachfolgend verwendet für qualitätssicherungsrelevante Daten, die über die Krankenkassen oder über das InEK erhoben werden können. Derzeit kann das AQUA-Institut (auf Antrag und nach Prüfung) die stationären Abrechnungsdaten nach §21 Abs. 3a KHEntgG, die Daten gemäß §299 Abs. 1a SGB V (Sozialdaten bei den Krankenkassen) und die sogenannten Morbi-RSA-Daten (§303 a–e SGB V) für Zwecke der Qualitätssicherung nutzen (Heller et al. 2014). Allerdings enthalten die letztgenannten Daten bisher keinen Leistungserbringerbezug und sind daher für Zwecke der Qualitätssicherung nur bedingt hilfreich.

Daten nach §21 KHEntgG

Die Nutzung von Daten nach §21 KHEntgG durch die Institution nach §137a SGB V ist seit April 2007 in Abs. 3a des §21 KHEntgG wie folgt geregelt:

> *Die nach §137a Abs. 1 des Fünften Buches Sozialgesetzbuch mit der Durchführung von Maßnahmen der einrichtungsübergreifenden Qualitätssicherung beauftragte Institution auf Bundesebene kann ausgewählte Leistungsdaten aus den Buchstaben a bis f des Absatzes 2 Nr. 2 anfordern, soweit diese nach Art und Umfang notwendig und geeignet sind, um Maßnahmen der Qualitätssicherung nach §137a Abs. 2 Nr. 1, 2 und 3 des Fünften Buches Sozialgesetzbuch durchführen zu können. Die Institution auf Bundesebene kann entsprechende Daten auch für Zwecke der einrichtungsübergreifenden Qualitätssicherung auf Landesebene anfor-*

dern und diese an die jeweils zuständige Institution auf Landesebene weitergeben. Die DRG-Datenstelle übermittelt die Daten, soweit die Notwendigkeit nach Satz 1 von der Institution auf Bundesebene glaubhaft dargelegt wurde. Absatz 3 Satz 9 und 10 gilt entsprechend.

Genereller Vorteil der Daten nach §21 KHEntgG ist, dass sie alle Krankenkassen, d.h. auch Leistungen an Privatversicherten umfassen. Demgegenüber steht der Nachteil, dass nur Daten der Krankenhäuser enthalten sind. Für sektorenübergreifende Analysen ist dieser Datenpool deshalb nicht geeignet. Vielmehr ist er im Wesentlichen hilfreich für die Weiterentwicklung der externen stationären Qualitätssicherung.

Bisher (Stand Oktober 2014) hat das AQUA-Institut als Institution nach §137a SGB V in acht Fällen Daten gemäß Abs. 3a des §21 KHEntgG beim InEK beantragt. Sofern eine Umsetzung der geplanten Projekte möglich war, haben sich die angeforderten Daten als überaus hilfreich für die Weiterentwicklung der Qualitätssicherung erwiesen. Beispielsweise konnte im Bereich Dekubitus die Zahl der für die Qualitätssicherung händisch zu dokumentierenden Dokumentationsbögen um etwa 900.000 reduziert werden bei gleichzeitiger Verbesserung der Aussagekraft der Qualitätsindikatoren durch Einbeziehung des gesamten Jahres anstatt nur des ersten Quartals und durch Einbeziehung einer größeren Patientengruppe (ab 20 Jahren anstatt zuvor ab 75 Jahren). Erste Ergebnisse aus dieser Weiterentwicklung zeigen, dass nunmehr bestehende Qualitätsprobleme wesentlich besser aufgedeckt werden können (Heller et al. 2014; AQUA 2013c; AQUA 2011b).

Für jedes Projekt, in dem Daten nach §21 KHEntgG genutzt werden sollen, ist ein Antragsverfahren zu starten. Je nach Umfang der angeforderten Daten dauert die Prüfung von drei Monaten bis zu über einem Jahr. Eine wesentliche Herausforderung zur Nutzung dieser Daten besteht daher darin, das Antragsverfahren möglichst früh zu beginnen, damit die geplanten Analysen und Operationalisierungen vor Beginn des RAM-Panels oder bereits für vorangehende Entwicklungsschritte vorliegen.

Daten gemäß §299 Abs. 1a SGB V

Mit der Neuformulierung bzw. dem Inkrafttreten des §299 SGB V (2012) im Rahmen des Versorgungsstrukturgesetzes (VStG) sind neben den Leistungserbringern erstmals auch Krankenkassen gesetzlich befugt und verpflichtet, erhobene und gespeicherte Versichertenstamm- und Abrechnungsdaten bei den Krankenkassen für Zwecke der Qualitätssicherung zu nutzen, soweit diese Rückschlüsse auf die Qualität der Leistungserbringung ermöglichen.

Grundsätzlich sind über den §299 Abs. 1a SGB V alle Datenbestände nutzbar, für die bundesweit einheitliche Übermittlungsformate vorliegen und die im Kontext des SGB V verwendet werden. Im Zuge der Erstellung einer allgemeinen Spezifikation für Sozialdaten bei den Krankenkassen wurden die Datenbestände konkretisiert (siehe Tabelle 7), Exportformate definiert und allgemeine Übermittlungsvorschriften festgelegt (AQUA 2014b; AQUA 2013a).

Tabelle 7: Für die QS nutzbare Datenbestände

Datenbestand nach Abrechnungskontext (jeweils SGB V)	Beschreibung	Verfügbare Informationen (Auszug)
§301	Stationäre Behandlung im Krankenhaus	Aufnahmedatum, Entlassungsdatum, Diagnosen (gem. ICD-10-GM), Prozeduren (gem. Operationen- und Prozedurenschlüssel)
§115b	Ambulantes Operieren im Krankenhaus	Diagnosen (gem. ICD-10-GM), Prozeduren (gem. Operationen- und Prozedurenschlüssel)
§116b	Ambulante Spezialfachärztliche Versorgung	Diagnosen (gem. ICD-10-GM), Prozeduren (gem. Operationen- und Prozedurenschlüssel)
§117	Hochschulambulanz	Diagnosen (gem. ICD-10-GM), Prozeduren (gem. Operationen- und Prozedurenschlüssel)

© 2015 AQUA-Institut GmbH

Datenbestand nach Abrechnungskontext (jeweils SGB V)	Beschreibung	Verfügbare Informationen (Auszug)
§118	Psychiatrische Institutsambulanz	Diagnosen (gem. ICD-10-GM), Prozeduren (gem. Operationen- und Prozedurenschlüssel)
§119	Sozialpädiatrisches Zentrum	Diagnosen (gem. ICD-10-GM), Prozeduren (gem. Operationen- und Prozedurenschlüssel)
§295 (kollektiv)	Vertragsärzte mit kollektivvertraglicher Abrechnung	Erstes Behandlungsdatum im Quartal, letztes Behandlungsdatum im Quartal, Gebührenordnungspositionen (gem. EBM, inkl. Datum), Diagnosen (gem. ICD-10-GM), Prozeduren (gem. Operationen- und Prozedurenschlüssel)
§295a (selektiv)	Vertragsärzte mit selektivvertraglicher Abrechnung (gemäß §§73b, 73c und 140a SGB V)	Gebührennummer (inkl. Datum), Diagnosen (gem. ICD-10-GM, inkl. Diagnosesicherheit, Lokalisation und Datum), Prozeduren (gem. Operationen- und Prozedurenschlüssel, inkl. Lokalisation und Datum)
§300	Arzneimittel/-verordnungen	Art des Leistungserbringers, Pharmazentralnummer, Verordnungsdatum des Arzneimittels
§302	Heilmittelverordnung nach §302	Art des Leistungserbringers, Art der abgegebenen Leistungen (Positionsnummer), Indikationsschlüssel, Verordnungsdatum des Heilmittels
§284	Versichertenstammdaten nach §284	Geschlecht, Geburtsdatum (Alter), Sterbedatum

Bezogen auf ein konkretes Thema ist es erforderlich, aus den in Tabelle 7 dargestellten Datenbeständen, die für Zwecke der Qualitätssicherung erforderlichen Daten auszuwählen und genau zu beschreiben. Hierin besteht die Aufgabe der Entwicklungsprojekte. Sie sollen eine Basis dafür liefern, dass der G-BA eine Richtline über die zu erhebenden Daten beschließen kann, und dafür, dass eine standardisierte Spezifikation der von den Krankenkassen zu erhebenden Daten erstellt werden kann (siehe Abschnitt 10.1.2).

Gegenwärtig bietet §299 SGB V allerdings keine Rechtsgrundlage, um Daten für die Entwicklung und Operationalisierung von Qualitätsindikatoren anzufordern. Vielmehr werden zu diesem Zweck freiwillige Forschungskooperationen zwischen unterschiedlichen Krankenkassen und dem AQUA-Institut etabliert. Bis dato (Stand: Oktober 2014) wurden über zehn Forschungskooperationen mit Krankenkassen im Kontext von Beauftragungen durch den G-BA etabliert. Auch hier besteht eine wesentliche Herausforderung darin, die Forschungskooperation so frühzeitig zu beginnen, dass geplante Analysen rechtzeitig für das RAM-Panel und vorangehende Entwicklungsschritte zur Verfügung stehen können.

7.2.3 Methodische Aspekte

Die Nutzung von Routinedaten im Zuge von Entwicklungsprozessen gliedert sich idealerweise in vier Phasen:

- Datenbeschaffung
- Datenaufbereitung
- Operationalisierung der Zielpopulation und Exploration
- Operationalisierung von Indikatoren

Die nachfolgenden Ausführungen fokussieren auf die Vorgehensweise zur Entwicklung fachlicher Grundlagen für eine Richtline des G-BA, um Daten gemäß §299 Abs. 1a SGB V zu übermitteln. Wie oben beschrieben ist zur Nutzung von Routinedaten im Zuge der Entwicklung zunächst die Etablierung einer Forschungskooperation mit einer Krankenkasse notwendig.

Datenbeschaffung

Im Zuge der Abstimmung mit einzelnen Krankenkassen sind zunächst unterschiedliche Datenstrukturen, Datenqualitäten und Zugangswege, sowie spezifische vertragliche und datenschutzrechtliche Anforderungen zu berücksichtigen (Swart et al. 2014: passim). Insbesondere folgende Schritte sind durchzuführen:

- Auswahl einer geeigneten Krankenkasse, die in Bezug auf das Thema die entsprechenden Daten vorhält (z.B. bei Selektivverträgen)
- Abstimmung und Erstellung eines Kooperationsvertrags
- Erstellen einer Anforderung an die anonymisierten Routinedaten inklusive einer umfassenden Feldspezifikation für die zu liefernden Daten; hierfür im Vorfeld Festlegung geeigneter Kriterien für die erforderlichen Fälle (Zielpopulation, Zeiträume, Diagnosen, Prozeduren, ggf. weitergehende Kriterien)
- Konsentieren der Feldspezifikation mit der Kasse sowie Verhandlung über Art und Umfang der Stichprobe
- im Vorfeld ggf. Klärung von Datenschutzaspekten

Datenaufbereitung

Nach erfolgter Datenlieferung werden die Daten aufbereitet bzw. qualitativ überprüft (Horenkamp-Sonntag et al. 2014). Dies umfasst folgende Schritte:

- Einlesen der Daten
- Datenvalidierung und weitere Aufbereitung (sind die ICD/OPS/EBM-Kodes korrekt und in einem verarbeitbaren Format?, stimmen die Datumangaben?, sind die einzelnen Teildatensätze verknüpfbar?, sind die Daten einheitlich pseudonymisiert?, stimmen die Behandlungsdaten? usw.
- Plausibilitätsprüfungen
- Eventuelle Korrektur-/Nachlieferung von Daten beim gleichen Datenlieferanten
- Eventuelle zusätzliche Beschaffung bei weiteren Lieferanten

Operationalisierung der Zielpopulation und Exploration[23]

In diesem Schritt werden folgende unterschiedliche Fragestellungen beantwortet:

- Wie kann die Zielpopulation des zu entwickelnden Verfahrens anhand von Routinedaten möglichst exakt beschrieben werden?
- Welche allgemeinen Eigenschaften und ggf. Einschränkungen weisen die verwendeten Routinedaten in Bezug auf das Thema auf?
- Können zu einem identifizierten Qualitätspotenzial Aussagen aus den Routinedaten abgeleitet werden?
- Können aus den Routinedaten weitere Qualitätspotenziale abgeleitet werden?

Zur Umsetzung dieser Schritte ist eine enge Zusammenarbeit zwischen Analysten, medizinischen Experten und Kodierexperten notwendig. Neben allgemeinen Explorationen (z.B. Betrachtung häufig abgerechneter OPS- und ICD-Kodierungen in Bezug auf das Thema) werden hierbei in einem iterativen Prozess Hypothesen generiert und überprüft. Diese wird in einem fortlaufend aktualisierten Auswertungsplan festgehalten.

Operationalisierung der Erhebungsinstrumente

Hierbei kann es unterschiedliche Ausgangspunkte geben, je nachdem, ob es sich um bereits vorhandene/recherchierte Indikatoren oder um neu entwickelte Indikatoren handelt (siehe Abschnitt 6.2). In beiden Fällen ist das Ziel, alle notwendigen Informationen für die zu erstellenden Indikatordatenblätter zu generieren.

Auch zur Umsetzung dieser Schritte sind eine enge Zusammenarbeit zwischen Analysten, medizinischen Experten und Kodierexperten und ein fortlaufend zu aktualisierender Auswertungsplan notwendig.

[23] Konzeptionell ist dieser Schritt Teil der Themenerschließung. Aus systematischen Gründen ist er jedoch an dieser Stelle dargestellt.

7.3 Datenerhebungen bei/über Patienten

Im Rahmen der einrichtungsübergreifenden Qualitätssicherung gem. §137 SGB V sollen verfahrensspezifische Patientenbefragungen als neues Erhebungsinstrument entwickelt werden, weil mit den derzeit zur Verfügung stehenden Erhebungsinstrumenten relevante Qualitätsaspekte aus Sicht der Patienten nicht ausreichend berücksichtigt werden. Dies gilt insbesondere für Aspekte der mittel- und langfristigen Ergebnisqualität, aber auch für Aspekte der Struktur- und Prozessqualität, inklusive der Indikationsqualität.

7.3.1 Ziele und Anforderungen von Patientenbefragungen

Ziel der Patientenbefragungen im Rahmen der Qualitätssicherung ist es, die Präferenzen, Sichtweisen und Erfahrungen sowie auch den Gesundheitszustand der Patienten (Patient-Reported Outcomes)[24] zu erfassen. Patienten können auf mögliche Qualitätsdefizite hinweisen, die sich einer Messung bzw. Darstellung durch Indikatoren, die auf Routinedaten oder QS-Dokumentation basieren, entziehen. Bei der Orientierung an Qualitätspotenzialen (Qualitätsziele und -defizite) (vgl. Abschnitt 4.2.7) ist der Fokus auf die Patientenperspektive insbesondere dann zu legen, wenn patientenrelevante Endpunkte nicht unbedingt kongruent zu einem medizinisch gesetzten Behandlungserfolg sind (z.B. Gelenkbeweglichkeit vs. Schmerzfreiheit) (Mühlhauser et al. 2009; Badura 1999).

Die Entwicklung des Erhebungsinstruments Patientenbefragung ist mit besonderen Herausforderungen verbunden, da sich die zugrunde liegende Datenquelle substanziell von anderen Datenquellen unterscheidet. Dies gilt insbesondere für die inhaltliche Entwicklung des Fragebogens. Entscheidend hierfür ist zunächst, dass das Instrument inhaltlich valide ist, d.h. dass es geeignet ist, die verfahrensspezifischen Ziele bzw. die Qualitätspotenziale abzubilden (Pfaff et al. 2009; Terwee et al. 2007; Rammstedt 2004). Um möglichst valide Antworten von Patienten zu erhalten, sind besondere Anforderungen an die Entwicklung des Erhebungsinstruments zu stellen. Um beispielsweise die Subjektivität von Patientenbewertungen möglichst gering zu halten, empfiehlt sich zur Darstellung der Prozess- bzw. Strukturqualität über Items der Einsatz von Reporting-Fragen (ereignis-/faktenorientierter Ansatz). Die Anwendung dieser Methodik hat sich bereits seit den 1990er Jahren etabliert, da sie im Gegensatz zu einer reinen Zufriedenheitsanalyse mithilfe von Rating-Fragen geeigneter erscheint, Qualitätspotenziale aufzuzeigen (Noest et al. 2014; Cleary et al. 1993). Darüber hinaus ist die Verständlichkeit der Fragen für den Patienten zentral, weshalb Patienten in die Entwicklung des Instruments einbezogen werden (Terwee et al. 2007), z.B. durch Durchführung von Fokusgruppen in der Frühphase der Entwicklung sowie auch in späteren Validierungsschritten im Rahmen von kognitiven Pretests. Zur fachlichen Validierung der Inhalte, aber auch zu konkreten Umsetzungsfragen (z.B. zum richtigen Befragungszeitpunkt), werden sowohl Fachexperten als auch mindestens ein Patientenvertreter hinzugezogen. Darüber hinaus werden konkrete methodische Schritte in Abstimmung mit dem G-BA vorgenommen.

Die Entwicklung des Instruments der Patientenbefragung ist zudem vor dem Hintergrund verfahrensspezifischer Anforderungen im Rahmen der gesetzlichen Qualitätssicherung zu sehen. Mit der Orientierung an den Zielsetzungen eines konkreten QS-Verfahrens sowie der Einschätzung der Abbildbarkeit der Qualitätspotenziale mit verfügbaren bzw. einer im Auftrag bereits vorgegebenen Nutzung von spezifischen Datenquellen soll ein Patientenfragebogen entwickelt werden, der dezidiert die versorgungsspezifischen Aspekte fokussiert

7.3.2 Entwicklung der Befragungsinstrumente

Die Methodik zur Entwicklung der Patientenbefragungsinstrumente (Fragebögen) und zur Durchführung der Befragungen orientiert sich unter anderem an folgenden Zielen:

- Die Befragungen sollen sich auf verfahrensspezifische Qualitätspotenziale (Qualitätsziele und -defizite) beziehen.

[24] Es existiert keine allgemeingültige Definition des Begriffs „Patient-Reported Outcomes“. Der Begriff wird je nach Zielrichtung der Anwendung entweder nur auf die Ergebnisqualität einer Behandlung oder auf die Prozess- *und* die Ergebnisqualität (aus Patientensicht) bezogen; immer jedoch geht es um „von Patienten selbst berichtete Erfahrungen und Angaben über ihren Gesundheitszustand“, nicht um eine Darstellung aus der Perspektive Anderer. Das AQUA-Institut verwendet den Begriff so, wie er auch in der Aufgabenbeschreibung des G-BA verwendet wird: Für alle von Patienten berichteten Behandlungserfahrungen (also bezogen auf Behandlungsprozesse und -ergebnisse), in Abgrenzung zu den auf anderem Wege erhobenen Daten wie z.B. Dokumentationen der Einrichtungen und Routinedaten.

- Bei der Entwicklung der Fragebögen werden die Erfahrungen und die Bedürfnisse der Patienten selbst einbezogen (z.B. im Rahmen der Durchführung von Fokusgruppen und kognitiven Pretests).
- Die Befragungsinstrumente sind im Prinzip sektorenübergreifend angelegt. Es wird verfahrensspezifisch entschieden, in welchen Bereichen ggf. Schwerpunkte gesetzt werden.
- Die entwickelten Befragungsinstrumente sollen für den routinemäßigen Einsatz geeignet sein und einen Einrichtungsbezug und -vergleich ermöglichen.
- Die Befragung soll zielgerichtet und aufwandsarm sein (insbesondere im Hinblick auf die Fragebogenlänge und die Logistik der Befragung).
- Die Einrichtungen sollen Einfluss auf die erfragten Qualitätsmerkmale und damit die Möglichkeit der Umsetzung von Qualitätsverbesserungen haben.
- Die Rücklaufquote sollte möglichst hoch sein (dies impliziert auch mindestens eine Nachfassaktion).
- Die Ergebnisse sollten möglichst repräsentativ für die betroffene Patientengruppe sein (geringstmöglicher Bias bei der Erhebung).
- Die Diskriminationsfähigkeit des Instruments sollte möglichst hoch sein.
- Die Ergebnisdarstellung (u.a. in den Rückmeldeberichten an die Einrichtungen) sollte eindeutig und verständlich sein.

Nachfolgend werden wesentliche methodische Schritte der Itemgenerierung in der zeitlichen Abfolge beschrieben. Aus jedem Schritt resultiert eine überarbeitete Version eines Fragenbogenentwurfs, der kontinuierlich weiterentwickelt wird.

7.3.2.1 Itemgenerierung

Der Prozess der Itemgenerierung, d.h. der Auswahl von Fragen und Antwortvorgaben aus anderen Instrumenten, oder der Neuentwicklung von Items sowie der Überprüfung der Verständlichkeit von Items, findet schrittweise vor dem Hintergrund der identifizierten Qualitätspotenziale statt:

- Recherche nach Befragungsinstrumenten, Operationalisierung
- Durchführung von Fokusgruppen
- Einbindung von Fachexperten und Patientenvertretern
- Kognitive Pretests
- ggf. Durchführung eines Panelprozesses

Falls das AQUA-Institut es als erforderlich erachtet, kann weitere themenspezifische Expertise eingebunden werden.

Die aus diesem Prozess generierten Items lassen sich hinsichtlich ihrer Erhebungsziele unterscheiden: Items zur Erfassung der Ergebnisqualität, der Prozessqualität, von Risikofaktoren und ggf. zur Identifizierung der Zielpopulation (evtl. Subpopulationen).

Die recherchierten Items zur Erfassung der Ergebnis- und Prozessqualität können nach inhaltlichen Kriterien zu Fragenblöcken zusammengefasst werden. Aus den Items/Fragen bzw. Fragenblöcken werden Indikatoren gebildet. Dabei kann ein Indikator aus einem/r einzelnen Item/Frage aber auch aus der Zusammenfassung mehrerer Items/Fragen oder auch Fragenblöcken gebildet werden. Im Allgemeinen ist davon auszugehen, dass durch die Abbildung eines Indikators durch mehrere Items, die Validität und Reliabilität erhöht werden kann (Schnell et al. 2013: 143ff). Ergänzend muss definiert werden, welche Antwortvorgabe(n) als Ausdruck eines Qualitätsdefizits bzw. welche als Ausdruck guter Versorgungsqualität betrachtet werden sollen. Werden Items aus bereits vorhandenen validierten Befragungsinstrumenten genutzt, sind die dort verwendeten Skalen zu berücksichtigen (Bortz et al. 2005). Um die Befragungsergebnisse auf Indikatorebene auswerten zu können, müssen konkrete Rechenregeln definiert werden, die auf den Indikatordatenblättern ausgewiesen werden (Definition von Zähler und Nenner). Des Weiteren ist zu bestimmen, wann ein Indikator eine rechnerische Auffälligkeit hervorbringt, also das gemessene Ergebnis außerhalb eines vorher festgelegten Referenzbereichs liegt. Wenn aus der Literatur kein ableitbarer Referenzbereich vorliegt, wird das 95. Perzentil als Referenzbereich

empfohlen. Dieser Wert wird auf Grundlage der ersten Befragungen im Regelbetrieb überprüft und ggf. modifiziert.

Da der Behandlungserfolg auch von patientenindividuellen Gegebenheiten (Alter, Geschlecht, Komorbiditäten usw.) abhängt (Schneeweiss et al. 2001), werden Items zur Erfassung von Risikofaktoren generiert. Eine Risikoadjustierung, insbesondere der Ergebnisindikatoren, ist die Voraussetzung für einen fairen Leistungserbringervergleich und damit letztlich die Grundlage für die Akzeptanz eines möglichen QS-Verfahrens. Das methodische Vorgehen zur Risikoadjustierung von Qualitätsindikatoren ist in Abschnitt 8.1.5 detailliert dargestellt.

Es kann darüber hinaus erforderlich sein, Items zur Identifizierung der Zielpopulation bzw. von Subpopulationen in den Fragebogen aufzunehmen, um die Möglichkeit zu haben, Indikatoren nach Subpopulationen differenziert betrachten zu können bzw. indikatorspezifische Ausschlusskriterien zu definieren.

7.3.2.2 Recherche nach Befragungsinstrumenten

Die Berücksichtigung der Patientenperspektive ist im Rahmen der Neuentwicklung eines QS-Verfahrens ein integraler Bestandteil der systematischen Literaturrecherche (vgl. Abschnitt 5.2), sodass diese neben der Suche nach Qualitätspotenzialen auch genutzt werden kann, geeignete und bereits verwendete verfahrensspezifische Befragungsinstrumente zur Ergebnis- und/oder Prozessqualität zu recherchieren. Zur Entwicklung einer Patientenbefragung bzw. zur Itemgenerierung sind zudem zusätzliche Recherchen nach bereits vorliegenden themenspezifischen Befragungsinstrumenten in speziellen Datenbanken notwendig (Bortz et al. 2005). Dabei sollten vorhandene Befragungsinstrumente bzw. einzelne Items nicht unkritisch hinsichtlich ihrer Anwendung und insbesondere ihrer Güteeigenschaften (Objektivität, Reliabilität und Validität) übernommen werden. In vielen Fällen ist zudem eine sprachliche Anpassung notwendig (Bortz et al. 2005).

Die Suche nach Befragungsinstrumenten erfolgt grundsätzlich über die:

- Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID; www.proqolid.org)
- Patient Reported Outcome Measurement Group (PROMs; phi.uhce.ox.ac.uk)

Die Suche in diesen speziellen Datenbanken wird durch eine auf themenspezifische Befragungsinstrumente und Validierungsstudien fokussierte Suche in Embase und der Cochrane-Library (Cochrane Collaboration) ergänzt.

Für viele Erkrankungen und Behandlungsmethoden liegen bereits Instrumente (oder Patientenfragebögen) zur Erfassung der Ergebnisqualität vor, selten jedoch zur Erfassung der Prozessqualität. Kann kein validiertes, verfahrensspezifisches Instrument zur Erfassung der Prozessqualität identifiziert werden, wird auf den vom AQUA-Institut entwickelten und validierten generischen Patientenfragebogen zur Erfassung der sektorenübergreifenden Prozessqualität zurückgegriffen (Noest et al. 2014; AQUA 2012d). Es werden für das jeweilige Verfahren zur Erfassung der Qualitätspotenziale geeignete Fragen ausgewählt, ggf. verfahrensspezifisch angepasst und im Rahmen der Verfahrensentwicklung validiert.

Sollte für ein Thema bereits ein validiertes, verfahrensspezifisches Instrument zur Bewertung der Prozessqualität vorliegen, so wird geprüft, ob dieses dazu geeignet ist, die für das QS-Verfahren identifizierten Qualitätspotenziale abzubilden. Ist dies nicht der Fall, wird ebenfalls auf den generischen Patientenfragebogen zurückgegriffen.

Ziel dieser Recherchen ist es, die mit einer vollständigen Neuentwicklung von Items für eine Patientenbefragung verbundenen Risiken zu minimieren. So kann die Adaption eines bereits entwickelten und in Anwendung befindlichen verfahrensspezifischen Instruments das Risiko der Aufnahme von nicht validen Items verringern. Darüber hinaus ist anzunehmen, dass die allgemeine Akzeptanz für validierte und in Anwendung befindliche Instrumente im Entwicklungsprozess höher ist. Es ist jedoch zu beachten, dass Änderungen bereits bestehender Befragungsinstrumente (entfernen oder ergänzen von Items, Umformulierungen) je nach Umfang Auswir-

kungen auf die Validität und Reliabilität haben können (Ruprecht 2001). Ein Nachteil der Nutzung von bereits entwickelten und validierten Instrumenten besteht in ihrer geringen Flexibilität.

Verfügbare Befragungsinstrumente werden nach folgenden Kriterien geprüft und bewertet:

- Orientierung an identifizierten Qualitätspotenzialen
- Anzahl der Items
- Inhaltliche Plausibilität der Items
- Vorliegen von Entwicklungs-, Validierungs- und Anwendungsstudien mit Informationen zu psychometrischen Eigenschaften des Instruments
- Vorliegen einer validierten deutschsprachigen Übersetzung
- Verfügbarkeit (Lizenzen)

Wenn durch die Recherchen kein geeignetes Befragungsinstrument identifiziert werden kann und keine Items abgeleitet werden können, bedarf es einer methodisch aufwendigeren Neuentwicklung eines Befragungsinstruments bzw. einzelner Items (Kilian et al. 2009; Wensing et al. 2008).

7.3.2.3 Durchführung von Fokusgruppen

Im Rahmen der Entwicklung einer Patientenbefragung ist die Einbindung von betroffenen Patienten und ihren Erfahrungen von zentraler Bedeutung. Dies erfolgt im Rahmen von Fokusgruppen.[25] Bei einer Fokusgruppe handelt es sich um eine moderierte Diskussion mit einer Kleingruppe über ein bestimmtes Thema. Strukturiert wird der Diskussionsprozess anhand eines Leitfadens, der als Orientierungshilfe für den Moderator fungiert und sicherstellt, dass in der Gruppe alle relevanten Aspekte angesprochen werden (Schulz et al. 2012).

Mit der Durchführung von Fokusgruppen als qualitative Methode im Rahmen von Fragebogenentwicklungen können unterschiedliche Ziele verfolgt werden (Rose et al. 2011; Morgan 2009; Geyer 2003):

Generierung eines Fragebogens:

- Identifizieren relevanter Themen aus Patientensicht, bezogen auf den Gegenstand der Befragung
- Verständnis von themenspezifischen Begriffen

Prüfung/Diskussion eines Fragebogenentwurfs:

- Diskussion und inhaltliche Validierung des Instruments in der Gruppe

Die Teilnehmerzahl in Fokusgruppen sollte zwischen vier und zwölf Teilnehmern liegen, ideal sind erfahrungsgemäß sechs bis zehn Teilnehmer. Die Auswahl der Teilnehmer sollte so erfolgen, dass die spätere Untersuchungsgruppe in ihren unterschiedlichen Merkmalen gut erfasst ist (z.B. im Hinblick auf Alter, Geschlecht, Bildungsstand u.a.). Es sollten mindestens zwei Fokusgruppen mit unterschiedlichen Teilnehmern mit einer Dauer von idealerweise ein bis zwei Stunden durchgeführt werden (Geyer 2003). Die Gruppensitzungen werden für die spätere Auswertung digital aufgezeichnet (Videoaufnahme) und anschließend transkribiert.

Bei der Entwicklung eines Patientenbefragungsinstruments im Rahmen der sektorenübergreifenden Qualitätssicherung hat die Durchführung der Fokusgruppen vor allem folgende Funktionen:

- Prüfung der Relevanz der im Fragebogenentwurf enthaltenen Themen/Fragen aus Patientensicht
- Prüfung, ob wichtige Qualitätspotenziale aus Patientensicht fehlen
- Ggf. Generierung ergänzender, verfahrensspezifisch relevanter Themen/Fragen

Die Rekrutierung der Teilnehmer der Fokusgruppen erfolgt über ein beauftragtes Unternehmen nach vom AQUA-Institut vorgegebenen Kriterien.

[25] Bei der Beauftragung der Entwicklung einer Patientenbefragung wird grundsätzlich für die Validierungsschritte „Einsatz von Fokusgruppen“ und „Testung des Instruments“ ein Ethikvotum eingeholt (vgl. auch Abschnitt 10.2.2). Zur Beurteilung der Studie werden u.a. der Ablauf und das Ziel der Studie berichtet sowie Patienteninformationen zur Studie und das entwickelte Befragungsinstrument eingereicht.

7.3.2.4 Einbindung von Fachexperten und Patientenvertretern

In den Entwicklungsprozess werden Fachexperten (möglichst mit zusätzlicher Expertise in der Patientenbefragung) und geeignete Patientenvertreter eingebunden, um zu einzelnen Fragebogeninhalten bzw. zum Gesamtinstrument (Fragebogenentwurf) Stellung zu nehmen.

Die Einbindung von Fachexperten und Patientenvertretern hat zum Ziel, die Fragebogenentwicklung bezüglich folgender Aspekte zu unterstützen:

- Prüfung, ob die Fragebogeninhalte zur Erfassung der für das Verfahren identifizierten Qualitätspotenziale (aus ärztlicher Sicht) zutreffen und ausreichend sind
- Prüfung, ob die Fragebogeninhalte und Fachbegriffe (aus ärztlicher Sicht) richtig und ausreichend präzise formuliert sind
- Prüfung der Eignung der zu erfassenden Items bzw. der vorgeschlagenen Indikatoren zum Zweck der Qualitätssicherung
- Diskussion von Umsetzungsfragen (z.B. Subpopulationen, Risikoadjustierung, Befragungszeitpunkt etc.)

Bei der Rekrutierung der Fachexperten werden in erster Linie Mitglieder bzw. Teilnehmer bereits bestehender verfahrensspezifischer Expertengruppen (z.B. Bundesfachgruppen, RAM-Panel) berücksichtigt. Hierbei wird eine angemessene Verteilung sowohl des stationären als auch des ambulanten Sektors angestrebt. Darüber hinaus werden Patientenvertreter gemäß §140f SGB V rekrutiert.

7.3.2.5 Kognitive Pretests

Kognitive Pretests (Interviews) betroffener Patienten im Rahmen der Fragebogenentwicklung dienen der Überprüfung (Prüfer et al. 2005):

- der Verständlichkeit und Eindeutigkeit der Fragebogeninhalte (Fragen und Antwortvorgaben, Ausfüllhinweise, einleitende Erläuterungen zu den unterschiedlichen Fragebogenabschnitten, u.a.), sowie der
- Handhabbarkeit des Fragebogens insgesamt (Fragebogenlänge, Nachvollziehbarkeit des Fragebogenaufbaus, u.a.).

Die Auswahl der Patienten für die kognitiven Pretests sollte so erfolgen, dass möglichst Patienten mit unterschiedlichen Charakteristika (z.B. im Hinblick auf Alter, Geschlecht, Bildungsstand, u.a.) vertreten sind. Des Weiteren sollten bei einer Bildung von Subpopulationen Patienten aus allen Subpopulationen an den kognitiven Pretests teilnehmen.

Für die Entwicklung von Fragebögen greift das AQUA-Institut auf die in Tabelle 8 dargestellten methodischen Schritte bei der Durchführung kognitiver Pretests zurück (Prüfer et al. 2005; Geyer 2003; Prüfer et al. 2000).

Tabelle 8: Methoden kognitiver Interviews

Methode	Vorgehensweise	Ziel
Probing	Direkte Frage an den Patienten nach Verständlichkeit einer bestimmten Frage oder eines bestimmten Begriffs	Wird die Frage/der Begriff so verstanden, wie vom Fragebogenentwickler beabsichtigt?
	Bitte an den Patienten um Auswahl der zutreffenden Antwortmöglichkeit (Antwortvorgabe) zu einer bestimmten Frage und Begründung der gewählten Antwort	Werden die Antwortvorgaben so verstanden, wie vom Fragebogenentwickler beabsichtigt? Sind die Antwortvorgaben aus Patientensicht zutreffend und ausreichend?
General Probing	Frage an den Patienten, ob es im Fragebogen (über das bereits Besprochene hinaus) etwas gibt, was nicht verständlich ist	Gibt es Unklarheiten, die bisher nicht aufgefallen waren.

Methode	Vorgehensweise	Ziel
Paraphrasing	Bitte an den Patienten, eine Frage im Fragebogen mit eigenen Worten zu formulieren bzw. einen bestimmten Begriff mit eigenen Worten zu umschreiben	Wird die Frage/der Begriff so verstanden, wie vom Fragebogenentwickler beabsichtigt?
Think Aloud	Bitte an den Patienten, bei der Beantwortung der Frage „laut zu denken“ bzw. nachträglich anzugeben, was ihm bei der Beantwortung durch den Kopf ging	Wie gehen Patienten bei der Beantwortung dieser Frage vor, d.h. wie „durchsuchen“ sie ihr Gedächtnis? Wird die Frage so verstanden, wie beabsichtigt?

Das AQUA-Institut wendet für die Durchführung der kognitiven Pretests überwiegend die Schritte des *Probing* und *General Probing* an. *Paraphrasing* oder *Think Aloud* kommen dann zur Anwendung, wenn gezielt überprüft werden soll, ob Patienten bestimmte Begriffe, Fragen oder Antwortvorgaben anders interpretieren, als vom Fragebogenentwickler beabsichtigt. Das Paraphrasieren durch die Patienten hilft, eine Frage oder einen Begriff für Patienten verständlicher zu formulieren. Die Rekrutierung der Teilnehmer der kognitiven Pretests erfolgt ebenfalls über ein beauftragtes Unternehmen nach vom AQUA-Institut vorgegebenen Kriterien.

7.3.2.6 Durchführung eines Panelprozesses

Ist die Entwicklung einer Patientenbefragung Bestandteil einer Neuentwicklung eines QS-Verfahrens zu einem bestimmten Thema, werden diejenigen Indikatoren, die über eine Patientenbefragung erhoben werden sollen, wie alle anderen entwickelten Indikatoren von den Teilnehmern des RAM-Panels hinsichtlich der Kategorien Relevanz und Praktikabilität bewertet (vgl. Abschnitt 8.2). Die Teilnehmer des RAM-Panels bewerten diese Indikatoren nicht isoliert, sondern vor dem Hintergrund des Gesamtverfahrens und den weiteren zur Verfügung stehenden Erhebungsinstrumenten. Darüber hinaus erfolgt keine Bewertung und Diskussion des Befragungsinstruments.

Im Panelprozess kann es sowohl zu Modifikationen als auch zum Ausschluss von Indikatoren kommen. Die dadurch evtl. beeinflussten Items werden in der aktuellen Fragebogenversion entsprechend geändert.

Wird das Instrument der Patientenbefragung einem bereits entwickelten QS-Verfahren (mit vorhandenen Instrumenten wie z.B. QS-Dokumentation und Routinedatennutzung) hinzugefügt, wird kein RAM-Prozess durchgeführt.

7.3.3 Festlegung des Befragungszeitpunktes

Bei der Entwicklung einer verfahrensspezifischen Patientenbefragung muss die Entscheidung für eine Methode der Veränderungsmessung bezüglich der Ergebnisqualität sowie die Festlegung eines geeigneten Befragungszeitpunktes (bzw. Befragungszeitraums) berücksichtigt werden. Für den Regelbetrieb der sektorenübergreifenden Qualitätssicherung sollten aufwandsarme Patientenbefragungen angestrebt werden.

7.3.3.1 Methoden der Veränderungsmessung für die Ergebnisqualität

Die Messung des Erfolgs eines Eingriffs bzw. einer Behandlung erfordert einen Vergleich des erkrankungsbezogenen Zustands des Patienten vor und nach der Intervention. Dabei gibt es die Möglichkeiten des direkten, indirekten und quasi-indirekten Vergleichs (Bitzer et al. 2011; Farin et al. 2009; Schmidt et al. 2004; Kohlmann et al. 1998b). Bei der direkten Methode der Veränderungsmessung wird die Veränderung retrospektiv erfasst, indem Patienten gebeten werden, eine Veränderung einzuschätzen. Bei dieser Methode weiß man aber weder genau, wie es dem Patienten vor dem Eingriff ging, noch wie es ihm danach ging, sondern nur, wie stark und in welche Richtung sich aus seiner Sicht etwas verändert hat (z.B. etwas verbessert, stark verbessert, gleich geblieben, etwas verschlechtert) (Bitzer et al. 2011; Kohlmann et al. 1998a). Eine Anwendung dieser Methode im Rahmen der sektorenübergreifenden Qualitätssicherung wird daher in der Regel als zu wenig aussagekräftig bewertet.

Im Rahmen der indirekten Methode werden die Patienten vor und nach dem Eingriff, also zu zwei Befragungszeitpunkten, nach ihrem aktuellen Gesundheitszustand gefragt. Dies bedeutet beispielsweise neben einem

organisatorischen Mehraufwand für zwei Befragungen (Bitzer et al. 2011; Farin et al. 2009) auch einen Mehraufwand hinsichtlich einer dann notwendigen fallbezogenen Verknüpfung der Daten (Befragungsergebnisse).

Bei einer quasi-indirekten Methode der retrospektiven Befragung hingegen sollen die Patienten ihren Gesundheitszustand vor dem Eingriff und ihren aktuellen Gesundheitszustand innerhalb ein und desselben Fragebogens berichten (Bitzer et al. 2011). Eine nachträgliche Verknüpfung von Befragungsergebnissen ist hier nicht notwendig. Der geeignete Befragungszeitpunkt zur Erfassung der Ergebnisqualität hängt jeweils von der eingriffsspezifischen Genesungsdauer ab.

Folgende Tabelle gibt einen Überblick über die Bewertungen der infrage kommenden Erhebungskonzepte der indirekten und der quasi-indirekten Methode durch verschiedene Autoren.

Tabelle 9: Unterschiedliche Studienbewertungen des quasi-indirekten Vergleichs als Methode der Veränderungsmessung

Bewertung	Begründung	Autor/Studie
Quasi-indirekter Vergleich ist nicht valide	Ergebnisse aus indirektem und quasi-indirektem Vergleich stimmen nicht genau überein; indirekter Vergleich ist valider, da kein Erinnerungs-Bias bezüglich Zustand vor Operation*	Mancuso et al. (1995) Pellise et al. (2005) Lingard et al. (2001)
Quasi-indirekter Vergleich ist valider als indirekter Vergleich	Ergebnisse aus indirektem und quasi-indirektem Vergleich stimmen nicht genau überein; quasi-indirekter Vergleich ist jedoch valider, da Patienten ihren Gesundheitszustand zu beiden Zeitpunkten von ein und demselben inneren Standpunkt (Bewertungsmaßstab) aus bewerten (bei zwei Befragungszeitpunkten: Problem der Response-Shift, d.h. unterschiedlicher Bewertungsmaßstab zu beiden Befragungszeitpunkten)	Nieuwkerk (2006)
Quasi-indirekter Vergleich ist ausreichend valide, wenn indirekter Vergleich nicht möglich ist (z.B. Notfall)	Quasi-indirekter Vergleich erzielt ähnliche und valide Ergebnisse wie indirekter Vergleich	Balain et al. (2009) Bitzer et al. (2011) Middel et al. (2006) Meyer et al. (2013) Meffert et al. (2010) Schmidt et al. (2004)

*Die Studien, die im Methodenvergleich zu sehr unterschiedlichen Ergebnissen kamen, wiesen jedoch methodische Einschränkungen auf: Vor und nach dem Eingriff wurde nicht das gleiche Befragungsinstrument eingesetzt, sondern in der Nachbefragung nur noch einzelne, ausgewählte Fragen (Lingard et al. 2001), vor dem Eingriff wurde schriftlich und nach dem Eingriff telefonisch befragt (sowie Zeitabstand zum Eingriff mindestens 1,5 Jahre) (Mancuso et al. 1995), der zeitliche Abstand zwischen Eingriff und Nachbefragung variierte zwischen 2 und 58 Monaten, bei einem Durchschnitt von 3 Jahren (Pellise et al. 2005).

Wenn auch aus wissenschaftlicher Sicht die indirekte Methode der Veränderungsmessung häufig bevorzugt wird, wird deutlich, dass sowohl der indirekte als auch der quasi-indirekte Vergleich mögliche Erhebungskonzepte für den Regelbetrieb der sektorenübergreifenden Qualitätssicherung sein können. Es ist verfahrensbezogen zu entscheiden, welche Methode geeignet ist; dabei sind Aufwand und Nutzen gegeneinander abzuwägen.

7.3.3.2 Methodik zur Ermittlung des Befragungszeitpunkts (bzw. Befragungszeitraums)

Nach der Entscheidung für eine Methode der Veränderungsmessung stellt sich die Frage, wann Patienten nach einer Indexleistung einen Fragebogen erhalten sollten. Dies muss sowohl bei Anwendung der indirekten als auch der quasi-indirekten Methode berücksichtigt werden.

Bei Patientenbefragungen zur Prozessqualität erfolgt nach Angaben in der Literatur die Fragebogenzusendung an die Patienten im Allgemeinen im Zeitabstand zwischen 2 und 8 Wochen nach der Behandlung (Saur et al. 2008; Freise 2003; Trojan et al. 2001; Kosinski et al. 1998). Mitunter ist der Zeitraum auch deutlich länger,

wenn dies aus eingriffsspezifischen Gründen sinnvoll erscheint (z.B. im Durchschnitt längerer Genesungszeitraum) (Schäfer et al. 2010; Freise 2003).

In der Literatur finden sich verschiedene Hinweise zur Rücklaufquote in Bezug auf den Befragungszeitpunkt, was ebenfalls bei der Wahl des Befragungszeitpunktes berücksichtigt werden sollte. Nach Freise (2003) haben Befragungszeitpunkte innerhalb von 3-17 Wochen nach Indexleistung weder Effekte auf die Rücklaufquoten noch Effekte auf die Ergebnisse der Patientenbefragungen, wohingegen Jensen et al. (2010) zeigen konnten, dass bei einem späten Befragungszeitpunkt die Rücklaufquote schlechter ist.

Der Befragungszeitpunkt wird mit den Fokusgruppen, mit den Fachexperten und den Patientenvertretern sowie mit den Teilnehmern des RAM-Panels diskutiert, die einen aus ihrer Sicht optimalen Befragungszeitpunkt empfehlen. Hier wird insbesondere die Länge des Genesungsprozesses berücksichtigt, um die Ergebnisqualität gut abbilden zu können. Darüber hinaus werden Methoden der Befragung aus Befragungsinstrumenten, die zur Itemgenerierung herangezogen werden, berücksichtigt und Erfahrungen aus der sich anschließenden Validierungsstudie in die Ermittlung des optimalen Befragungszeitpunkts, bzw. -zeitraums, integriert.

Die Umsetzung dieser Empfehlungen hängt von weiteren Rahmenbedingungen (wie z.B. Datenflüssen im Regelbetrieb) ab. Im Anschluss an die inhaltliche Entwicklung des Fragebogens schließt sich eine statistische, psychometrische Validierung des Gesamtinstruments an (vgl. Abschnitt 10.2.2).

© 2015 AQUA-Institut GmbH

8 Auswahl und Bewertung von Indikatoren

Ein „guter“ Indikator soll (medizinisch) relevant, klar formuliert und praktisch, d.h. unter Alltagsbedingungen und im gegebenen versorgungspolitischen Rahmen umsetzbar sein. Zudem soll er aussagekräftig sein, d.h. er ist (ggf.) risikoadjustiert, verfügt über gute Messeigenschaften und eignet sich deswegen zur öffentlichen Berichterstattung. Angesichts dieser zahlreichen Anforderungen fällt den Auswahl- und Bewertungsprozessen von Indikatoren eines Qualitätssicherungsverfahrens – unabhängig davon, ob dieses praktisch etabliert oder neu zu entwickeln ist – eine große versorgungspolitische Bedeutung zu.

Während weitgehend Einigkeit darüber besteht, dass die Qualität der Indikatoren selbst zu definieren und sicherzustellen ist (Sens et al. 2007), mangelt es letztlich an einer verbindlichen und alle Erfordernisse vollständig abdeckenden Methode zur Bewertung von Indikatoren. Dies gilt umso mehr, wenn Qualitätsindikatoren unter Beachtung versorgungspolitischer Rahmenbedingungen entwickelt werden sollen, für die es keine wissenschaftliche Begründung im engeren Sinne gibt. Insofern ist die transparente Darlegung jener Kriterien, nach denen Indikatoren zu unterschiedlichen Zeitpunkten und in unterschiedlichen Entwicklungsstadien ausgewählt und bewertet werden, der einzig verbindliche Maßstab.

Im Folgenden werden zunächst die Gütekriterien unterschiedlicher Bewertungskategorien von Indikatoren sowie ihre Operationalisierung in den unterschiedlichen Phasen der Indikatorenentwicklung erläutert.[26] Hier sind auch jene Bewertungskategorien eingeschlossen, die über den Rahmen des RAM-Prozesses – der als Kernelement der wissenschaftlichen Bewertungsmethode im Abschnitt 8.2 erläutert wird – hinausgehen.

8.1 Bewertungskategorien und Gütekriterien

Der erst im nächsten Abschnitt beschriebene RAM-Prozess ist zwardie zentrale Methode der Qualitätsbewertung von Indikatoren – und angesichts konkurrierender Auswahl- und Bewertungsprozesse auch die Methode der Wahl des AQUA-Instituts. Angesichts der einleitend genannten weiteren Anforderungen an Qualitätsindikatoren und Qualitätssicherungsverfahren ist aber ersichtlich, dass die Durchführung eines Expertenpanels zur Auswahl und Bewertung von Indikatoren nicht erschöpfend ist. Zum einen sind durch den zu beachtenden versorgungspolitischen Rahmen Grenzen gesetzt, die es opportun scheinen lassen, einzelne, möglicherweise relevante Indikatoren prospektiv auszuschließen (vgl. Abschnitt 6.1). Zum anderen lassen sich einige Qualitätsaspekte von Indikatoren erst retrospektiv, d. h. in der praktischen Erprobung bzw. im Regelbetrieb und der Analyse der Ergebnisse bewerten.

Die folgende Tabelle vermittelt einen Überblick über alle zur Anwendung kommenden Bewertungen von Indikatoren im Zusammenhang mit abgrenzbaren Phasen der Entwicklung vom RAM-Prozess über die Gesamtwürdigung des abgestimmten Indikatorensets bis zu einem (angenommenen) Regelbetrieb.

Tabelle 10: Kategorien und Phasen der Auswahl und Bewertung von Indikatoren

Bewertungskategorie	Zeitpunkt der Beurteilung
Relevanz (medizinische Effektivität und Patientennutzen)	RAM-Prozess (Bewertungsrunde 1)
Klarheit/Verständlichkeit	RAM-Prozess (Bewertungsrunde 1)
Praktikabilität (Umsetzbarkeit)	Im RAM-Prozess (Bewertungsrunde 2) erfolgt eine verbindliche Einschätzung. Die tatsächliche Praktikabilität kann nur in der Anwendung bzw. mit den Anwendern und den Softwareherstellern überprüft werden. D.h. die Praktikabilität wird in spezifisch zu bestimmenden Erprobungen bzw. dem Regelbetrieb beurteilt.

[26] Die folgenden Ausführungen zur Prüfung der Güte von Qualitätsindikatoren bauen auf dem RAM-Prozess und damit einer Neuentwicklung von Qualitätsindikatoren auf. Für die Indikatoren der bestehenden Verfahren der externen stationären Qualitätssicherung gelten die dort ebenfalls zutreffenden Kriterien im Regelfall analog.

© 2015 AQUA-Institut GmbH

Bewertungskategorie	Zeitpunkt der Beurteilung
(Erhebungs-)Aufwand	Würdigung des Indikatorensets.
Risikoadjustierung	Die Risikoadjustierung ist eine Frage des Bedarfs. Dieser Bedarf oder auch Nicht-Bedarf kann schon während des RAM-Prozesses, aber auch im Verlauf von Erprobungen oder im Regelbetrieb festgestellt werden. Die abschließende Beurteilung kann erst im Regelbetrieb erfolgen, insbesondere wenn hierzu Daten aus verschiedenen Datenquellen erforderlich sind.
Messeigenschaften (insb. Diskriminationsfähigkeit)	Die konkreten Messeigenschaften können erst im Regelbetrieb abschließend beurteilt werden.
Eignung der Indikatoren für die öffentliche Berichterstattung	Die Überprüfung findet im Regelbetrieb statt. Beurteilungsgrundlage sind Ergebnisse der Datenvalidierung, der Prüfung von Messeigenschaften und ggf. der Risikoadjustierung.

Die Indikatoren werden in verschiedenen Phasen der Entwicklung eines QS-Verfahrens (bis hin zum Regelbetrieb) immer wieder unter Anlegung verschiedener Bewertungskategorien überprüft. Auf die Umsetzung der Gütekriterien im Prozess der Indikatorenentwicklung und ihrer Implementierung sowie auf die Methoden zur empirischen Überprüfung wird im Folgenden genauer eingegangen.

Das nachfolgend dargestellte System der wissenschaftlichen Bewertungskategorien und Gütekriterien von Qualitätsindikatoren greift auf folgende Arbeiten zurück, deren wesentliche Kategorien und Kriterien vom AQUA-Institut in einem eigenen System synthetisiert sind:

- die RAND Appropriateness Method (RAM) (Fitch et al. 2001)
- das QUALIFY-Instrument (Reiter et al. 2008)
- die Systematik der Qualitätsindikatoren für die ambulante Versorgung (Szecsenyi et al. 2009)
- die von Terwee et al. (2007) vorgeschlagenen Kriterien zur Messgüte
- die von Campbell et al. (2003) vorgeschlagenen Qualitätskriterien für Indikatoren
- das niederländische AIRE-Instument (de Koning 2007; de Koning et al. 2007)
- die von Wollersheim et al. (2007) dargelegten Bewertungskriterien für Indikatoren
- die von Ouwens et al. (2007) dargelegte Systematik zur Entwicklung sektorenübergreifender Qualitätsindikatoren und zur Einbeziehung der Patientenperspektive

Bei der Systematisierung und Umsetzung von den in der Literatur zu findenden Gütekriterien wurden projektspezifische Zielstellungen und Methoden der sektorenübergreifenden Qualitätssicherung berücksichtigt.

8.1.1 Relevanz

Definition: Die Bewertungskategorie *Relevanz* umfasst verschiedene Kriterien zur Beurteilung der Bedeutung eines Indikators.

In der Literatur zur Anwendung des RAM-Prozesses (vgl. Abschnitt 8.2) finden sich verschiedene Operationalisierungen dieser wichtigen Bewertungskategorie, die auch allgemeiner als „Validität“ oder „Angemessenheit“ bezeichnet oder auch in anderen ähnlichen Ausprägungen kategorisiert wird.

Informationsgrundlage: Grundlage für die Bewertung ist das Indikatordatenblatt (siehe Abschnitt 6.3).

Gütekriterien: Folgende Kriterien sollen von den Experten zur Beurteilung der Relevanz herangezogen werden:

Die Relevanz ist eine globale Einschätzung und Beurteilung zur Bedeutung des Indikators.[27] Grundsätzlich ist ein Indikator umso relevanter,

- je stärker die vorliegende Evidenz für einen Zusammenhang zwischen der Zielstellung des Indikators und dem Versorgungsaspekt, z.B. zum klinischen Outcome, ist.
- je besser der Indikator zwischen schlechter und guter Leistung einer Einrichtung differenzieren kann.
- je stärker die Bedeutung für Patienten ist (patientenrelevante Endpunkte).
- je stärker die Bedeutung des Indikators für die betrachtete Population ist (epidemiologische Relevanz).
- je stärker der Nutzen bei Erfüllung des Indikators mögliche Risiken überwiegt.
- je besser das Ergebnis des Indikators durch die betrachtete Einrichtung beeinflussbar ist.

Operationalisierung: Die Bewertung der Relevanz erfolgt durch das RAM-Panel in der ersten Bewertungsrunde auf einer 9-stufigen Skala (1 = überhaupt nicht relevant bis 9 = sehr relevant) – und bezüglich der genannten Kriterien summarisch.

8.1.2 Klarheit/Verständlichkeit

Definition: Die Bewertungskategorie *Klarheit/Verständlichkeit* umfasst Gütekriterien zum sprachlichen und logischen Aufbau des Indikators. Die Bewertung erfolgt hier subjektiv aus Sicht der Experten des RAM-Panels. Die Experten sollen an dieser Stelle keine „stellvertretende“ Perspektive (etwa die eines Patienten) einnehmen.

Informationsgrundlage: Grundlage für die Bewertung der Verständlichkeit sind die Angaben zu den Indikatoren auf den Indikatordatenblättern.

Gütekriterien: Folgende Kriterien werden zur Beurteilung von Klarheit und Verständlichkeit herangezogen:

- Zähler und Nenner sind klar definiert.
- Keine Interpretationsspielräume: Der Indikator ist eindeutig, unmissverständlich, reproduzierbar.
- Die Formulierung ist überregional gültig.
- Die Formulierung ist institutionsunabhängig.

Operationalisierung: Die Bewertung erfolgt summarisch durch die Experten des Panels, die während der ersten Bewertungsrunde die Klarheit/Verständlichkeit auf einer 2-stufigen Skala (1 = ja, 2 = nein) beurteilen. Sind die Angaben oder Formulierungen zu einem Indikator nicht verständlich, können die unverständlichen Abschnitte benannt und alternative Formulierungen vorgeschlagen werden. Nach entsprechender Überarbeitung der Formulierungen werden die Indikatoren ein zweites Mal beurteilt. Sofern weiterhin Verständnisprobleme auftreten, werden die Formulierungen durch das Expertenpanel modifiziert und konsentiert.

8.1.3 Praktikabilität (Umsetzbarkeit)

Definition: Die Gütekriterien der Bewertungskategorie *Praktikabilität* überprüfen die praktische Anwendbarkeit und Umsetzbarkeit vor dem Hintergrund der vorgeschlagenen Instrumente zur Abbildung des Qualitätsindikators. Die Bewertung erfolgt mit unterschiedlichen Schwerpunkten zu *zwei verschiedenen Zeitpunkten*:

1. im Rahmen des RAM-Prozesses durch das Panel (zweite Bewertungsrunde)
2. im Rahmen möglicher Erprobungen (vgl. Abschnitt 10.2)

Zu 1: Prüfung der Praktikabilität (Umsetzbarkeit) durch das Expertenpanel
Informationsgrundlage: Anschließend an die Relevanz-Bewertung durch das RAM-Panel erfolgt die Beurteilung der Umsetzbarkeit der Indikatoren unter Real- bzw. Alltagsbedingungen. Bezogen auf Datenerhebungen über Einrichtungen steht die praktische Expertise der Teilnehmer des RAM-Panels hierbei im Vordergrund. Es soll

[27] Die Vorteile einer summarisch-subjektiven Bewertung von „Relevanz“ gegenüber der Bewertung einzelner Items und der Verrechnung in einem Relevanz-Score können dem Abschnitt 8.3. des Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al. 2011) entnommen werden.

einschätzen, ob die zu erhebenden Daten im medizinischen Routinebetrieb schon dokumentiert werden bzw. wie sich der zusätzliche Erhebungsaufwand darstellt. Weiterhin soll beurteilt werden, ob die vorgeschlagenen Prüfmechanismen geeignet sind, Dokumentationsfehler zu vermeiden. Bezogen auf Datenerhebungen über Routinedaten oder Patienten wird die Praktikabilität solcher Erhebungen prinzipiell in separaten Prozessen überprüft. Dem Panel werden die Ergebnisse dieser Prüfprozesse vor der Beurteilung der Praktikabilität zur Diskussion dargelegt.

Gütekriterien: Folgende Gütekriterien werden zur Beurteilung der Praktikabilität herangezogen:

- Angemessenheit des Erhebungsinstruments (kein alternatives Erhebungsinstrument vorhanden, das mit geringerem Aufwand mindestens gleichwertige Ergebnisse liefert)
- Angemessenheit des Zeitaufwands zur Erhebung der geforderten Daten (nur für QS-Dokumentation relevant)
- Implementierungsbarrieren (z.B. personell-fachlicher Aufwand für die Umsetzung der Instrumente) sind nicht vorhanden bzw. sie wurden angemessen berücksichtigt (nur für QS-Dokumentation relevant)
- Die geplanten Plausibilitätskontrollen sind angemessen zur Vermeidung von Dokumentationsfehlern

Operationalisierung: Die Bewertung erfolgt summarisch durch das RAM-Panel, das die Praktikabilität auf einer 9-stufigen Skala (1 = überhaupt nicht praktikabel bis 9 = sehr praktikabel) beurteilt. Indikatoren, die vom RAM-Panel als „nicht praktikabel“ bewertet wurden, sind von der weiteren Umsetzung ausgeschlossen.

Zu 2: Prüfung der Praktikabilität (Umsetzbarkeit) im Rahmen von Erprobungen
Informationsgrundlage: Grundlage sind die vom AQUA-Institut zur Umsetzung für ein QS-Verfahren empfohlenen Indikatoren bzw. die entsprechenden Angaben auf den Indikatordatenblättern und in den Entwicklungsberichten.

Gütekriterien: Technische Funktionsfähigkeit und Praktikabilität gemäß der Protokolle/Berichte von Erprobungen

Operationalisierung: Je nach thematischen Kontext und Erhebungsinstrument kann der notwendige Umfang von Prüfungen sehr unterschiedlich sein. In Abschnitt 10.2 sind mögliche Erprobungen dargelegt. Inwiefern solche Erprobungen notwendig sind, wird themenspezifisch entschieden.

8.1.4 Erhebungsaufwand

Definition: Die Bewertungskategorie *Erhebungsaufwand* bezieht sich auf den prinzipiellen, d.h. ggf. zusätzlich erforderlichen Dokumentationsaufwand der Leistungserbringer.

Informationsgrundlage: Grundlage zur Bewertung des Erhebungsaufwands ist die Analyse verfügbarer Datenquellen und Erhebungsinstrumente.

Gütekriterien: Zur Einschätzung des Erhebungsaufwands bei der Verwendung unterschiedlicher Datenquellen ist die Perspektive der Leistungserbringer das Gütekriterium:

- Niedriger Erhebungsaufwand besteht für den Leistungserbringer bei Verwendung folgender Datenquellen:
  - Sozialdaten bei den Krankenkassen (Spezifikation)
  - Abrechungsdaten nach §21 KHEntgG (Antrag auf Nutzung ausgewählter Datenfelder)
  - Abrechnungsdaten der Einrichtungen, automatisiert ausgeleitet
  - Patientenbefragung (sofern die Versendung der Befragungen nicht durch die Leistungserbringer erfolgt)

- Mittlerer Erhebungsaufwand besteht für den Leistungserbringer bei Verwendung folgender Datenquellen:
  - QS-Dokumentation, sofern sich die erhobenen Informationen aus der Patientendokumentation ableiten lassen (z.B. Medikation im stationären Bereich)
  - QS-Dokumentation, sofern sie nicht fallbezogen erhoben werden muss, sondern (z.B.) zu ausgewählten Stichtagen (Einrichtungsbefragung)

- Hoher Erhebungsaufwand besteht für den Leistungserbringer bei Verwendung folgender Datenquellen:
  - QS-Dokumentation, sofern die erhobenen Informationen sich nicht aus der Patientendokumentation ableiten lassen (z.B. manuelle Erfassung des Langzeitüberlebens, wie gegenwärtig in den Leistungsbereich der Transplantationsmedizin)

8.1.5 Risikoadjustierung

Definition: Die *Risikoadjustierung* von Qualitätsindikatoren ist dann erforderlich, wenn spezifische Merkmale von Patienten, wie z.B. Alter oder Schweregrad der Erkrankung, die Ausprägung eines Qualitätsindikators in einer Einrichtung beeinflussen. Die Auswahl von Qualitätsindikatoren für eine Risikoadjustierung sowie die Auswahl relevanter Einflussfaktoren und deren Modellierung ist ein mehrstufiger Prozess, weshalb nachfolgend von dem üblichen Beschreibungsschema abgewichen wird.[28]

1. Schritt: Suche nach möglichen Einflussfaktoren
Um Qualitätsindikatoren zu erkennen, für die eine Risikoadjustierung relevant ist, wird nach potenziell relevanten Einflussfaktoren für das Indikatorergebnis recherchiert. Die im Rahmen der Recherchen der Themenerschließung und der Indikatorenrecherche gefundenen Risikofaktoren und bereits vorhandene Risikoadjustierungsmodelle werden ggf. durch die Experten des Panels ergänzt und in den jeweiligen Indikatordatenblättern zitiert.

2. Schritt: Ergänzende Suche nach möglichen Einflussfaktoren auf Basis von Routinedaten
Sofern dem AQUA-Institut Routinedaten zur Verfügung stehen (z.B. Daten nach §21 KHEntgG), erfolgt eine ergänzende Suche nach möglichen Einflussfaktoren. Auch diese Einflussfaktoren werden in den Indikatordatenblättern dargelegt. Schritt 2 kann z.B. aus Gründen der Datenverfügbarkeit auch im Zuge der initialen Modellentwicklung (Schritt 5) durchgeführt werden.

3. Schritt: Modifikation und Anpassung durch das RAM-Panel
Im Rahmen der Bewertung und der Diskussion der vorgestellten Qualitätsindikatoren werden die Risikofaktoren zu den Qualitätsindikatoren ggf. ergänzt und dem nationalen Kontext angepasst.

4. Schritt: Entwicklung von Instrumenten zur Abbildung der Risikofaktoren
Nach Auswahl der relevanten Qualitätsindikatoren werden die Erhebungsinstrumente zur Abbildung der Qualitätsindikatoren und der dazugehörigen Risikofaktoren entwickelt.

5. Schritt: Initiale Modellentwicklung
Auf Basis der ersten Datenerfassungen nach Umsetzung der Qualitätsindikatoren erfolgt die initiale Entwicklung des Risikoadjustierungsmodells.

6. Schritt: Empfehlung zur Umsetzung
Die Ergebnisse der initialen Modellentwicklung werden mit Wissenschaftlern und Praktikern (für die externe stationäre Qualitätssicherung: die Bundesfachgruppen) diskutiert. Aus der Diskussion erfolgt eine Empfehlung zur weiteren Ausarbeitung bzw. Umsetzung des Risikoadjustierungsmodells.

7. Schritt: Umsetzung und Aktualisierung
Sofern keine weitere Ausarbeitung des Modells notwendig ist, wird es umgesetzt. Bei festgestelltem Bedarf erfolgt eine Aktualisierung des Modells.

[28] Für Details zur Risikoadjustierung sei an dieser Stelle auf den Abschnitt 11.2 verwiesen.

8.1.6 Messeigenschaften

Definition: Die Gütekriterien der Bewertungskategorie *Messeigenschaften* folgen den empirisch-statistischen Gütekriterien Vollzähligkeit (der Dokumentation), statistische Unterscheidungsfähigkeit (Diskriminationsfähigkeit) und Reliabilität der Messmethode.

Informationsgrundlage: Grundlage sind die gelieferten Daten der QS-Dokumentationen und Ergebnisse zur Vollzähligkeit der Daten (z.B. Soll-Ist-Vergleich).

Gütekriterien:

- Die **Vollzähligkeit** der Dokumentation beschreibt das Verhältnis von erwarteten Dokumentationsfällen zu vorhandenen Dokumentationsfällen. Für die Daten der QS-Dokumentation wird stets die Vollzähligkeit der Dokumentationen einzelner Einrichtungen angestrebt.
- Die **statistische Unterscheidungsfähigkeit** (Diskriminationsfähigkeit) spezifiziert die notwendige Mindestfallzahl, um gute von schlechter Qualität zu unterscheiden. Anschließend wird die Anzahl der Einrichtungen berechnet, die diese notwendige Fallzahl erreichen (AQUA 2011a; Heller 2010; Dimick et al. 2004).
- **Decken- und Bodeneffekte** beschreiben, ob mehr als 15 % der Einrichtungen den bestmöglichen bzw. schlechtesten Wert des Indikators erreichen. Wenn diese Effekte festgellt werden, können ein schlechtes Design des Indikators oder ein ausgeschöpftes Qualitätspotenzial die Ursachen dafür sein.
- **Reliabilität** beschreibt die Zuverlässigkeit der Messmethode. Sie wird üblicherweise ermittelt durch Prüfung der Übereinstimmung von Messwerten nach Wiederholung des gleichen Messverfahrens am selben Objekt (Test-Retest-Verfahren) oder durch den Vergleich von Messungen unterschiedlicher Bewerter (Inter-Rater-Verfahren). So wurde im Rahmen der „QI-Prüfung“ die Reliabilität der Messwerte über vier Quartale eines Jahres analysiert. Anhand der bisher durchgeführten Analysen (AQUA 2014i; AQUA 2013d; AQUA 2011a; AQUA 2011b) hat sich gezeigt, dass die Ergebnisse zur Test-Retest-Reliabilität und zur Diskriminationsfähigkeit zu faktisch identischen Ergebnissen führen.

Operationalisierung: Die Ergebnisse der empirisch-statistischen Analysen werden mit Wissenschaftlern und Praktikern (für die externe stationäre Qualitätssicherung: die Bundesfachgruppen) diskutiert. Aus der Diskussion erfolgt eine Empfehlung an den G-BA zur weiteren Ausarbeitung bzw. Umsetzung des jeweiligen Qualitätsindikators.

8.1.7 Eignung für die öffentliche Berichterstattung

Definition: Unter *Eignung der Indikatoren für die öffentliche Berichterstattung* wird im vorliegenden Zusammenhang verstanden, dass ein Qualitätsindikator in der Lage ist, ausreichend faire Qualitätsvergleiche zwischen einzelnen Einrichtungen zu ermöglichen, um die Ergebnisse der interessierten Öffentlichkeit zur Verfügung zu stellen.

Informationsgrundlage:

- Risikoadjustierung: Ergebnisse der Modellierung und Diskussion der Risikoadjustierung (sofern für den jeweiligen Indikator relevant)
- Überprüfung der Messeigenschaften: Ergebnisse der Analyse und Diskussion der Messeigenschaften
- Datenvalidierung: Ergebnisse der Analyse und der Diskussion aus dem Datenvalidierungsverfahren (Für die Datenvalidierung in der sektorenübergreifenden Qualitätssicherung hat das AQUA-Institut ein Konzept vorgelegt (vgl. (AQUA 2011e) und Abschnitt 9.2).

Gütekriterien: Folgende Kriterien zur Bewertung der Eignung eines Qualitätsindikators für die öffentliche Berichterstattung werden angewendet:

- Der Indikator ist auch für die interessierte Öffentlichkeit verständlich bzw. seine Ergebnisse sind richtig interpretierbar.
- Der Indikator kann zwischen guten und schlechten Ergebnissen unterscheiden (Messeigenschaften)
- Einflussgrößen für eine Risikoadjustierung werden, sofern erforderlich, angemessen berücksichtigt.
- Es besteht kein Risiko für Fehlsteuerungen der Versorgung bei einer einrichtungsbezogenen öffentlichen Berichterstattung.

Operationalisierung: Die Prüfung findet im Regelbetrieb statt. Nachdem auch die statistischen Analysen (Risikoadjustierung, Messeigenschaften, Datenvalidierung) und Erfahrungen aus der Umsetzung des Indikators vorliegen, wird die Eignung des Indikators für eine öffentliche Berichterstattung empirisch und unter Berücksichtigung der Einschätzung von Wissenschaftlern und Praktikern (für die externe stationäre Qualitätssicherung: die Bundesfachgruppen) in einem standardisierten mehrstufigen Prozess bewertet. Zu allen erhobenen Indikatoren wird eine Empfehlung an den G-BA übermittelt (AQUA 2014i).

8.1.8 Nicht berücksichtigte Gütekriterien

Die bisher genannten Gütekriterien decken ein weites Spektrum möglicher Kriterien ab, die in der Literatur genannt werden. Kriterien der Konstruktvalidität werden in eine routinemäßige Überprüfung der Qualitätsindikatoren aus Zeitgründen nicht einbezogen, weil solche Analysen spezifische, auf einen Indikator ausgerichtete Evaluationen beinhalten und einen Goldstandard voraussetzen.

8.2 RAND Appropriateness Method (RAM)

Das vom AQUA-Institut in einer spezifisch angepassten Form verwendete methodische Kernelement zur Auswahl und Bewertung von Indikatoren ist die von der Arbeitsgruppe „Research and Developement“ (RAND) an der Universität von Los Angeles (UCLA) entwickelte „Appropriateness Method“ (RAM) (Fitch et al. 2001). Im RAM-Prozess werden im Wesentlichen Relevanz und Praktikabilität bewertet.

Grundlage der Indikatorbewertung ist das in Kapitel 6 beschriebene Register bzw. die entsprechenden Operationalisierungen auf den Indikatordatenblättern (vgl. Abschnitt 6.3). Für viele der dort hinterlegten, recherchierten Indikatoren kann keine Evidenzbasierung auf der Grundlage randomisierter kontrollierter Studien vorausgesetzt werden (Hearnshaw et al. 2001; McColl et al. 1998). Dies gilt in besonderem Maße für neu entwickelte Indikatoren. Aus diesem Grund wird für das weitere Auswahl- und Bewertungsverfahren eine externe Sachverständigengruppe (RAM-Panel) gebildet, die diese Aufgabe übernimmt.

8.2.1 Grundlagen

Zur Einbeziehung von Experten in die Entwicklung von Indikatoren werden verschiedene Methoden verwendet (Campbell et al. 2003). Gebräuchlich sind Konsensus-Konferenzen (Murphy et al. 1998), Delphi-Befragungen (Fink et al. 1984) und der RAM-Prozess (Fitch et al. 2001). In der folgenden Tabelle sind die genannten Verfahren hinsichtlich der Art der Entscheidungsfindung und der Frage, ob persönliche Treffen der Experten oder schriftliche Expertenbefragungen (explizites Verfahren) stattfinden, unterschieden.

Tabelle 11: Unterscheidung gebräuchlicher Verfahren zur Einbeziehung von Expertenmeinungen

	Konsensus-Konferenzen	Delphi-Befragungen	RAM-Prozess
Explizites Verfahren zur Entscheidungsfindung?	Nein	Ja	Ja
Persönliche Treffen der Experten?	Ja	Nein	Ja

Explizite Verfahren zeichnen sich dadurch aus, dass die Expertenmeinungen über schriftliche Befragungen erhoben werden und die Auswertung der Befragung an die Experten zurückgemeldet wird. Die Entscheidung zum Ein- oder Ausschluss von Indikatoren erfolgt über ein vorab festgelegtes Verfahren.

Bei impliziten Verfahren erfolgt diese Entscheidung in der Gruppe, z.B. durch Abstimmung (Mehrheitsentscheidungen). Eine Gefahr impliziter Verfahren ist, dass dominante Persönlichkeiten in der Gruppe die Entscheidungsfindung stark beeinflussen können. Deshalb fordert Grol (1993) die konsquentte Anwendung expliziter Verfahren im Rahmen einer Entscheidungsfindung.

Ob persönliche Treffen der Experten stattfinden, ist das zweite dargestellte Unterscheidungsmerkmal der Verfahren. Die Vorteile persönlicher Treffen sind: Differenzierter Austausch unterschiedlicher Standpunkte und mögliche Einarbeitung der Diskussionsergebnisse in eine veränderte Fassung der Indikatoren sowie bessere

Vermittlung des Gesamtkonzepts. Diesen Vorteilen stehen als Nachteile ein höherer finanzieller Aufwand, eine eingeschränkte Gruppengröße und eine mögliche Beeinflussung der Ergebnisse durch dominante Panelteilnehmer im Rahmen der Arbeitstreffen gegenüber(Campbell et al. 2003).

Der RAM-Prozess – bzw. die inzwischen zahlreichen an diesen angelehnten Varianten – wird inzwischen weltweit zur Entwicklung von Indikatoren angewendet (Mourad et al. 2007; Campbell et al. 2003; Marshall et al. 2003). Er verbindet als einziger den expliziten Ansatz mit persönlichen Treffen der Experten in einer sehr strukturierten Arbeitsweise. Aufgrund der breiten Erfahrungsbasis und der Tatsache, dass diese Methode nachweislich in der Lage ist, gute Ergebnisse hervorzubringen (Campbell et al. 2003), ist sie eine wesentliche Grundlage für die Entwicklungsarbeit des AQUA-Instituts.

Inzwischen kann der RAM-Prozess als internationaler Goldstandard für Entwicklung, Auswahl und Bewertung von Qualitätsindikatoren im Gesundheitswesen angesehen werden – unabhängig davon, ob es sich um recherchierte, neu entwickelte oder aus Leitlinien abgeleitete Indikatoren handelt, und unabhängig davon, auf welche Art und Weise die Methode an spezifische Erfordernisse angepasst wurde. Neben der Anwendung bei den deutschen Qualitätsindikatorensystemen im ambulanten Bereich (wie AQUIK und QiSA) sind dazu allein in den letzten vier Jahren mehr als 30 internationale Publikationen aus so unterschiedlichen Themenbereichen wie z.B. Knie- und Hüftgelenkersatz (SooHoo et al. 2011), ältere chirurgische Patienten (McGory et al. 2009), Multiple Sklerose (Cheng et al. 2010), Epilepsie (Bokhour et al. 2009), Lupus Erythematodes (Gillis et al. 2011), gastrointestinale Blutungen (Kanwal et al. 2010) oder Patientensicherheit (Kristensen et al. 2009) erschienen.

Die konkrete Umsetzung der Methode kann – wie der Blick in die auf den RAM-Prozess verweisende Literatur zeigt – in Details und hinsichtlich unterschiedlicher Projektziele variieren. Dies betrifft die Bewertungskategorien bzw. die diesen zugrunde liegenden Gütekriterien, aber auch die Methode der Ergebnisauswertung. Im Kern bleiben jedoch immer die oben genannten Basiselemente erhalten: Ein explizites Verfahren zur Abstimmung und ein persönliches Treffen des Panels. Das RAND-Handbuch zur Methode vermerkt hierzu:

> *At its centre, the RAM is a modified Delphi method that, unlike the original Delphi, provides panellists with the opportunity to discuss their judgements between the rating rounds.* (Fitch et al. 2001: 7)

Auch im Kontext der Entwicklung von Qualitätsindikatoren durch das AQUA-Institut wird der ursprüngliche RAM-Prozess an verschiedenen Stellen variiert. Die wichtigsten Änderungen sind:

- Das dem Panel zur Bewertung vorgestellte Indikatorenregister enthält nicht alle recherchierten Indikatoren, sondern nur eine auf die recherchierten Qualitätspotenziale und den Versorgungspfad fokussierte Auswahl.
- Zwei Teilnehmer des Panels dürfen als Patientenvertreter durch die in §2 der Patientenbeteiligungsverordnung (PatBeteiligungsV) „für die Wahrnehmung der Interessen der Patientinnen und Patienten und der Selbsthilfe chronisch kranker und behinderter Menschen maßgeblichen Organisationen auf Bundesebene“ benannt werden. Sie nehmen als gleichberechtigte Experten am RAM-Prozess teil.
- Es wird ein zusätzliches Treffen der Experten vor Beginn der Bewertung durchgeführt, um die Teilnehmer mit der Methodik des Entwicklungsprozesses und den Zielen des jeweiligen QS-Verfahrens vertraut zu machen.
- Es wird ein abschließendes (optionales) Treffen am Ende der Bewertung durchgeführt, um den Experten die Chance einer Gesamtwürdigung zu geben
- Im Unterschied zur üblichen Vorgehensweise anderer Institutionen wird die Praktikabilität der Indikatoren erst in einer zweiten Bewertungsrunde beurteilt. Dies ist erforderlich, weil ein Teil der zur Bewertung der Praktikabilität notwendigen Angaben zur Operationalisierung (insbesondere die Spezifizierung von relevanten Kodes und Datenfeldern zur Erhebung) aufwendig ist und daher erst nach der Auswahl der relevanten Indikatoren erfolgt.

8.2.2 Ablauf des RAM-Prozesses

Der konkrete Ablauf des RAM-Prozesses ist in Anlehnung an die ursprüngliche Methode formalisiert und beinhaltet in der Adaption durch das AQUA-Institut fünf bedeutsame Elemente. Konkret sind dies eine gemeinsame Auftaktveranstaltung mit dem Projektteam des AQUA-Instituts und allen Teilnehmern des Panels sowie insgesamt vier, immer schriftlich abzugebenden Bewertungen zu unterschiedlichen Zeitpunkten des Bewer-

tungsprozesses. Die schriftlichen Bewertungen teilen sich in zwei Bewertungsrunden auf, in denen unterschiedliche Eigenschaften der Indikatoren bewertet werden. Jede Bewertungsrunde besteht aus einer rein postalischen bzw. einer Online-Bewertung und einer Bewertung während eines gemeinsamen Treffens.

	Indikatorenbewertung	
	postalisch/online	**Treffen**
Auftaktveranstaltung		Einführung in die Bewertungsmethodik
Bewertungsrunde 1	Relevanz + Klarheit/Verständlichkeit	Relevanz
Bewertungsrunde 2	Praktikabilität	Praktikabilität (ggf. Gesamtwürdigung)
(optionales Treffen)		Gesamtwürdigung, sofern nicht bereits durchgeführt

Abbildung 9: Ablauf des RAM-Prozesses mit Zuordnung zu den einzelnen Bewertungsrunden

Auftaktveranstaltung

Dem AQUA-Institut ist aus Erfahrung bei der Entwicklung von Indikatoren bekannt, dass es von Vorteil ist, eine gemeinsame Auftaktveranstaltung mit den Experten durchzuführen. Diese Vorgehensweise ist auch in anderen internationalen Studien erprobt (Fitch et al. 2001).

Im Rahmen der obligatorischen Auftaktveranstaltung werden die Panelteilnehmer mit dem Thema bzw. dem Versorgungspfad und den Qualitätspotenzialen sowie dem Ablauf des RAM-Prozesses vertraut gemacht. Inhaltliche Grundlage des Treffens sind die Ergebnisse der Themenerschließung. Die Teilnehmer erhalten dazu ein ausführliches Manual, das diese Ergebnisse zusammenfasst und einen Überblick über die konkreten Aufgaben und Vorgehensweisen im RAM-Prozess gibt.

In einer Gruppenarbeit während der Auftaktveranstaltung werden die Bewertungsmethodik und die Anwendung der Bewertungskategorien an einem konkreten Beispiel geübt. Ferner besteht die Gelegenheit, grundlegende Fragen zum Verfahren zu diskutieren. Für alle Teilnehmer des Panels bzw. besonders für die Patientenvertreter besteht auch während des laufenden Verfahrens jederzeit die Möglichkeit, beim AQUA-Institut detailliertere Informationen zu medizinischen Fragen per Telefon oder E-Mail einzuholen.

Die Teilnahme an der Auftaktveranstaltung ist verpflichtend, weil nur so sichergestellt werden kann, dass es nicht zu prinzipiellen Missverständnissen aufgrund mangelnder Vertrautheit mit der Aufgabenstellung und der Methode kommt. Eine Gruppenarbeit schon zu Beginn verbessert überdies die Qualität des Prozesses der Bearbeitung von Indikatoren. Eine Nichtteilnahme an der Auftaktveranstaltung führt deshalb zum Ausschluss aus dem Panel.

Am Ende der Auftaktveranstaltung erhalten alle Teilnehmer des RAM-Panels die Unterlagen für den ersten Teil der Bewertungsrunde 1. Dabei kann es sich – je nach dem, auf welche Art und Weise die Bewertung durchgeführt wird – um schriftliche Unterlagen oder persönliche Zugangskodes zu einer (anonymisierten) Online-Befragung handeln.

Bewertungsrunde 1

Der erste Teil der Bewertungsrunde 1 verfolgt das Ziel, die Relevanz der Indikatoren sowie ihre sprachlich-logische Klarheit/Verständlichkeit zu bewerten. Hierzu ist jeder Panelteilnehmer zunächst angewiesen, eine *erste individuelle und selbstständige Bewertung* ohne Absprachen abzugeben.

Im zweiten Teil der Bewertungsrunde 1, der im Rahmen eines Paneltreffens stattfindet, werden alle Indikatoren auf Grundlage der statistisch ausgewerteten Ergebnisse der postalischen Vorbewertung und der Kommentare zu den einzelnen Indikatoren diskutiert, ggf. modifiziert und hinsichtlich der Relevanz erneut bewertet.

Die Entscheidungsgrundlage für die Bewertung der Relevanz von Indikatoren ist einerseits die auf den Indikatordatenblättern ausgewiesene Evidenz und andererseits die Fachexpertise der Panelteilnehmer. Die den Bewertungskategorien „Relevanz“ und „Klarheit/Verständlichkeit“ unterliegenden Gütekriterien wurden bereits im vorherigen Abschnitt (8.1) im Zusammenhang mit den anderen Bewertungskategorien näher erläutert.

Die formale Arbeitsgrundlage für die Bewertungen sind die Indikatordatenblätter des Registers, denen hier gesonderte Skalen zur Bewertung von „Relevanz“ und „Klarheit/Verständlichkeit“ sowie ein Freitextfeld für Anmerkungen hinzugefügt sind.

Bewertung des Indikators (bitte ankreuzen bzw. ausfüllen):	
Relevanz	1=überhaupt nicht relevant 9=sehr relevant
Ist der Indikator für das Thema relevant?	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ 1 2 3 4 5 6 7 8 9
Klarheit/Verständlichkeit	
Ist der Indikator klar und verständlich formuliert?	Ja ☐ Nein ☐
Wenn nein, haben Sie Vorschläge für eine klarere und verständlichere Formulierung des Indikators?	

Abbildung 10: Bewertungsfeld für die Bewertungsrunde 1

Bewertungsrunde 2

In der Bewertungsrunde 2 bewertet das Panel die Indikatoren hinsichtlich ihrer Praktikabilität. Davor werden die verbliebenen, relevanten Indikatoren um weitere Informationen zu ihrer Umsetzung ergänzt. Dabei handelt es sich überwiegend um weitere Operationalisierungen der Erhebungsinstrumente bzw. Details zur Berechnung der Indikatoren (vgl. Kapitel 7).

Bei der Bewertung der *Praktikabilität* geht es um eine Beurteilung der Umsetzbarkeit der Indikatoren. Vor dem Hintergrund der eigenen praktischen Erfahrungen sollen die Teilnehmer des RAM-Panels einschätzen, ob die vorgeschlagenen Erhebungsinstrumente in der Praxis zur Erhebung der benötigten Informationen geeignet sind.

Alle Teilnehmer des RAM-Panels erhalten die Unterlagen für den ersten Teil der Bewertungsrunde 2; erneut kann es sich dabei entweder um schriftliche Unterlagen oder persönliche Zugangskodes zu einer (anonymisierten) Online-Befragung handeln.

Auf Basis der Ergebnisse des ersten Teils der Bewertungsrunde 2 werden die Indikatoren bei einem weiteren Paneltreffen erneut bewertet und ggf. nochmals modifiziert.

Die Bewertung der Praktikabilität durch das Panel kann aus naheliegenden Gründen – z.B. aufgrund der erforderlichen informationstechnischen Kenntnisse und sonstiger Umsetzungsvoraussetzungen (z.B. vorhandene Richtlinien, Datenschutz) – nicht abschließend sein. Eine diesbezügliche Evaluation erfolgt im Rahmen der weiteren Umsetzung der entwickelten Indikatoren (siehe Kapitel 10).

Bewertung des Indikators (bitte ankreuzen bzw. ausfüllen):	
Praktikabilität/Umsetzbarkeit	1=überhaupt nicht praktikabel/umsetzbar 9=sehr praktikabel/umsetzbar
Ist der Indikator in der Praxis gut umsetzbar?	☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ 1 2 3 4 5 6 7 8 9

Abbildung 11: Bewertungsfeld für die Bewertungsrunde 2

Die Gütekriterien zur Bewertung der Praktikabilität sind – wie im vorigen Abschnitt (8.1.3) erläutert – prozessübergreifend bzw. nicht auf den Zeitraum des RAM-Prozesses beschränkt.

8.2.3 Auswertung

Insgesamt werden die Indikatoren im Panel nach drei Kategorien bewertet:

- Relevanz (Bewertungsrunde 1)
- Klarheit/Verständlichkeit (Bewertungsrunde 1)
- Praktikabilität (Bewertungsrunde 2)

Relevanz

Die zentrale, weil abschließende Bewertungskategorie zur Auswahl der Indikatoren im RAM-Prozess ist die der Relevanz. Hier ist das Panelvotum insoweit abschließend, als dass die Relevanz von Indikatoren im weiteren Umsetzungsprozess keiner empirisch-statistischen Überprüfung ausgesetzt ist.

Die zur Anwendung kommenden Kriterien sind detailliert im Abschnitt 8.1.1 erläutert.

Die Panelteilnehmer bewerten die Relevanz des Indikators anhand einer 9-stufigen Skala (1 = überhaupt nicht relevant bis 9 = sehr relevant).

Für die Auswertung wird der Median als Lagemaß der Verteilung der Bewertungen genutzt:

- Median in [7–9]: relevant
- Median in [4–6]: fraglich relevant
- Median in [1–3]: nicht relevant

Innerhalb des RAM-Prozesses kommen nur Indikatoren in die Bewertungsrunde 2, die nach Abschluss der Bewertungsrunde 1 als relevant konsentiert sind. Konsens darüber, dass der Indikator relevant ist, liegt vor, wenn mindestens 75 % der Bewertungen in der Klasse [7–9] liegen. Je nach Größe des Panels und Zahl der Panelteilnehmer (gerade/ungerade) werden geringfügige Abweichungen dieser Grenze nach unten toleriert.

Klarheit/Verständlichkeit

Die Bewertung der Klarheit/Verständlichkeit des Indikators erfolgt auf einer 2-stufigen Skala (1 = ja, 2 = nein). Wenn die Teilnehmer einen Indikator als nicht verständlich einstufen, werden sie gebeten, alternative Formulierungsvorschläge zu machen.

Praktikabilität

Die Panelteilnehmer bewerten die Praktikabilität des Indikators anhand einer 9-stufigen Skala (1 = überhaupt nicht praktikabel bis 9 = sehr praktikabel).

Ein Indikator gilt als praktikabel, wenn der Median einen Wert ≥ 4 einnimmt. Im Rahmen der durch das AQUA-Institut angewendeten Methodik werden Indikatoren, die als „nicht praktikabel“ bewertet werden, von der weiteren Umsetzung ausgeschlossen. Eine ergänzende Einschätzung erfolgt zu einem späteren Zeitpunkt im Rahmen von etwaigen Erprobungen.

© 2015 AQUA-Institut GmbH

8.3 Würdigung des Indikatorensets

Aus dem Ergebnis der beiden Bewertungsrunden resultiert ein abgestimmtes und vollständig operationalisiertes Indikatorenset, das vom Panel und vom AQUA-Institut vor einer Empfehlung zur Umsetzung oder Erprobung abschließend gewürdigt wird.

Die abschließende Würdigung erfolgt mit Blick auf die Gesamtzielsetzung des QS-Verfahrens. Die Bewertungsgrundlage ist die Beantwortung der Frage, inwieweit die Ziele bzw. die konkretisierten Qualitätspotenziale des einzuführenden QS-Verfahrens mit den praktisch umsetzbaren Indikatoren voraussichtlich erreicht werden können. Dabei ist die Zielerreichung dezidiert nicht mit einer bestimmten Anzahl („Mindestanzahl") umsetzbarer Indikatoren verbunden.

Im Einzelnen werden drei Fragenkomplexe adressiert:

- Werden die im Rahmen der Themenerschließung identifizierten Qualitätspotenziale durch die entwickelten bzw. ausgewählten relevanten Indikatoren hinreichend abgebildet?

  Die Würdigung der einzelnen Indikatoren erfolgt mit Blick auf ihren unmittelbaren Nutzen in Bezug auf die Erreichung/Messung verfahrensimmanent gesetzter Ziele. In der Regel ist davon auszugehen, dass es sich dabei um Qualitätsdefizite bzw. um Verbesserungspotenziale handelt (ÄZQ 2009). Sollte die Summe der einzelnen Indikatoren in der Gesamtschau wesentliche Ziele verfehlen, behält sich das AQUA-Institut in Ausnahmefällen vor, einzelne Indikatoren im Set zu belassen, für die im Panel kein positiver Konsens bestand, wenn damit erreicht werden kann, das wesentliche, mit vertretbarem Aufwand messbare Qualitätspotenziale mithilfe dieser Indikatoren abgedeckt werden können.

- Steht der Aufwand zur Erhebung einzelner Qualitätsindikatoren in einem adäquaten Verhältnis zum Gesamtaufwand bzw. kann angesichts des Aufwands ein Regelbetrieb empfohlen werden?

  Auch hier wird die qualitative Relevanz des Indikatorensets mit Blick auf die in den Qualitätspotenzialen niedergelegten Ziele des Verfahrens überprüft. Wenn die Erreichung der Ziele durch die unter den gegebenen Rahmenbedingungen umsetzbaren Indikatoren nicht gewährleistet scheint oder die Aussicht auf Qualitätsverbesserungen zu gering ist oder der Erhebungsaufwand einzelner Indikatoren in keiner Relation mehr zum Nutzen des Verfahrens steht, werden dem G-BA entsprechende Empfehlungen zur Beendigung des Umsetzungsprozesses oder zu alternativen Umsetzungsmöglichkeiten unterbreitet. Dies gilt auch für weitere Umsetzungsschritte, d.h. das über den Panel-Prozess hinaus mitgeführte Indikatorenset wird nach jedem weiteren Entwicklungsschritt (Erprobung) neu und mit Blick auf diese Aufwand-/Nutzenrelation bewertet.

- Wie ist die Umsetzung der Indikatoren unter den allgemeinen rechtlichen (z.B. Datenschutz) und versorgungspolitisch gesetzten Rahmenbedingungen, wie z.B. der Datensparsamkeit, zu würdigen?

  Unter Berücksichtigung der zur Erhebung und Auswertung notwendigen Ressourcen und der allgemeinen Forderung nach ausgewogenen und „schlanken" Qualitätssicherungsverfahren behält sich das AQUA-Institut vor, einzelne, im RAM-Prozess als relevant und praktikabel bewertete Indikatoren nicht zur Umsetzung zu empfehlen, wenn:

  - zwei oder mehrere ähnliche Indikatoren den gleichen Aspekt des Versorgungspfads abdecken. Die Priorisierung erfolgt anhand der besseren/schlechteren Relevanz-Bewertung durch das RAM-Panel.
  - ein Indikator nicht mit dem Gebot der Datensparsamkeit vereinbar ist bzw. die rechtlichen Rahmenbedingungen für seine Implementierung auf absehbare Zeit nicht gegeben sind.

Das in dieser Weise vom AQUA-Institut unter Beteiligung des RAM-Panels abschließend gewürdigte Indikatorenset wird zur Begutachtung an die gemäß §137a Abs. 2 (3) SGB V an der Entwicklung zu beteiligenden Institutionen übergeben. In einem folgenden Stellungnahmeverfahren können die beteiligten Institutionen Bedenken, Einwände und Verbesserungsvorschläge zu den zur Umsetzung empfohlenen Indikatoren äußern. Soweit die vorgebrachten Argumente sachgerecht und begründet sind, kann das Institut von der Möglichkeit Gebrauch machen, das Indikatorenset letztmalig zu korrigieren.

© 2015 AQUA-Institut GmbH

8.4 „Verkürzte“ Auswahl- und Bewertungsprozesse

Je nach Leistungsbereich und spezifischer Fragestellung kann für Aufgaben im Rahmen der sektorspezifischen Weiterentwicklung ggf. auf vorhandene Strukturen (z.B. Fachgruppen) zurückgegriffen werden.

Für diese Fälle ist es denkbar – vorbehaltlich eines entsprechend konkretisierten Auftrags – ein inhaltlich, zeitlich und personell verkürztes Auswahl- und Bewertungsverfahren durchzuführen. Dabei bleibt der methodische Kern des im vorigen Abschnitt beschriebenen RAM-Prozesses weitgehend erhalten. Je nach Thema und erforderlichem, sachgerechtem Aufwand kann

- der Aufwand der Recherchen (Themenerschließung und Indikatorenrecherche) verkürzt werden, wenn sie nur in Bezug auf eine klar definierte Zielgruppe/Subpopulation, eine konkrete einzelne Versorgungsleistung (bspw. ein Follow-up) oder einen Indikator durchgeführt werden,
- der zeitliche Aufwand zur Bildung eines Expertenpanels reduziert werden, wenn auf die personellen Ressourcen einer bestehenden, adäquaten Fachgruppe zurückgegriffen werden kann und/oder
- die Anzahl der notwendigen Bewertungsrunden bei gegebenen Erhebungsinstrumenten halbiert werden, wenn das Expertengremium nur Relevanz sowie Klarheit/Verständlichkeit bewerten muss.

Insofern ergibt sich im Vorfeld aller Überlegungen zur Neuentwicklung einzelner Indikatoren ein grundsätzlicher Prüfbedarf:

- In Bezug auf die Möglichkeit, ein verkürztes RAM-Verfahren durchzuführen, muss geprüft werden, ob eine adäquate Fachgruppe zur Bewertung der Indikatoren vom AQUA-Institut, ggf. unter Berücksichtigung bestehender Fachgruppen (z.B. Bundesfachgruppen), ausgewählt werden kann.
- Wenn für einen Indikator neue/andere Erhebungsinstrumente erforderlich sind, muss geprüft werden, ob dafür eine Erprobung durchgeführt werden muss.
- Wenn für eine geplante Neuentwicklung Routinedaten verwendet werden (sollen), muss dies mit Blick auf die notwendigen Antragsverfahren zur Datennutzung (z.B. Nutzung der Daten nach §21 KHEntgG) im Zusammenhang mit dem Auftrag abgestimmt werden.

Ob der beschriebene erhebliche Aufwand (Erstellen von Indikatordatenblättern, Koordinierung und Auswertung der Bewertungen usw.) stets erforderlich ist, wird vom AQUA-Institut für jeden Fall einzeln überprüft.

© 2015 AQUA-Institut GmbH

9 Übergreifende Aspekte von Entwicklungsleistungen

Neben den auf einzelne Indikatoren bezogenen Empfehlungen enthalten die Abschlussberichte zur Entwicklung von Indikatoren, Instrumenten und der notwendigen Dokumentation übergreifende Empfehlungen zur Umsetzung eines QS-Verfahrens. Ein Teil der Umsetzungsvoraussetzungen wird zusammen mit den Indikatoren entwickelt (z.B. Fragen in Bezug auf die Auslösung einer Dokumentationspflicht, siehe Abschnitt 7.1.1), teilweise handelt es sich aber um eigenständige Fragestellungen. Dies gilt insbesondere für die nachfolgend thematisierten Aspekte.

9.1 Stichproben

Am Ende der Abschlussberichte zu den Entwicklungsleistungen stehen Empfehlungen zur Umsetzung an den G-BA. Die Empfehlungen nehmen insbesondere Bezug auf die Inhalte der themenspezifischen Bestimmungen (§3 Qesü-RL, vgl. auch Abschnitt 2.3.3), die der G-BA im Falle eines Beschlusses zur Umsetzung trifft. Eine gewisse Sonderstellung haben hier die Empfehlungen zu „Art und Umfang der Stichprobe und Festlegung der zu erhebenden Daten“ bzw. zur „Begründung einer eventuell notwendigen Vollerhebung“.

Eine ähnliche Regelung findet sich im §299 Abs. 1 Satz 4 und 5 SGB V:

> *(1) (...) Die Richtlinien und Beschlüsse sowie Vereinbarungen nach Satz 1 haben darüber hinaus sicherzustellen, dass*
> *1. in der Regel die Datenerhebung auf eine Stichprobe der betroffenen Patienten begrenzt wird (...)*
> *(...) Abweichend von Satz 1 Nr. 1 können die Richtlinien, Beschlüsse und Vereinbarungen auch eine Vollerhebung der Daten aller betroffenen Patienten vorsehen, sofern dieses aus gewichtigen medizinisch fachlichen oder gewichtigen methodischen Gründen, die als Bestandteil der Richtlinien, Beschlüsse und Vereinbarungen dargelegt werden müssen, erforderlich ist. Auch Auswahl, Umfang und Verfahren der Stichprobe sind in den Richtlinien und Beschlüssen sowie den Vereinbarungen nach Satz 1 festzulegen und von den an der vertragsärztlichen Versorgung teilnehmenden Ärzten und den übrigen Leistungserbringern zu erheben und zu übermitteln.*

Vor dem Hintergrund, dass eine solche Empfehlung die Darlegung von Entscheidungskriterien erforderlich macht,[29] die ihrerseits von weiteren Rahmenbedingungen – Ziele des Verfahrens, Gebot der Datensparsamkeit usw. – abhängen, wird im Folgenden eine kurze, aber grundsätzliche Erörterung des Themas gegeben.

9.1.1 Bestandsaufnahme und grundsätzliche Kriterien zur Nutzung von Stichproben

Die derzeitigen Verfahren der externen stationären Qualitätssicherung werden oft fälschlich als Vollerhebungen angesehen. Die Auswahl der Leistungsbereiche oder die Auswahl von Tracer-Verfahren (Kessner et al. 1973) für die Qualitätssicherung sind jedoch gezielt und damit Stichproben (Schnell et al. 2013: 257f). Auch innerhalb der bestehenden QS-Verfahren der externen stationären Qualitätssicherung werden gezielt Fälle bzw. Patienten ausgewählt, für die eine vergleichende Betrachtung der Qualität sinnvoll ist. So sind beispielsweise im Leistungsbereich *Cholezystektomie* die Begleit-Cholezystektomien sowie auch eine Reihe von Krebserkrankungen von der vergleichenden Betrachtung ausgeschlossen. Auch im ambulanten Sektor ist eine überwiegend auf Stichproben basierende Qualitätssicherung gemäß §136(2) SGB V etabliert.

Für neue sektorenübergreifende QS-Verfahren wird dieses grundlegende Prozedere weiterentwickelt. Ausgehend von der in einem Auftrag definierten Grundgesamtheit (Population) werden im Rahmen der Themenerschließung geeignete Stichproben von Patienten (Subpopulationen) definiert. Kriterien der Auswahl sind zu-

[29] Vom Folgenden unberührt sind die etablierten Stichprobenprüfungen im Rahmen der Datenvalidierung der bestehenden (stationären) QS-Verfahren sowie Stichproben im Rahmen der Patientenbefragung.

nächst Qualitätspotenziale und Versorgungsrelevanz in Bezug auf die Fragestellung (vgl. Abschnitt 4.1). Im weiteren Entwicklungsprozess werden auch Aspekte der Umsetzbarkeit bzw. das Aufwand-/Nutzenverhältnis in die Bildung einer themenspezifischen Stichprobe einbezogen.

Eine darüber hinausgehende Bildung von Zufallsstichproben innerhalb der zuvor definierten Stichproben (Subpopulationen) ist dann zielführend, wenn bereits kleinere Datenmengen mit ausreichender Sicherheit auf Qualitätsprobleme hinweisen. Wie die Daten des AQUA-Instituts sowie die Erfahrung aus anderen Qualitätssicherungsprogrammen (Bardach et al. 2010; Welke et al. 2010; Dimick et al. 2004; Rollins 2004) zeigen, ist dies für zahlreiche Qualitätsmessverfahren nicht gegeben bzw. wird nur äußerst selten erreicht. Wenn aus einer Zufallsstichprobe keine hinreichend aussagefähigen Daten gewonnen werden können, ist dies ein methodisch hinreichender Grund, eine Vollerhebung im Rahmen der definierten Grundgesamtheit zu empfehlen (Eichenhofer et al. 2013, S. 1854). Dennoch sollte aufgrund von Effizienzüberlegungen grundsätzlich versucht werden, wann immer aus methodischer Sicht möglich, Zufallsstichproben innerhalb definierter QS-Verfahren zu bilden.

9.1.2 Methodische Grundlagen zur Bildung von Zufallsstichproben

Wie vorangehend dargestellt, werden Stichproben von Patienten in der Qualitätssicherung häufig über eine fachliche Auswahl von definierten Erkrankungen oder Behandlungen durchgeführt. Nachfolgende Ausführungen beziehen sich auf darüber hinausgehende Möglichkeiten zu Bildung von Zufallsstichproben.

Unter einer Grundgesamtheit (Population) versteht man die Menge aller Elemente, über die Aussagen getroffen werden sollen. Im Allgemeinen spricht man von einer Vollerhebung, wenn alle Elemente einer Grundgesamtheit in die Erhebung einbezogen werden. Werden hingegen nur von einer Teilmenge der Grundgesamtheit Daten erfasst, spricht man von einer Teil- oder Stichprobenerhebung.

Unterhalb der Ebene der Population befindet sich der „sampling frame“. Dieser stellt die Operationalisierung der Population, also eine Liste von Elementen, die die Population enumerieren soll, dar (Lohr 2010: 4). Beispielsweise kann dies eine Liste von Datensätzen sein, in denen bestimmte OPS-Kodes vorhanden sind, die Patienten mit bestimmten operativen Eingriffen erfassen sollen.

Die Frage, ob eine Vollerhebung der definierten Grundgesamtheit oder eine Stichprobe vorzuziehen ist, hängt von der jeweiligen Fragestellung, der Definition und der Größe der Grundgesamtheit sowie dem erforderlichen Aufwand ab. Der größte Nachteil einer Vollerhebung ist dabei der in der Regel hohe Ressourcenaufwand. Vollerhebungen sollten also nur dann durchgeführt werden, wenn der erhöhte Ressourcenverbrauch aufgrund von methodischen Überlegungen zwingend erforderlich ist.

Stichprobenuntersuchungen sind weniger aufwendig als Vollerhebungen und gehen daher mit geringeren Kosten der Erhebung und Aufbereitung einher. Sie können zudem schnellere Ergebnisse liefern, insbesondere bei rasch veränderlichen Merkmalen. Für Stichprobenuntersuchungen spricht außerdem die Möglichkeit einer gründlicheren Durchführung, wodurch genauere Ergebnisse als bei Vollerhebungen erzielt werden können (Lohr 2010: 18). Dies gilt jedoch nur, wenn die Ergebnisse aus der Stichprobe auf eine Grundgesamtheit hochgerechnet werden, d.h. es muss sich um eine echte Zufallsstichprobe handeln.

Von einer Zufallsstichprobe wird gesprochen, wenn die Ziehungswahrscheinlichkeiten für die Elemente der Stichprobe angegeben werden können bzw. gleich sind (einfache Zufallsstichprobe). Um eine Zufallsstichprobe ziehen zu können, müssen die Grundgesamtheit klar definiert und ein zufälliger Prozess der Ziehung vorliegen. In QS-Verfahren ist der „sampling frame“ eindeutig durch den QS-Filter definiert. Es sind stichprobentheoretisch nicht lösbare Fragen der Dokumentationsqualität sowie medizinisch-inhaltliche Fragen, ob der „sampling frame“ die Zielpopulation adäquat abbildet. Ist dies der Fall, kann ein „sampling bias“ nur entstehen, wenn Datensätze nicht vollzählig sind, was allerdings in der Qualitätssicherung empirisch gesehen eine untergeordnete Rolle spielt: So betrug beispielsweise die Vollzähligkeit im Leistungsbereich Geburtshilfe für das Erfassungsjahr 2013 99,9 % (AQUA 2014e).

Zufallsstichproben ermöglichen Rückschlüsse auf die Grundgesamtheit in Bezug auf einen schätzbaren zufälligen Fehler. Der zufällige Fehler einer Stichprobe hängt dabei ganz wesentlich von der Größe der Stichprobe, aber auch von der Streuung der betrachteten Merkmale (z.B. Qualitätsindikatoren, Risikoadjustierungsvariablen) ab. Die Streuung der betrachteten Merkmale wird dabei in aller Regel aus eigenen Pilotdaten (z.B. den Daten der Vorjahreserhebung) oder externen Erhebungen weitgehend bekannt sein, und die Größe der Stichprobe sollte so gewählt werden, dass der zufällige Fehler akzeptabel gering oder, umgekehrt formuliert, die Präzision der Qualitätsaussage für die betrachtete Grundgesamtheit ausreichend groß ist.

Geschichtete, mehrstufige Stichproben sind Zufallsstichproben und bilden eine Möglichkeit, den Erhebungs- und Auswertungsaufwand bei vergleichbarer Aussagekraft gering zu halten (Lohr 2010: 73ff). Dabei werden Stichproben gezogen, bei denen die Verteilung der relevanten Merkmalsausprägungen denen in der Gesamtpopulation entspricht oder aber bestimmte Merkmalsausprägungen bewusst überrepräsentiert sind, um sichere Aussagen über bestimmte Subpopulationen machen zu können (Lohr 2010: 73ff). In beiden Fällen besteht jedoch das Problem, dass die Verteilung in der Population bekannt sein muss (Lee et al. 2006: 6).

In jedem Fall ist es bei Stichprobenziehungen notwendig, Poweranalysen bzw. Fallzahlberechnungen durchzuführen. Mit ihrer Hilfe kann abgeschätzt werden, wie groß die benötige Fallzahl der Stichprobe bei gegebenen Merkmalen sein muss, um unter definierten Randbedingungen tatsächlich existierende Qualitätsunterschiede erkennen zu können (AQUA 2011a; Heller 2010; Pfandzelter et al. 2009; Dimick et al. 2004).

In der sektorenübergreifenden Qualitätssicherung wird zumeist ein statistisch tragfähiger Vergleich zwischen einzelnen Leistungserbringern angestrebt. Dies erfordert, dass die Einrichtungen die oben definierte Minimalfallzahl erbringen (Dimick et al. 2004). Dies liegt daran, dass (die vorwiegend verwendeten) binären Qualitätsindikatoren aufgrund ihrer Beschaffenheit in der Regel besonders niedrige oder besonders hohe Prävalenzen aufweisen (Mor 2005: 342; Credé et al. 1990: 199; Luft et al. 1986). Daraus folgt, dass zumeist eher große Fallzahlen pro Leistungserbringer benötigt werden. Dabei weisen einige Analysen aus dem stationären Bereich darauf hin, dass dieses Ziel auch bei den aktuellen und bewussten Auswahlen (Stichproben) oft nicht in befriedigendem Maße erreicht wird (AQUA 2013d; AQUA 2011a; Heller 2010; Dimick et al. 2004). Solange ein Einrichtungsvergleich angestrebt wird, ist daher nicht zu erwarten, dass künftig Zufallsstichproben von Patienten in breitem Umfang in der Qualitätssicherung eingesetzt werden können.

Sofern im Rahmen der sektorenübergreifenden Qualitätssicherung Stichproben von Patienten gebildet werden können, ist darauf zu achten, ob die Abbildung mehrerer Versorger zur Ermittlung des Indikators notwendig ist. Zu erwarten ist in einem solchen Fall, dass die resultierenden Stichproben auf Leistungserbringerebene ihren Zweck nicht erfüllen, denn die in einem Verfahren eingeschlossenen Indexpatienten werden im Zeitverlauf in immer mehr Einrichtungen behandelt. Die von einem Leistungserbringer minimal benötigte Fallzahl würde dann ggf. nicht mehr erreicht werden, weil alle beteiligten Einrichtungen berücksichtigt werden müssen.

Wenn keine Vergleiche von einzelnen Leistungserbringern, sondern von mehreren Leistungserbringern angestrebt werden (Regionen, Area, Netzwerke von Versorgern), gilt die minimal benötigte Fallzahl für diese Gruppe von Leistungserbringern.

Das Ziehen einer Zufallsstichprobe durch eine Einrichtung setzt allerdings voraus, dass dem Leistungserbringer mitgeteilt werden kann, welche Fälle er zu erheben hat. Die Ziehungswahrscheinlichkeit müsste aus Gründen der Praktikabilität auf der Basis von Vorjahresdaten geschätzt und in die QS-Software ein zufälliger Ziehungsalgorithmus integriert werden, weil bei einer den Einrichtungen überlassenen Auswahl von Patienten mit systematischen Verzerrungen zu rechnen ist. Dies könnte auch über die Angabe eines (zufällig) bestimmten Intervalls erfolgen, in dem Fälle zu dokumentieren sind. In jedem Fall ist dafür Sorge zu tragen, dass keine verzerrte Auswahl von zu dokumentierenden Fällen oder Patienten erfolgt.

Eine solche Zufallsstichprobe von Fällen oder Patienten innerhalb der Einrichtung käme für alle Leistungserbringer infrage, die mit ihren Fallzahlen über der anhand der Vorjahresdaten ermittelten minimalen Fallzahl für die betrachteten Merkmale des jeweiligen Verfahrens liegen.

In Bezug auf die Umsetzung von Patientenbefragungen wird gegenwärtig diskutiert, ob die Bildung einer Stichprobe von Patienten besser durch eine neutrale Stelle durchgeführt werden sollte.

Eine weitere Möglichkeit für die fernere Zukunft bzw. für spezielle Anwendungsgebiete ist, dass – falls eine Identifikation von Patienten mit auffälligen Qualitätsergebnissen über Routinedaten möglich ist – nur für diese Patienten, sowie auch für eine Reihe von Kontrollpatienten in einem Verfahren, Dokumentationen anzufertigen sind. Anschließend wären Analysen im Rahmen eines eingebetteten Fall-Kontroll-Designs möglich (Rothman et al. 2008). Allerdings müssen hier mindestens die gleichen Voraussetzungen bezüglich einer eindeutigen Zuordnung von QS-Dokumentationen zu Routinedaten gegeben sein, wie oben beschrieben.

Theoretisch wäre denkbar, dass man Auffälligkeiten in Qualitätsindikatoren zunächst anhand von Routinedaten ermittelt, und in einem zweiten Durchgang alle auffälligen, aber nur eine Zufallsstichprobe der unauffälligen Leistungserbringer zur Nachdokumentation verpflichtet. Tatsächlich ist ein solches Vorgehen im Rahmen der Datenvalidierung, die auf Krankenakten zurückgreift, vorgesehen (AQUA 2011e: 34f). In der Praxis ist eine solche Nachdokumentation in den eigentlichen QS-Verfahren aufgrund der Zeitdifferenz zwischen der Verfügbarkeit der Routinedaten (inklusive ihrer Auswertung) und dem eigentlichen Eingriff – die bei den InEK-Daten beispielsweise mehr als anderthalb Jahre betragen kann – jedoch kaum praktikabel.

Zusammenfassend lässt sich sagen, dass nahezu immer Stichproben in der Qualitätssicherung genutzt werden. Zunächst wird eine willkürliche Auswahl von betroffenen Patienten innerhalb eines QS-Verfahrens getroffen, für die eine vergleichende Analyse sinnvoll ist. Abhängig vom Ziel der Qualitätssicherung und von der Verteilung der zugrunde liegenden Merkmale können darüber hinaus weitere Zufallsstichproben gezogen werden. Wenn das Ziel nicht der Qualitätsvergleich von Einrichtungen ist, können (gegebenenfalls geschichtete) Stichproben von Einrichtungen und anschließend innerhalb der Einrichtungen Stichproben von Fällen bzw. Patienten gezogen werden. Andernfalls hängt die Möglichkeit einer Stichprobenerhebung davon ab, ob die Fallzahlen in den einzelnen Einrichtungen groß genug sind, um statistisch zuverlässige Aussagen zu erlauben.

9.2 Datenvalidität und Dokumentationsqualität

Valide Qualitätsmessungen können nur auf der Basis von Daten mit hoher Dokumentationsqualität erhoben werden. Im Zentrum der Qualitätsprüfungen standen bisher Überprüfung der Korrektheit (jetzt: Konkordanz), Vollständigkeit und Vollzähligkeit der Daten (AQUA 2011c: 17). Um eine noch aussagekräftigere und transparentere Prüfung der Daten- und Dokumentationsqualität durchführen zu können, hat das AQUA-Institut im Zuge des Konzepts zur sektorenübergreifenden Datenvalidierung diese Merkmale noch weiter systematisiert und ihre Legitimation evaluiert (AQUA 2014a; Schubert et al. 2014; AQUA 2013b; AQUA 2012b). Allgemein ist allerdings zu beachten, dass gute Dokumentationsqualität zwar eine Voraussetzung für eine gute Reliabilität und Validität der Daten ist, diese jedoch nicht impliziert.

9.2.1 Allgemeine Prüfkriterien der Dokumentationsqualität

Für die Festlegung von Prüfkriterien für die Dokumentationsqualität werden folgende Eigenschaften geprüft:

- bisheriger Einsatz im nationalen Kontext
- Einsatz im internationalen Kontext
- Meinungen in der Fachliteratur zur Methodik
- geringer Ressourcenverbrauch
- geringe Invasivität (Arbeit mit Daten, die möglichst nicht zusätzlich erhoben werden müssen, z.B. Zeitreihendaten)
- möglichst hohe Diskriminationsfähigkeit (Wie deutlich kann „gute“ von „schlechter“ Dokumentationsqualität unterschieden werden?)
- möglichst hohe Fehlerursachenerkennung:
  - Möglichkeit, systematische Fehler von Zufallsfehlern zu unterscheiden
  - Möglichkeit, Rückschlüsse auf mögliche Fehlerursachen (Datenmanipulation, Softwarefehler etc.) zu ziehen

Mithilfe dieser Kriterien wurden in Anlehnung an die Typologie von Nonnemacher et al. (2007) folgende Prüfkriterien für die Überprüfung der Dokumentationsqualität ausgewählt:

Kriterien der Akkuratesse

Die Kriterien der Akkuratesse sind *Vollzähligkeit*, *Vollständigkeit* und *Konkordanz (Korrektheit)*. Sie wurden auch bisher immer in den Verfahren der externen stationären Qualitätssicherung geprüft.

Die Vollzähligkeit bezieht sich auf Erfassung möglichst aller Fälle in einem Leistungsbereich und wird durch einen Soll-Ist-Vergleich geprüft.

Das Vollständigkeitskriterium (z.B. Batini et al. 2006: 26) bezieht sich auf die Frage, inwieweit die erhobenen Datensätze gültige Werte in allen zu erhebenden Datenfeldern aufweisen. Dies wird primär durch die QS-Software beim Leistungserbringer überprüft und sichergestellt.

Ein besonders aussagekräftiges Prüfkriterium ist die Konkordanz, d.h. die Übereinstimmung der Daten mit einer Referenzquelle, deren Angaben als Abbildung der Realität angenommen werden (Nonnemacher et al. 2007: 15f; Batini et al. 2006: 4ff). Die übliche Referenzquelle für medizinische Daten ist dabei die Krankenakte (Arts et al. 2002: 605). Da jedoch ein Vergleich der im Rahmen des QS-Verfahrens erfassten Daten mit denen der Krankenakte, besonders dann, wenn diese nur in Papierform vorliegt, mit einem hohen Aufwand verbunden ist (Besuch des Leistungserbringers durch die auswertende Stelle, Gewährleistung des Datenschutzes, organisatorischer Aufwand des Leistungserbringers), wird dieses Kriterium nur stichprobenartig überprüft.

Kriterien der Plausibilität

An der Grenze zwischen Akkuratesse- und Plausibilitätskriterien liegt der Anteil des *Auftretens unzulässiger Werte*, wie beispielsweise eine negative Altersangabe oder ein nicht-existierender ICD-Kode (Nonnemacher et al. 2007: 23; Olsen 2003: 155ff). Da die Prüfung auf unerlaubte Werte durch einfache Softwarelösungen vorgenommen werden kann, wird ein Großteil dieser Prüfung bereits durch die Software beim Leistungserbringer durchgeführt.

Auch bei der Prüfung des Plausibilitätskriteriums der *Widerspruchsfreiheit* (Chin et al. 2008: 304; Nonnemacher et al. 2007: 17f; Körner et al. 2006: 112) übernimmt die Erhebungssoftware beim Leistungserbringer einen Großteil dieser Prüfung bereits vor der Datenübermittlung.

Ein weiteres gängiges Plausibilitätskriterium ist die Untersuchung der *Werteverteilung* (Maydanchik 2007: 73; Batini et al. 2006: 86). Darunter kann sowohl die Analyse von Ausreißern in der Werteverteilung fallen als auch die Form der Verteilung selbst, d.h. die Überprüfung, ob die Verteilung der Werte der zur erwartenden Verteilung entspricht. Beispielsweise deuten überdurchschnittlich häufig durch 4 teilbare Atemfrequenzdaten (z.B. im Leistungsbereich *Ambulante Pneumonie*) auf Messungen im 15-Sekunden-Intervall und anschließende Multiplikation mit 4 hin. Insofern müssen die Daten als ungenau betrachtet werden.

Das ambitionierteste Plausiblitätskriterium ist die *Replikation bekannter Korrelationen* (Nonnemacher et al. 2007: 26; Stausberg et al. 2006: 480; Hasford et al. 1994). Dieses Kriterium eignet sich besonders dafür, Täuschungen aufzudecken. Auf der Kostenseite stehen hier zwar hohe Anforderungen an die Entwicklung geeigneter Korrelationsmatrizen zur Identifizierung bekannter Korrelationen, solche Matrizen können aber über Jahre hinweg verwendet werden.

Strukturell dem Kriterium der bekannten Korrelationen ähnlich ist das Kriterium des *Anteils der fehlenden Werte bei optionalen Datenfeldeinträgen* (Nonnemacher et al. 2007: 21). Hier werden bekannte empirische Ausfallraten mit den tatsächlich auftretenden Ausfallraten verglichen, um statistisch auffällige Abweichungen (typischerweise um mehr als zwei Standardabweichungen bei Daten, bei denen keine a-priori-Standards existieren) zu identifizieren.

Kriterium der Organisation

Ein Organisationskriterium ist die Anzahl der Dubletten. Unter Dubletten versteht man hier zwei oder mehr Datensätze, die sich mit hinreichender Wahrscheinlichkeit auf denselben Fall beziehen (Nonnemacher et al. 2007: 29; Batini et al. 2006: 24; Winkler 2004). Hier könnten durch bisher weitgehend ungenutzte Techniken der „fuzzy searches" im Data Mining auch nicht-exakte Dubletten gefunden werden (Supekar et al. 2002), die eventuell Rückschlüsse auf Manipulationen – beispielsweise das „Auffüllen" von Problemdaten mit unverdächtigen Daten aus anderen Fällen – zulassen würden. In der leistungserbringerübergreifenden Prüfung könnte dies bei widersprüchlichen Daten zwischen zwei Datenlieferanten sogar den Fehlerverursacher „identifizieren" (Pseudonym), weil dieser annähernde Dubletten liefern würde.

9.2.2 Bewertung der Prüfergebnisse

Zur Bewertung der Ergebnisse schlagen Nonnemacher et al. (2007: 57) ein gleichsam sehr elaboriertes und starres Score-Schema vor. Dieses sieht Schwellenwerte vor, die auf Empfehlungen des Instituts für Medizinische Informatik, Biometrie und Epidemiologie (IMIBE) basieren. Allerdings sind diese Schwellenwerte keinesfalls international bewährt. Der NHS arbeitet beispielsweise im Fall der Vollzähligkeit mit vier Qualitätsstufen, die durch die Perzentile bei 97 %, 90 % und 80 % Vollzähligkeit definiert sind (NHS 2007: 14), während die schwedische Gesundheits- und Sozialbehörde für zwei wichtige Datenbanken Vollzähligkeiten von 89 % und 92 % angibt (Socialstyrelsen 2010: 4). Insofern ist die vom IMIBE vorgeschlagene Marke von 90 % durchaus plausibel. Aufgrund der organisatorisch sehr unterschiedlichen Strukturen der verschiedenen Gesundheitssysteme empfiehlt das Institut dennoch zunächst ein Benchmarking im ersten Jahr durchzuführen, anstatt derartige Werte 1:1 auf das deutsche Gesundheitssystem zu übertragen.

Noch problematischer ist die Gewichtung von Prüfergebnissen aus der Anwendung unterschiedlicher Kriterien. Keine der betrachteten Datenbanken nimmt eine derartige Gewichtung vor, sondern bewertet die Ergebnisse anhand der einzelnen Kriterien separat. Da es mit zusammengeführten Maßen bezüglich der Dokumentationsqualität noch keine ausreichenden Erfahrungen gibt, schlägt das AQUA-Institut deshalb vor, die Auswertung der verschiedenen Kriterien zunächst unabhängig voneinander durchzuführen und dabei auf allgemein übliche Auffälligkeitsmerkmale, beispielsweise Abweichung um mehr als zwei Standardabweichungen, zurückzugreifen. Mit Blick auf den notwendigen Aufwand werden die konkreten Bedingungen verfahrensspezifisch zu entscheiden sein.

9.2.3 Datenvalidierung

Im Unterschied zu kontinuierlich stattfindenden Datenprüfungen bei der Dateneingabe oder im Zuge des Datenaustauschs beschreiben Datenvalidierungsverfahren im engeren Sinne Prüfungen, die auf Grundlage zusammengeführter Daten eines definierten Zeitraums durchgeführt werden. Innerhalb des Datenvalidierungsverfahrens werden statistische Basisprüfung und Stichprobenprüfung unterschieden. Erstere beziehen sich auf den gesamten Datenpool ausgewählter QS-Verfahren, Letztere auf eine Auswahl von Leistungserbringern und/oder Patienten innerhalb ausgewählter QS-Verfahren.

Um die Dokumentationsqualität der über den Leistungserbringer dokumentierten QS-Daten zu prüfen, geht das AQUA-Institut grundsätzlich in zwei Schritten vor: einer *statistischen Basisprüfung* folgt eine *Stichprobenprüfung*.

Statistische Basisprüfung

Die statistische Basisprüfung hat zum Ziel, jene rechnerischen Auffälligkeiten zu identifizieren, die auf eine fehlerhafte Dokumentation hinweisen. Datenbasis ist der Bundesdatenpool für ein ausgewähltes QS-Verfahren. Um ein transparentes Vorgehen zu gewährleisten, sollen systematische Auffälligkeitskriterien formuliert werden. Diese sollen so konstruiert sein, dass Auffälligkeiten auf Ebene der Leistungserbringer bestimmt werden. Diesem Vorgehen unterliegt die Annahme, dass die Dokumentationsqualität zwischen den Leistungserbringern variiert und sich eine mangelnde Dokumentationsqualität folglich auch bei einzelnen Leistungserbringern niederschlägt. Für die rechnerisch auffälligen Leistungserbringer wird ein Stellungnahmeverfahren eingeleitet (vgl. §17 Abs. 2 Qesü-RL), dessen Ziel es ist, Ursachen zu eruieren und gegebenenfalls Verbesserungsmaßnahmen einzuleiten.

Stichprobenprüfung

Die Stichprobenprüfung ist ein Verfahren zur externen Validierung. Sie zielt darauf ab, quantitative Aussagen darüber treffen zu können, inwiefern die Angaben in den Datenfeldern der QS-Dokumentation die tatsächliche Versorgung abbilden. Um dies zu ermöglichen, wird ein externes Kriterium verwendet, von dem angenommen wird, dass es die reale Versorgungsqualität besser abbildet als die QS-Dokumentation. Sowohl in der externen stationären Qualitätssicherung als auch in der ambulanten Qualitätssicherung ist es üblich, dafür Angaben aus der Patientenakte heranzuziehen. Insbesondere können hier auch objektive Messwerte, wie sie z.B. aus Laborbefunden und bildgebenden Verfahren generiert werden, genutzt werden. Für eine Stichprobe von Leistungserbringern und Patienten, zu denen es eine QS-Dokumentation gibt, erfolgt dann eine Zweiterfassung ausgewählter Datenfelder der QS-Dokumentation anhand der Patientenakte. Anschließend werden die fallbezogenen Angaben aus der Patientenakte mit denen aus der QS-Dokumentation abgeglichen.

Zentrales Prüfkriterium der Stichprobenprüfung ist die *Konkordanz*, das einzige Prüfkriterium, das aufgrund seiner Beschaffenheit *ausschließlich* in der Stichprobenprüfung validiert werden kann und das gleichzeitig auch das wichtigste Kriterium der Akkuratesse ist. Außerdem können im Rahmen der Stichprobenprüfung ebenfalls mögliche Dubletten überprüft und fehlende Werte nachgetragen werden.

Der Umfang der Stichprobe hängt dabei von der ausgewählten Verfahrensart ab. In den Vorgaben zum Konzept der sektorenübergreifenden Sicherung der Dokumentationsqualität hat der G-BA beispielsweise festgelegt, dass 5 % der Leistungserbringer in die Stichprobenprüfung einbezogen werden sollen. Dabei soll das Konzept einer einrichtungsbezogenen Klumpenstichprobe angewendet werden. Der Hauptvorteil einer solchen Vorgehensweise ist die Kostenersparnis gegenüber einer patientenbezogenen einfachen Zufallsstichprobe, da wesentlich weniger Einrichtungen besucht werden müssen (Singh 2003: 765).

10 Vorbereitung der Umsetzung von Entwicklungsleistungen

In den vorangegangenen Kapiteln wurden methodische Vorgehensweisen dargelegt, wie sie im Allgemeinen bei Beauftragungen des AQUA-Instituts zur Entwicklung von Indikatoren, Instrumenten und der notwendigen Dokumentation durch den G-BA (vgl. Abschnitt 2.3.2) angewendet werden. Die Abschlussberichte zu solchen Entwicklungsleistungen liefern wesentliche fachliche Grundlagen, damit der G-BA Richtlinien bzw. themenspezifische Bestimmungen erstellen kann (vgl. Abschnitt 2.3.3), um so die Basis für den Regelbetrieb eines QS-Verfahrens zu schaffen.

Da die Methoden einem stetigen Wandel bzw. der Weiterentwicklung unterliegen, kann es sein, dass einzelne Bestandteile im Zuge einer Beauftragung (noch) nicht umgesetzt werden konnten. Dies gilt insbesondere bezüglich der Methoden zur Einbindung von Routinedaten und zur Entwicklung von Patientenbefragungen. In solchen Fällen werden diese Schritte zumeist nachträglich umgesetzt. Beispiele hierfür sind die Entwicklung einer Patientenbefragung für das Thema PCI[30] oder die empirischen Prüfungen mit Routinedaten für die entwickelten Qualitätsindikatoren der Themen „Vermeidung nosokomialer Infektionen: Postoperative Wundinfektionen“ und „Vermeidung nosokomialer Infektionen: Gefäßkatheter-assoziierte Infektionen“[31].

Sofern eine Erhebung der Daten über Einrichtungen (QS-Dokumentation) oder Patienten (Patientenbefragungen) geplant ist, besteht vor Einführung eines Regelbetriebs zumeist Klärungsbedarf dahingehend, ob die Erhebungsinstrumente in der gewünschten Art und Weise funktionieren und valide Aussagen liefern. Dies kann in Form von spezifischen Erprobungsprojekten (QS-Dokumentation) oder Validierungsstudien (Patientenbefragung) beauftragt werden. Bei einer angestrebten Verwendung von Sozialdaten können empirische Überprüfungen anhand von anonymisierten Abrechnungsdaten bzw. Routinedaten von Krankenkassen erfolgen.

Weil im Prozess der Beschlussfassung durch den G-BA noch Abweichungen von den Empfehlungen des AQUA-Instituts auftreten können, ist die informationstechnische Aufbereitung der Datenerhebungen in Form von Spezifikationen erst dann vorgesehen, wenn ein entsprechender Beschluss des G-BA zur Umsetzung eines QS-Verfahrens vorliegt.

10.1 Informationstechnische Aufbereitung

Gegenstand der informationstechnischen Aufbereitung ist die Übertragung der im Rahmen der Indikatorenentwicklung vorgeschlagenen Erhebungsinstrumente bzw. der für die QS zu übermittelnden Informationen in eine Spezifikation, d.h. in eine Form, die verständlich und eindeutig für die Softwareentwickler ist, die die konkrete Software für die Leistungserbringer oder Krankenkassen erstellen.

Die Spezifikation ist die Gesamtheit aller Vorgaben, nach denen die Sollfestlegung eines QS-Verfahrens bzw. die Selektion der benötigten Daten und der Export sowie die Übermittlung der Daten erfolgen soll. Sie wird jeweils als Spezifikationspaket zum Download zur Verfügung gestellt, das aus einzelnen Komponenten besteht. Komponenten können dabei die Spezifikationsdatenbanken, die unter anderem die exakte Definition von Datenfeldern, Ausfüllhinweise für die Anwender und Hilfestellungen bezüglich der Auslösekriterien für die Dokumentationspflicht bzw. Definitionen zur Datenselektion enthalten, eine zielgruppengerichtete technische Dokumentation, XML-Schemata, Verschlüsselungsprogramme und ein Datenprüfprogramm sein.

Jede Komponente dieser Spezifikationspakete kann unabhängig und parallel zu anderen aktualisiert werden.

[30] „Beschluss des Gemeinsamen Bundesausschusses über eine Beauftragung der Institution nach §137a SGB V: Entwicklung von Patientenbefragungen im Rahmen der Qualitätssicherungsverfahren Perkutane Koronarintervention und Koronarangiographie (PCI) und Arthroskopie am Kniegelenk“ vom 21. März 2013. Vgl.: https://www.g-ba.de/downloads/39-261-1682/2013-03-21_AQUA_Patientenbefragung-PCI-ASK.pdf (Zugriff am 13.02.2015).

[31] Beauftragungen vom 19. Dezember 2013 und 20. März 2014 (die Beschlüsse dazu sind nicht veröffentlicht).

Gegenwärtig werden zwei Arten von Spezifikationen, entsprechend den verwendeten Erhebungsinstrumenten, unterschieden:

- Spezifikation für die QS-Dokumentation
- Spezifikation für Sozialdaten bei den Krankenkassen

Patientenbefragungen sind zzt. noch nicht spezifiziert, da die exakten Datenflüsse und Verantwortlichkeiten für eine Umsetzung in Regelbetrieben noch nicht festgelegt sind.

10.1.1 Spezifikation für die QS-Dokumentation

Gemäß §13 Abs. 2 der Richtlinie zur einrichtungs- und sektorenübergreifenden Qualitätssicherung (Qesü-RL) werden die zu erstellenden Anforderungen und Datenschnittstellen mit den Beteiligten abgestimmt. Die derzeit bestehenden Regelungen der QSKH-RL beziehen sich vorrangig auf die QS-Dokumentation durch Leistungserbringer, weil die anderen für die Qualitätssicherung avisierten Erhebungsinstrumente noch in Entwicklung sind. In die Qesü-RL wurden im Jahr 2014 die Ergänzungen für die Nutzung von Sozialdaten bereits aufgenommen.

Für die Spezifikation für die Leistungsbereiche der externen stationären Qualitätssicherung ist in Anlage 1 der QSKH-RL Folgendes festgelegt:

> *Die spezifische Darstellung der einbezogenen Leistungen ist definiert durch die Listen der Ein- und/oder Ausschlusskriterien in der Spezifikation für QS-Filter-Software. Die spezifische Darstellung der Dokumentationsinhalte ist definiert durch die Anforderungen der Spezifikation für QS-Dokumentationssoftware. Diese werden durch den G-BA beschlossen und in der jeweils geltenden Fassung durch die Institution nach §137a SGB V auf ihrer Homepage im Internet öffentlich bekannt gemacht.*

Die Veröffentlichung dieser Spezifikation erfolgt auf der SQG-Webseite im Bereich Datenservice. Aktuell wird die Spezifikation für die stationäre Qualitätssicherung einmal jährlich geändert und jeweils Ende Juni des Jahres vor der Datenerfassung veröffentlicht. Ein Service-Release hierzu erfolgt dann im Herbst, nachdem das Deutsche Institut für Medizinische Dokumentation und Information (DIMDI) die offiziellen ICD- und OPS-Schlüssel für das Folgejahr herausgegeben hat.

Regelungen zur Erstellung von Spezifikationen für die sektorenübergreifende Qualitätssicherung sind in §13 Abs. 2 der Qesü-RL festgehalten. Dort heißt es:

> *[1]Der G-BA legt in den themenspezifischen Bestimmungen jeweils die erforderlichen Daten fest. [2]Er gibt die Softwarespezifikationen für die Erfassung der Daten nach §14 vor. [3]Dabei kann er sich der Institution nach §137a SGB V oder einer anderen Stelle auf Bundesebene bedienen. [4]Die jeweils beauftragte Stelle auf Bundesebene bezieht zur Definition der erforderlichen Datenschnittstellen und Anforderungen an die Dokumentationssoftware Vertreterinnen und Vertreter aller Parteien ein, die durch den Datenfluss berührt sind. [5]Dies sind insbesondere die KBV, die KZBV, die DKG, der GKV-SV, die PKV, die LAGen und Anbieter von Praxis- und Krankenhaussoftware. [6]Im Rahmen der nach Satz 2 vorgegebenen Spezifikation übernehmen die KBV für den vertragsärztlichen Bereich und die KZBV für den vertragszahnärztlichen Bereich die Softwarezertifizierung innerhalb von zwei Monaten nach Antragsstellung. [7]Die Institution nach §137a SGB V erstellt und pflegt ein Datenprüfprogramm gemäß §4 Abs. 2 und macht es öffentlich zugänglich.*

Voraussichtlich Ende Februar 2015 wird erstmals ein Qesü-Modul veröffentlicht werden. Weil es sich um die erstmalige Umsetzung einer sektorenübergreifenden Spezifikation handelt, soll den Beteiligten mehr Zeit zur Umsetzung gegeben werden. Künftig ist vorgesehen, auch die Qesü-Module Ende Juni eines jeden Jahres zu veröffentlichen.

Mit der Einführung von Datenübermittlungen mit personenidentifizierenden Merkmalen (PID) sind die Spezifikationen in der stationären und der sektorenübergreifenden Qualitätssicherung durch komplexere Datenflüsse, höhere Anforderungen an den Datenschutz (vgl. §299 SGB V) und die Notwendigkeit eines Datenprüfpro-

gramms gekennzeichnet. Abbildung 12 stellt das serielle Datenflussmodell dar, welches für die sektorenübergreifende Qualitätssicherung angewendet werden soll. Es zeigt, dass in die Datenflüsse der sektorenübergreifenden Qualitätssicherung verschiedenste Institutionen eingebunden sind. Gänzlich neu ist beispielsweise die Vertrauensstelle, die einerseits patientenbezogene Informationen (lebenslange Versichertennummer) pseudonymisieren und andererseits als Datenannahmestelle für den ambulant-selektivvertraglichen Bereich fungieren soll.

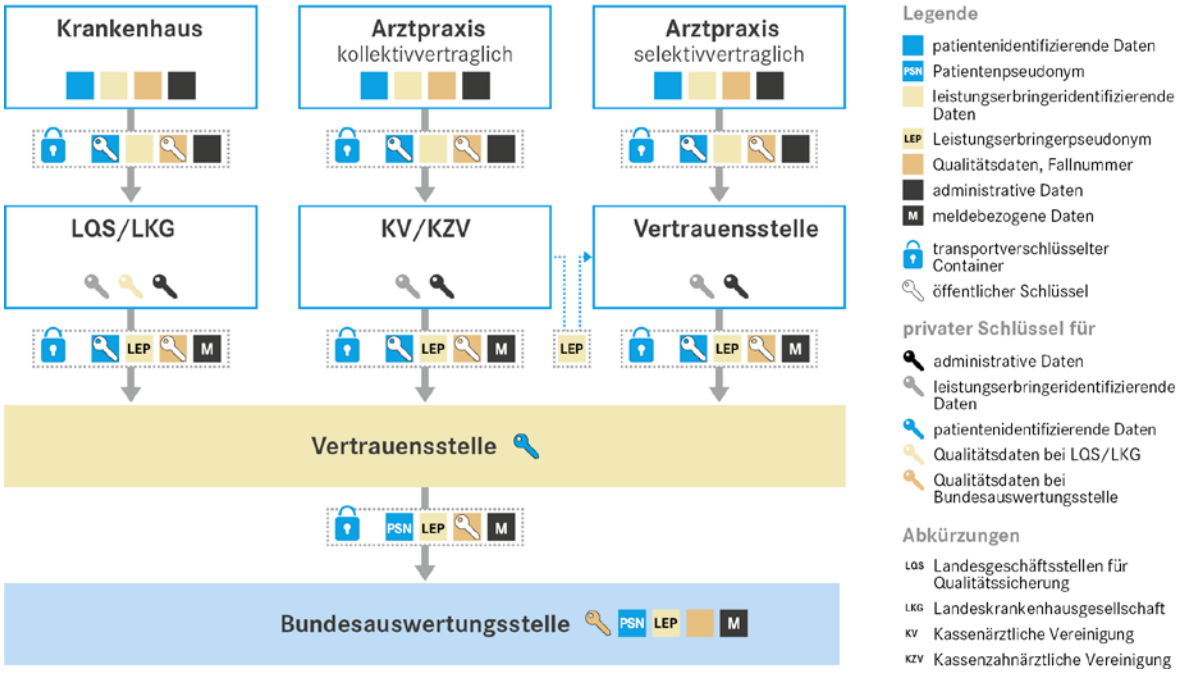

Abbildung 12: Serielles Datenflussmodell für die sektorenübergreifende Qualitätssicherung

Jede der in die Datenflüsse der sektorenübergreifenden Qualitätssicherung eingebundenen Institutionen darf nur die für sie relevanten Informationen einsehen. Hierfür ist eine differenzierte Verschlüsselung unterschiedlicher Inhalte der Datenlieferungen notwendig. Weiterhin ist es möglich, dass ein Leistungserbringer je nach Abrechnungskontext unterschiedliche Datenannahmestellen beliefern muss. Auch hierdurch erhöht sich die Komplexität der Datenflüsse enorm. Aus diesen Gründen wird die Spezifikation im „Baukastenprinzip“ aus den oben erwähnten Komponenten erstellt, die jeweils auf die spezifischen Stellen im Datenfluss zugeschnitten sind. Seit dem Erfassungsjahr 2014 ist die Spezifikation zudem als „Basisspezifikation“ konzipiert, die die Regelungsbereiche beider Richtlinien (QSKH-RL und Qesü-RL) und beider Erhebungsinstrumente abdeckt. Abbildung 13 stellt die einzelnen Komponenten übersichtlich dar.

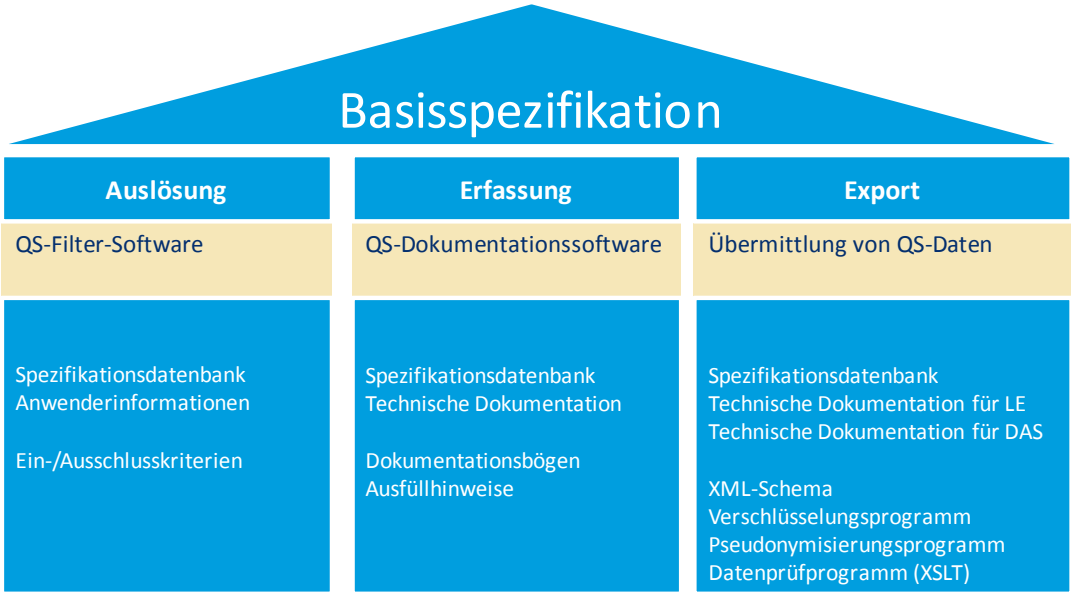

Abbildung 13: Bereiche der Spezifikation der sektorenübergreifenden Qualitätssicherung

Gemäß diesem Modell bildet die Basispezifikation die drei Hauptprozesse der QS-Dokumentation ab (AQUA 2014g):

- Auslösung (Spezifikation für den QS-Filter)
- Erfassung (Spezifikation für die QS-Dokumentation)
- Datenexport und -übermittlung

Die Spezifikation für den QS-Filter enthält alle Regeln für die Auslösung einer Dokumentation und die Übermittlung der Vollzähligkeit. Die Spezifikation für die QS-Dokumentationssoftware enthält alle Regeln für die Datenerfassung in den Einrichtungen. Die Spezifikation für den Datenexport enthält alle Regeln für den Datenexport und die Datenübermittlungen zwischen Leistungserbringer (LE), Datenannahmestelle (DAS), Vertrauensstelle (VST) und Bundesauswertungsstelle (BAS), inklusive der Rückprotokollierung und der Übermittlung der Auswertungsberichte. Für die Durchführung der Prozesse Datenexport und -übermittlung enthält die Spezifikation eine Reihe von weiteren Hilfskomponenten, wie die XML-Schemata (XML: Extensible Markup Language) oder Datenprüf- und Verschlüsselungsprogramme.

XML als Metasprache im Rahmen der Datenflüsse

Für die Gestaltung der Datenflüsse in der sektorenübergreifenden Qualitätssicherung dient XML (W3C 2008) als eine der wesentlichen technischen Grundlagen für die Spezifikation. In XML werden komplexe Datenbeziehungen durch die Datenstruktur übersichtlich und schnell nachvollziehbar dargestellt. Die Bezeichnung der Daten erfolgt nicht über Tabellen- und Feld-Definitionen, sondern durch XML-Elemente, sog. <Tags>, direkt in den Daten selbst (Vonhoegen 2009).

Statt eine Tabelle mit Vorgangsnummer, Aufnahmedatum, Entlassungsdatum zu definieren und folgendes CSV auszugeben:

```
123456,2010-01-01,2010-02-01
123457,2010-01-02,2010-02-02
123458,2010-01-03,2010-02-03
```

sieht dieses in XML folgendermaßen aus:

```
<datensätze Anzahl=„3“>
   <datensatz vorgangsnummer=„123456“>
      <Aufnahmedatum>2010-01-01</Aufnahmedatum>
      <Entlassungsdatum>2010-02-01</Entlassungsdatum>
   </datensatz>
   <datensatz vorgangsnummer=„123457“>
```

```
      <Aufnahmedatum>2010-01-02</Aufnahmedatum>
      <Entlassungsdatum>2010-02-02</Entlassungsdatum>
   </datensatz>
   <datensatz vorgangsnummer=„123458“>
      <Aufnahmedatum>2010-01-03</Aufnahmedatum>
      <Entlassungsdatum>2010-02-03</Entlassungsdatum>
   </datensatz>
</datensätze>
```

Während dieselben Inhalte in XML zwar aufgrund der redundanten Bezeichner-Informationen deutlich mehr Speicherplatz erfordern, wird die Lesbarkeit erhöht. Die Datenstruktur ist mit den Inhalten verbunden. Je komplexer die Datenstrukturen der Erfassungsbögen werden (beispielsweise PCI mit vier Teildatensätzen), umso lesbarer gestalten sie sich in XML im Vergleich zu CSV. Zusätzlich schafft dies die Möglichkeit, mehrere Dateien für einen Basisbogen in Form einer hierarchischen Baumstruktur in einer Datei zusammenzufassen, anstatt diese und die dazugehörigen Teildatensätze zu exportieren:

```
<datensätze Anzahl=„3“>
   <datensatz vorgangsnummer=„123456“>
      <Aufnahmedatum>2010-01-01</Aufnahmedatum>
      <Entlassungsdatum>2010-02-01</Entlassungsdatum>
      <tds name=„diagnosen“>
         <diagnose>Code1</diagnose>
         <diagnose>Code2</diagnose>
         <diagnose>Code3</diagnose>
      </tds>
   </datensatz>
</datensätze>
```

Der prinzipielle Vorteil von XML als Metasprache ist, dass Standardtechnologien und frei verfügbare Programmen genutzt werden können. Zum Beispiel bietet XML verschiedene Möglichkeiten, die Validität von Daten zu überprüfen und Bedingungen für diese vorzugeben. Neben der „Wohlgeformtheit“ von XML (dem Einhalten grundlegender syntaktischer Vorgaben) kann mit einem XML-Schema vorgegeben werden, welchem Aufbau das XML-Dokument zu folgen hat, um im Sinne des Schemas valide zu sein (Sebastian 2009).

Das vom World Wide Web Consortium herausgegebene XML-Schema beschreibt in einer standardisierten Schemasprache neben der Datenstruktur alle Datentypen der einzelnen Felder und Bögen. Um Abweichungen zwischen dem XML-Schema und der Spezifikationsdatenbank zu vermeiden, werden die XML-Schemata automatisch aus der zentralen Spezifikationsdatenbank generiert.

Ein XML-Dokument, welches validiert werden soll, gibt dieses Schema in seinem Root-Element bekannt. Im Beispiel oben ist das Root-Element das Element <datensätze>. Hier würde z.B. stehen:

```
<datensätze Anzahl=„3“ xmlns:xsi=„http://www.w3.org/2001/XMLSchema-instance“
xsi:schemaLocation="urn:gba:sqg interface_le.xsd>
```

Damit wüsste die validierende Instanz sofort, wo die Schemainformationen zu finden sind. Hier nämlich in der Schemadatei „interface_le.xsd“, die im selben Verzeichnis wie die XML-Datei liegt.

In dieser Schemadatei finden sich nun Informationen darüber, wie ein zulässiger Aufbau der XML-Datei aussehen darf. Dabei wird eine Struktur und Typisierung der Daten vorgegeben. Nur wenn diese Vorgaben erfüllt sind, ist der XML-Datensatz valide (Harold et al. 2004). Ein Schema für den obigen XML-Datensatz würde in der Beispielansicht in etwa aussehen wie in Abbildung 14.

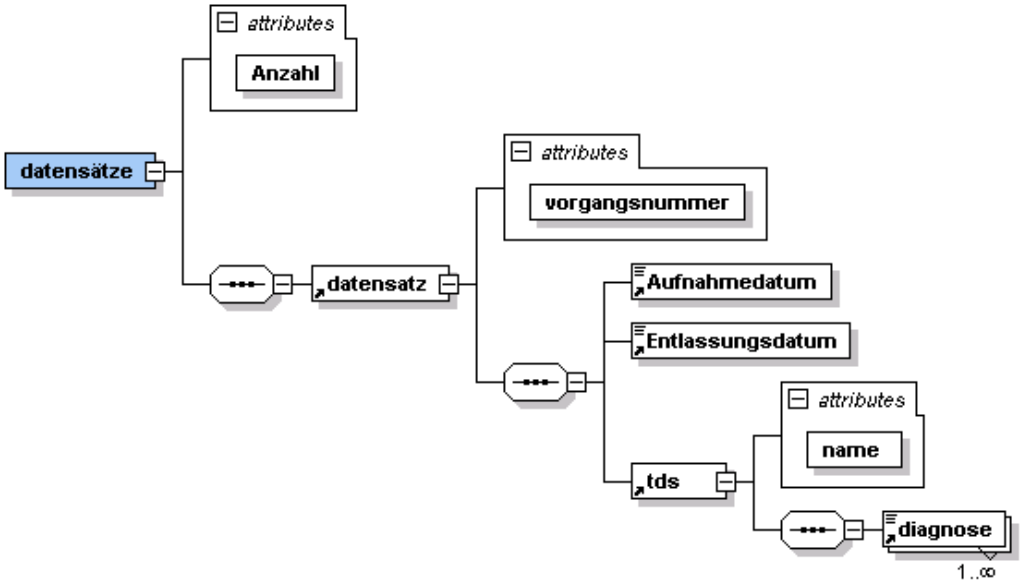

Abbildung 14: Beispielansicht eines XML-Schemas

Weitergehende Möglichkeiten zur Datenprüfung ergeben sich durch die Verwendung von XSLT (Extensible Stylesheet Language Transformation) (Kay 2008). Mittels XSLT-Skripts können inhaltliche, in der Spezifikation für die Datenerfassungen festgelegte Plausibilitätsregeln vor Ort durch ein Datenprüfprogramm angewendet werden. Die Nutzung dieser Technologie erfordert demnach keine separate Prüfung durch die eingesetzte Dokumentationssoftware. Es reicht allein eine dem Schema entsprechende Formatierung der XML-Daten aus.

Datenprüfprogramm

Aufgrund der Anforderung der Qesü-RL, bereits vor Ort beim Leistungserbringer Datenprüfungen durchzuführen, wurde ein Prüfprogramm eingeführt. Dieses Prüfprogramm bezieht die Plausibilitätsregeln direkt aus der Spezifikationsdatenbank und testet die zu übermittelnden Datensätze vor deren Verschlüsselung. Es kann sowohl von den Datenannahmestellen in ihre Datenservices integriert als auch von den Softwareanbietern für die Durchführung des Softwaretests verwendet werden.

Die eigentliche Prüfung erfolgt auf der XML-Exportdatei, deren Struktur der entsprechenden Dokumentation bzw. dem gültigen XML-Schema zu entnehmen ist (AQUA 2014g).

Verschlüsselungsprogramme

Aufgrund der Anforderungen der Qesü-RL, mehrere Arten von Daten (QS-Daten, personenidentifizierende Daten, administrative Daten) gezielt an verschiedene Akteure im Gesamtprozess (Leistungserbringer, Datenannahmestellen, Vertrauensstelle, Bundesauswertungsstelle) zu übermitteln, wurde ein Verschlüsselungsprogramm entwickelt, das alle Datensicherheitsanforderungen erfüllt. Das AQUA-Institut stellt ein Ver- und Entschlüsselungspaket für die Anwendung in der externen Qualitätssicherung entsprechend QSKH- und Qesü-RL zur Verfügung. Die Ver- und Entschlüsselungsalgorithmen sind entsprechend ausgewählter W3C-Standards implementiert. Das bedeutet, dass unabhängig von der Implementierung des AQUA-Instituts an jeder beliebigen Stelle im Workflow auch jede andere W3C-Standard-Implementierung verwendet werden kann.

Das Ver- und Entschlüsselungspaket besteht aus den Programmen TPacker für die Transportverschlüsselung und XPacker für die XML-Verschlüsselung. Eine interaktive Alternative zur Verwendung dieser Programme stellt der GPacker dar. Er integriert den XPacker und TPacker und ermöglicht mithilfe einer einfachen grafischen Oberfläche für die Endanwender die Verschlüsselung und Komprimierung von QS-Dateien (Abbildung *15*).

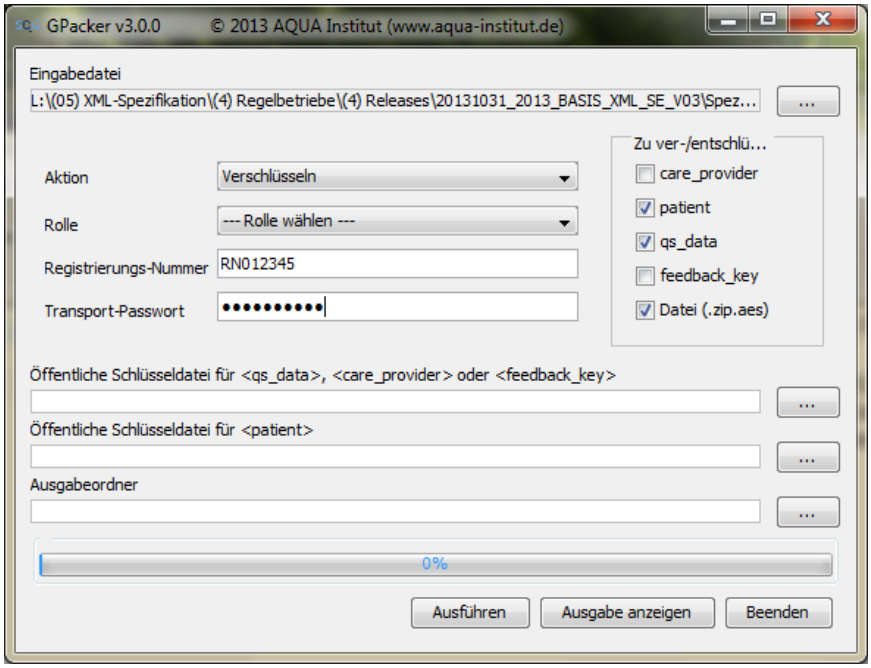

Abbildung 15: GPacker, interaktive Oberfläche zur Verschlüsselung der XML-Daten

Verschlüsselung

Das vom AQUA-Institut entwickelte Verschlüsselungsprogramm verwendet den W3C-Standard „XML Encryption Syntax and Processing“ (XML-Encryption). XML-Encryption bietet eine adäquate Verschlüsselungslösung an, um komplexe Anforderungen an die Sicherheit des Datenaustauschs zu erfüllen: Mit XML-Encryption können unterschiedliche Dokumentabschnitte (XML-Knoten) für unterschiedliche Datenempfänger mit unterschiedlichen Schlüsseln verschlüsselt werden, anstatt diese Dokumentabschnitte, die logisch miteinander verbunden sind, in verschiedenen Dokumenten an die jeweiligen Datenempfänger senden zu müssen. Auf diese Weise bleibt der Zusammenhang der Daten erhalten.

Für die Verschlüsselung der XML-Knoten wird die hybride Verschlüsselung nach XML Encryption verwendet, da diese die Vorteile symmetrischer und asymmetrischer Verschlüsselungsverfahren –Schnelligkeit und Sicherheit – kombiniert. Als Verschlüsselungsalgorithmen werden „AES128“ für die symmetrische Verschlüsselung der XML-Elemente und „RSA mit 2048-Bit“ für die asymmetrische Verschlüsselung des generierten symmetrischen Schlüssels verwendet.

Abschließend wird für die Datenübermittlung noch eine Transportverschlüsselung benötigt, welche die gesamte Datei verschlüsselt und vor unberechtigten Zugriffen schützt. Die Transportverschlüsselung ist nur dann notwendig, wenn kein sicherer Übertragungskanal (z.B. VPN) genutzt wird, also insbesondere bei der Übermittlung der Datei als E-Mail-Anhang.

10.1.2 Spezifikation für Sozialdaten bei den Krankenkassen

Nach §299 Absatz 1a SGB V sind die Krankenkassen befugt und verpflichtet, einen zweckgebundenen Ausschnitt aus den von ihnen gem. §284 SGB V erhobenen Daten („Sozialdaten bei den Krankenkassen“) für die gesetzliche Qualitätssicherung zur Verfügung zu stellen (AQUA 2014b; AQUA 2013a).

Die technische Umsetzung der Anforderungen an die Verfahren hinsichtlich Datenselektion, Plausibilitätsprüfung und Datenübermittlung werden analog zur Basisspezifikation in einer Spezifikationsdatenbank definiert. Sie enthält zusammen mit der technischen Dokumentation alle Informationen, die die Krankenkassen für die Selektion und Übermittlung der Sozialdaten brauchen. Für bestimmte, prozessspezifische Anwendungen werden weitere Anwendersichten automatisch aus der Spezifikationsdatenbank generiert und bereitgestellt (Abbildung 16). Diese verfahrensübergreifenden Informationen bilden den allgemeinen Teil der Spezifikation.

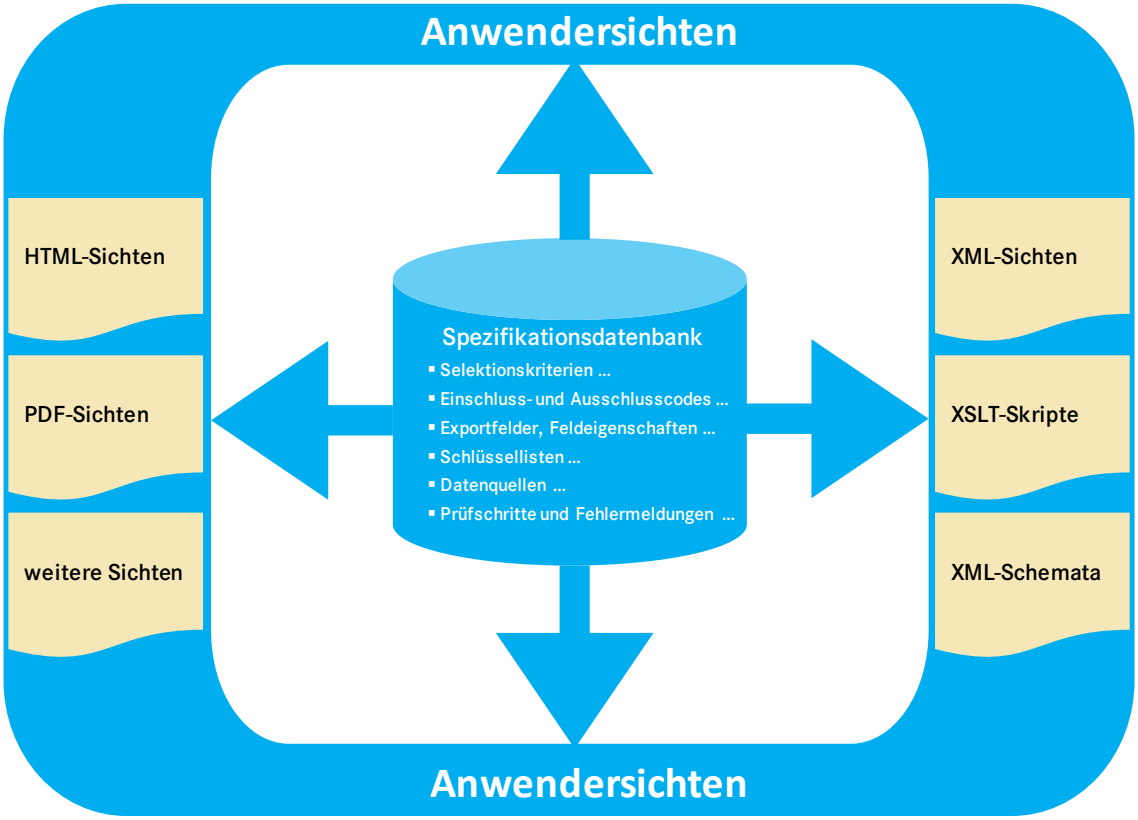

Abbildung 16: Spezifikationsdatenbank und die Ableitung der Anwendersichten

Darüber hinaus kann die Spezifikationsdatenbank um spezifische Angaben zu bestimmten QS-Verfahren erweitert werden, die in einzelne Module gegliedert sind und für die ebenfalls Anwendersichten generiert werden können.

Datenflüsse der Sozialdaten

Für die Sozialdaten bei den Krankenkassen wurde der Qesü-RL ein weiterer Datenfluss hinzugefügt. Dieser entspricht in seinen Grundzügen dem Datenfluss der von den Leistungserbringern übermittelten Daten. Abweichende technische Festlegungen werden in der Technischen Dokumentation beschrieben. Die Datenflüsse unterscheiden sich im Wesentlichen hinsichtlich der zuständigen Datenannahmestelle, der technischen Beschaffenheit der Daten und deren Transport- und Verschlüsselungsverfahren. Die Einbindung einer pseudonymisierenden Vertrauensstelle ist nur in Verfahren notwendig, welche PID bzw. die daraus erzeugten Pseudonyme zur Zusammenführung von Datensätzen benötigen. Bei Verfahren ohne PID erfolgt der Transport der QS-Daten direkt von der Datenannahmestelle für die Krankenkassen (DAS-KK) zur Bundesauswertungsstelle (BAS). Sozialdaten bei den Krankenkassen werden auch für die QS im stationären Bereich (esQS) zum Einsatz kommen, sodass aktuell entsprechende Anpassungen der QSKH-RL in Vorbereitung sind.

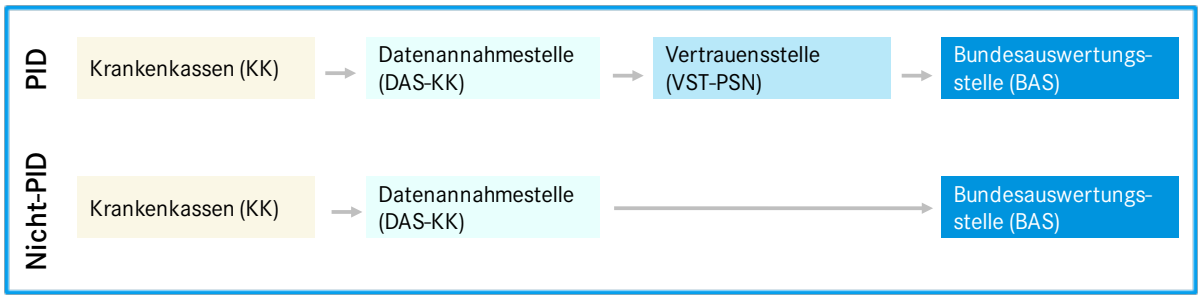

Abbildung 17: Datenfluss bei den Krankenkassen im Rahmen der Qesü-RL bzw. QSKH-RL

Für die Datenflüsse in der sektorenübergreifenden QS wurden weitere Details bereits per Richtlinie beschlossen, siehe Abb. Abbildung 18.

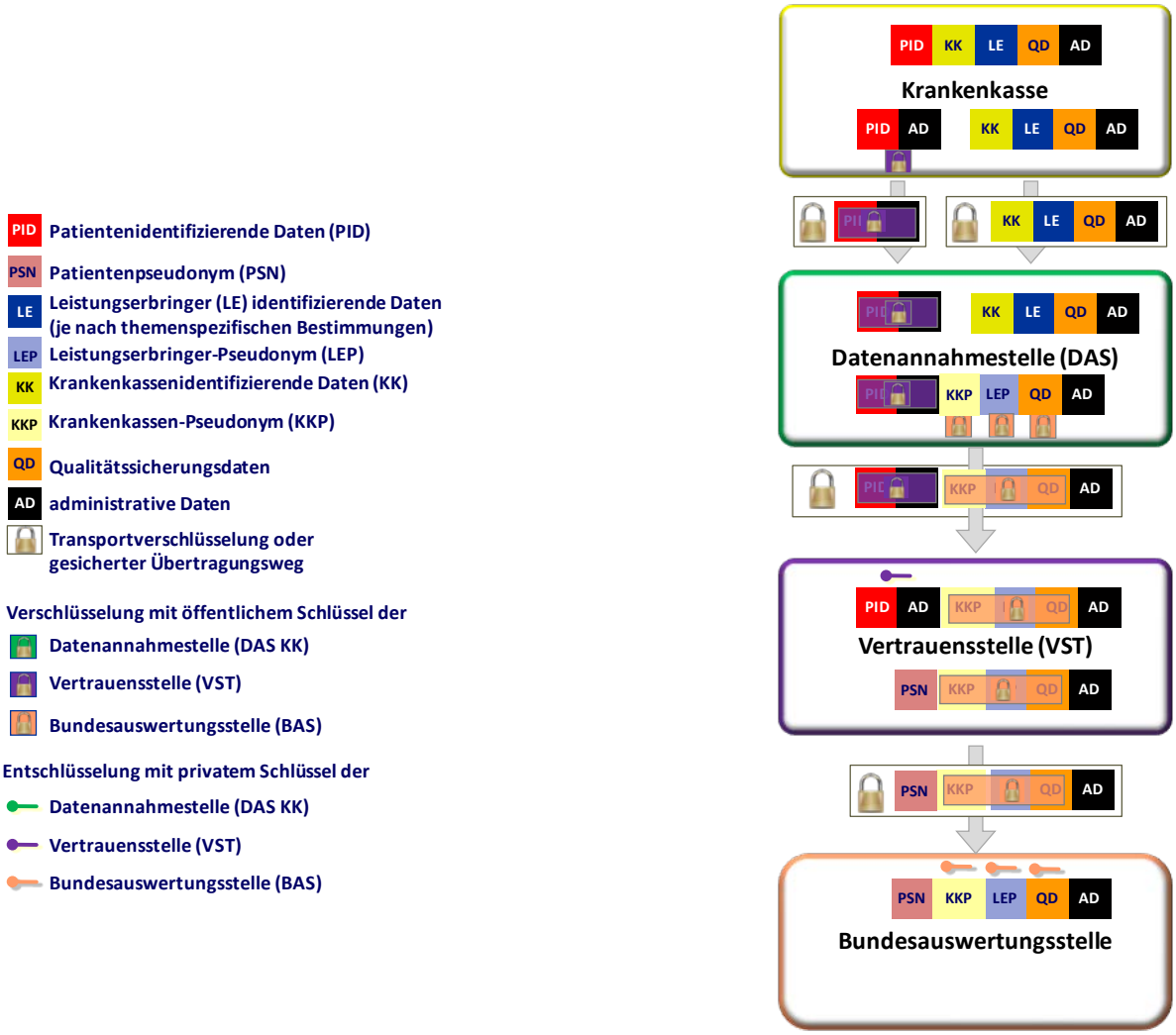

Abbildung 18: Serielles Datenflussmodell Sozialdaten bei den Krankenkassen (Qesü-RL, Anlage zu Teil 1, Abb. 2)

Prozesse bei den Krankenkassen

Im folgenden Abschnitt werden die Prozesse beschrieben, die notwendig sind, um aus den Sozialdaten bei den Krankenkassen nach Vorgaben des G-BA die Daten zu selektieren und anschließend an die Institution nach §137a SGB V zu übermitteln.

Die Prozesse können in folgendem seriellen Modell dargestellt werden.[32]

[32] XSLT ist nicht Teil der allgemeinen Spezifikation, sondern wird verfahrensspezifisch erstellt.

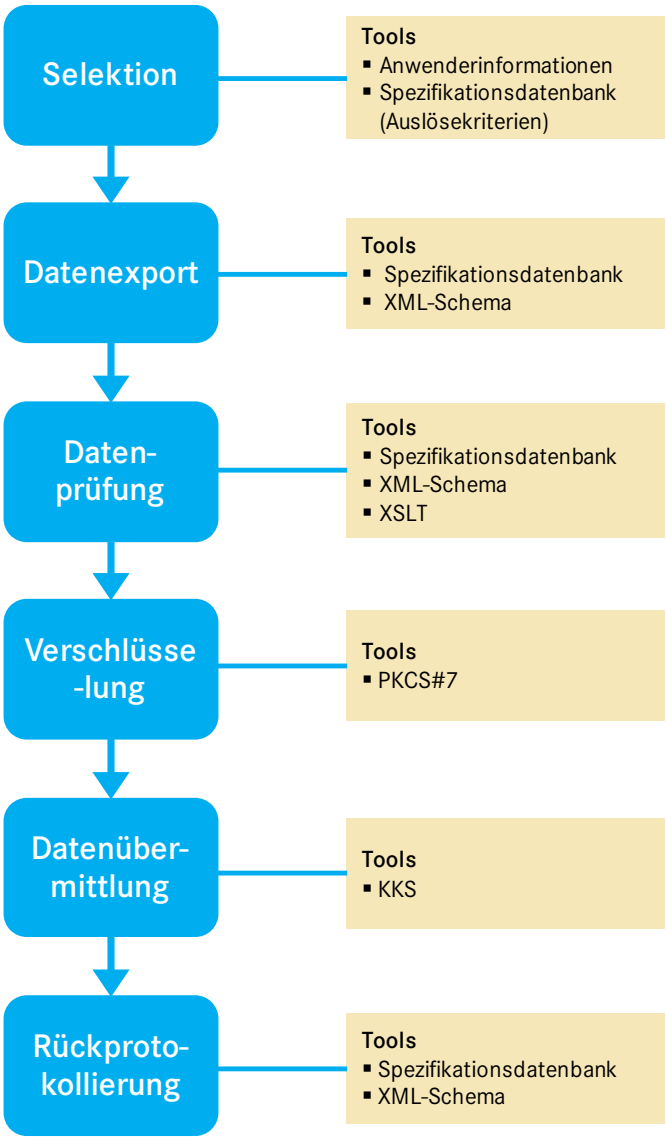

Abbildung 19: Prozesse bei den Krankenkassen

Datenselektion

Dem Grundprinzip der Datensparsamkeit folgend, dürfen gezielt Daten nur solcher Fälle selektiert und übermittelt werden, die für das jeweilige QS-Verfahren von Relevanz sind bzw. für die eine Zweckbestimmung ausgewiesen werden kann.

Um die Sozialdaten annehmen und verarbeiten zu können, muss sichergestellt sein, dass alle Krankenkassen die gleichen Inhalte in gleicher Form übermitteln. Da die Datenbestände der Krankenkassen unterschiedlich strukturiert sind muss die Spezifikation es ermöglichen, die benötigten Datenfelder unabhängig von der konkreten Struktur der Datenbestände einer Krankenkasse eindeutig zu identifizieren und das Format für die weitere Verarbeitung festzulegen.

Bedingt durch die Heterogenität der Datenstruktur der Krankenkassen ist es nicht möglich, im Rahmen der Allgemeinen Spezifikation einen direkten Bezug zu den Datenfeldern herzustellen. Stattdessen wird zur eindeutigen Beschreibung der Datenfelder die Tatsache genutzt, dass in den entsprechenden technischen Anlagen der einzelnen Datenaustauschverfahren zwischen Leistungserbringern (LE) und Krankenkassen alle dort übermittelten Datenfelder (Eingangsdaten) eindeutig benannt und beschrieben werden.[33] An diesem Punkt sind die Datenbestände noch einheitlich, die Unterschiede entstehen erst im weiteren Verarbeitungsprozess, wenn die

[33] Siehe http://www.gkv-datenaustausch.de/leistungserbringer/leistungserbringer.jsp

Krankenkassen die Daten in ihre jeweiligen Systeme einspielen. Diese Beschreibungen des Datenaustauschverfahrens werden in der Allgemeinen Spezifikation für Sozialdaten verwendet. Auf dieser Grundlage ist es den Krankenkassen möglich, den Bezug eines referenzierten Datenfeldes zu ihren eigenen Datenbeständen herzustellen.

Die Selektion von Daten für den Export an die Bundesauswertungsstelle geschieht immer für ein konkretes QS-Verfahren.[34] Die Grundlage der Selektion ist eine Referenztabelle, die für jedes Datenfeld eine eindeutige Kennung und die genaue Herkunft in den Eingangsdaten spezifiziert, also jener Daten, die die Krankenkasse von den LE erhalten hat. Weiterhin wird in dieser Tabelle das genaue Format des Datenfeldes spezifiziert, das für die weitere Verarbeitung erforderlich ist. Gegebenenfalls müssen Daten aus im Datenbestand der Krankenkasse verwendeten Formaten in das spezifizierte Format konvertiert werden. Die Referenztabelle ist in der Spezifikationsdatenbank hinterlegt.

Datenexport

Der Datenexport hat zum Ziel, die im vorherigen Schritt selektierten Daten der Krankenkassen im zuvor von der BAS spezifizierten Format in Exportdateien für die weitere Verarbeitung und Übermittlung abzulegen. Die Datenlieferfristen werden in der QSKH-RL bzw. in den themenspezifischen Bestimmungen eines Verfahrens der Qesü-RL festgelegt.

Um die Daten für die weitere Verarbeitung exportieren zu können, müssen Art, Format und Struktur der zu erzeugenden Dateien klar definiert sein. Wie schon im Zusammenhang mit der Datenselektion bemerkt, kann sich die Spezifikation nur auf die Datenstruktur der von den Leistungserbringern an die Krankenkassen übermittelten Daten (Eingangsdaten) beziehen und nicht auf die individuell verschiedenen Datenspeichersysteme der Krankenkassen. Für die Umsetzung des Datenexports muss die einzelne Krankenkasse in der Lage sein, zu identifizieren, wo und in welcher Form die in ihrem System selektierten Daten innerhalb der Datenstruktur der Exportdateien abzulegen sind. Dazu ist eine Überleitungstabelle (Mapping-Tabelle) notwendig, mit deren Hilfe die Daten der Leistungserbringer zu den Exportdaten übergleitet werden können. Mithilfe dieser Mapping-Tabelle können die einzelnen Datenfelder der Eingangsdaten eindeutig einem entsprechenden Datenfeld in den Exportdateien zugeordnet werden.

Art und Anzahl der Exportdateien in einer Datenlieferung ergeben sich aus den verschiedenen Datenkategorien und ihrem Empfänger im Datenfluss. In der Datenlieferung sind fünf Kategorien zu unterscheiden:

Tabelle 12: Datenkategorien in der Datenlieferung

Datenkategorie	Beschreibung
Admin-Daten/ Kopfdaten	Enthalten Meta-Informationen zum Datenlieferanten, über das Datenpaket, Informationen zur Organisation des Datenflusses und zur Identifizierung der Transaktion. Pro Datei nur einmal vorhanden. Jede Stelle im Datenfluss muss diese Daten einsehen können.
Krankenkassenidentifizierende Daten (KID)	Enthält die Haupt-IK-Nummer der Krankenkasse. Anforderungen an die Pseudonymisierung der KID werden durch die G-BA-Richtlinien festgelegt.
Patientenidentifizierende Daten (PID)	Die PID müssen mit dem öffentlichen Schlüssel der VST verschlüsselt werden. Nur die VST kann diesen Bereich einsehen und ersetzt die PID durch ein (verfahrensspezifisches) Pseudonym.

[34] Definitionen in der Spezifikationsdatenbank sind meist modulbezogen. In der Regel entspricht ein Modul einem bestimmten QS-Verfahren. Es gibt jedoch auch QS-Verfahren, die sich über mehrere Module erstrecken. In diesem Fall ist jedes Modul von den Krankenkassen separat zu bearbeiten. Die Zusammenführung der Module zu einem QS-Verfahren findet erst bei den im Datenfluss folgenden Institutionen statt.

Datenkategorie	Beschreibung
Leistungserbringeridentifizierende Daten (LID)	Die LID müssen gemeinsam mit den QS-Daten mit dem öffentlichen Schlüssel der DAS-KK verschlüsselt werden. Nur die DAS-KK kann diesen Bereich einsehen und ersetzt die LID durch ein Pseudonym.
QS-Daten	Dies sind die Fall- und Versichertendaten ohne PID. Diese Daten dürfen von der DAS-KK und der BAS eingesehen werden. Sie müssen von der absendenden Krankenkasse mit dem öffentlichen Schlüssel der DAS-KK und von der DAS-KK mit dem öffentlichen Schlüssel der BAS verschlüsselt werden.

Zur Datenübertragung zwischen Krankenkassen und der Datenannahmestelle wird das Krankenkassenkommunikationssystem (KKS)[35] verwendet. Es ist ein bei allen gesetzlichen Krankenkassen vorhandener Standard, der die gesetzlichen Vorgaben zur Sicherheit und Datenschutz erfüllt. Das KKS unterscheidet bei jeder Datenübertragung zwischen einem Nutzdatensatz und einem zugehörigen Auftragsdatensatz mit Metainformationen. Für jede Datenübertragung sind also zwei separate Dateien erforderlich.

Als Exportformat wird XML verwendet. Die Struktur der XML-Dateien wird jeweils durch ein XML-Schema aus der Spezifikationsdatenbank eindeutig definiert. Da die Datenbestände der Krankenkassensysteme nicht einheitlich gestaltet sind, kann die Erzeugung der Exportdatei nicht im Detail beschrieben werden und muss individuell durch die Krankenkassen erfolgen. In der Allgemeinen Spezifikation wird stattdessen die Struktur der zu erzeugenden Dateien erläutert. Das Mapping der selektierten Daten in diese Struktur muss jeweils in Abhängigkeit von der Struktur der Datenbestände der Krankenkassen erfolgen. Das Format des Auftragsdatensatzes ist durch das KKS vorgegeben.

Datenprüfung

Bei der Datenprüfung sind drei Arten von Prüfungen zu unterscheiden:

1. Struktur- und Formatprüfungen
   Stellen sicher, dass die zu überprüfenden Felder im richtigen Format und ihre Werte im richtigen Gültigkeitsbereich liegen.
2. Plausibilitätsregeln
   Beziehen sich auf QS- und leistungserbringeridentifizierende Daten und sind verfahrensbezogen. Fehlende oder fehlerhafte Angaben in den QS-Datensätzen werden dadurch verhindert.
3. Administrative Prüfungen
   Sind verfahrensunabhängig und definieren, welche Prüfungen im Rahmen des Exports (von der Krankenkasse bis zur Bundesauswertungsstelle) innerhalb einer Einrichtung im Datenfluss in welcher zeitlichen Reihenfolge durchzuführen sind. Administrative Prüfungen beziehen sich nicht auf Datenfelder innerhalb der Nutzdatensätze, sondern auf übergeordnete Prüfungen, wie die Feststellung korrekter Dateinamen, Einhaltung des Schemas, Prüfung auf fehlende Segmente etc.

Verschlüsselung

Die Übertragung von Sozialdaten zur Qualitätssicherung enthält schützenswerte Datenbestände, welche nur für den berechtigten Empfänger bzw. die auswertende Stelle lesbar sein dürfen. Zu diesem Zweck sind, unter Berücksichtigung vorhandener Infrastrukturen, geeignete sichere Übertragungswege zu wählen und Verschlüsslungen einzusetzen. Bei der Dateiverschlüsselung und -übermittlung sind die im Folgenden genannten Standards im Bereich der gesetzlichen Krankenversicherung bereits etabliert, auf die im Rahmen der Spezifikation für Sozialdaten zurückgegriffen wird.

[35] Siehe http://www.gkv-datenaustausch.de/standards_und_normen/richtlinien_fuer_den_datenaustausch_kks/richtlinien_fuer_den_datenaustausch_kks.jsp (abgerufen am 13.11.2014).

Die für die Verwendung im KKS vorgeschriebene Security-Schnittstelle für den Datenaustausch im Gesundheits- und Sozialwesen (SECON)[36] sieht, anders als die bislang von der BAS spezifizierte Verschlüsselung, ausschließlich die Verschlüsselung vollständiger Dateien – nicht einzelner Bereiche der Dateien – nach dem Standard PKCS#7 vor. Aus den Möglichkeiten der in der Security-Schnittstelle genutzten Verschlüsselungsverfahren und den der externen gesetzlichen QS zugrunde liegenden Richtlinien ergibt sich die Notwendigkeit, die patientenidentifizierenden Daten in einer separaten Datei gesondert von den QS-Daten zu versenden. Daraus resultiert, dass Datenlieferungen, welche patientenidentifizierende Daten enthalten, stets aus zwei Nutzdateien bestehen (zuzüglich Auftragsdatensätze):

Tabelle 13: Teildateien der Datenlieferungen

Datei	verarbeitende Stelle	Inhalt
PID-Datei	VST	▪ Admin-Daten ▪ patientenidentifizierende Daten
QS-Datei	DAS-KK	▪ Admin-Daten ▪ leistungserbringeridentifizierende Daten ▪ krankenkassenidentifizierende Daten ▪ QS-Daten

Bei QS-Verfahren, die keine PID verwenden (Nicht-PID-Verfahren), entfällt die Beteiligung der Vertrauensstelle im Datenfluss. In diesem Fall wird keine separate PID-Datei erzeugt, sondern nur eine QS-Datei zuzüglich Auftragsdatensatz.

Übermittlung

Zur Übermittlung der verschlüsselten Dateien an die DAS-KK wird das Krankenkassenkommunikationssystem genutzt. Dazu stellt die Krankenkasse eine sFTP-Verbindung[37] zum Server der DAS-KK her. Der Verbindungsaufbau erfolgt mittels Benutzername und Passwort bzw. Schlüsseln in der Public-Key-Infrastruktur (PKI) der Vertrauensstelle.

Rückprotokollierung

Nach erfolgreicher Eingangsverarbeitung durch die DAS-KK erhält die Krankenkasse von der DAS-KK eine Empfangsbestätigung, die den Erhalt aller Segmente einer Transaktion bestätigt. Ist eine Datenlieferung nicht bearbeitbar, wird anstelle einer Empfangsbestätigung ein Datenflussprotokoll an die Krankenkasse übermittelt.

10.2 Erprobung von Erhebungsinstrumenten

10.2.1 Erprobung der QS-Dokumentation

Die Erprobung einer QS-Dokumentation mithilfe standardisierter Verfahren (Machbarkeitsprüfung und Probebetrieb mit freiwilligen Teilnehmern) im Rahmen der Qesü-RL hat sich in der Vergangenheit nur teilweise als zielführend erwiesen. Daher wird nunmehr durch den G-BA der Ansatz verfolgt, spezifische Projekte mit engem Bezug zu den unterschiedlichen, klärungsbedürftigen Fragen vor der Umsetzung eines Regelbetriebs durchzuführen.

Die konkreten Fragestellungen hinsichtlich der Erprobung einer QS-Dokumentation vor der Umsetzung eines Regelbetriebs können sehr unterschiedlich sein. Typische klärungsbedürftige Aspekte in Zusammenhang mit der QS-Dokumentation können z.B. sein:

[36] Siehe http://www.gkv-datenaustausch.de/standards_und_normen/sicherheitsverfahren/sicherheitsverfahren.jsp (Stand 27.03.2014, abgerufen am 13.11.2014).
[37] Secure File Transfer Protocol, ein Netzwerkprotokoll

- Kriterien der Auslösung
- Verfügbarkeit der Daten
- Verständlichkeit der Datenfelder
- Verständlichkeit der Ausfüllhinweise
- Dokumentationsaufwand
- inhaltliche Überprüfung des Indikatorensets
- Daten- und Schemaprüfung
- Entwicklung der Rechenregeln und Probeberechnung (soweit ohne PID-Verknüpfung möglich)
- Verwertbarkeit der Daten für die Durchführung der QS
- Zielgruppenspezifisches Auswertungskonzept
- Spezifikation der QS-Dokumentation
- Spezifikation des QS-Filters für die QS-Dokumentation
- Erkennen von Implementierungsbarrieren

Die Prüfung der entwickelten QS-Dokumentation erfolgt mittels strukturierter Interviews, die von AQUA-Mitarbeitern mit freiwilligen Leistungserbringern geführt werden. Ziel ist es, sowohl ein Feedback zum Verfahren insgesamt als auch zur Verständlichkeit der Dokumentation von den Leistungserbringern einzuholen. Diese erhalten zur Vorbereitung den geplanten Dokumentationsbogen, inkl. der Ausfüllhinweise. Anhand eines allgemeinen Fragebogens werden verschiedene Aspekte der Umsetzbarkeit der Dokumentation systematisch geprüft.

Im strukturierten Interview wird zu jedem Datenfeld der QS-Dokumentation erfragt:

- Sind die Daten bekannt und/oder werden sie in Ihrer medizinischen Dokumentation regelmäßig erhoben?
- Sind die Angaben (elektronisch oder in der Patientenakte) vorhanden?
- Wie hoch schätzen Sie die Zuverlässigkeit/Aussagekraft der Daten ein?
- Sind die Formulierungen der einzelnen Felder klar, verständlich und eindeutig?
- Sind die Ausfüllhinweise verständlich?
- Sind die Ausfüllhinweise ausreichend?

Zusätzlich können weitere allgemeine Informationen, die für das Verfahren relevant sind, erfragt werden, z.B.:

- wer für die medizinische Dokumentation zuständig ist,
- wer für die Dokumentation der Abrechnung zuständig ist,
- wie die Auslösekriterien eingeschätzt werden,
- mit welcher Genauigkeit das entsprechende Patientenkollektiv erfasst wird,
- wie der Dokumentationsaufwand eingeschätzt wird,
- ob es Überschneidungen zu anderen, ggf. unberücksichtigten Dokumentationsanforderungen gibt sowie
- ob das Verfahren in der vorgeschlagenen Form grundsätzlich praktikabel ist.

Im Rahmen der qualitativen Protokollierung der Interviews werden identifizierte Probleme sowie Verbesserungs- und Ergänzungsvorschläge aufgenommen. Auf dieser Grundlage können bei Bedarf erforderliche Modifikationen (z.B. einzelner Indikatoren, Datenfelder, Ausfüllhinweise) vorgenommen werden.

© 2015 AQUA-Institut GmbH

10.2.2 Erprobung von Patientenbefragungen

Im Anschluss an die Entwicklung von Fragebögen für Patienten (vgl. Abschnitt 7.3) schließt sich eine statistische, psychometrische Validierung des Gesamtinstruments an.

10.2.2.1 Deskriptive Analysen

Im Rahmen der statistischen und psychometrischen Validierung sollen zunächst deskriptive Analysen durchgeführt werden, um Anhaltspunkte für z.B. Lagemaße (Mittelwert), Streuwerte und Boden-Decken-Effekte zu erhalten und daraus Schlussfolgerungen für eine eventuell notwendige Änderung von Items ziehen zu können. Darüber hinaus geben deskriptive Analysen der Items Aufschluss über zukünftige Stichprobengrößen. Ein weiteres Ziel ist es zu überprüfen, ob alle Fragen von den Patienten beantwortet werden und ob der Umgang mit *Missing Values* angemessen ist.

10.2.2.2 Analysen zur Aussagekraft der Stichprobe (Non-Responder-Analyse)

Es soll geprüft werden, ob sich die Patienten, die den Fragebogen ausgefüllt haben, systematisch von der Grundgesamtheit hinsichtlich bestimmter Merkmale unterscheiden. Dabei können nur patientenbezogene Eigenschaften, die seitens des Leistungserbringers in der Grundgesamtheit auch identifizierbar und bei Patienten abfragbar sind (z.B. Alter und Geschlecht), in der Analyse berücksichtigt werden. So soll untersucht werden, ob die Befragungsergebnisse einen Stichproben-Bias aufweisen (Schnell et al. 2013; Diekmann 2006).

10.2.2.3 Faktorenanalyse

Im Allgemeinen wird angestrebt, Indikatoren durch mehrere Items abzubilden, um ihre Validität und Reliabilität zu erhöhen. Mit einer Faktorenanalyse kann geprüft werden, ob die einem Indikator zugeordneten Items diesen messtechnisch gut abbilden. Dies ist der Fall, wenn Items auf einen gemeinsamen Indikatorfaktor laden (Schnell et al. 2013: 151; Diekmann 2006: 378f). Sollte dies nicht der Fall sein, werden auf inhaltlicher Ebene begründete Argumente abgewogen, die ein Item dennoch als relevant erscheinen lassen. Andernfalls kann es zur Änderung oder zum Wegfallen von Items kommen.

10.2.2.4 Prüfung der Diskriminationsfähigkeit

Um statistisch signifikante Unterschiede zwischen den Einrichtungen messen zu können, muss eine Stichprobe ausreichend groß sein (Pfandzelter et al. 2009; Dimick et al. 2004). Stichprobengrößen für Einrichtungsvergleiche erfordern die ungefähre Kenntnis über die Merkmalsausprägungen der Items im Fragebogen. Insbesondere Kennwerte wie der Mittelwert und die Streuung der Items sind Voraussetzungen für die Abschätzung von Stichprobengrößen. In Studien über bereits validierte Befragungsinstrumente finden sich häufig Informationen zu den benötigten Kennwerten, die zur Abschätzung der erforderlichen Stichprobengröße genutzt werden können (Abdallah et al. 2013; Roos 2012; Cohen et al. 2011; Roos et al. 2003). Gibt es dazu keine zuverlässigen Anhaltspunkte, kann die statistische und psychometrische Validierung des Fragebogens Aufschluss über die benötigten Kennwerte der Items geben, um für den nachfolgenden Prozess die notwendige Stichprobengröße unter Berücksichtigung von Designeffekten (Schnell et al. 2013: 271ff) für den Regelbetrieb festlegen zu können.

10.2.2.5 Prüfung der Reliabilität

Um die Reliabilität des Fragebogens, d.h. die Zuverlässigkeit des Instruments, schätzen zu können, kann die Test-Retest-Reliabilität geprüft werden. Dazu wird einem Teil der Befragten der Fragebogen nach ca. 2-3 Wochen ein zweites Mal vorgelegt. Ziel ist eine möglichst hohe Korrelation zwischen den Antworten beider Messzeitpunkte. Ist diese nicht vorhanden, kann es im Ergebnis ggf. zur Umformulierung oder Löschung von Items kommen (Schnell et al. 2013: 141; Diekmann 2006: 217f). Einen weiteren Aspekt der Reliabilität stellt die Prüfung der internen Konsistenz dar, bei der die Items eines Indikators auf ihre Messgenauigkeit untersucht werden (Bortz et al. 2005).

10.2.2.6 Empirische Überprüfung der Risikofaktoren

Im Rahmen der statistischen und psychometrischen Validierung soll ebenfalls eine empirische Überprüfung der potenziellen Risikofaktoren durchgeführt werden, indem der Zusammenhang zwischen diesen und den Indikatoren zur Ergebnisqualität analysiert wird (vgl. auch Abschnitt 11.2). Für einen späteren Regelbetrieb sollen dann ggf. Items gelöscht werden, bei denen keine Effekte festgestellt wurden.

11 Auswertungen

11.1 Rechenregeln

Rechenregeln sind im Kontext der stationären und sektorenübergreifenden Qualitätssicherung die vom AQUA-Institut zu erstellenden bundeseinheitlichen Vorschriften zur Berechnung der entwickelten Qualitätsindikatoren.

In der externen stationären Qualitätssicherung gelten die Regelungen nach §8 Abs. 1der QSKH-RL:

> *Die Qualitätsindikatoren sind nach bundesweit einheitlichen Rechenvorschriften, die von der Institution nach §137a SGB V veröffentlicht werden, auszuwerten. Die Rechenvorschriften sollen bis spätestens 28. Februar des Jahres vorliegen, welches der Datenerhebung im Krankenhaus folgt.*

Die Veröffentlichung der Rechenregeln ist für die auf Landesebene beauftragten Stellen zur Erstellung und Übersendung von Auswertungen wichtig. Diese sind zuständig für die mindestens jährliche Erstellung von Auswertungsberichten in den indirekten Verfahren der stationären Qualitätssicherung. Zur Erfüllung dieser Aufgaben erhalten diese Institutionen detaillierte Berechnungsvorschriften. Für die allgemeine Öffentlichkeit erfolgt die Veröffentlichung der Rechenregeln in vereinfachter Form, wobei für jeden Qualitätsindikator Zähler und Nenner beschrieben werden.

Für die sektorenübergreifende Qualitätssicherung liegen bisher noch keine konkreten Fristen zur Veröffentlichung von Rechenregeln vor. Relevant sind die Rechenregeln grundsätzlich für Institutionen, die Auswertungen der Qualitätssicherungsdaten nach bundeseinheitlichen Kriterien vornehmen (§10 Abs. 1 Qesü-RL).

11.2 Risikoadjustierung

Der für einen Qualitätsindikator erhobene Wert für einen Patienten hängt nicht nur von der Qualität der behandelnden Einrichtung ab, sondern auch von individuellen patientenbezogenen Risikofaktoren wie z.B. dem Schweregrad der Erkrankung, den Komorbiditäten und dem Alter. Hinzu kommt eine Vielzahl weiterer Faktoren, die sich auf den Behandlungserfolg auswirken können. Gemeinsam mit Eigenschaften der Versorgungsstruktur sind Patienteneigenschaften neben der ärztlichen Leistung die wichtigsten Prädiktoren medizinischer Ergebnisse (Schneeweiss et al. 2001). Im Falle einer zufälligen Zuweisung der Patienten zu den zu vergleichenden Institutionen wären die Einrichtungsdaten bezüglich der Patientencharakteristika strukturgleich. Dadurch wäre ein Vergleich auf Basis einfacher deskriptiver Maßzahlen der empirischen Verteilung des jeweiligen Qualitätsindikators unproblematisch. Tatsächlich erfolgt die Zuteilung meist jedoch nicht zufällig, sondern eine ganze Reihe von Faktoren trägt dazu bei, dass ein Patient eine spezifische Einrichtung aufsucht. Die Faktoren selbst können wiederum mit den intrinsischen Risikofaktoren der Patienten zusammenhängen (z.B. Gyrd-Hansen et al. 2012; Phibbs et al. 1993; Cohen et al. 1985; Hebel et al. 1982). Als Konsequenz ergeben sich häufig Unterschiede zwischen den zu vergleichenden Institutionen hinsichtlich des Risikoprofils ihrer Patienten. Dabei ist zu berücksichtigen, dass Patienten mit bestimmten Konstellationen von Risikofaktoren möglicherweise systematisch schlechtere Behandlungen erfahren (Damman et al. 2009: 503).

Die zentrale Herausforderung bei der statistischen Analyse von Qualitätsindikatoren mit dem Ziel einer Gegenüberstellung der Ergebnisse verschiedener Einrichtungen besteht deshalb darin, durch die Anwendung sogenannter *Risikoadjustierungsverfahren* einen fairen Vergleich zu ermöglichen. Durch die Anwendung adäquater statistischer Verfahren werden dabei Unterschiede in den Ausgangsbedingungen hinsichtlich patientenbezogener Einflussfaktoren ausgeglichen.

11.2.1 Adjustierungsvariablen

Die Risikoadjustierung beinhaltet zunächst die Auswahl aller relevanten Risikofaktoren, welche zuverlässig und möglichst vollständig erhoben werden müssen. Die Auswahl der Adjustierungsvariablen sollte nach inhaltlichen und empirischen Kriterien erfolgen und alle für den jeweiligen Qualitätsindikator relevanten Einflussfaktoren

berücksichtigen, die von der Einrichtung selbst nicht beeinflussbar sind. Bei der Festlegung der Adjustierungsvariablen werden Erkenntnisse aus der Literatur ebenso berücksichtigt wie die Ergebnisse aus Analysen vergleichbarer Datensätze. Notwendige Voraussetzung für die Durchführung einer sachgemäßen Risikoadjustierung ist die Verfügbarkeit der hierfür notwendigen Daten.

Ziel des AQUA-Instituts ist es, langfristig für alle Ergebnisindikatoren auf Basis der Erkenntnisse evidenzbasierter Medizin eine Risikoadjustierung durchzuführen. Dies bedeutet grundsätzlich einen Zuwachs an komplexeren Modellen, insbesondere von logistischen Regressionsmodellen. Allerdings kann es sinnvoll sein, darüber hinaus auf weniger komplexe, aber leichter zu vermittelnde Verfahren zurückzugreifen. Die verschiedenen Adjustierungsverfahren, die im Folgenden vorgestellt werden, sind im Hinblick auf die voranstehend erläuterten Anforderungen unterschiedlich leistungsfähig.

11.2.2 Adjustierungsverfahren

Grundsätzlich werden international derzeit drei verschiedene statistische Methoden zur Risikoadjustierung für binäre Endpunkte angewandt, nämlich die *Einschränkung der Population*, die *Stratifizierung* und die *logistische Regression* (Ding 2009: 555). Daneben etablieren sich zunehmend auch *Mehrebenenmodelle* (Clark et al. 2010; Gatsonis et al. 1995), die wiederum zumeist auf logistischen Regressionen aufbauen.

Stratifizierung

Der einfachste differenzielle Ansatz zur Risikoadjustierung ist die sogenannte *stratifizierte Analyse* (Schumacher et al. 2008: 66ff.). Hier werden die Ergebnisse für den jeweiligen Qualitätsindikator innerhalb definierter *Strata* (Schichten) betrachtet, d.h. für Kombinationen von Ausprägungen der Einflussfaktoren. Werden alle relevanten patientenbezogenen Einflussfaktoren berücksichtigt, so wird durch die Bildung von Schichten mit entsprechendem Risikoprofil eine Vergleichbarkeit der Ergebnisse für die Qualitätsindikatoren erreicht. Ein Nachteil dieses Auswertungsansatzes ist, dass dieses Vorgehen nur handhabbar ist, wenn nicht zu viele Adjustierungsvariablen betrachtet werden. Auch können Interaktionseffekte zwischen mehreren Risikofaktoren bestenfalls rudimentär modelliert werden. Weiterhin müssen stetige Einflussfaktoren für die Bildung der Strata in Kategorien eingeteilt werden, was einer gewissen Willkür unterliegt. Insgesamt ist somit die Anwendbarkeit der Methode begrenzt.

Ein Anwendungsbeispiel für eine einfache Stratifizierung ist die Adjustierung der perioperativen Sterblichkeit anhand von ASA-Klassen. Das von der *American Society of Anesthesiologists* (ASA) entwickelte Klassifikationsschema beurteilt den körperlichen Zustand von Patienten. In Anlehnung an dieses Schema unterscheidet das AQUA-Institut fünf Kategorien:

- ASA 1: Normaler, gesunder Patient
- ASA 2: Patient mit leichter Allgemeinerkrankung
- ASA 3: Patient mit schwerer Allgemeinerkrankung
- ASA 4: Patient mit schwerer Allgemeinerkrankung, die eine ständige Lebensbedrohung darstellt
- ASA 5: Moribunder Patient, von dem nicht erwartet wird, dass er ohne Operation überlebt

Tabelle 14 schlüsselt (an einem Beispiel) die Daten der Patienten des Leistungsbereichs *Hüftgelenknahe Femurfraktur* nach den ASA-Kategorien 1 bis 5 auf. Die Sterblichkeit bei allen Patienten (Gesamtrate) beträgt 5,27 %. Die Stratifizierung nach ASA führt zur Aufdeckung erheblicher Unterschiede innerhalb der Grundgesamtheit: Die Raten bewegen sich zwischen 1,24 % für die zusammengefassten ASA-Kategorien 1 und 2 (deutlich unterhalb der Gesamtrate) und 48,63 % für ASA 5.

Tabelle 14: Qualitätsindikator „Sterblichkeit" für den Leistungsbereich Hüftgelenknahe Femurfraktur (Ergebnis 2009)

	behandelte Patienten	verstorbene Patienten	Anteil
bei allen Patienten	99.671	5.252	5,3 %
Stratifizierung nach ASA-Klassifikation			
ASA 1–2	28.044	349	1,2 %
ASA 3	63.414	3.200	5,0 %
ASA 4	7.921	1.561	19,7 %
ASA 5	292	142	48,6 %

Es wird deutlich, dass die Stratifizierung von Qualitätsindikatoren nach bekannten Risikofaktoren (hier: die ASA-Klassifikation) dabei helfen kann, fairere Vergleiche zwischen Leistungserbringern durchzuführen. Allerdings stößt diese Methode an ihre Grenzen, wenn mehrere Risikofaktoren gleichzeitig berücksichtigt werden müssen. Dieses Problem kann jedoch, wie eingangs erwähnt, beispielsweise durch den Einsatz regressionsanalytischer Adjustierungsmethoden behoben werden. Ein Grundvoraussetzung aller Adjustierungsverfahren ist jedoch, dass die genutzten Risikofaktoren in homogener Weise verstanden und kodiert werden, weil sonst auch eine Erhöhung der Verzerrung resultieren kann (Heller et al. 2007). In der Vergangenheit wurde jedoch mehrfach angezweifelt, ob dies für das ASA-Klassifikation gegeben ist (vgl. die zusammenfassende Diskussion in Sankar et al. 2014)

Regressionsanalysen

Regressionsanalysen sind statistische Standardverfahren für die Modellierung des Einflusses verschiedener Größen auf Zielvariablen. Bei stetigen Qualitätsindikatoren kommt für die Risikoadjustierung die multiple lineare Regression infrage (Weisberg 2005; Seber et al. 2003). In der Qualitätssicherung handelt es sich jedoch zumeist um binäre Zielvariablen, wie etwa die Sterblichkeit im Krankenhaus, bei denen im Bereich der Medizin sehr häufig die multiple logistische Regression (Hosmer et al. 2000) zum Einsatz kommt. Dabei wird die Wahrscheinlichkeit *p* für das Eintreten eines Ereignisses (Y = a) bei Vorliegen mehrerer Einflussgrößen X wie folgt modelliert:

$$p\,(Y = a|X_1, X_2, \dots, X_n) = \frac{e^{\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n}}{1 + e^{\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n}}$$

Formel 1: Logistische Regressionsgleichung

In dieser Gleichung repräsentiert $ß_0$ die Regressionskonstante, $ß_1$, ..., $ß_n$ die Regressionskoeffizienten und X_1, ..., X_n die ausgewählten Einflussgrößen. Die Regressionskoeffizienten $ß_i$ stellen die in dem Regressionsmodell berechneten Gewichtungen der einzelnen Risikofaktoren dar und dienen dazu, den Einfluss der Risikofaktoren auf den Qualitätsindikator zu quantifizieren. Aus den Regressionskoeffizienten $ß_i$ einer logistischen Regression lässt sich die (nach allen anderen Risikofaktoren adjustierte) jeweilige Odds Ratio (OR) als OR = $e^{ßi}$ berechnen. Die OR ist als das Verhältnis zweier Chancen (Odds) definiert. Mithilfe des OR lässt sich angeben, um welchen Faktor die einzelnen Einflussvariablen die Chance erhöhen, dass ein Ereignis (z.B. Versterben im Krankenhaus) eintritt. Ein Wert von $e^{ßi}$=1 bedeutet, dass Patienten, bei denen der Risikofaktor vorliegt, mit der gleichen Wahrscheinlichkeit versterben wie Patienten, bei denen der Risikofaktor nicht vorliegt. Eine OR größer als 1 besagt, dass bei Vorliegen des Risikofaktors die Chance erhöht ist, dass ein Ereignis eintritt. Bei einem Wert kleiner als 1 handelt es sich um einen Faktor, der diese Chance verringert. In der Praxis werden die Risikofaktoren im AQUA-Institut so kodiert, dass die OR stets über 1 liegt.

In den bisherigen Analysen gingen stetige Variablen linear in das Modell ein oder wurden kategorisiert. Bei kategorialen Variablen wurde die Kategorie mit dem geringsten Risiko als Referenzkategorie definiert. Die ORs aller anderen Kategorien wurden dann in Bezug auf diese Referenzkategorie berechnet. Die Risikofaktoren, die für die Adjustierung verwendet wurden, werden bei den Ergebnissen für die einzelnen Qualitätsindikatoren ausgewiesen. Zur Beurteilung, ob der beobachtete Effekt statistisch signifikant und klinisch relevant ist, müssen bei allen Effektschätzern zusätzlich der p-Wert und das Konfidenzintervall betrachtet werden.

Berechnung individueller Patientenrisiken
Mithilfe der multiplen logistischen Regressionsgleichung (Formel 1) kann für jeden Patienten unter Berücksichtigung seines individuellen Risikoprofils die erwartete Wahrscheinlichkeit des vorhergesagten Ereignisses (hier: Versterben im Krankenhaus) berechnet werden.

Berechnung risikoadjustierter Indikatorwerte
Im nächsten Schritt werden die erwarteten (E = expected) Ereignisse (z.B. Todesfälle) pro Einrichtung berechnet und mit der Zahl der beobachteten Ereignisse (O = observed) in dieser Einrichtung in Beziehung gesetzt. Die Anzahl der erwarteten Ereignisse unter Berücksichtigung der Risikostruktur der Einrichtung ergibt sich aus den summierten Ereigniswahrscheinlichkeiten p über alle Patienten dieser Einrichtung. Wird die Anzahl der erwarteten bzw. beobachteten Ereignisse jeweils durch die Anzahl der Patienten in der Einrichtung dividiert, so ergeben sich die entsprechenden erwarteten und beobachteten Ereignisraten.

Ein Vergleich der erwarteten mit den tatsächlich beobachteten Ereignisraten kann über das Verhältnis O/E oder über die Differenz von O und E (O minus E) erfolgen. Diese Differenz kann als absolute Abweichung der beobachteten von der erwarteten Ereignisrate betrachtet werden und zeigt im Prinzip auch an, wie gut die Ereignisrate mit dem vorliegenden Modell vorhergesagt werden kann. Über das Verhältnis O/E lässt sich erkennen, um welchen Faktor die beobachtete Ereignisrate von der erwarteten Ereignisrate abweicht.

Sind O und E gleich groß, ergibt sich eine Differenz O minus E = 0 und ein Verhältnis von O/E = 1. Eine positive Differenz von O minus E bzw. ein Verhältnis O/E > 1 deutet darauf hin, dass das Risiko für das betreffende Ereignis in der betrachteten Einrichtung höher ist, als es nach dem Risikoprofil der Patienten zu erwarten wäre. Eine negative Differenz von O minus E bzw. ein Verhältnis O/E < 1 zeigt dagegen an, dass das Risiko für das Ereignis in der betrachteten Einrichtung niedriger ist, als es nach dem Risikoprofil der Patienten zu erwarten wäre.

Die Darstellung der Ergebnisse umfasst die Angabe der Fallzahl, aussagekräftige Maßzahlen der Lage und Streuung der empirischen Verteilung (die dem Skalierungsniveau des jeweiligen Qualitätsindikators angemessen sind) sowie 95 %-Konfidenzintervalle für die Lagemaße. Diese Zahlen werden sowohl für die nicht-adjustierte Analyse als auch für die risikoadjustierte Analyse angegeben. Ergänzt werden diese Angaben durch geeignete Grafiken, die die wesentlichen Charakteristika der empirischen Verteilung und die geschätzten Werte mit der zugehörigen Schätzgenauigkeit – jeweils im Vergleich mit dem zugehörigen Referenzbereich – komprimiert und aussagekräftig darstellen. Beispiele für entsprechende Abbildungen sind Kreis- und Säulendiagramme, Box-and-Whisker-Plots, grafische Darstellungen der Schätzwerte mit 95 %-Konfidenzintervallen sowie eine vergleichende Gegenüberstellung der nicht-adjustierten Ergebnisse und der Prädiktionsintervalle für die bei der Risikoadjustierung erwarteten Ergebnisse (Ash et al. 2003) (vgl. auch Abschnitt 12.4).

Mehrebenenmodelle

Auch die Anwendung „klassischer“ Regressionsmodelle beim Vergleich von Einrichtungen wird zunehmend kritisch betrachtet. Sie berücksichtigen nämlich nicht, dass sich Patienten aus derselben Einrichtung meist untereinander ähnlicher sind als Patienten, die aus unterschiedlichen Einrichtungen stammen („patient clustering“). Weiterhin ist bekannt, dass die Schätzungen für Einrichtungen mit kleinen Fallzahlen ungenau und instabil sein können, woraus insbesondere für diese Einrichtungen falsche Schlussfolgerungen über deren Qualität resultieren können.

Zukünftig sollten daher auch Verfahren, die dieser Abhängigkeit der Daten von äußeren Faktoren auf Einrichtungsebene besonders Rechnung tragen, eingesetzt werden. Diese Modellanforderung wird insbesondere von Mehrebenenmodellen erfüllt, welche sowohl die Ebene der Patienten als auch die Ebene der Einrichtungen adäquat abbilden (Alexandrescu et al. 2011). Solche hierarchischen Modelle berücksichtigen die Mehrebenenstruktur der Daten und die damit verbundene mögliche Korrelation der Beobachtungen sowie die verschiedenen Quellen der Variation, insbesondere solche innerhalb und zwischen den Einrichtungen (Raudenbush et al. 2002; Greenland 2000; Snijders 1999; Sullivan et al. 1999). Hierdurch ist eine sachgerechte Modellierung der Einflussfaktoren auf verschiedenen hierarchisch geordneten Ebenen (z.B. Zwei-Ebenen-Modell: Einrichtung – Patient; oder Drei-Ebenen-Modell: Einrichtung – Abteilung – Patient) mit der Spezifizie-

rung von festen und zufälligen Effekten möglich. Aufgrund der zahlreichen Vorteile gegenüber anderen Risikoadjustierungsverfahren wird diese Methode zunehmend favorisiert (Shahian et al. 2001; Christiansen et al. 1997; Goldstein et al. 1996), allerdings noch selten angewandt. Damit ist die Anwendung hierarchischer linearer Modelle die bevorzugte Methode zur Risikoadjustierung bei Einrichtungsvergleichen. Allerdings ist sicherzustellen, dass die notwendigen Voraussetzungen für die Validität des Verfahrens in der jeweiligen Anwendungssituation erfüllt sind, z.B. hinsichtlich der Fallzahl (Dimick et al. 2010a; Maas et al. 2005; Cohen et al. 2003; Hox et al. 2001).

Für hierarchische lineare Modelle existieren Verfahren zur Parameter- und Intervallschätzung, sodass die methodischen Voraussetzungen für die Anwendung bei Einrichtungsvergleichen gegeben sind. Die rechentechnische Umsetzung dieser Methoden wird schnell außerordentlich komplex, weshalb bei der statistischen Auswertung häufig Simulationsverfahren wie das sogenannte Markov Chain Monte Carlo bzw. *MCMC-Verfahren* (WinBUGS) zum Einsatz kommen. Entsprechende Algorithmen sind inzwischen in den einschlägigen Statistik-Software-Paketen implementiert (Spiegelhalter et al. 2003; Rabash et al. 2001; Singer 1998). Damit ist die praktische Durchführung von Auswertungen auch an großen Datensätzen realisierbar.

Allerdings sind die Schätzalgorithmen für diese Verfahren deutlich komplexer und führen je nach Softwarepaket zu merklichen Ergebnisunterschieden. Auch sollte angemerkt werden, dass die Anwendung derartiger Modelle keine unverzerrten Schätzer mehr liefert. Darüber hinaus werden sehr gute oder sehr schlechte Ergebnisse bei kleinen Einrichtungen zu mäßig guten oder mäßig schlechten Ergebnissen transferiert (sogenanntes „shrinkage“; vgl. z.B. Greenland 2000). Diese Tatsachen dürften eine Akzeptanz innerhalb des G-BA, aber auch in der Kommunikation zu Leistungserbringern, Patienten oder sonstigen Beteiligten, ggf. deutlich einschränken.

11.2.3 Stand der Risikoadjustierung in der Qualitätssicherung

Die Adjustierungsmethoden, die das AQUA-Institut im Erfassungsjahr 2009 einsetzte (AQUA 2010), entsprachen weitgehend noch den bisher von der BQS verwendeten Methoden. Die Qualitätsindikatoren wurden für Risikofaktoren adjustiert, die in der Literatur oder in Vorjahresauswertungen als prognostisch relevant identifiziert wurden.

Folgende Methoden zur Risikoadjustierung von Qualitätsindikatoren werden vom AQUA-Institut verwendet:

- Stratifizierung: Bildung von Schichten nach Ausprägung des Risikofaktors;
- Additive Scores: Vereinfachung komplexer Modellen (z.B. multipler logistischer Regression) in Form von Scores;
- Multiple logistische Regression: Untersuchung des Einflusses von Risikofaktoren auf einen binären Qualitätsindikator.

Bis 2009 wurde – neben Überlegungen zur Definition und damit Restriktion der Population – hauptsächlich auf Stratifizierungstechniken zurückgegriffen. Seit 2010 werden multiple logistische Regressionen vermehrt eingesetzt, weil logistische Regressionen eine tiefgreifende Risikoadjustierung unter Berücksichtigung zahlreicher Risikofaktoren erlauben. Der Begriff „Risikostandardisierte Fallkonstellation“ wurde fallen gelassen, weil alle Qualitätsindikatoren der Ergebnisqualität mit Referenzbereich und ohne zusätzliche risikoadjustierte Qualitätsindikatoren nach Einschätzung der Bundesfachgruppen und des AQUA-Instituts als ausreichend risikoadjustiert gelten.

Deshalb hat das AQUA-Institut die Risikoadjustierung einer Reihe von Qualitätsindikatoren grundlegend überarbeitet und so die relative Häufigkeit der Adjustierungsmethoden untereinander verschoben. In Einzelfällen werden darüber hinaus auch bei bereits bestehenden Adjustierungsmodellen Prädiktoren, die für die aktuellen Daten nicht mehr statistisch signifikant sind, aus Modellen entfernt.

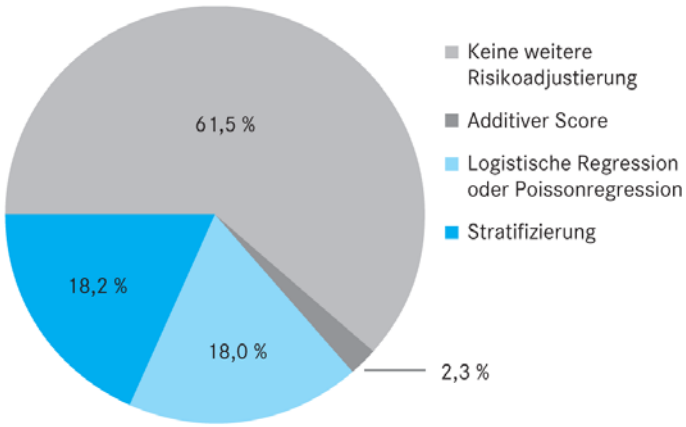

Abbildung 20: Risikoadjustierungsmethode der Qualitätsindikatoren (König et al. 2014)

Abbildung 20 visualisiert die Verteilung der Risikoadjustierungsmethoden über alle 434 Indikatoren des Erfassungsjahres 2013: 79 dieser Indikatoren werden stratifiziert, 76 mittels logistischer Regression, 2 mittels einer Poisson-Regression und 10 mittels additiver Scores risikoadjustiert. Damit verbleiben nunmehr 267 Indikatoren ohne weitere Risikoadjustierung.

Bemerkenswert ist zunächst, dass sich der Anteil der mit Regressionsmethoden risikoadjustierten Variablen seit der Übernahme der externen stationären Qualitätssicherung durch das AQUA-Institut von rund 6 % auf rund 18 % fast verdreifacht hat. Berücksichtigt man außerdem, dass Indikatoren zur Indikationsstellung und Prozessindikatoren aus methodischen Gründen nicht risikoadjustiert werden sollen, stehen lediglich 334 Ergebnisindikatoren zur Risikoadjustierung zur Verfügung. Da man aus den Ergebnissen der risikoadjustierten Rate zudem nicht die für die öffentliche Berichterstattung wichtige Prävalenz ermitteln kann, wird außerdem jedem mithilfe von Regressionsanalysen adjustierten Indikator ein Pendant der rohen Rate an die Seite gestellt, was ebenfalls per definitionem nicht risikoadjustiert werden darf. Das bedeutet, dass nunmehr 65,2 %, also fast zwei Drittel, der theoretisch einer Risikoadjustierung zugänglichen Indikatoren auch risikoadjustiert werden. Der ganz überwiegende Teil der 89 verbleibenden Ergebnisindikatoren weist – insbesondere in den Leistungsbereichen der Transplantationsmedizin – so geringe Fallzahlen auf, dass eine elaborierte Risikoadjustierung aus statistischen Gründen kaum möglich ist.

11.2.4 Weiterentwicklung der Risikoadjustierung

Mittelfristig wird die Risikoadjustierung aller Ergebnisindikatoren überarbeitet. Für jeden Indikator wird überprüft, ob eine Risikoadjustierung durchgeführt werden sollte und welche Methode ggf. infrage kommt. Dieses – auch schon 2011 praktizierte – Vorgehen wird im folgenden Abschnitt beschrieben. Das Vorgehen des AQUA-Instituts bei der Evaluation der Risikoadjustierung folgt dabei stets dem gleichen systematischen Schema. Dieses beinhaltet die folgenden Schritte:

- Auswahl der Adjustierungsvariablen
  - Inhaltliche Kriterien
  - Datenquellen
  - Statistische Voraussetzungen

- Modellspezifikation
  - Schätzung des vollständigen Modells
  - Entfernen nicht-signifikanter Variablen
  - Schätzung des reduzierten Modells

- Modellevaluation
  - Evaluation der Anpassungsgüte des Modells
  - Überprüfung der statistischen Voraussetzungen für eine Regression bei den ausgewählten Variablen (ggf. Modellanpassung)
  - Scatterplot der Einrichtungsergebnisse von rohen Raten mit standardisierten Mortalitätsraten (SMR berechnet als O/E)

Die drei Schritte (Auswahl der Adjustierungsvariablen, Modellspezifikation, Modellevaluation) werden im Folgenden beschrieben:

1. Auswahl der Adjustierungsvariablen

Wie oben bereits ausgeführt, ist für einen fairen Einrichtungsvergleich die Auswahl der Adjustierungsvariablen entscheidend. Aus praktischen Erwägungen basiert die Auswahl zunächst auf den Variablen, die mithilfe der bereits erhobenen Datenfelder gemessen werden können.

Inhaltliche Kriterien

Als Voraussetzung für die Identifizierung und Auswahl der für die Fragestellung relevanten Adjustierungsvariablen sollten zunächst die folgenden vier Fragen beantwortet werden:

- In Bezug auf welchen Qualitätsindikator,
- in welchem Zeitraum,
- in welcher Population und
- zu welchem Zweck

... soll für potenzielle Risikofaktoren adjustiert werden? (Iezzoni 2003: 20)

Entscheidend bei der Auswahl von Adjustierungsvariablen ist, dass nur solche Variablen berücksichtigt werden sollten, die von der Einrichtung selbst nicht beeinflussbar sind, so z.B. alle Patienteneigenschaften zum Aufnahmezeitpunkt. Die von der Einrichtung veränderbaren Variablen dürfen hingegen nicht adjustiert werden, da gerade diese die Qualitätsunterschiede zwischen den Einrichtungen abbilden. Potenzielle Einflussfaktoren können in die folgenden Kategorien eingeteilt werden (Iezzoni 2003):

- demografische Faktoren (z.B. Alter, Geschlecht)
- klinische Faktoren (z.B. Komorbiditäten)
- sozioökonomische Faktoren (z.B. Bildung)
- gesundheitsbezogenes Verhalten (z.B. Rauchen)
- Einstellungen und Präferenzen von Patienten (z.B. hinsichtlich Lebensqualität)

Die Auswahl der Adjustierungsvariablen anhand dieser Kriterien hängt weitestgehend vom Inhalt der jeweiligen Qualitätsindikatoren ab. Eine Ausnahme bilden die beiden demografischen Faktoren Alter und Geschlecht, die sich bei fast allen Indikatoren auf das Risiko auswirken. Da diese Variablen auch grundsätzlich in jedem QS-Dokumentationsbogen vorhanden sind, werden sie in allen Anfangsmodellen mit geschätzt, um eventuelle Effekte zu erfassen. Die restlichen Adjustierungsvariablen werden indikatorspezifisch ausgewählt. Neben Einflussfaktoren auf Patientenebene können hier auch relevante Variablen auf Einrichtungsebene verwendet werden.

Datenquellen

Ziel der Auswahl eines geeigneten Modells ist es, für die Qualitätsindikatoren eine möglichst umfassende Risikoadjustierung auf einer Datenbasis zu erreichen, deren Erfassung mit einem vertretbaren Aufwand einhergeht. Der Zugang zu GKV-Routinedaten bzw. zu den Sozialdaten bei den Krankenkassen spielt in diesem Zusammenhang eine Schlüsselrolle, da er die Berücksichtigung zahlreicher potenziell relevanter Einflussfaktoren ermöglicht und auch die Erkrankungsvorgeschichte von Patienten mit vertretbarem Aufwand in der Risikoadjustierung Eingang finden könnte. So können risikoadjustierte Analysen mit Routinedaten formal problemlos durchgeführt werden (Heller 2006). Inwieweit jedoch die zur Verfügung stehenden Daten für die jeweilige Risikoadjustierung geeignet sind, muss in der entsprechenden Situation näher untersucht werden. Die Weiterentwicklung der Risikoadjustierung hat das Ziel zu gewährleisten, dass die verwendeten Variablen einheitlich verstanden, zuver-

lässig erfasst und übermittelt werden sowie valide sind. Andernfalls kann eine Risikoadjustierung trotz einer Verbesserung der Modellanpassung auch zu einer Verminderung eines fairen Leistungserbringervergleichs führen (Heller et al. 2007).

Insgesamt hat die Risikoadjustierung eine zentrale Bedeutung bei der Zuordnung von Ergebnisverantwortung und hier insbesondere bei längsschnittlichen Analysen. Dabei wurde vom AQUA-Institut in verschiedenen Verfahren vorgeschlagen, neben Routinedaten weitere Erhebungsinstrumente, z.B. QS-Daten zu verwenden, um so mit möglichst geringem Aufwand valide Qualitätsanalysen im Allgemeinen und risikoadjustierte Analysen im Speziellen durchführen zu können. Allerdings können derartige Risikoadjustierungsmodelle zumeist erst entwickelt werden, wenn Daten aus diesen Erhebungsinstrumenten vorliegen. Dabei setzt die Entwicklung eines Risikoadjustierungsmodells, neben der Modellbildung, eine Abstimmung mit Kodier- und Fachexperten in einem ggf. mehrstufigen, iterativen Prozess voraus.

Statistische Voraussetzungen

Bei der Entwicklung geeigneter Risikoadjustierungsmodelle sind vor der eigentlichen Modellierung statistische Untersuchungen der Daten notwendig. Erste Schritte umfassen eine einfache deskriptive Analyse der Adjustierungsvariablen und einen Scatterplot zwischen Adjustierungsvariable und Qualitätsindikator, um den Einfluss einzelner Beobachtungen zu evaluieren sowie mögliche Linearitätsverletzungen und/oder Heteroskedastizität des (putativen) Zusammenhangs im Vorfeld zu erkennen (Pearson 2010: 290f.).

Weiterhin müssen die Beziehungen der Adjustierungsvariablen untereinander auf Multikollinearität geprüft werden (Pearson 2010: 289). Dies kann unter Zuhilfenahme des Varianzinflationsfaktors geschehen (Rawlings et al. 2001), welcher bei keinem Regressor einen Wert über 4 erreichen sollte. Dieser Wert ist ein Kompromiss aus der Tatsache, dass der Wert 10 am häufigsten als maximal tolerierbarer Wert (*cut-off point)* genannt wird (O'Brien 2007), bei logistischen Regressionen aber oft mit Vorsicht zu genießen ist (Elashoff et al. 2005: 591). Dies führt wiederum dazu, dass bei logistischen Regressionen mitunter sogar ein Wert kleiner 2½ empfohlen wird (Katz 2011: 91; Tarling 2009: 34f).

Schließlich ist insbesondere eine Untersuchung von Datensätzen mit fehlenden Werten wichtig, denn Einrichtungsvergleiche führen nur dann zu gültigen Aussagen, wenn die Daten, auf denen sie basieren, die Versorgungswirklichkeit innerhalb der verglichenen Einrichtungen zuverlässig abbilden. Spiegeln die vorliegenden Daten die Versorgungsqualität der verglichenen Einrichtung nicht ausreichend wider, so kann auch die Risikoadjustierung nicht sicherstellen, dass die Bewertungen der Einrichtungen fair sind. In den Analysen der Qualitätsindikatoren wurden bisher nur Fälle mit vollständigen Daten berücksichtigt, während Fälle mit fehlenden Daten aus den Analysen ausgeschlossen wurden. Diese sogenannte *complete cases analysis* wird durch die softwareseitige Überprüfung der Vollständigkeit der Datensätze, welche zu extrem hohen Vollständigkeitsraten führt, erleichtert. Fehlen dennoch Datenfelder, weil diese zum Beispiel nur optional ausgefüllt werden müssen und eine geringe Vollzähligkeit vorliegt, sind die auf solchen Daten basierenden Analysen nur unter bestimmten Annahmen gültig.

Die Weiterentwicklung der Risikoadjustierung wird in diesem Zusammenhang bedeuten, Mechanismen für das Fehlen der Daten zu erforschen und adäquate Methoden zum Umgang mit fehlenden Werten einzusetzen. Ersteres wird dabei in der Datenvalidierung und von der Systempflege geleistet; Letzteres findet sich derzeit in einem (aufgrund der algorithmisch gesicherten, sehr selten vorkommenden) listenweisen Ausschluss unvollständiger Datensätze wieder.

2. Modellspezifikation

Sind a priori die möglichen Risikoadjustierungsvariablen bestimmt, wird zunächst ein vollständiges Modell geschätzt. Aus diesem Modell werden alle nicht signifikanten Variablen entfernt, um das (zunächst) endgültige Modell zu erreichen, in dem alle Regressoren auf dem 95 %-Konfidenzniveau signifikant sind. Das so erhaltene Modell wird dann in der Folge evaluiert.

3. Modellevaluation

Die Anpassungsgüte des Modells wird, wie beispielsweise von Chertow et al. (2006: 1125) vorgeschlagen, über den Hosmer-Lemeshow-Test (Hosmer et al. 2000: 147ff) und ROC-Kurven ermittelt. Zu einem geringeren Grad

gehen auch Pseudo-R^2 in die Modellbewertung ein. Sind diese Werte in der Gesamtschau zufriedenstellend, werden nochmals die statistischen Voraussetzungen geprüft. Falls einer dieser beiden Schritte Probleme aufzeigt, wird das Modell erneut spezifiziert. Andernfalls wird ein Scatterplot der Einrichtungsergebnisse von rohen Raten mit SMRs (= O/E) produziert, um die Auswirkungen der Risikoadjustierung zu visualisieren. Zusammengefasst wird diese Visualisierung im Korrelationskoeffizienten Pearsons *r*.

11.2.5 Anwendungsfälle

Mit den obigen Ausführungen ist ein allgemeiner Rahmen für das Vorgehen bei der risikoadjustierten Auswertung von Qualitätsindikatoren abgesteckt. Die Anwendung eines einheitlichen Verfahrens für alle Fragestellungen ist weder sachgerecht noch angemessen. Welches Verfahren in der jeweiligen Situation adäquat ist, hängt von einer Vielzahl von Faktoren ab, wie z.B. der Fallzahl oder der Verfügbarkeit und Qualität der für die jeweilige Analyse notwendigen Daten. Deshalb sind für jeden Qualitätsindikator die jeweils spezifischen Einflussfaktoren zu definieren und zu dokumentieren. Zudem ist ein jeweils problemadäquates Auswertungsverfahren festzulegen und zu validieren. Ein Vergleich verschiedener Risikoadjustierungsverfahren sowie eine Auswahl der für den jeweiligen Qualitätsindikator relevanten Einflussfaktoren kann beispielsweise auf Basis der vergleichbaren Daten aus vorangegangenen Jahren durchgeführt werden. Hannan et al. (2005) beschreiben, wie die prädiktive Wertigkeit verschiedener Modelle mit einem solchen Vorgehen vergleichend evaluiert werden kann. Entsprechende Methoden sollen angewendet werden, um den in der jeweiligen Situation sachgerechten Modellansatz zu identifizieren. Die verwendeten statistischen Verfahren sind dabei nicht auf die oben genannten beschränkt. Sofern die für einen bestimmten Modellannsatz jeweils notwendigen Daten erhoben werden können, sollen auch alternative Methoden zum Einsatz kommen. Ein Beispiel sind die sogenannten *Control Charts*, die u.a. die Analyse und Darstellung von Lernprozessen bei Qualitätsvergleichen erlauben (Tekkis et al. 2003).

11.3 Area-Indikatoren

Seit den nunmehr über vier Jahrzehnte zurückliegenden ersten Arbeiten von John Wennberg (Wennberg et al. 1973) hat die Betrachtung von regionalen (engl. „area“) Unterschieden in der Gesundheitsversorgung neue Möglichkeiten der Analyse und Darstellung von Daten zur Qualität der Versorgung eröffnet. Bekannteste internationale Beispiele sind der *Dartmouth Atlas of Healthcare*[38] aus den USA oder der *NHS Atlas of Variation*[39] aus England.

Der Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen hat darauf hingewiesen, dass „die meisten Qualitätsindikatoren in Deutschland [...] eine institutionelle, auf die Leistungserbringer (insbesondere Krankenhäuser) gerichtete Perspektive“ auszeichne. Deswegen seien „Area-Indikatoren“, die „nicht eine Institution, sondern die Gesamtversorgung in einer Region“ beschreiben, für die „Diskussion der Qualität der Krankenhausversorgung unter dem gegenwärtigen Strukturwandel“ eine besonders wichtige Gruppe von Qualitätsindikatoren, weil sie meist eine „transsektorale Perspektive“ einnehmen (SVR-G 2007: 501).

Area-Indikatoren sind in Anlehnung an die Aussagen des Sachverständigenrats nicht per se ein völlig neuer Indikatortyp, sondern zunächst eine andere Auswertungsperspektive. Allerdings ist für Area-Indikatoren die Bildung einer festen Bezugsgröße oder Population wichtig. Dies können Regionen (z.B. Bundesländer und Landkreise), Einwohner, Versicherte oder Patienten in einer Region oder Gruppen von Einrichtungen der Gesundheitsversorgung (z.B. Krankenhäuser und Praxen) sein.

In Deutschland bekannte Darstellungen von Versorgungsdaten als ratenbasierte Indikatoren auf regionaler Ebene sind beispielsweise der *Faktencheck Gesundheit*[40] der Bertelsmann-Stiftung (Verwendung von Daten aus der externen stationären Qualitätssicherung und des Statistischen Bundesamtes) und der *Versorgungsatlas*[41] des Zentralinstituts für die Kassenärztliche Versorgung, in dem Abrechnungsdaten aus der vertragsärztlichen Versorgung verwendet werden (z.B. zur Grippeschutzimpfung).

[38] www.dartmouthatlas.org; letzter Zugriff am 13. November 2014.
[39] www.rightcare.nhs.uk/index.php/nhs-atlas; letzter Zugriff am 13. November 2014.
[40] www.faktencheck-gesundheit.de; letzter Zugriff am 13. November 2014.
[41] www.versorgungsatlas.de; letzter Zugriff am 13. November 2014.

Im Februar 2012 hat der G-BA eine erste Länderauswertung des AQUA-Instituts mit Geodarstellung für Indikatoren mit besonderem Handlungsbedarf („C-Indikatoren") freigegeben (AQUA 2012a). Diese Berichtsform wurde weiter ausgebaut und richtet sich sowohl an die interessierte Öffentlichkeit und die Leistungserbringer als insbesondere auch an die LQS.

In folgenden Bereichen erscheinen aus Sicht der Qualitätssicherung außerdem die Messung, Berechnung und Darstellung von Indikatoren auf regionaler Ebene besonders interessant:

- Leistungsmenge und Inanspruchnahme (Utilisation). Diese Betrachtungsweise kann hilfreich für das Verständnis von Indikationsstellungen in einer Region sein. Ein Beispiel ist die Durchführung von Prostataentfernungen pro 100.000 (männliche) Einwohner.
- Vermeidbare oder verringerbare stationäre Einweisungen (*Ambulatory Care Sensitive Conditions*). Dieser Indikatortyp kann Hinweise auf die Qualität des ambulanten Versorgungsystems geben (hausärztlich, fachärztlich und vertragsärztlicher Notdienst). Ein Beispiel dafür sind stationäre Einweisungen aufgrund Blutzucker-Entgleisungen bei Patienten mit Diabetes mellitus Typ 2 pro 100.000 Einwohner.
- Medizinische Prozesse auf Ebene der Leistungserbringer, für die eine Zielerreichung aus Leitlinienempfehlungen abgeleitet werden kann. Beispiele dafür sind der Anteil der gegen Grippe geimpften über 65-jährigen Patienten oder der Anteil der mit Antibiotika behandelten Kinder in einem bestimmten Zeitraum in einer Region.
- Ergebnisse, für die eine sektorenübergreifende, regionale Betrachtung sinnvoll ist. Ein Beispiel ist der Anteil der Patienten mit frühen Stadien von Darmkrebs (sog. pT1-Karzinome bei kolorektalem Karzinom) an allen neu entdeckten Darmkrebserkrankungen in einer Region. Dieser Indikator gibt einen Hinweis auf das Zusammenspiel aller Leistungserbringer im Hinblick auf Prävention und Früherkennung (AQUA 2011d).

Bezogen auf die Qualitätssicherung ist es besonders entscheidend, dass Qualitätsindikatoren für die Leistungserbringer handlungsrelevant sind und beeinflusst werden können. Das AQUA-Institut versteht unter Area-Indikatoren solche Qualitätsindikatoren, für die eine Darstellung auf regionaler Ebene und bezogen auf eine eingrenzbare Gruppe von Leistungserbringern sowie von ihnen versorgte Patienten oder Einwohner besonders sinnvoll erscheint. Eine wichtige Entwicklungsperspektive sind dabei die Indikatoren, die die Qualität der Versorgung möglichst sektorenübergreifend auf der Ebene einer Region oder eines regionalen Versorgungsmodells abbilden. Sie spiegeln damit die Verantwortung der Leistungserbringer gegenüber einer Population von Versicherten wider. Zweck dieser Indikatoren kann dabei einerseits die Abbildung von Versorgungsketten sein, andererseits aber auch die Darstellung geografischer Effekte, welche beispielsweise aufgrund der Versorgungs- oder der Sozialstruktur einen Einfluss auf Behandlungserfolge ausüben können (Galea 2007: 11ff.). Dies ist bei einem Vergleich von Regionen ggf. zu berücksichtigen. Eine besondere Herausforderung bei den Area-Indikatoren ist die Zuschreibung von Ergebnissen der Qualitätsmessung zu einzelnen Einrichtungen oder Teilen der Versorgungskette.

Bei der Umsetzung von Area-Indikatoren ist jeweils zu berücksichtigen, auf welcher Berichtsebene die Indikatorergebnisse zurückgemeldet werden sollen. Gerade für die Darstellung der Versorgungsqualität über mehrere Abschnitte des Versorgungspfads hinweg fehlen teilweise noch Erfahrungen über geeignete Darstellungsarten. Ein erster Schritt wird dabei die Darstellung auf der Ebene von Bundesländern sein, wobei zu berücksichtigen ist, dass die Gesundheitsversorgung auch über Landesgrenzen hinweg stattfindet. Allerdings sind auf diese Weise Leistungserbringer und Landesgeschäftsstellen sowie weitere Akteure (Landeskrankenhausgesellschaft, Kassenärztliche Vereinigung, ggf. Selektivvertragspartner) klar benennbar, und entsprechende Auswertungskonzepte sind vorhanden und erprobt.

Eine wesentlich größere Herausforderung wird die sektorenübergreifende Darstellung auf der Ebene von sektorenübergreifenden Versorgungsnetzen sein, die entweder formal-vertraglich zusammenarbeiten oder die sich faktisch durch eine stattfindende Zusammenarbeit ergeben. Für die formal-vertragliche Zusammenarbeit, für die es erst wenige Beispiele in der Bundesrepublik gibt, würde die Rückmeldung der Ergebnisse der Qualitätsmessung für alle Vertragspartner zusammen in einem gesonderten Bericht erfolgen. Für Kooperationen ohne besondere vertragliche Bindung müssen noch geeignete Auswertungs- und Berichtskonzepte erprobt werden. Ein innovativer und vielversprechender Ansatz ergibt sich aus dem Prinzip der sozialen Netzwerkanalyse, die

z.B. für die Behandlung von Parkinson-Patienten in den Niederlanden (Wensing et al. 2011) oder im Rahmen des EUREGIO-Netzes zur Vermeidung von Infektionen mit antibiotikaresistenten Erregern bereits beschrieben wurde (Friedrich 2011). Entsprechende Software-Programme sind vorhanden. Unklar ist jedoch noch, ab welchem Grad oder ab welcher Häufigkeit der Zusammenarbeit eine Integration von Leistungserbringern in eine derartige Auswertung bzw. in einem Bericht sinnvoll und interpretierbar ist oder von diesen akzeptiert wird. Gerade in Ballungsgebieten können sich sehr komplexe Netzwerke von Zu- und Überweisern sowie auch eine durch Patienten selbst gesteuerte Inanspruchnahme ergeben, sodass sich die Frage stellt, welche maximale Anzahl von Einrichtungen eine solche Netzwerkanalyse beinhalten sollte. Diese Aspekte müssen für jeden Area-Indikator im Rahmen von Erprobungsprojekten weiter bearbeitet werden.

11.4 Qualitätsindizes

Zahlreiche Qualitätsindikatoren, insbesondere Indikatoren der Ergebnisqualität, besitzen „ungünstige“ statistische Eigenschaften. Wenn beispielsweise die Anzahl bestimmter Behandlungen in einer Einrichtung (Fallzahl) sehr gering ist, erhöht sich die Wahrscheinlichkeit, dass ein gemessenes seltenes Ereignis (z.B. eine Komplikation) hier nur zufällig aufgetreten ist. Kurz gesagt: Geringe Fallzahlen und niedrige Häufigkeiten („Fallzahl-Prävalenz-Problem“) von Qualitätsindikatoren erschweren statistisch zuverlässige Aussagen über die Ergebnisqualität (König et al. 2014; Heller 2010).

11.4.1 Fallzahl-Prävalenz-Problem

Dies sei am Beispiel des Qualitätsindikators „Sterblichkeit bei Risiko-Lebendgeburten“ aus dem Leistungsbereich *Neonatologie* illustriert. Die durchschnittliche Sterberate betrug im Erfassungsjahr 2011 0,8 % bei einer durchschnittlichen Anzahl von 195 Fällen pro Krankenhaus. Abbildung 21 zeigt die Verteilung der krankenhausspezifischen Sterblichkeitsraten. 303 der insgesamt 511 datenliefernden Krankenhäuser (60,4 %) wiesen überhaupt keine Sterbefälle auf. Dabei kann schon die Tatsache, dass ein Indikator in über 60 % der Untersuchungseinheiten eine Nullmessung aufweist, als deutlicher Hinweis auf mögliche Messprobleme verstanden werden.

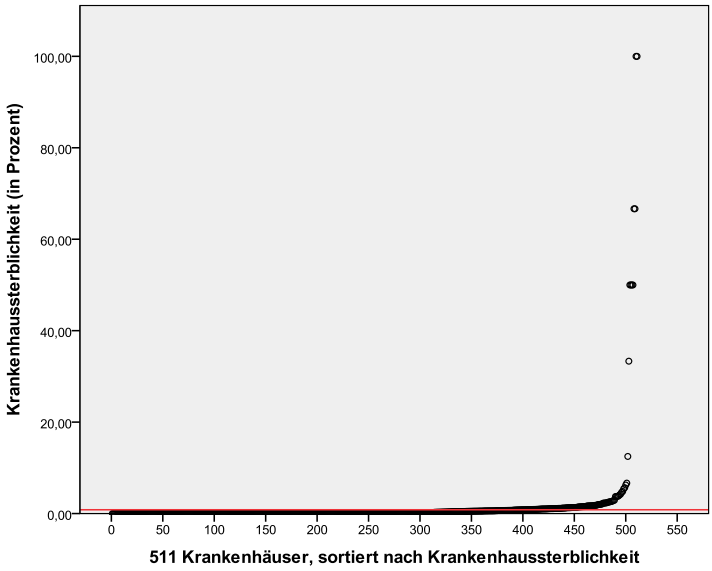

Abbildung 21: Verteilung krankenhausspezifischer Sterblichkeitsraten

Die Frage, ob ein Indikator tatsächliche Abweichungen oder nur zufällige Schwankungen misst, kann mit Methoden der sogenannten schließenden Statistik beantwortet werden. Dimick et al. (2004) analysierten die Unterscheidungsfähigkeit (Diskriminationsfähigkeit) von Qualitätsindikatoren, indem sie von folgender Annahme ausgingen: Die Verdopplung des Wertes eines Qualitätsindikators für eine Einrichtung gegenüber dem Durchschnittswert in der Population soll auf einem 95 %-Signifikanzniveau mit einer Wahrscheinlichkeit von 80 % erkannt werden. Auf Basis dieser Annahme können minimal benötige Fallzahlen berechnet werden und anschließend kann ermittelt werden, wie viele Einrichtungen diese Fallzahl tatsächlich erbringen (Dimick et al.

2004).[42] Für den angeführten Beispielindikator „Sterblichkeit bei Risiko-Lebendgeburten“ ergibt sich eine minimal benötigte Fallzahl von 1.049. Im Jahr 2011 lag diese Fallzahl aber nur in 2 von 511 Krankenhäusern vor. Insofern lässt sich sagen, dass die Unterscheidungsfähigkeit dieses Qualitätsindikators unbefriedigend ist.

Dieses Problem ist auch in anderen Leistungsbereichen vorhanden, was beispielsweise (2004) für sieben operative Leistungsbereiche in den USA nachgewiesen haben. Nur für einen dieser Bereiche (Aortokoronare Bypassoperation) wurde die Unterscheidungsfähigkeit der Indikatoren als ausreichend beurteilt. Vor einigen Jahren wurden analog auch in Deutschland Indikatoren der Ergebnisqualität aus verschiedenen stationären Leistungsbereichen untersucht.[43] Auch hier wurden für die große Mehrheit der Indikatoren deutliche Probleme hinsichtlich ihrer Unterscheidungsfähigkeit festgestellt (Heller 2010).

Das AQUA-Institut hat im Rahmen der sogenannten „Schnellprüfung und Bewertung der Indikatoren der externen stationären Qualitätssicherung hinsichtlich ihrer Eignung für die einrichtungsbezogene öffentliche Berichterstattung“ insgesamt 302 Indikatoren der externen stationären Qualitätssicherung in der beschriebenen Art und Weise analysiert (AQUA 2011a). Abbildung 22 illustriert, dass 44 % der Indikatoren der datenliefernden Krankenhäuser über keine hinreichende Unterscheidungsfähigkeit verfügten. Für 87 % der Indikatoren war die Unterscheidungsfähigkeit für weniger als die Hälfte der Krankenhäuser gegeben. Nur für 7 % der analysierten Indikatoren konnte nach den oben beschriebenen Kriterien eine Unterscheidungsfähigkeit für mehr als 75 % der Krankenhäuser errechnet werden.

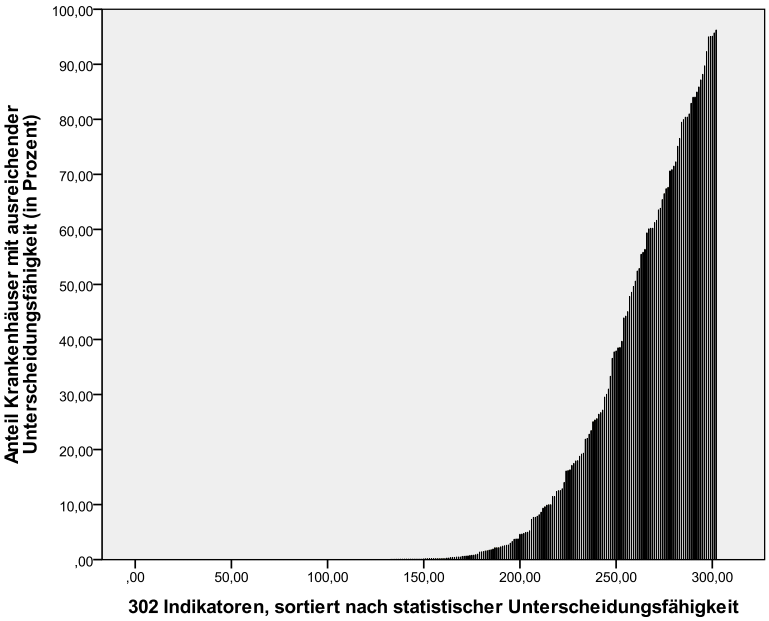

Abbildung 22: Unterscheidungsfähigkeit der Qualitätsindikatoren der externen stationären Qualitätssicherung

[42] Ein grundsätzlich ähnliches Verfahren zur Ermittlung der statistischen Unterscheidungsfähigkeit von Indikatoren wurde von der BQS im Rahmen des QUALIFY Verfahrens entwickelt (vglhttp://www.bqs-qualify.com/online/public/leistungen/qualify, letzter Zugriff am 13. November 2012). Allerdings wurde eine minimal benötigte Fallzahl bestimmt, um einen Unterschied zu den beobachteten Raten der 5% besten Krankenhäuser und der 5% schlechtesten Krankenhäuser statistisch nachweisen zu können. Dieses Vorgehen setzt aber voraus, dass die tatsächliche Verteilung (ohne zufällige Fehler) für einen untersuchten Qualitätsindikator bekannt ist. Dies ist allerdings nicht der Fall, weil die empirisch beobachteten Werte auch die zufälligen Fehler beinhalten. Besteht die Verteilung eines Qualitätsindikators aus einem fixen Wert (alle Krankenhäuser sind gleich gut) und einem zufälligen Fehler, wird gemäß diesem Verfahren dennoch eine minimal benötigte Fallzahl ermittelt, die alleine von der Streuung des zufälligen Fehlers abhängig ist, damit sehr unterschiedliche Ergebnisse haben kann und fälschlicherweise als eine Verteilung von echten Werten interpretiert wird. Dabei würde dieser Indikator keinerlei Unterscheidungsfähigkeit aufweisen. Die minimal benötigte Fallzahl für diesen Indikator wäre demnach unendlich. Dieses Gedankenbeispiel zeigt, dass das statistische Verfahren der BQS im QUALIFY-Verfahren zur statistischen Prüfung der Unterscheidungsfähigkeit von Qualitätsindikatoren grundsätzlich fehlerhaft konzeptioniert ist. Das von Dimick et al. (2004) gewählte Vorgehen orientiert sich dagegen am Mittelwert der Verteilung, sodass die (Verteilung der) zufälligen Fehler keinen (nennenswerten) Einfluss auf die Berechnung der minimal benötigten Fallzahlen nach diesem Verfahren haben sollte.

[43] Untersucht wurden: Helios-Indikatoren der zweiten Generation, ausgewählte BQS-Indikatoren sowie A-Indikatoren des Projektes Qualitätssicherung mit Routinedaten (QSR).

11.4.2 Lösungsansätze

Für das offenbar weitverbreitete Problem der Unterscheidungsfähigkeit von Indikatoren existieren mehrere Lösungsvorschläge (Heller et al. 2004):

- Auswerten längerer Zeiträume bzw. Zusammenfassen mehrerer Untersuchungsjahrgänge: Die Analyse von gleitenden Durchschnittswerten (*moving averages*) und die Erstellung von sogenannten „kumulativen Summen Plots“ bzw. *Quality Control Charts* gehören ebenfalls in diese Kategorie (Tekkis et al. 2003).
- Zusammenfassen von einzelnen Leistungserbringern zu Gruppen von Leistungserbringern: Volume-Outcome-Analysen, regionale Analysen oder auch Analysen nach Trägern oder weiteren Eigenschaften von Leistungserbringern wären hier zu subsummieren.
- Die Anwendung von Mehrebenenanalysen/empirischen Bayes-Analysen kann das Fallzahl-Prävalenz-Problem insofern in Ansätzen lösen, als dass, neben dem errechneten Wert für eine Einrichtung, angenommen wird, dass eine Vorinformation (*informative prior*) existiert, die bei einer empirischen Bayes-Analyse der Gesamtmittelwert ist. Unter Berücksichtigung dieser Information werden Verteilungen berechnet, die zu einer Verminderung der Streuung und ggf. zu einer besseren Prognosefähigkeit der leistungserbringerbezogenen Ergebnisse führen können (Dimick et al. 2010b; Ash et al. 2003).
- Zusammenfassung mehrerer Qualitätsindikatoren zu einem Qualitätsindex (Heller et al. 2011; Heller 2010; Heller 2008)[44]

Die letztgenannte Lösung soll im Folgenden am Beispiel des 2011 im Leistungsbereich *Neonatologie* neu eingeführten „Qualitätsindex der Frühgeborenenversorgung“ erläutert werden.

Der „Qualitätsindex der Frühgeborenenversorgung“ setzt sich aus folgenden Einzelindikatoren zusammen:

- Tod des Kindes
- Intra- und Periventrikuläre Hirnblutung
- Nekrotisierende Enterokolitis mit Operation
- Zystische Periventrikuläre Leukomalazie
- Bronchopulmonale Dysplasie
- Höhergradige Frühgeborenenretinopathie

Für die Berechnung des Qualitätsindex wurde jeder der genannten Endpunkte mittels sequenzieller logistischer Regression risikoadjustiert.

[44] In früheren Publikationen wurde auch der Begriff „Komplikationsindex“ verwendet. Wenn aber außer Komplikationen ggf. auch das Ziel einer Behandlung oder die Indikationsqualität zum Gegenstand der Qualitätsbetrachtung gemacht werden soll, ist der Begriff zu eng gefasst. Im internationalen Kontext wird diesbezüglich auch der Terminus „*composite measure*“ verwendet (Dimick et al. 2012; AHRQ 2011; Staiger et al. 2009). Allerdings werden hier oft auch über mehrere Leistungsbereiche zusammengefasste Maßzahlen beschrieben.

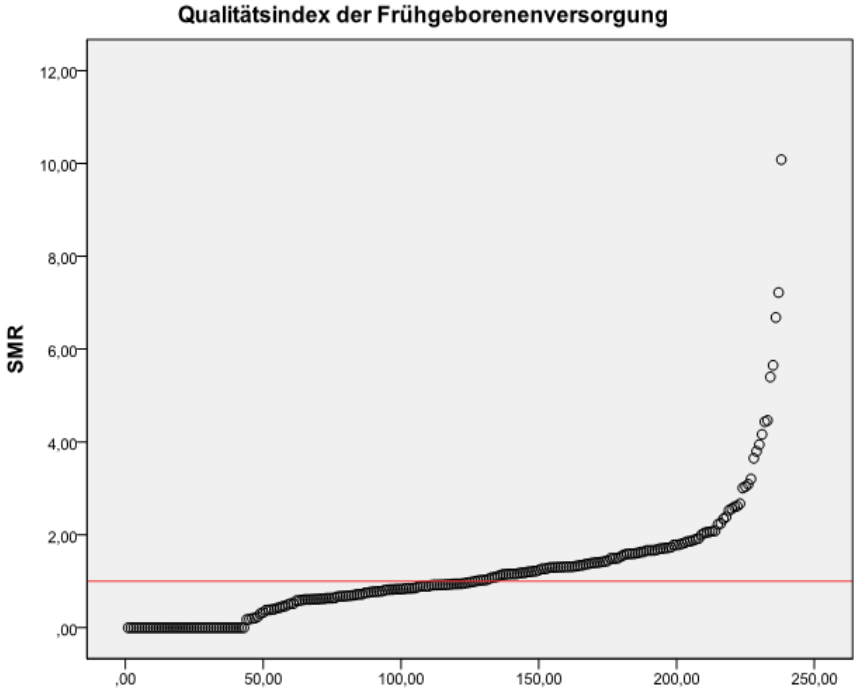

Abbildung 23: Krankenhausspezifischer Prävalenz des Qualitätsindex

Abbildung 23 zeigt die Verteilung der krankenhausspezifischen Häufigkeiten für den Qualitätsindex der Frühgeborenenversorgung aus dem Erhebungsjahr 2013 für Krankenhäuser, die mindestens 5 derartige Frühgeborene versorgt hatten.

Im Ergebnis fand sich bei 15,8 % der analysierten Fälle einer der oben genannten Endpunkte. Hieraus ergibt sich anhand der beschriebenen Methode zur Ermittlung einer Mindestfallzahl eine minimal benötigte Fallzahl von 40. Diese wird immerhin von 45 % der datenliefernden Krankenhäuser erreicht. Im Vergleich zur alleinigen Betrachtung der Sterblichkeit von Frühgeborenen ergibt sich bei Anwendung eines Index also eine erhebliche Verbesserung der statistischen Unterscheidungsfähigkeit. Anders formuliert: Der Einfluss von zufälligen Fehlern bei der Qualitätsmessung wird durch die Verwendung von Qualitätsindizes erheblich abgeschwächt. Es ist demnach eine deutlich bessere, zuverlässigere Qualitätsmessung auf Basis des Qualitätsindex zu erwarten.

11.4.3 Vor- und Nachteile der Verwendung von Indizes

Darüber hinaus ermöglichen zu einer Maßzahl zusammengefasste Qualitätsindizes eine detailliertere Qualitätsmessung als einzelne Qualitätsindikatoren. Grundsätzlich kann eine entsprechend umfassende Messung der Qualität auch durch die einzelnen zugrunde liegenden Indikatoren ermöglicht werden, wie z.B. in sogenannten Dashboard- oder Cockpit-Verfahren. Diese Vorgehensweisen sind aber komplex und bei kritischer Betrachtung wenig objektiv, solange keine einheitlichen Regeln vorgegeben sind. Bei der Berechnung eines Qualitätsindex liegt dagegen ein einheitliches und transparentes Verfahren zur ersten Einstufung eines Leistungserbringers vor, das eine objektivere und zuverlässigere Vergleichbarkeit von Einrichtungsergebnissen ermöglicht. Da zusätzlich die Ergebnisse einzelner Indikatoren dargestellt werden, kann zudem ermittelt werden, aufgrund welcher Indikatoren es zu einer Auffälligkeit gekommen ist. Um dem Problem zu begegnen, dass hier ggf. eng zusammenhängende oder aufeinanderfolgende Komplikationen gemessen werden, wurde eine Hierarchisierung der Qualitätsindikatoren durchgeführt: So wird beispielsweise eine Hirnblutung nur bei überlebenden Kindern und eine nekrotisierende Enterokolitis (eine Darmerkrankung) nur bei überlebenden Kindern ohne Hirnblutung gezählt.

Tatsächlich kann auch bezweifelt werden, dass die Berücksichtigung von „Vorstufen“ späterer Komplikationen bei der vorherigen Einzelbetrachtung von Indikatoren immer berücksichtigt wurde. Es ist also fraglich, ob bisher beispielsweise bei der Beurteilung des Qualitätsindikators „Höhergradige Frühgeborenenretinopathie“ (eine Erkrankung der Netzhaut) ausreichend berücksichtigt wurde, dass die Ausprägung dieses Indikators davon

abhängig ist, ob und in welchem Ausmaß sehr unreife Frühgeborene in einem Krankenhaus überlebt haben. Im Qualitätsindex wird dieser Sachverhalt implizit berücksichtigt, weil hier sowohl verstorbene Kinder als auch Kinder mit höhergradiger Frühgeborenenretinopathie ermittelt werden. Würde dabei ein Krankenhaus bei einem Qualitätsindex aufgrund von schlechten Werten zur höhergradigen Frühgeborenenretinopathie auffällig und wiesen die hierarchisch darüberliegenden Einzelindikatoren (Sterblichkeit, Hirnblutung, Enterokolitis usw.) überdurchschnittlich gute Werte auf, so könnte dies ein Hinweis darauf sein, dass die Auffälligkeit für die höhergradige Frühgeborenenretinopathie erst durch die gute Ergebnisqualität der darüberliegenden Qualitätsindikatoren zur Frühgeborenenversorgung entstanden ist.

Weiterhin wird gelegentlich das Problem der Mehrdimensionalität von Indizes thematisiert. Es wird nicht nur die Sterblichkeit, sondern die Sterblichkeit und eine Hirnblutung etc. abgebildet. Allerdings ist dieser Sachverhalt kein auf Indizes beschränktes, sondern auch bei Qualitätsindikatoren existentes Phänomen. Bei der Berechnung von Qualitätsindizes wird dieses „Problem“ allerdings im Gegensatz zu den bisher bestehenden mehrdimensionalen Qualitätsindikatoren adressiert, indem diese als einzelne Dimensionen ausgewiesen und mit einer eigenen Risikoadjustierung versehen werden. Umgekehrt kann auch formuliert werden, dass die Versorgungsqualität tatsächlich mehrdimensional ist und ein Index dies als „Indikatorprofil“ berücksichtigt (Kopp et al. 2007). Sowohl bei Indikatoren als auch bei Indizes dämmt eine geeignete Risikoadjustierung das Mehrdimensionalitätsproblem zusätzlich ein.

Für die Einführungsphase wurde mit der Bundesfachgruppe Perinatalmedizin konsentiert, alle Indikatoren des Index gleich zu gewichten. Dabei kann kritisiert werden, dass hier Indikatoren mit sehr unterschiedlicher Tragweite (z.B. Versterben oder eine bronchopulmonale Dysplasie) gleichberechtigt in den Index einfließen. Dieses Problem kann formal durch eine Gewichtung der einzelnen Indikatoren relativ einfach gelöst werden. Aufwendig ist allerdings eine an inhaltlichen Kriterien orientierte Ermittlung bzw. Konsentierung der Gewichtungsfaktoren mit Fachexperten. Diese Problematik ist in gleicher Weise z.B. auch bei der Konstruktion von „disabilty adjusted life years“ zu beobachten (Murray 1994).

Bei der Indexbildung selbst gibt es zahlreiche unterschiedliche Herangehensweisen bzw. Möglichkeiten der Konstruktion (AHRQ 2011; Staiger et al. 2009; Heller et al. 2008; Reeves et al. 2007; Dimick et al. 2004). Die Berechnungen können je nach den Erfahrungen, die mit dem Umgang mit Qualitätsindizes gemacht werden, und auch in Bezug auf wechselnde Anwendungskontexte angepasst werden.

Das AQUA-Institut empfiehlt den Qualitätsindex auf Auffälligkeitskriterium zu nutzen und sich in einem ersten Schritt anhand der Ausprägung der darunterliegenden Einzelindikatoren zu orientieren, wo entsprechende Qualitätsprobleme in einer Einrichtung ggf. vorliegen. Anschließend sollten für alle diejenigen Fälle, die zur Auffälligkeit im Qualitätsindex beitrugen, ein intensiver Dialog geführt und die Ursachen für das Auftreten entsprechender Ereignisse ermittelt werden. Wird dabei ein Qualitätsproblem identifiziert, sollte der Qualitätsindex als auffällig bewertet werden.

11.5 Referenzbereiche

Grundsätzlich werden die Qualitätsindikatoren hinsichtlich ihrer Auslösung im Hinblick auf statistische Auffälligkeiten in zwei Gruppen systematisiert, nämlich in auf aggregierten Daten basierende Indikatoren und in Sentinel-Event-Indikatoren (Nadzam et al. 1993). Sentinel-Events sind sehr seltene Ereignisse (Michaels et al. 2007: 829), deren Folgen so gravierend sind, dass jedes einzelne dieser Ereignisse qualitativ untersucht werden soll. Diese Indikatoren benötigen daher keinen einrichtungsbezogenen Nenner. Demgegenüber stehen Indikatoren auf der Basis von Aggregatdaten (Jäckel 2009). Zum einen sind dies Indikatoren, bei denen pro Patient eine kontinuierliche Variable erhoben wird (z. B eine Strahlenbelastung) und anschließend Werte der zentralen Tendenz (oft der Median) über alle in einer Einrichtung behandelten Patienten ermittelt werden, die innerhalb eines bestimmten Bereichs liegen sollten. Zum anderen sind dies ratenbasierte Indikatoren, bei denen eine bestimmte Rate des Auftretens unabhängig von jeglichen Qualitätsstandards erwartet wird. Bei Indikatoren auf der Basis von Aggregatdaten können die Einrichtungsergebnisse über die Fallzahl oder – bei risikoadjustierten Qualitätsindikatoren – die erwartete Wahrscheinlichkeit standardisiert werden (Mainz 2003: 524). Mit Rückgriff auf wissenschaftliche Untersuchungen kann bisweilen ein sogenannter Zielbereich festgelegt werden, innerhalb dessen ein Einrichtungsergebnis bei angemessener Qualität liegen sollte (Methode 4 in

Haller et al. 2009: 1161). Für solche Indikatoren wird zumeist ein fester Wert als Referenzbereich festgelegt. Demgegenüber gibt es bei zahlreichen anderen Indikatoren keine klare evidenzbasierte Grenze. Bei diesen Qualitätsindikatoren werden auf Basis des Vorjahresgesamtergebnisses (oder vergleichbaren Datenkörpern) Toleranzbereiche für unauffällige Ergebnisse definiert (Methode 3 in Haller et al. 2009: 1160f.). Häufig werden in diesen Fällen auch Perzentile als Toleranzbereiche verwendet. Ergebnisse, die außerhalb des Toleranzbereichs liegen, müssen genauer analysiert werden.

Grundsätzlich werden Sentinel-Events und Referenzbereiche auf Grundlage eines systematischen Literaturreviews in Verbindung mit Expertengesprächen in den Bundesfachgruppen im Vorjahr des Erhebungszeitraums festgelegt. Lässt sich auf der Grundlage des systematischen Literaturreviews ein eindeutiger evidenzbasierter Grenzwert ermitteln, so wird dieser zur Grundlage des Zielbereichs genommen. Andernfalls wird auf Basis der wissenschaftlichen Diskussion, wie sie in den Expertengesprächen zum Tragen kommt, ein Toleranzbereich als Referenzbereich gesetzt.

Auf Basis dieser Methode wurde in der externen stationären Qualitätssicherung für das Erfassungsjahr 2013 die in der folgenden Tabelle dargestellte Verteilung der verschiedenen Arten von Referenzbereichen festgelegt.

Tabelle 15: Systematik der Referenzbereiche in der externen stationären Qualitätssicherung im Erfassungsjahr 2013

Indikatorart	**Referenzbereichtyp**	**Anzahl**		
		gesamt	**grenzwertbasiert**	**perzentilbasiert**
Sentinel-Event	Sentinel-Event	40	Sentinel-Event	
ratenbasiert (ohne weitere Risikoadjustierung)	Zielbereich	33	33	0
	Toleranzbereich	64	36	28
	Nicht definiert	135		
ratenbasiert (mit Risikoadjustierung)	Zielbereich	26	26	0
	Toleranzbereich	108	50	58
	Nicht definiert	26		
mengenbasiert	Toleranzbereich	2	0	2

12 Berichterstattung

Die Institution nach §137a SGB V ist verpflichtet, die Ergebnisse der Qualitätssicherungsmaßnahmen in einer für die Allgemeinheit verständlichen Form zu veröffentlichen. Die Veröffentlichungspflicht umfasst die übersichtliche Darstellung der in der Qualitätssicherung verwendeten Indikatoren, Instrumente und Dokumentationen[45] sowie die laufende Berichterstattung über die Ergebnisse, Projekte und abgeschlossene Aufträge der Qualitätssicherung.

Das zentrale Medium zur Veröffentlichung sind spezielle Webseiten des AQUA-Instituts (www.sqg.de). Dort werden alle vom G-BA abgenommenen und mit Beschlussfassung freigegebenen Dokumente in einer mindestens barrierearmen Form[46] der Öffentlichkeit zur Verfügung gestellt.

12.1 Hintergrund

Der versorgungspolitische Hintergrund bzw. das wesentliche Ziel der öffentlichen Berichterstattung im Gesundheitswesen ist die Stärkung des Qualitätswettbewerbs zwischen Leistungserbringern (Schrappe 2011). Dem liegt die Annahme zugrunde, dass veröffentlichte Informationen von Leistungserbringern bessere Wahlentscheidungen für Patienten ermöglichen und die Leistungserbringer auf diese Weise motiviert werden, die Qualität der Versorgung weiter zu verbessern. In praktischer Hinsicht spiegelt der angenommene Zusammenhang von verfügbaren Qualitätsinformationen und individuellen Wahlentscheidungen ein komplexes Problemfeld wider. So konnten Studien zeigen, dass die Veröffentlichung von qualitätsrelevanten Informationen allein oftmals nur wenig Einfluss auf die Auswahl eines gesundheitlichen Versorgungsangebotes durch Patienten hat (Hibbard 2008). Ob und inwieweit dies damit zusammenhängt, dass die meisten Qualitätsberichte deutscher Krankenhäuser für Patienten oft nur schwer verständlich (Verbraucherzentrale NRW 2014; Friedemann et al. 2009) und bisher wenig bekannt sind (Verbraucherzentrale NRW 2014; de Cruppé et al. 2011b; Geraedts et al. 2011; Geraedts et al. 2010), ist unklar. Generell ist der Effekt einer öffentlichen Qualitätsberichterstattung auf die Effektivität der Versorgung bislang eher ungeklärt (Ketelaar et al. 2011; Geraedts et al. 2009; Fung et al. 2008). Demgegenüber steht die mittlerweile starke und weiter zunehmende Frequentierung von Gesundheitsportalen, die Qualitätsinformationen auch aus der externen stationären Qualitätssicherung mit dem Anspruch der Laienverständlichkeit darbieten.

Die Problematik einer adäquaten und für die Allgemeinheit verständlichen Darstellung von Qualitätsinformationen betrifft das AQUA-Institut in besonderer Weise, weil sich die Ergebnisse mehrheitlich an das Fachpublikum der Qualitätssicherung bzw. an Leistungserbringer richten. Gleichwohl sieht das AQUA-Institut keine Alternative zur bisherigen, umfassenden Veröffentlichung seiner Ergebnisse und damit einer möglichst transparenten Darlegung seiner Arbeitsweise: Das AQUA-Institut stellt alle relevanten Informationen über seine Arbeit, die zur Anwendung kommenden Methoden und die Ergebnisse der Qualitätssicherung grundsätzlich der Allgemeinheit auf der Webseite www.sqg.de zur Verfügung.

12.2 Adressaten der Berichterstattung

Insgesamt kann nicht davon ausgegangen werden, dass die Bereitstellung von Informationen mit der Vermittlung von Wissen gleichzusetzen ist. Besonders bei der Befragung von Patienten zeigt sich, dass diese oft Schwierigkeiten haben, statistische Angaben wie Mittelwerte und Prozentangaben zu verstehen. Um diesem Problem gerecht zu werden, ist eine möglichst einfache und „lebensnahe“ Form der Datenaufbereitung erforderlich, denn die Art und Weise der Darstellung beeinflusst die Interpretation der Daten erheblich (Gigerenzer 2009; Gigerenzer et al. 2007; Grande et al. 2005). Andererseits mangelt es patientengerechten aufgearbeiteten Informationen zumeist an der notwendigen Präzision, um von professionellen Akteuren aus den Bereichen der Qualitätssicherung oder von medizinischen Fachexperten gleichermaßen genutzt werden zu können.

[45] Die bisherige Berichterstattung zu Qualitätssicherungsmaßnahmen beschränkt sich auf die Leistungsbereiche der stationären Qualitätssicherung, weil sich die bisher entwickelten sektorenübergreifenden QS-Verfahren noch in der Erprobung befinden. Einzelheiten zur Darstellung der Ergebnisse im Rahmen der sektorenübergreifenden Qualitätssicherung werden in den themenspezifischen Bestimmungen festgelegt.
[46] Die Webseite www.sqg.de selbst ist barrierefrei nach den Kriterien der BITV 2.0.

Das AQUA-Institut unterscheidet bei den Adressaten in Bezug auf die Berichterstattung deswegen grundsätzlich zwischen:

- Laien (Menschen ohne medizinische und qualitätssicherungsrelevante Vorkenntnisse)
- Interessierter Öffentlichkeit (Menschen und Institutionen mit medizinischen und/oder qualitätssicherungsrelevanten Interessen, z.B. Patienten, Beschäftige aus Gesundheitsberufen usw.)
- Betroffene Leistungserbringer und Fachpublikum der Qualitätssicherung (Menschen und Institutionen aus den Arbeitsbereichen der Qualitätssicherung, z.B. QS-Beauftragte der Krankenhäuser, Qualitätsinitiativen, G-BA, LQS usw.)
- Medizinischen Experten (Ärzte und andere medizinisch-therapeutische Leistungserbringer)

Die genannten Zielgruppen unterscheiden sich erheblich durch den Grad ihrer Sach- und Fachkenntnisse. Sie haben nicht nur unterschiedliche Informationsbedürfnisse und Zugangsmöglichkeiten, sondern auch teilweise ganz unterschiedliche Erwartungshaltungen bezüglich der Art und Weise der Informationsaufbereitung.. Das betrifft sowohl die Form der Darstellung von Datenmaterial (Zahlen, Tabellen, Grafiken, Text) als auch die Art und die Ausführlichkeit der gegebenen Informationen wie auch das zu wählende Medium der Darstellung (gedruckt oder computergestützt).

12.2.1 Laien

Laien (im oben genannten Sinne) sind eine eher selten adressierte Zielgruppe der Berichterstattung des AQUA-Instituts. Selbst der Qualitätsreport, der sich weitgehend um Allgemeinverständlichkeit, im Sinne der Reduktion des Abstraktions- und Komplexitätsniveaus, bemüht, kann die Komplexität der medizinischen Leistungen und der diesbezüglich stattfindenden Qualitätssicherung nicht auf „echte“ Laienverständlichkeit reduzieren.

Eine patientengerechte Darstellung, d.h. eine „Übersetzung“ der Ergebnisse der Qualitätssicherung ist derzeit in weiten Teilen Aufgabe der Leistungserbringer und bspw. in den Qualitätsberichten der Krankenhäuser zu veröffentlichen. Dass diese Berichte diesen Anforderungen nur selten gerecht werden (Verbraucherzentrale NRW 2014; Geraedts et al. 2010), ist ein Mangel, der im Rahmen der Regelungen zu den Qualitätsberichten der Krankenhäuser behoben werden müsste.

Durch den G-BA wurde bislang nur das sogenannte NICU-Projekt (www.perinatalzentren.org) im Sinne einer patientengerechten Darstellung der Ergebnisse beauftragt (siehe Abschnitt 12.4.3). Hierfür hat sich das AQUA-Institut eingehend mit Experten des Max-Planck-Instituts für Bildungsforschung (Berlin) beraten

12.2.2 Interessierte Öffentlichkeit

Das AQUA-Institut interpretiert die auftragsgemäß zu adressierende „Allgemeinheit“ als interessierte Öffentlichkeit im oben genannten Sinne und geht davon aus, dass Patienten (auch potenzielle Patienten und Angehörige) die primäre Zielgruppe sind. Auch wenn noch weitgehend unklar ist, ob und inwieweit Patienten bereitgestellte Qualitätsinformationen tatsächlich zur Wahlentscheidung nutzen (de Cruppé et al. 2011a; Hibbard et al. 2005), so ist das Ziel der Berichterstattung zunächst die Herstellung von Transparenz, um überhaupt Einrichtungen vergleichen zu können und so Wahlentscheidungen zu ermöglichen (Murphy et al. 2010).

Paradigmatisch hierfür sind nicht nur Patienteninformationen, sondern auch der Qualitätsreport. In Letzterem wird versucht, die Ergebnisse der hochkomplexen Materie der Qualitätssicherung auch der interessierten Öffentlichkeit zugänglich zu machen. Der Report verfügt über eine Leseanleitung und ein Glossar, die Ziele, Inhalte und Methoden der Qualitätssicherung werden in einem Hintergrundkapitel erläutert, medizinisch-therapeutische und pflegerische Sachverhalte sind vereinfacht dargestellt und die Diskussion einzelner Qualitätsindikatoren ist auf exemplarische und/oder Fälle mit besonderen Merkmalen beschränkt (AQUA 2014e).

12.2.3 Leistungserbringer und Fachpublikum der Qualitätssicherung

Die betroffenen Leistungserbringer sowie alle Personen und Einrichtungen, die im weitesten Sinne professionell mit der Qualitätssicherung im Gesundheitswesen befasst sind, gehören zur primären Zielgruppe der meisten Formen der Berichterstattung des AQUA-Instituts.

Fachpublikum und „Allgemeinheit“ bzw. interessierten Öffentlichkeit werden voneinander unterscheiden, da es sich sowohl bei den Leistungserbringern als auch bei den Akteuren der „Qualitätssicherungsbranche“ um Experten handelt, deren Interesse sich sehr stark auf präzise dargelegte formale, methodische und technische Details richtet. In den speziell an dieses Publikum gerichteten Berichten wird deshalb weitgehend auf Kommentierungen verzichtet, weil medizinische Experten stärker komprimierte und an Zahlen orientierte Darstellungen bevorzugen (Geraedts et al. 2011).

In Bezug auf das Fachpublikum liegt die Anforderung nicht so sehr darin, allgemeinverständliche Texte zu formulieren. Vielmehr sind hier die Art und Weise der Bereitstellung der Daten sowie eine gute Kommunikationsstruktur wesentlich entscheidender als die Allgemeinverständlichkeit dieser Texte. Das AQUA-Institut trägt diesen Erfordernissen durch einen komfortablen Bereitstellungsweg aller Daten über ein Internetportal, der Aufbereitung von Informationen in verschiedenen Dateiformaten (XLS/CSV) und eine Hotline, die alle Fragen rund um den „Datenservice“ beantwortet, Rechnung.

Auch die vom AQUA-Institut individuell erstellten Rückmeldeberichte für die Einrichtungen richten sich naturgemäß an das Fachpublikum bzw. die jeweils adressierten Leistungserbringer. Hier ist das AQUA-Institut an der öffentlichen Berichterstattung nur indirekt beteiligt. Zum einen erhalten die Einrichtungen Rückmeldeberichte in Form von Benchmarkreports. Die öffentliche Berichterstattung obliegt aber den Einrichtungen (derzeit ausschließlich Krankenhäuser) selbst, die ihre Ergebnisse in einem nach Vorgaben des G-BA zu erstellenden Strukturierten Qualitätsbericht veröffentlichen müssen. Das AQUA-Institut hat für den G-BA dazu im Jahr 2013 eine Empfehlungsliste mit allgemeinverständlichen „Übersetzungen“ für alle veröffentlichungspflichtigen Indikatoren der externen Qualitätssicherung angefertigt. Diese Liste ist vom G-BA als (unverbindliche) Empfehlung an die Einrichtungen weitergegeben worden.

Mit Blick auf die Studienlage geht das AQUA-Institut davon aus, dass sich die Veröffentlichung möglichst vieler Indikatoren positiv auf die Umsetzung von Qualitätsverbesserungsmaßnahmen auswirkt (Fung et al. 2008). So setzten etwa Krankenhäuser nach veröffentlichter hoher Mortalitätsrate bei koronarer Bypass-Chirurgie vermehrt Qualitätsverbesserungsmaßnahmen um (Chassin 2002). Ebenso scheint allein der Vorgang der Veröffentlichung von Ergebnisdaten die Bereitschaft zur Qualitätsverbesserung zu erhöhen (Hibbard et al. 2003). Gleichsam sind dem Institut auch die Gefahren bewusst, die hiermit verbunden sein können – vor allem etwaige Strategien zur patientenbezogenen Risikoselektion (Geraedts et al. 2009; Smith 1995). Solche Strategien können dazu führen, dass beispielsweise multimorbide Patienten an andere Einrichtungen verwiesen werden, um die eigene Bewertung nicht zu gefährden (Moscucci et al. 2005). Das AQUA-Institut begegnet diesem Problem mit umfassenden Ansätzen zur Risikoadjustierung (vgl. Abschnitt 11.2). Vor allem der Anteil der mit Regressionsmethoden adjustierten Variablen hat sich seit 2010 von ca. 6 auf fast 18 Prozent fast verdreifacht ((König et al. 2014).

Zum anderen prüft das AQUA-Institut seit 2011 im Auftrag des G-BA alle Indikatoren der externen stationären Qualitätssicherung mit Blick auf ihre Eignung für eine einrichtungsbezogene öffentliche Berichterstattung. Als Ergebnis dieser Prüfungen stieg die Anzahl der ab dem Erfassungsjahr 2011 verpflichtend zu veröffentlichenden Indikatoren von 28 auf mittlerweile 295 (für das Erfassungsjahr 2013).

12.2.4 Wissenschaftliche Experten

Wissenschaftliche Experten (z.B. akademisch arbeitende Mediziner und andere vorwiegend forschend tätige Experten) sind eine wichtige und vielfältig in die Qualitätssicherungsarbeit des AQUA-Instituts einbezogene Gruppe, z.B. als Teilnehmer an Scoping-Workshops, als Teilnehmer des RAM-Panels oder der Bundesfachgruppen. Damit verbunden ist das Anliegen des AQUA-Instituts, z.B. qualitätssicherungsrelevante medizinisch-therapeutische Sachverhalte korrekt darzustellen.

Diese Experten sind vor allem in Bezug auf die Ergebnisse und die Erläuterung der Maßnahmen der Qualitätssicherung Adressaten der Berichterstattung des AQUA-Instituts, nicht aber hinsichtlich wissenschaftlich-medizinischer Hintergründe.[47]

Da die primären Adressaten der Berichterstattung des Instituts die interessierte Öffentlichkeit und das Fachpublikum der Qualitätssicherung (inkl. der von der Qualitätssicherung betroffenen Mediziner) sind, werden im Qualitätsreport und den Entwicklungsberichten medizinische Kontexte möglichst verständlich dargelegt. Wenn Berichte primär die interessierte Öffentlichkeit und das Fachpublikum der Qualitätssicherung adressieren, beschränkt sich die Bereitstellung allgemeinerer medizinischer Hintergrundinformationen auf die Darlegung qualitätsrelevanter medizinischer Sachverhalte, die überdies im sprachlichen Duktus an einer möglichst weitgehenden Allgemeinverständlichkeit orientiert sind. Gleiches gilt für Abbildungen und Tabellenformate, die möglichst auch ohne spezifisches medizinisches Hintergrundwissen verstanden werden sollen.

12.3 Berichtsarten

12.3.1 Berichte im Rahmen der gegenwärtigen einrichtungsübergreifenden Qualitätssicherung

Bundesauswertung

Das AQUA-Institut veröffentlicht jährlich zum 30. Juni des Folgejahres alle Ergebnisse zu den Leistungsbereichen der bundesweiten gesetzlich verpflichtenden Qualitätssicherung eines Erfassungsjahres (Bundesauswertung). Für jeden bestehenden Leistungsbereich wird – sofern Daten für das Erfassungsjahr vorliegen – jährlich ein separater Bericht erstellt, der die Ergebnisse aller durch die Rechenregeln definierten Indikatoren eines Verfahrens auf Bundesebene aggregiert darstellt.

Jeder Bericht enthält

- Angaben zur Datengrundlage,
- grafisch aufbereitete Auswertungen zu jedem Indikator, inklusive eines Vergleichs zu den Ergebnissen des Vorjahres,
- Detailergebnisse zum Indikator sowie
- Erläuterungen zur Berechnung einzelner Indikatoren (Adjustierungsvariablen, Scores und sonstige Interpretationshilfen).

Jeder Bericht wird ergänzt um die sogenannte Basisauswertung dieses Leistungsbereichs, die u.a. grundsätzliche Angaben zu Fallzahlen, Aufenthaltsdauer, Patienten, Diagnostik und Therapie enthält.

Die Bundesauswertungen repräsentieren die grundlegendste Analyse der erfassten Daten zu einem Leistungsbereich. Sie werden in Form von PDF-Dokumenten geordnet nach Leistungsbereichen auf den Webseiten des AQUA-Instituts veröffentlicht. Obwohl die Berichte auf diese Weise auch der Allgemeinheit zur Verfügung stehen, richten sie sich in ihrer detaillierten, aber unkommentierten Form an das Fachpublikum der Qualitätssicherung.

Qualitätsreport

Für die interessierte Öffentlichkeit werden die Ergebnisse der Bundesauswertung (in möglichst allgemeinverständlicher Form) in einem Qualitätsreport veröffentlicht, der in einer PDF-Fassung und in gedruckter Form erscheint.

Der jährlich im September veröffentlichte Qualitätsreport enthält die Ergebnisse aller Leistungsbereiche und Indikatoren, stellt sie aber in einer stärker zusammengefassten Form dar. Um die Allgemeinverständlichkeit zu erhöhen, enthält der Qualitätsreport eine ausführliche Leseanleitung in Bezug auf die Formen der Ergebnisdarstellung sowie ein Glossar. Die Ziele, Inhalte und Methoden der Qualitätssicherung werden in gesonderten

[47] Da die primären Adressaten der Berichterstattung des AQUA-Instituts die interessierte Öffentlichkeit und das Fachpublikum der Qualitätssicherung (inkl. der von der Qualitätssicherung betroffenen Mediziner) sind, beschränkt sich die Bereitstellung allgemeinerer medizinischer Hintergrundinformationen auf eine Darlegung qualitätsrelevanter medizinischer Sachverhalte, die überdies im sprachlichen Duktus an einer möglichst weitgehenden Allgemeinverständlichkeit orientiert sind. Gleiches gilt für Abbildungen und Tabellenformate, die möglichst auch ohne spezifisches medizinisches Hintergrundwissen verstanden werden sollen

Hintergrundkapiteln erläutert, medizinisch-therapeutische Sachverhalte sind vereinfacht dargestellt und die Diskussion einzelner Qualitätsindikatoren ist auf exemplarische und/oder Fälle mit besonderen Merkmalen beschränkt.

Rückmeldeberichte (Benchmarkreports) für Leistungserbringer
Alle datenliefernden Leistungserbringer erhalten Rückmeldung zu ihren Ergebnissen. Struktur und Form der Berichte für die Leistungserbringer, die durch das AQUA-Institut erstellt werden, sind im Grundaufbau identisch mit dem Aufbau der Bundesauswertungen. Zusätzlich enthalten die Berichte individuelle Analysen des Leistungserbringers und seiner Ergebnisse. Damit die Leistungserbringer ihre Ergebnisse mit denen anderer Einrichtungen des Verfahrens vergleichen können, erfolgt eine grafische, anonyme Gegenüberstellung der Ergebnisse aller Leistungserbringer. Nur dem Empfänger eines Berichts ist dabei bekannt, welchen Wert in der Abbildung seine Einrichtung repräsentiert. Die Berichte sind damit Grundlage für eine vergleichende Qualitätssicherung.

Wie für die Bundesauswertung gilt auch für den Benchmarkreport, dass die Lesbarkeit der Dateien allein durch den Umfang der Informationen schon maßgeblich erschwert wird. Deshalb stellt das AQUA-Institut den Einrichtungen die Ergebnisse auf einer gesicherten Internetplattform zur Verfügung.

Berichte zum Strukturierten Dialog
Das AQUA-Institut erstellt für die direkten Verfahren Berichte der Ergebnisbewertungen und der durchgeführten qualitätssichernden Maßnahmen. Die Länder erstellen diese Berichte analog für die indirekten Verfahren auf Basis von bundeseinheitlichen, durch das AQUA-Institut spezifizierten Vorgaben (§15 Abs. 1 QSKH-RL) und übermitteln sie an das AQUA-Institut. Das AQUA-Institut wertet die Ergebnisse des Strukturierten Dialogs aus den indirekten Verfahren auf Bundesebene aus und erstellt einen abschließenden Bericht.

Die Berichterstattung über die Ergebnisse des Strukturierten Dialogs ist auf zwei Dokumente verteilt.

- Zum 15. März: Bericht zum Strukturierten Dialog aus den direkten Leistungsbereichen. In diesem Bericht sind auch die Ergebnisse der Datenvalidierung aus den direkten Leistungsbereichen enthalten.
- Zum 15. Mai: Zusammenfassender Bericht zum Strukturierten Dialog aus den indirekten und direkten Leistungsbereichen.

Berichte zur Datenvalidierung
Nach Abschluss des Datenvalidierungsverfahrens werden die Ergebnisse des Strukturierten Dialogs zur Datenvalidierung für die indirekten Verfahren von den LQS und die Ergebnisse des Stichprobenverfahrens mit Datenabgleich entsprechend der Spezifikation an das AQUA-Institut übersandt.

Die Ergebnisse des Datenvalidierungsverfahrens werden gemäß §15 Abs. 1 der QSKH-RL in den zusammenfassenden Berichten dargestellt. Darüber hinaus werden die Ergebnisse gemäß §15 Abs. 2 der QSKH-RL in einem Abschlussbericht präsentiert.

- Zum 15. März: Bericht über die statistische Basisprüfung mit Strukturiertem Dialog und dem Stichprobenverfahren mit Datenabgleich aus den direkten Leistungsbereichen. In diesem Bericht sind auch die Ergebnisse des Strukturierten Dialogs aus den direkten Leistungsbereichen dargelegt.
- Zum 15. Mai: Zusammenfassender Bericht zur Datenvalidierung aus den indirekten und direkten Leistungsbereichen.

Berichte über Entwicklungsleistungen
Zu den Ergebnissen jeder beauftragten Entwicklungs- oder Weiterentwicklungsleistung erstellt das AQUA-Institut einen ausführlichen Bericht. Dieser enthält eine Dokumentation aller durchgeführten Prozessschritte und Teilergebnisse sowie Art und Umfang der Einbeziehung von Fachexperten.

Nach Abnahme der Abschlussberichte zu den Entwicklungsleistungen durch das Plenum des G-BA werden diese inklusive der Anhänge und der Stellungnahmen der gesetzlich zur Stellungnahme berechtigen Organisati-

onen auf den SQG-Webseiten des AQUA-Instituts veröffentlicht. An gleicher Stelle wird auch über den Stand laufender Projekte berichtet.

Tätigkeitsbericht
Ein jährlich für den G-BA bis zum 31. März zu erstellender Tätigkeitsbericht fasst die Arbeitsschwerpunkte des AQUA-Instituts sowie die wesentlichen Ergebnisse im Rahmen der Beauftragung nach §137 a SGB V zusammen. Auch er wird auf den Webseiten des Instituts veröffentlicht.

Darstellung der Leistungsbereiche und Indikatoren
Alle dokumentationspflichtigen Leistungsbereiche der externen (bisher ausschließlich stationären) Qualitätssicherung werden zusammen mit den jeweiligen Indikatoren auf den Webseiten des Instituts ausführlich und in allgemeinverständlicher Form erläutert.

12.3.2 Auswertungskonzepte für sektorenübergreifende QS-Verfahren

Die Bundesauswertung, die Rückmeldeberichte für Leistungserbringer sowie die Berichte zum Strukturierten Dialog und zur Datenvalidierung sind derzeit (Stand November 2014) auf die Darlegung der Ergebnisse im Rahmen der bestehenden, d.h. der externen stationären und damit der sektorspezifischen Qualitätssicherung beschränkt. Mit der Berichterstattung zu den in Entwicklung befindlichen sektorenübergreifenden QS-Verfahren wird an die genannten Berichtsformate eine Reihe neuer Anforderungen gestellt, die sich nicht zuletzt aus neu zum Tragen kommenden Auswertungskonzepten ergeben.

Grundsätzlich orientiert sich jedes Auswertungskonzept an den Zielen eines QS-Verfahrens und damit an den Inhalten der Entwicklungsberichte bzw. den konkreten Angaben zu den Indikatoren (Zähler, Nenner, Risikoadjustierung, Ergebnis- und Dokumentationsverantwortlichkeit, Erhebungsinstrumente). Darüber hinaus sind immer die empirisch-statistischen Gegebenheiten der zu analysierenden Qualitätsindikatoren sowie auch die Besonderheiten einer sektorenübergreifenden Herangehensweise in Betracht zu ziehen.

Bei Auswertungskonzepten von sektorenübergreifenden QS-Verfahren sind folgende Besonderheiten zu berücksichtigen:

- Es existiert üblicherweise ein Follow-up.
- Es sind ggf. mehrere Leistungserbringer am Versorgungsprozess beteiligt.
- Es liegen ggf. Angaben aus Patientenbefragungen vor.
- Einige Indikatoren können andere Indikatoren bedingen, z.B. indem sie eine Vorstufe eines späteren Outcome abbilden.
- Ggf. ist eine regionale Auswertung in Betracht zu ziehen

Damit sind grundsätzlich die folgenden Fragen zu beantworten:

- Wie sind die Informationen im Follow-up aufzubereiten?
- Für welche Leistungserbringer sind die Informationen aufzubereiten?
- Wie fließen Angaben aus Patientenbefragungen in diese Auswertungen ein?
- Wie sind voneinander abhängige Indikatoren zu beurteilen?
- Wie wird die Region definiert und wie werden die Daten erhoben?

Damit stellt sich im Kontext der sektorenübergreifenden Qualitätssicherung die Frage nach der Ergebnisdarstellung und damit nach der Form der Rückmeldeberichte an die Leistungserbringer noch einmal neu, da hier eine konkrete Zuweisung von Verantwortlichkeiten oft schwierig ist. D.h. hier ist vor allem zu klären, welche Leistungserbringer beteiligt sind bzw. welche Ergebnisse welchen am Behandlungsprozess beteiligten Akteuren übermittelt werden können und dürfen.

Vorbehaltlich der mit dem G-BA abzustimmenden Einzelheiten sollen die folgenden Festlegungen gelten:

- Ein Leistungserbringer, der eine initiale Indexleistung erbringt, soll die Ergebnisse im Follow-up (typischerweise Indikatoren der Ergebnisqualität) für alle seine Patienten dargestellt bekommen, ggf. getrennt danach, ob er sie selbst nachbehandelt/nachuntersucht hat oder nicht.
- Ein Leistungserbringer, der eine Nachbehandlung/-untersuchung durchführt, soll die Ergebnisse für alle seine Nachbehandlungen/-untersuchungen dargestellt bekommen, ggf. getrennt danach, ob er sie selbst primär behandelt hat oder nicht. Bei dieser Art der Darstellung wird es sich oft nicht um Indikatoren der Ergebnisqualität, sondern um Indikatoren der Prozessqualität oder eine Information zu den durchgeführten Nachuntersuchungen handeln.
- Angaben aus Patientenbefragungen sind typischerweise Indikatoren der Prozessqualität (ob eine Patientenaufklärung sachgerecht durchgeführt oder der Patient in die Therapieentscheidung mit einbezogen wurde) und Ergebnisqualität (z.B. Zufriedenheit mit der Therapie bzw. funktionelles Outcome einer Therapie im Follow-up). Diese Angaben sollen den Leistungserbringern, die die initiale Indexleistung erbracht haben, für die von ihnen initial behandelten Patienten zurückgespiegelt werden. Leistungserbringer, die eine Nachbehandlung/-untersuchung durchgeführt haben, sollen die Ergebnisse von Patientenbefragungen für die von ihnen behandelten/untersuchten Patienten erhalten.
- Ggf. sollen Indikatoren, die eine Vorstufe zu anderen abbilden, als Subgruppen der späteren Indikatoren ausgewiesen werden. So könnten beispielsweise intraoperative Komplikationen das funktionelle oder allgemeine spätere Outcome deutlich beeinflussen. In einem solchen Fall ist es sinnvoll, den späteren Follow-up-Indikator auch getrennt danach auszuweisen, ob zuvor intraoperative Komplikationen vorlagen oder nicht.
- Die Auswertung von sogenannten Area-Indikatoren erfordert nicht nur eine frühzeitige Definition der auszuwertenden Region/Population, sondern auch der Erhebungsmethoden (ggf. müssen gesonderte Regions- oder Institutionskennzeichen mit erhoben oder neu geschaffen werden) sowie eines Konzepts zur Rückspiegelung von Ergebnissen.

Insgesamt sollen die Fristen im Krankenhaus und in der Arztpraxis für alle zukünftigen sektorenübergreifenden Verfahren einheitlich festgelegt werden, um zusätzlichen Aufwand oder Unsicherheiten bei den Leistungserbringern zu vermeiden. Gleiches gilt für die Datenannahmestellen und für Berichte der zukünftigen Landesarbeitsgemeinschaften sowie den Bericht des AQUA-Instituts an den G-BA. Eine Abstimmung der Fristen erfolgt im Zuge der Erstellung themenspezifischer Bestimmungen.

Über die adäquaten Formen der Ergebnisdarstellung (siehe Abschnitt 12.4) wird entsprechend dem Auswertungskonzept entschieden, das mit jedem Abschlussbericht als Muster/Entwurf mitgeliefert wird.

12.4 Formen der Ergebnisdarstellung

Je nach adressierter Zielgruppe und Ebene der Berichterstattung verwendet das AQUA-Institut unterschiedliche tabellarische und grafische Ergebnisdarstellungen. Das bedeutet einerseits, dass grundsätzlich gleiche Daten und Auswertungen für unterschiedliche Zielgruppen individuell aufbereitet werden. Andererseits sind bei sektorenübergreifenden Ergebnisdarstellungen oder auch bei einrichtungsübergreifenden Follow-up-Auswertungen differenzierende Konzepte mit Blick auf die Rückmeldeformen zu beachten.

12.4.1 Aktuell verwendete Darstellungen

Adressaten der Ergebnisdarstellung der externen stationären Qualitätssicherung (Bundesauswertung) sind die Leistungserbringer sowie in Form des Qualitätsreports auch die interessierte Öffentlichkeit. Die Kommunikation der Einzelergebnisse über die „Qualität der Krankenhäuser“ obliegt den Leistungserbringern in den jährlichen Qualitätsberichten.

Die Ergebnisse von Auswertungen werden je nach Sachverhalt als absolute Zahlen, als Anteile (Prozent) oder Raten sowie teilweise auch als dimensionslose Verhältniszahl (O/E) dargelegt.

In der Regel werden Ergebnisse, Konfidenzintervalle (Vertrauensbereiche), Verteilungswerte und Referenzbereiche mit einer Nachkommastelle ausgewiesen. Eine Ausnahme bilden dabei Werte < 9,995, die aufgrund der

Genauigkeit mit zwei Nachkommastellen ausgewiesen werden. Die rechnerische Auffälligkeit von Krankenhäusern sowie die Tendenz werden entsprechend diesem Schema ermittelt. Bei einer Fallzahl kleiner als 4 werden aus datenschutzrechtlichen Gründen keine Ergebnisse ausgewiesen.

Das Ergebnis für O/E-Indikatoren ermittelt sich anhand der systembedingten maximalen Genauigkeit (30 Stellen) der ungerundeten Werte.

Zur Veranschaulichung der Ergebnisse wählt das AQUA-Institut derzeit die folgenden Grafiken.

Linien- bzw. Fehlerbalkendiagramm
Linien- bzw. Fehlerbalkendiagramme zeigen die Lage eines Mittelwertes einer Stichprobe bzw. die wahrscheinliche Lage eines Mittelwertes einer Grundgesamtheit im Zusammenhang mit einem zugehörigen Vertrauensbereich (Konfidenzintervall). Der „Fehlerbalken“ (der Vertrauensbereich) gibt dabei den Bereich an, in dem der tatsächliche Wert eines Qualitätsindikators unter Berücksichtigung aller zufälligen Ereignisse (z. B. Dokumentationsfehler) mit einer bestimmten, zuvor festgelegten Wahrscheinlichkeit liegt.

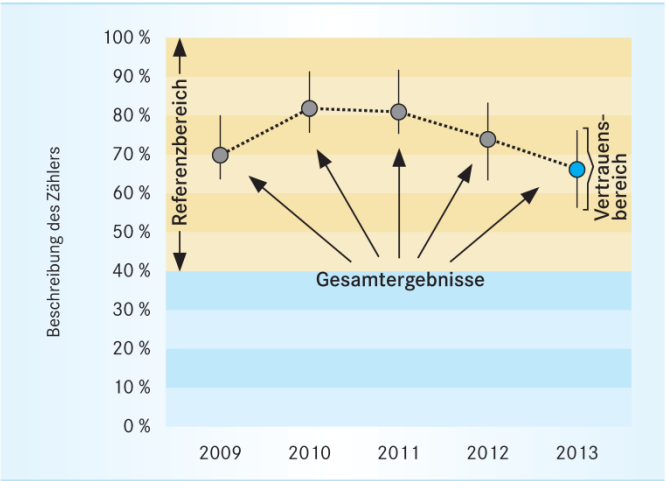

Abbildung 24: Beispiel für ein Fehlerbalkendiagramm (Stichprobe)

Im Qualitätsreport und in der Bundesauswertung wird das Fehlerbalkendiagramm zur Darstellung der Gesamtergebnisse auf Basis der Fälle (Patienten) verwendet. Die durch den Indikator abgebildete Versorgungsqualität wird hier zusätzlich im Verlauf aufeinander folgender Jahre dargestellt und kann dementsprechend auf einen Blick verglichen werden.[48] Sofern für den Qualitätsindikator ein Referenzbereich definiert ist, erscheint dieser als Ziel- oder Toleranzbereich entsprechend gelb unterlegt.

Liniendiagramme werden in Bundesauswertungen und Rückmeldeberichten verwendet. In der Bundesauswertung stellt das Liniendiagramm vergleichend die Bundeswerte (in der Regel Anteilswerte) eines Qualitätsindikators für das jeweilige Erfassungsjahr und das vorangehende Jahr grafisch dar.

[48] Abhängig vom Maßstab der y-Achse ist es möglich, dass ein Vertrauensbereich nicht erkennbar ist, wenn seine Grenzen innerhalb des Kreises liegen, der das Gesamtergebnis markiert.

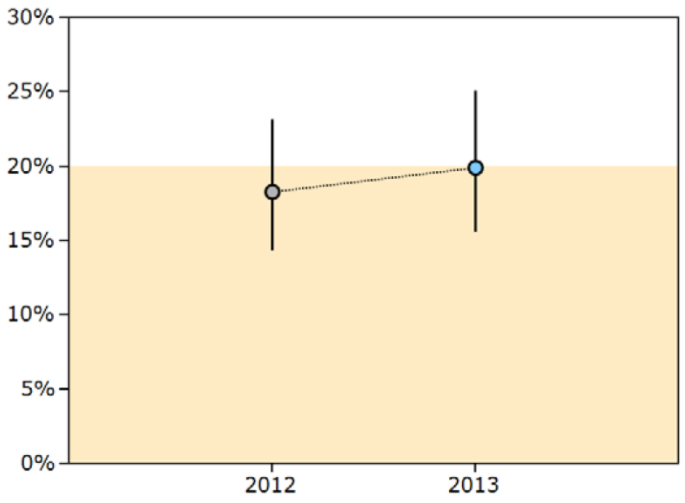

Abbildung 25: Beispiel für ein Liniendiagramm (Bundesergebnisse 2013: N = 276 Fälle und 2012: N = 295 Fälle)

Auf diese Weise kann die Versorgungsqualität zweier aufeinanderfolgender Jahre direkt anhand der Werte der beiden Jahre verglichen werden. Die Konfidenzintervalle der Jahreswerte werden als senkrechte Linien gezeigt.[49] Die Basis der Berechnungen (Anzahl der Fälle, resp. Patienten) kann der zweiten Zeile des Grafiktitels entnommen werden.

In den Rückmeldeberichten wird das Fehlerbalkendiagramm zum einen zur Darstellung der Krankenhäuser im Vergleich verwendet, wobei das jeweils eigene Krankenhaus gesondert gekennzeichnet ist, zum anderen wird das Ergebnis der eigenen Einrichtung im Vergleich zum Bundesergebnis dargestellt.

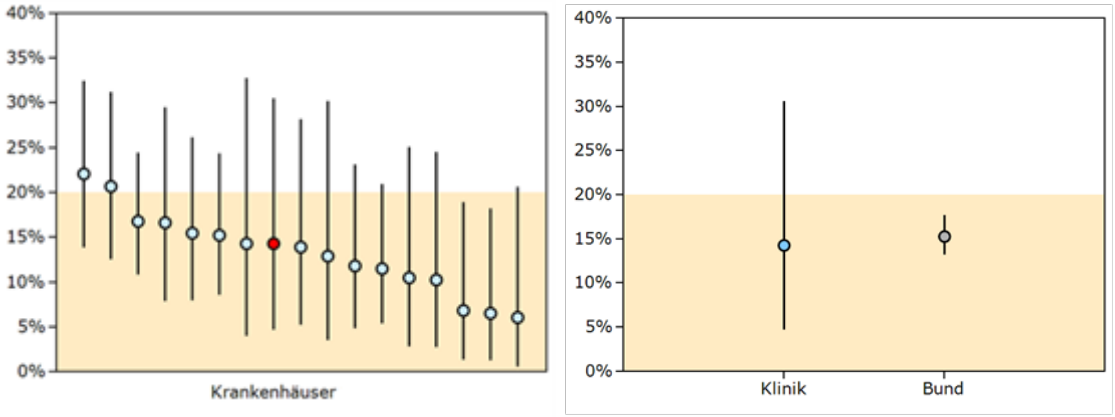

Abbildung 26: Beispiele für ein Fehlerbalkendiagramm (Benchmarkreport)

Benchmarkdiagramm

Benchmarkdiagramme werden derzeit in den Bundesauswertungen verwendet. Ein Benchmarkdiagramm zeigt dort die Indikatorwerte der einzelnen Krankenhäuser. Für jedes Krankenhaus wird das Indikatorergebnis durch einen Balken dargestellt. Die Balken werden so angeordnet, dass Krankenhäuser mit guten Ergebnissen immer rechts stehen. Die Beschriftung der x-Achse zeigt die absolute Anzahl der Krankenhäuser.[50]

Ein Benchmarking ist prinzipiell möglich, weil der durch einen Balken angezeigte QI-Wert eines einzelnen Krankenhauses unmittelbar mit den Indikatorwerten der anderen Krankenhäuser verglichen werden kann. Der Referenzbereich des Indikators (sofern definiert) wird als farbig unterlegte Fläche (gelb) kenntlich gemacht.

[49] Es ist allerdings möglich, dass ein Konfidenzintervall nicht erkennbar ist, weil seine Grenzen – abhängig vom Maßstab der y-Achse – sehr eng beieinander liegen. Dies ist dann der Fall, wenn sie innerhalb des Kreises liegen, der den Bundeswert markiert.
[50] Wegen der begrenzten Größe der Grafiken entsteht ein „flächiger“ Eindruck, wenn die Balken sehr vieler Krankenhäuser auf engem Raum nebeneinander gestellt werden.

Derzeit werden separate Diagramme für Häuser mit 1 bis 19 Fällen bzw. 20 Fällen und mehr bereitgestellt.

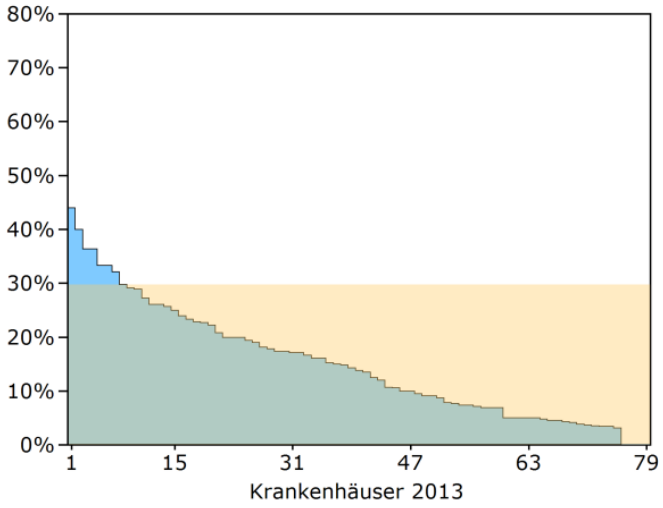

Abbildung 27: Benchmarkdiagramm (Beispiel für Krankenhäuser mit mindestens 20 Fällen)

Box-and-Whisker-Plots

Ein Box-and-Whisker-Plot bzw. Boxplot (Kastengrafik) fasst verschiedene Streuungs- und Lagemaße mindestens intervallskalierter Ergebnisdaten in einer Grafik zusammen. Seine Verwendung ermöglicht einen vereinfachten Überblick über die Verteilung der wichtigsten Werte: über den sogenannten Interquartilbereich, in dem 50 % aller Werte einer Verteilung liegen (gekennzeichnet als „Box“), einen variabel bestimmten Verteilungswert außerhalb des Interquartils (gekennzeichnet durch die über die Box hinausgehenden Linien bzw. „Whiskers“), den minimalen und maximalen Wert der Verteilung (gekennzeichnet durch das Sternsymbol) sowie den Median, also den mittleren Wert der Verteilung (gekennzeichnet als Strich innerhalb der Box).

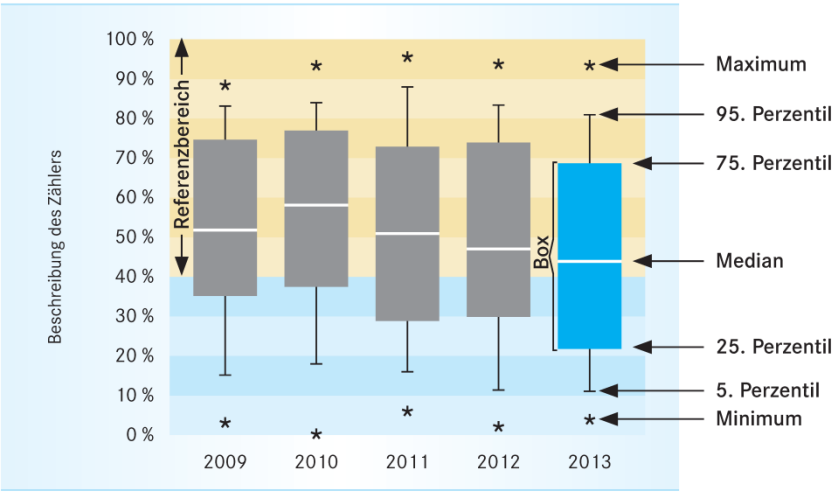

Abbildung 28: Beispiel für einen Box-and-Whisker-Plot (Kastengrafik)

Das AQUA-Institut verwendet diese Darstellung in den Bundesauswertungen, den Qualitätsreporten und in den Rückmeldeberichten für die Leistungserbringer in den direkten bzw. auch in den künftigen sektorenübergreifenden Verfahren. Sie werden als sogenannte gruppierte Boxplots zum Vergleich unterschiedlicher Analyseeinheiten genutzt; entweder als vergleichende Darstellung der Indikatorenergebnisse für aufeinanderfolgende Berichtsjahre oder als Vergleich des Einrichtungsgebnisses und des Bundeswertes eines Qualitätsindikators. In der Darstellung der Bundesauswertung ist zusätzlich der Bundeswert des Indikators – berechnet auf Fallbasis – eingezeichnet.

Säulendiagramm

Säulendiagramme werden verwendet, um Häufigkeitsverteilung diskreter Werte darzustellen. Das AQUA-Institut verwendet Säulendiagramme derzeit im Qualitätsreport. Zum einen finden Säulendiagramme Verwendung um die (absolute) Anzahl der Sentinel-Events (auf Patientenebene) bzw. die (absolute) Anzahl der Krankenhäuser mit Sentinel-Events darzustellen. Zum anderen werden sie zur Darstellung risikoadjustierter Indikatoren verwendet, bei denen das Verhältnis der beobachteten (O) und der statistisch erwarteten Ereignisse (E) unter Berücksichtigung der Risikostruktur abgebildet wird.

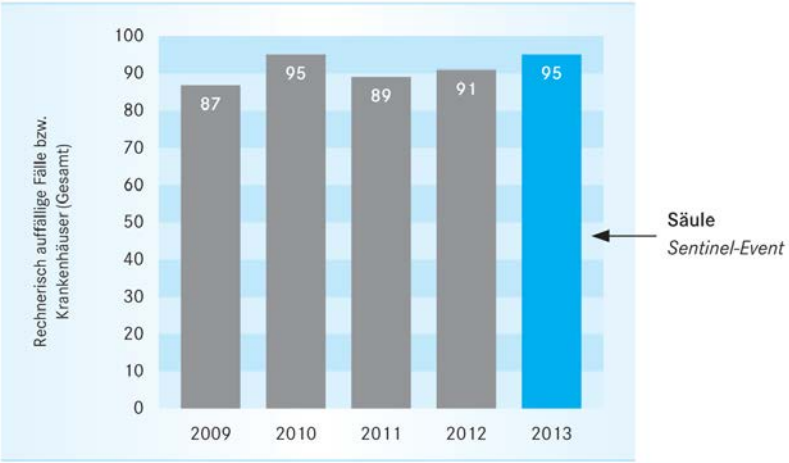

Abbildung 29: Beispiel für ein Säulendiagramm (Sentinel-Events)

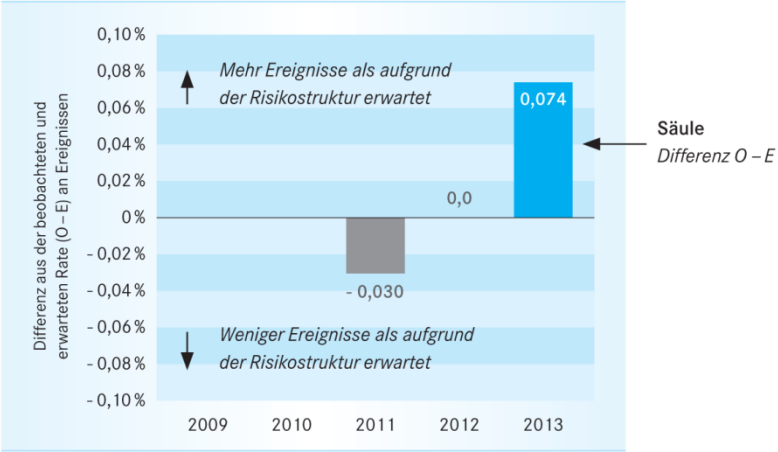

Abbildung 30: Beispiel für ein Säulendiagramm (Differenz O – E)

Geodarstellung

Im Zusammenhang mit der Aufgabe der bundesweiten Auswertung der Qualitätssicherungsdaten wurde das AQUA-Institut vom G-BA auch damit beauftragt, mittelfristig neue Darstellungsformen für die Ergebnisse der Qualitätssicherung zu entwickeln. Die seit 2012 jährlich veröffentlichte „Länderauswertung mit Geodarstellung für veröffentlichungspflichtige Indikatoren“ trägt diesem Auftrag Rechnung. Die Ergebnisse der veröffentlichungspflichtigen Indikatoren (Erfassungsjahr 2013 = 295 Indikatoren) sind hier in Form eines grafisch aufbereiteten Ländervergleichs und im zeitlichen Verlauf dargestellt.

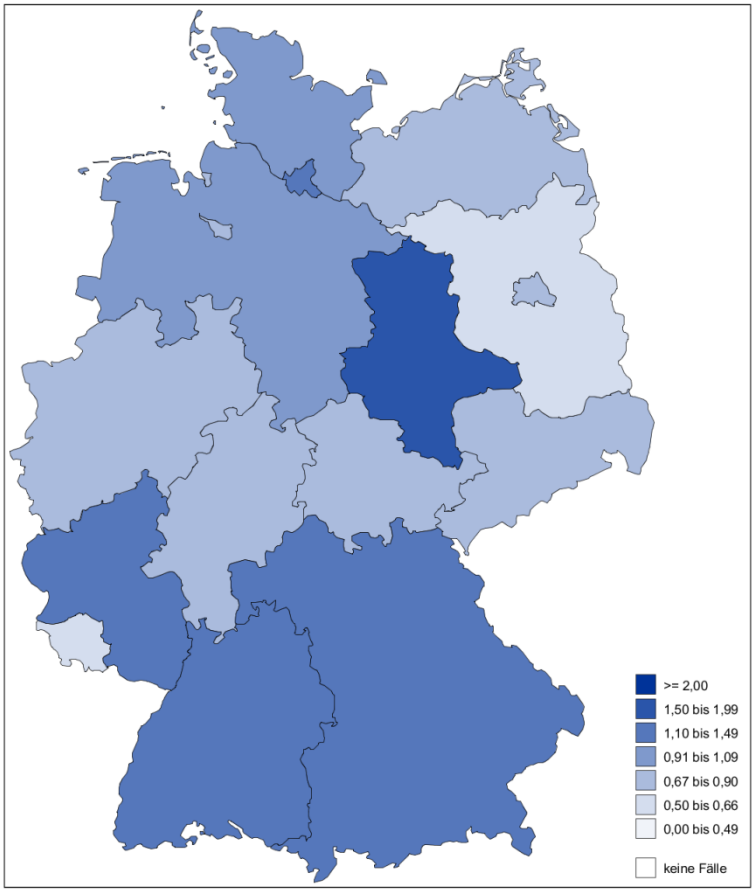

Abbildung 31: Beispiel Geodarstellung

12.4.2 Weiterentwicklung der Qualitätsanalysen und -darstellungen

Im Rahmen der Weiterentwicklung sind einerseits einige grundsätzliche Überlegungen mit Blick auf Auswertungskonzepte von sektorenübergreifenden Verfahren anzustellen und andererseits weitere, im Folgenden dargestellte grafische Darstellungsformen von Qualitätsanalysen zu berücksichtigen.

Trichterdiagramme (Funnel Plots)

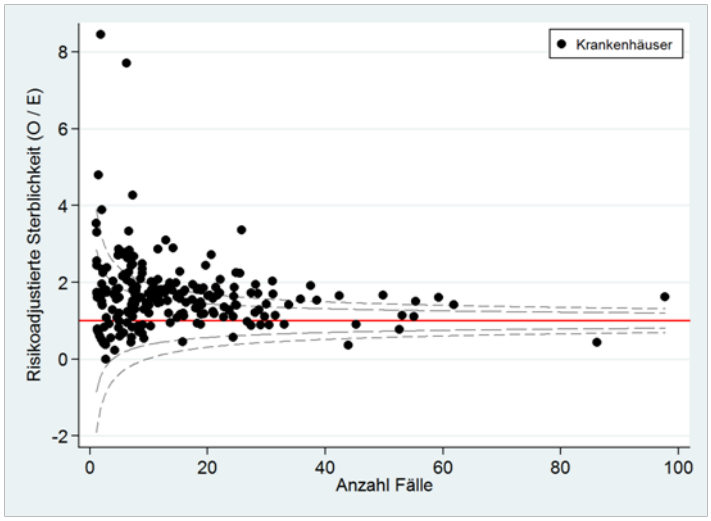

Abbildung 32: Beispiel eines Trichterdiagramms (Funnel Plot)

Trichterdiagramme (Funnel Plots) sind als Darstellungsformen zur Weiterentwicklung der Bundesauswertungen gedacht und stellen Streudiagramme von einrichtungsspezifischen Ergebnissen (Y-Achse) mit den erbrachten

Fallzahlen (X-Achse) dar. Zusätzlich kann anhand der abgebildeten Konfidenzintervalle, die den Trichter bilden, beurteilt werden, ob eine statistisch bedeutsame Abweichung vom Mittel oder ggf. Referenzwerten vorliegt (Lack et al. 2009). Damit können die Werte von signifikant abweichenden Einrichtungen für einen bestimmten Indikator, auch für kleinere Fallzahlen, identifiziert werden. Alle Einrichtungen außerhalb des Trichters sind außerhalb des Konfidenzintervalls und somit auffällig.

Kumulative Summenplots (CUSUM)

Daneben existieren seit Längerem weitere Anwendungen des Konzepts von verlängerten Analyseintervallen in Form von „kumulativen Summenplots“ bzw. *Quality Control Charts* (Rogers et al. 2004; Tekkis et al. 2003; Poloniecki et al. 1998).

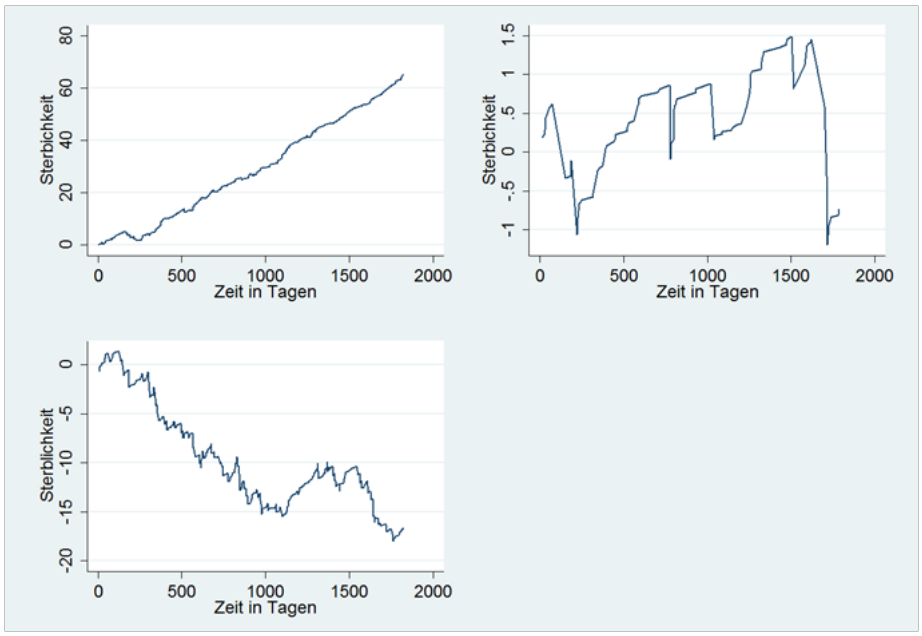

Abbildung 33: Beispiele eines Summenplots

Das AQUA-Institut verwendet die kumulierten Summenplots bisher noch nicht, sie sind aber als Weiterentwicklung der Benchmarkreports vorgesehen. In der statistischen Prozess- und Qualitätskontrolle ist die kumulative Summe oder CUSUM (***cu**mulative **sum***) eine sequenzielle Analysemethode zur Entdeckung von Änderungen in einer sequenziellen Datenreihe. In den obigen Beispielen sind die Abweichungen zwischen erwarteten und beobachteten Werten eines Indikators dargestellt.

In den dargestellten Grafiken wurde die Summe der Differenz zwischen der erwarteten Sterblichkeit und der beobachteten Sterblichkeit im Zeitverlauf abgebildet. So generiert die erste dargestellte Klinik mehr als 60 „potenziell vermiedene Todesfälle“. Die zweite Klinik zeigt absolut deutlich geringere Schwankungen ohne eindeutige Tendenz. Durch die größere Schwankung bei gleichzeitig kleinerer Metrik auf der Y-Achse lässt sich ableiten, dass diese Klinik weniger Fälle behandelt hat. Die letzte Klinik weist in etwa 17 Todesfälle mehr auf, als aufgrund der behandelten Erkrankungsschwere der Patienten zu erwarten gewesen wäre.

Analyse von gleitenden Durchschnittswerten

Ein typisches Problem der Ergebnisdarstellung betrifft das Fallzahl-Prävalenz-Problem (König et al. 2014). Eine Möglichkeit diesem zu begegnen, ist die Verwendung von gleitenden Durchschnittswerten (*Moving Averages*). Allerdings ist ggf. nachteilig, dass in das aktuelle Indikatorergebnis Prozesse einfließen, die schon länger zurückliegen und möglicherweise auf Probleme verweisen, die nicht mehr aktuell sind. Dies führt dazu, dass derartig konstruierte Indikatoren für ein Qualitätsmanagement nur eingeschränkt genutzt werden können, möglicherweise aber zum Zweck einer öffentlichen Berichterstattung hilfreich sein können (AQUA 2012c).

Das Problem der kleinen Fallzahlen verschärft sich – bezogen auf alle verwendeten Indikatoren – noch einmal bei unterjährigen (vierteljährlichen) Auswertungen. Beim gleitenden Durchschnitt wird eine kumulierte Periode

definiert (2, 4 oder 6 Quartale) und für die nachfolgenden Auswertungen fortgeschrieben. Da die Fallzahlen je nach Leistungsbereich und Indikator sehr unterschiedlich ausfallen, wäre die Möglichkeit, unterschiedlich lange Perioden zu parametrisieren, eine sinnvolle Ergänzung.

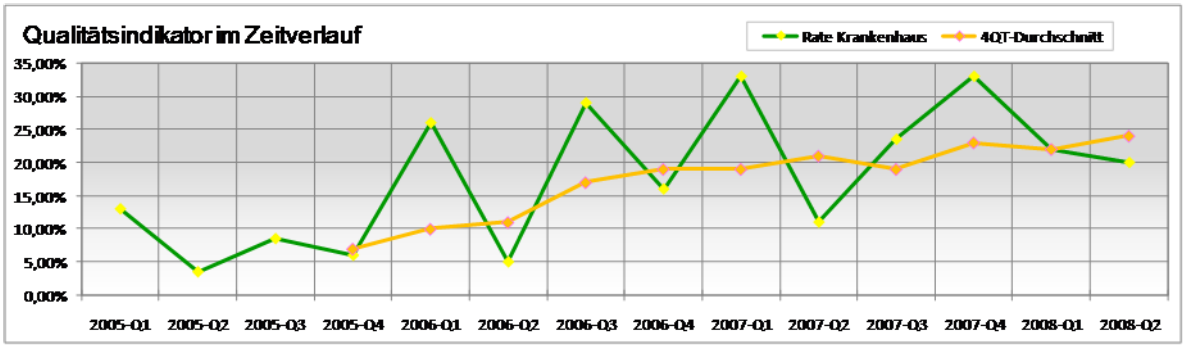

Abbildung 34: Indikatorverlauf bei vierteljährlicher Auswertung und Verwendung eines gleitenden Durchschnitts von 4 Quartalen (Eckardt 2008)

Abbildung 34 zeigt den beschriebenen Effekt: Bei niedriger Prävalenz führen Einzelergebnisse zu großen Schwankungen im zeitlichen Verlauf. Die orangefarbene Kurve entsteht bei Verwendung von gleitenden Durchschnittswerten. Wie im Beispiel gezeigt, funktioniert diese Methode auch jahresübergreifend, wenn man die Gleitphase mit jedem Quartal fortführt.

Kaplan-Meyer-Kurven

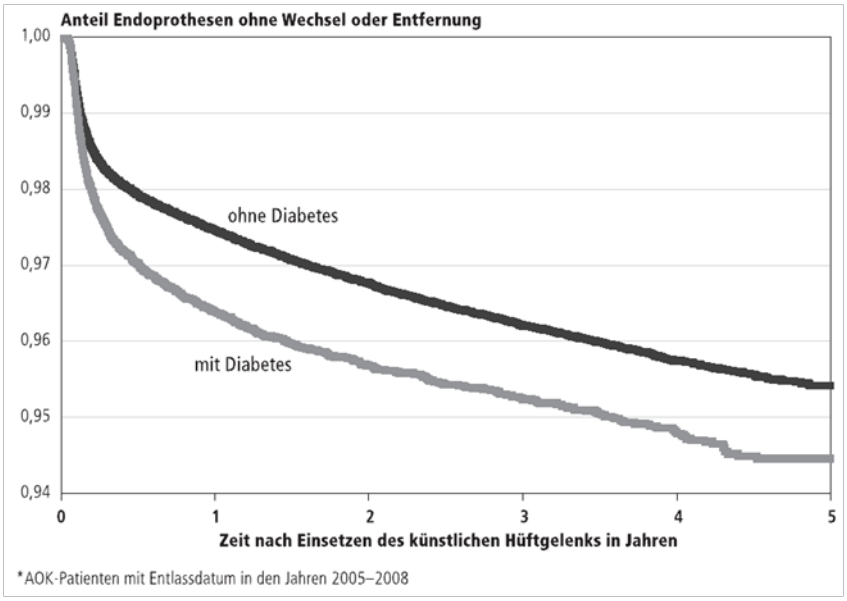

Abbildung 35: Überlebenszeitanalyse nach Implantation von Hüftendoprothesen bei Patienten mit und ohne Diabetes

Wenn Daten mit langen Nachbeobachtungszeiten (z.B. bei Transplantationen) analysiert werden, sind sogenannten Kaplan-Meyer-Kurven üblich. In Abbildung 35 sind die Standzeiten von Hüftgelenksimplantaten bei über 170.000 AOK-Patienten mit Entlassungsdatum zwischen 2005 und 2008, einer Gesamtbeobachtungsdauer von mehr als 480.000 Personenjahren und einer Nachbeobachtungsdauer von bis zu fünf Jahren getrennt nach Vorliegen oder Fehlen einer Diabetes-Erkrankung dargestellt.

Mit dieser Darstellungsform können Unterschiede von Behandlungen im Follow-up unter Berücksichtigung von sogenannten Zensierungen (neutrale Ausfälle) besonders gut und anschaulich dargestellt werden. Wenn zusätzlich Variablen zur Risikoadjustierung verwendet werden sollen, sind z.B. Cox Regressionen oder aber marginale strukturelle Modelle zu erwägen.

Verteilungsdiagramm
Verteilungsdiagramme bieten insbesondere die Möglichkeit, Ergebnisse aus Patientenbefragungen einrichtungsindividuell darzustellen. Folgende Grafik wurde am Beispiel der Frage: „Hatten Sie die Möglichkeit einer Ärztin/einem Arzt oder jemandem vom Pflege- bzw. Praxispersonal Fragen zu stellen, die Ihre Entlassung betrafen?“ erstellt.

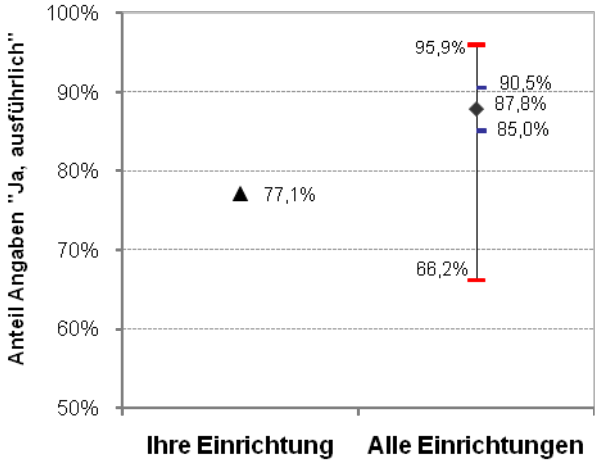

Abbildung 36: Beispiel für eine mögliche Darstellung der Patientenbefragungsergebnisse

Legende:
— Maximum
– 75 %
♦ Median
– 25 %
— Minimum

▲ Ihr Wert

Anzahl der antwortenden Patienten
Ihre Einrichtung: n = 105
Alle Einrichtungen n = 1.958

In der Abbildung ist sowohl der Wert für die einzelne Einrichtung auf der linken Seite wie auch die Verteilung der Werte in der Grundgesamtheit mit Median, Quartilsgrenzen, Minimum und Maximum dargestellt.

12.4.3 Allgemeinverständliche Aufbereitung von Qualitätsinformationen im Internet

Mit dem Auftrag zur „Aufbereitung einer aussagefähigen Internetdarstellung der Ergebnisse der Maßnahmen zur Versorgung von Früh- und Neugeborenen (NICU-Projekt)“ wurde ein neuer Weg der Qualitätsdarstellung in der externen Qualitätssicherung unternommen. Seit 2014 werden hier einrichtungsbezogene Auswertungen in einer vergleichenden Darstellung veröffentlicht (Thomas et al. 2014). Dabei richtet sich die Webseite sowohl an Laien (Eltern mit zu erwartender Geburt eines sehr kleinen Frühgeborenen) als auch an medizinische Experten (Fachärzte und Zuweiser). Darüber hinaus haben natürlich auch die Krankenhäuser die Möglichkeit, sich hinsichtlich ihrer Versorgungsergebnisse untereinander zu vergleichen. Auf Basis der ausgewerteten Daten der Neonatalerhebung werden die Informationen für die Zielgruppen unterschiedlich aufbereitet bzw. dargestellt.

Die Konzeption der sich an werdende Eltern, d.h. an im Regelfall an medizinische Laien, richtenden Teile der Webseite berücksichtigt dabei, dass die Standard-Indikatoren der Qualitätsberichterstattung – Mortalitätsraten, Überlebenszeiten, Komplikationen – oft nur wenig Einfluss auf die Auswahl eines bestimmten Krankenhauses durch Patienten haben (Simansiki et al. 2006; Robinson et al. 1997). Darüber hinaus ist bei der Darstellung von umfangreichen und unterschiedlichen Informationen die individuelle Nutzung der Informationen erschwert, was zu einer Verunsicherung der Patienten führen kann (Kelly et al. 2008).

Die Ergebnisse der Qualitätssicherung der Neonatalerhebung sind in einer praktisch anwendbaren Form aufbereitet. Über eine Suchmaske können entsprechende Perinatalzentren durch die Eingabe von Postleitzahlen und Umkreisentfernungen gesucht werden. Für die gefundenen Zentren werden auf der Hauptseite die Entfernungen zum eigenen bzw. zum eingegebenen Standort, die Anzahl behandelter Fälle sowie Ergebnisse zum Überleben ausgewiesen (vgl. Abbildung 37). Anhand dieser Ergebnisse sollen werdende Eltern unterstützt werden, informierte Entscheidungen zu treffen.

Abbildung 37: Patientengerechte Darstellung der Krankenhaussuche

Alle Ergebnisse werden auf einer tieferen Seitenebene für das medizinische Fachpublikum auch als detaillierte Fachinformationen ausgewiesen (vgl. Abbildung 38). Für diese stehen die ausgewerteten Daten der Neonatalerhebung inklusive der Angaben zur Berechnung in fachlich-adäquater Darstellung zur Verfügung. Diese Darstellung richtet sich vor allem an Ärzte und potenzielle Zuweiser. Letztlich sind die Ärzte für viele Patienten die wichtigsten Ansprechpartner und deren klinischer, sozialer und kommunikativer Kompetenz wird im Allgemeinen eine große Bedeutung für die eigene Entscheidung beigemessen (Singer et al. 2009; Epstein et al. 2005; Grande et al. 2005).

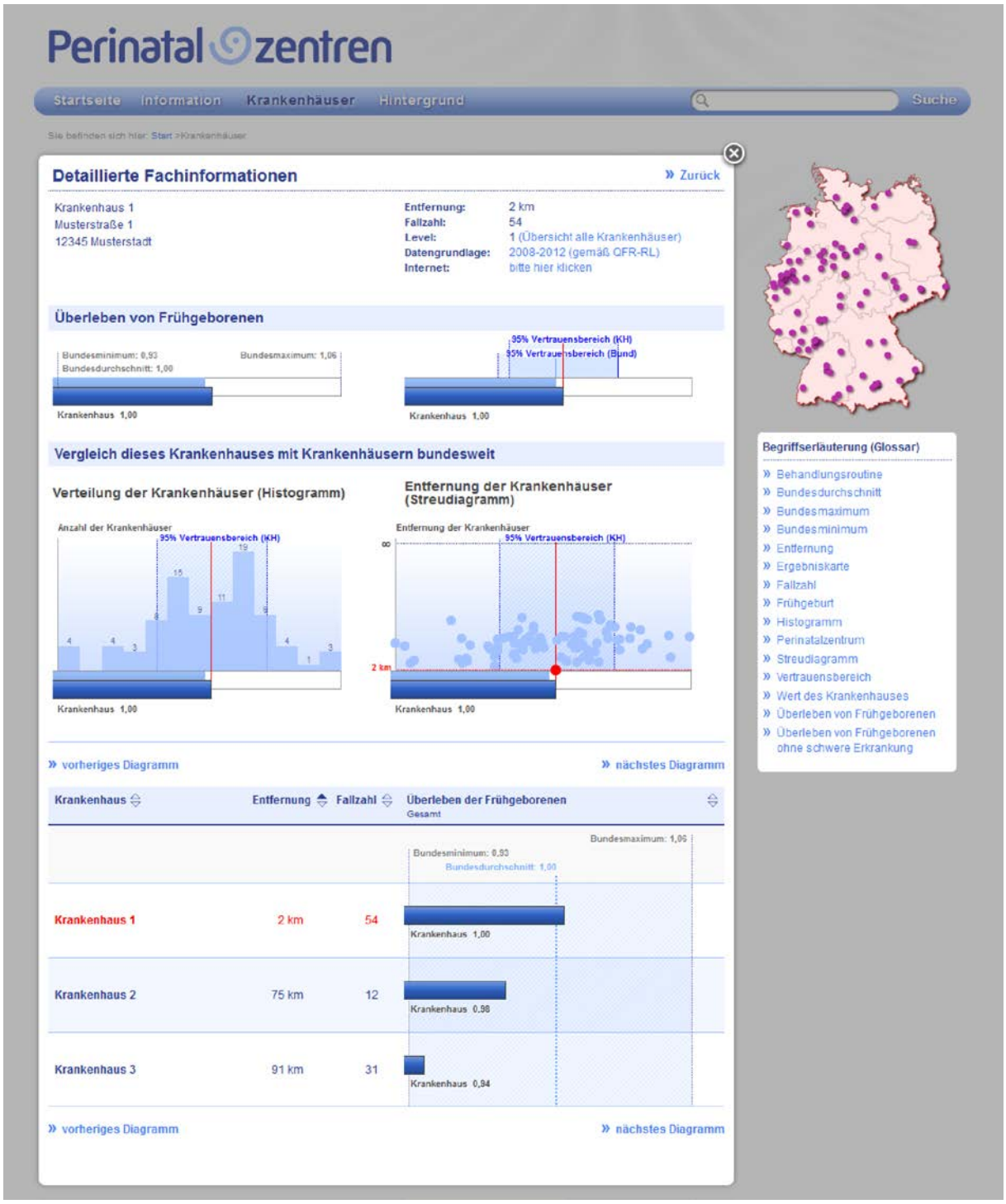

Abbildung 38: Darstellung der Fachinformationen

Glossar

Begriff	Beschreibung
Abrechnungsdaten	Daten, die von Leistungserbringern über von ihnen erbrachte Leistungen zum Zweck der Abrechnung mit den Kostenträgern dokumentiert und übermittelt wurden.
Additiver Score	Methode der Risikoadjustierung, bei der das Risiko der Patienten für den untersuchten Endpunkt, z.B. Versterben, durch die Addition von Punkten für unterschiedliche Merkmale (z.B. Diabetes = 2, koronare Herzerkrankung = 3) ermittelt wird.
Akkuratesse	engl.: accuracy. Dokumentationsqualitätskriterien, die Rückschlüsse auf die Übereinstimmung von Datensätzen mit den Referenzdaten zulassen.
Anonymisierung	Verändern personenbezogener Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können. (§3 Abs. 6a BDSG)
Area-Indikator	Regionales und/oder geografisch begründetes Auswertungskonzept von Indikatoren. Ein Area-Indikator wertet aggregierte Daten für ein definiertes Gebiet aus (z.B. Bundesland, Kassenbezirk).
ASA-Klassifikation	In der Medizin weitverbreitetes Schema der American Society of Anesthesiologists (ASA) zur Stratifizierung von Patienten nach ihrem allgemeinen Gesundheitszustand (ASA 1 bis 6). Anwendung findet das Schema vor allem zur Einschätzung von individuellen Operationsrisiken.
Auffälligkeit, qualitative	Teilmenge der rechnerischen Auffälligkeiten, die in einem qualitativen Beurteilungsprozess (z.B. Strukturierter Dialog) als Qualitätsproblem einer Einrichtung klassifiziert wurden.
Auffälligkeit, rechnerische	Das gemessene Ergebnis eines Indikators liegt außerhalb eines vorher festgelegten Referenzbereichs und/oder weicht im Verhältnis zu den gemessenen Ergebnissen anderer Einrichtungen signifikant ab. Durch rechnerische Auffälligkeiten werden Maßnahmen der Qualitätssicherung ausgelöst (siehe: Strukturierter Dialog).
Auslösekriterien	Algorithmus zur Auslösung der Dokumentationspflicht (QS-Filter).
Auswertungsstellen	Nehmen die Auswertung der QS-Daten nach bundeseinheitlichen Kriterien vor. Zu den Aufgaben der Auswertungsstellen gehören die Prüfung der an sie übermittelten Datensätze auf Vollzähligkeit, Vollständigkeit und Plausibilität, die patientenbezogene Zusammenführung von Datensätzen, die Auswertung von Datensätzen nach bundeseinheitlichen Kriterien, die Vorhaltung der Daten für Landesauswertungen (Nutzungsmöglichkeiten der Länder für länderbezogene Verfahren) und die Übermittlung angeforderter anonymisierter Datenauswertungen an den G-BA. Der G-BA beauftragt die Institution nach §137a SGB V oder eine andere Einrichtung als Bundesauswertungsstelle.
Basisauswertung	Auswertung der grundlegenden Angaben zu einem Leistungsbereich (Angaben zu Fallzahlen, Patienten, Aufenthaltsdauer, Diagnostik, Therapie usw.).

Begriff	Beschreibung
Basisprüfung, statistische	Statistische Analyse der gelieferten QS-Daten der Leistungserbringer nach vorab festgelegten Auffälligkeitskriterien (Plausibilität, Vollständigkeit, Vollzähligkeit). Zielt zusammen mit dem Strukturierten Dialog darauf ab, fehlerhafte, unvollständige und unvollzählige Dokumentationen in den QS-Daten zu identifizieren und gezielte Maßnahmen zur Verbesserung der Dokumentationsqualität zu initiieren.
Basisspezifikation	Gemeinsame Spezifikation für stationäre und sektorenübergreifende QS-Dokumentationen.
Benchmarking	Datengestützter Vergleich von Ergebnissen, Prozessen und Strukturen einzelner Institutionen mit definierten Zielgrößen (benchmark) bzw. mit aggregierten Daten anderer Institutionen.
Bundesauswertung	Auswertung und Darstellung der aggregierten Ergebnisse der bundesweiten Qualitätssicherung.
Bundesdatenpool	Zusammenführung aller bundesweit dokumentierten QS-Daten.
Bundesfachgruppen	Beratende Expertengruppen für die Verfahren der externen stationären Qualitätssicherung.
Bundesstelle	Vom G-BA nach §137a SGB V bestimmte Stelle zur Auswertung der QS-Daten auf Bundesebene.
Clusterstichprobe	siehe: Klumpenstichprobe
Content Validity	siehe: Inhaltsvalidität
Data mining	Algorithmen, die Muster in semi- oder unstrukturierten Daten identifizieren.
Daten	
~annahmestellen	Stellen, an die die Leistungserbringer oder andere Daten liefernde Stellen (z.B. Krankenkassen) die erhobenen Daten übermitteln. Sie prüfen die übermittelten Daten auf Plausibilität, Vollständigkeit und Vollzähligkeit (§9 Qesü-RL).

Begriff	Beschreibung
~ arten	1. Daten, deren formale Unterscheidung im Rahmen der bundesweiten externen Qualitätssicherung gesetzlich geregelt ist. Abhängig von ihrem spezifischen Zweck im Rahmen der gesetzlichen Qualitätssicherung wird zwischen a) personenbezogenen Daten (patienten- und/oder leistungserbringer-identifizierende Daten), b) Qualitätssicherungsdaten (für die Zwecke eines QS-Verfahrens ausgewiesene Behandlungsdaten) und c) administrativen Daten (Daten, die zu spezifischen Prüfzwecken oder zur technischen Organisation von Datenflüssen erhoben werden) unterschieden (§14 Qesü-RL). 2. Daten, die entsprechend ihres Inhalts unterschieden werden, z.B. a) Abrechnungsdaten (der Leistungserbringer), b) Routinedaten (siehe dort), c) Sozialdaten (der Krankenkassen; bestehend aus Versichertenstammdaten und Abrechungsdaten der Leistungserbringer), d) Patientenbefragungsdaten, e) Fall- oder Behandlungsdaten (der medizinischen QS-Dokumentation über Behandlungen) oder f) Strukturdaten (der Einrichtungen).
~basis	Im Einzelfall festzulegende bzw. festgelegte Gesamtmenge von auszuwertenden bzw. ausgewerteten Daten.
~extraktion	Standardisierte Vorgehensweise, um mithilfe strukturierter Formulare eine systematische Extraktion relevanter Daten aus Studien vorzunehmen.
~feld	Kleinste Einheit eines Datensatzes (z.B. Angabe des Geschlechts im Dokumentationsbogen).
~fluss	Übermittlung der Daten der QS-Verfahren in einem festgelegten Format und Inhalt, die vom Leistungserbringer über eine Datenannahmestelle, Vertrauensstelle (nur sektorenübergreifende PID-Verfahren) bis zur Datenauswertungsstelle gelangen. Die Datenflüsse sind grundsätzlich in der QSKH-RL und der Qesü-RL des G-BA festgelegt.
~quelle	hier: Der Ursprungsort, an dem Daten für die Qualitätssicherung generiert oder erhoben werden können, z.B. medizinische Einrichtungen, Krankenkassen, Patienten usw. Für die Erhebung stehen verschiedene Erhebungsinstrumente zur Verfügung.
~satz	Eine zusammenhängende Menge von QS-Daten, die einem Fall (beispielsweise einem Patienten) zugeordnet wird.
~sparsamkeit	Grundsatz, dass die Erhebung, Verarbeitung und Nutzung personenbezogener Daten nur streng zweckbestimmt erfolgen darf und für jeden Zweck möglichst wenige personenbezogene Daten erhoben werden dürfen (§3a Bundesdatenschutzgesetz).
~validierung	Verfahren zur Überprüfung von QS-Daten einerseits auf Vollzähligkeit, Vollständigkeit und Plausibilität (statistische Basisprüfung), andererseits ihre Übereinstimmung (Konkordanz) mit einer Referenzquelle wie bspw. der Krankenakte (Stichprobenverfahren mit Datenabgleich).

Begriff	Beschreibung
Diagnosis Related Groups (DRG)	Diagnosebezogene Fallgruppen zur pauschalierten Abrechnung stationärer Behandlungsfälle. Wesentliche Grundlage für die Ermittlung einer DRG, das sogenannte Grouping, sind Diagnosen- (ICD) und Prozeduren-Kodes (OPS).
Dialog, Strukturierter	siehe: Strukturierter Dialog
Diskriminationsfähigkeit	(Statistische) Unterscheidungsfähigkeit. Aussagekraft von Daten hinsichtlich tatsächlich bestehender Unterschiede.
Dokumentation	siehe: QS-Dokumentation
Dokumentationsentwicklung	„Entwicklung der notwendigen Dokumentation“ (§137a SGB V): Gestaltung der Erhebungsinstrumente durch die Konkretisierung von Datenfeldern zur Erhebung von QS-Daten und der Formulierung diesbezüglicher Ausfüllhinweise für Leistungserbringer.
Dokumentationsqualität	Güte der (medizinischen) QS-Dokumentation eines Leistungsbereiches, die durch die Größen Korrektheit (Übereinstimmung mit den Daten der Patientenakte), Vollständigkeit (einzelner Datensätze) und Vollzähligkeit (der Datensätze insgesamt) abgebildet wird.
Dokumentationsrate	Das Verhältnis dokumentierter Fälle im Verhältnis zu allen dokumentationspflichtigen Fällen. Siehe auch: Vollzähligkeit.
Einheitlicher Bewertungsmaßstab (EBM)	Vergütungssystem für die Abrechnung von Leistungen im vertragsärztlichen Bereich.
Empfehlungsgrad	auch: Empfehlungsstärke, engl.: Grade of Recommendation (GoR). Auf der Basis von Evidenzklassen abgegebene Behandlungsempfehlung in Leitlinien. Allgemein werden die Grade A („soll“), B („sollte“) und C („kann“) unterschieden. Die Grade werden von „allgemein üblicher medizinischer Praxis“ (GCP = „Good Clinical Practice) ergänzt, sofern keine Evidenzklassen für eine medizinische Behandlung vorliegen.
Effektivität	Zielerreichung bzw. Grad der Zielerreichung. Wirksamkeit (einer Behandlung).
Effizienz	Verhältnis zwischen einem erzielten Ergebnis und den dafür eingesetzten Ressourcen.
Einrichtung	siehe: Leistungserbringer
Endpunkte	Vorab definierte Ziele einer Behandlung.
patientenrelevante~	Endpunkte einer Behandlung, die aus Sicht der Patienten wesentlich sind.
therapierelevante~	Endpunkte einer Behandlung, die aus Sicht der Behandler wesentlich sind.
Erfassungsjahr	Das Jahr, in dem die Daten erhoben werden. Hierauf beruhen die Ergebnisse der Indikatoren. Die Kriterien für die Abgrenzung des Erfassungsjahres sind in der Spezifikation zur Sollstatistik definiert.
Erhebungsinstrument	Spezifizierung der Art und Weise der Erhebung von Daten (für die externe Qualitätssicherung) bzw. der Erschließung/Nutzbarmachung von Datenquellen. Beispiele: ▪ Dokumentationsbögen zur Datenerhebung durch die Leistungserbringer (QS-Dokumentation) ▪ Fragebögen zur Datenerhebung bei Patienten ▪ Technische Spezifikationen zur Erhebung von Routinedaten (z.B. aus den Sozialdaten bei den Krankenkassen)

Begriff	Beschreibung
Ethikvotum	In einem Ethikvotum zertifiziert eine Ethikkommission, ob ein geplantes Forschungsprojekt im Einklang mit ethischen Richtlinien steht. Ein eingereichter Antrag wird nach ethischen, rechtlichen (insbesondere auch datenschutzrechtlichen) und medizinisch-wissenschaftlichen Gesichtspunkten unter Berücksichtigung der Deklaration von Helsinki sowie internationalen Richtlinien von einer unparteilichen Sachverständigenkommission beraten.
Evidence-based Health Care	Evidenzbasierte Gesundheitsversorgung. Übertragung der Prinzipien der evidenzbasierten Medizin von der Ebene der individuellen Patientenbehandlung auf die Ebene der Versorgung von Patientengruppen und Populationen sowie auf das Gesundheitssystem (Management, Gesundheitspolitik).
Evidenzbasierte Medizin (EbM)	Abgeleitet von engl. „evidence-based medicine"; auf empirische Beweiskraft gestützte Medizin. Schulbildende medizinische Richtung, die die bestmögliche Versorgung von Patienten durch die Kombination von bester verfügbarer externer Evidenz (klinische Studien), individueller medizinisch-therapeutischer Expertise und Patientenpräferenzen zu erreichen sucht.
Evidenzklassen	Medizinische Studien können hinsichtlich ihres Studiendesigns in verschiedene Güteklassen (Evidenzklassen) eingeteilt werden.
Face Validity	„Augenscheinliche" Objektivität. Dabei handelt es sich um keine Validität im technischen Sinn. Der Begriff bezieht sich nicht darauf, was bspw. ein Test misst, sondern darauf, was er oberflächlich zu messen scheint, d.h. ob es „so aussieht", als würde er die relevanten Aspekte messen. Face Validity ist nicht objektiv im engeren Wortsinn und daher von der Inhaltsvalidität abzugrenzen; sie ist jedoch als wichtiger Faktor für die Wirkung von Tests anzusehen.
Fallzahl	Anzahl durchgeführter Behandlungen, z.B. in einem bestimmten Leistungsbereich in einem Krankenhaus.
Follow-up	auch: Mehrpunktmessung, Mehrzeitpunktmessung. Versorgungsleistungen, deren Qualität (z.B. ein Behandlungserfolg) nicht einmalig, sondern zu unterschiedlichen weiteren Zeitpunkten gemessen wird. Die Messungen können in einem bestimmten oder auch in unterschiedlichen Sektoren (vgl. QSKH-RL) stattfinden.
Fuzzy Search	Datenbanksuche, die neben identischen Werten oder Wertemengen auch solche Werte und Wertemengen findet, die sich nur in wenigen Teilen unterscheiden.
Gemeinsamer Bundesausschuss (G-BA)	Oberstes Beschlussgremium der gemeinsamen Selbstverwaltung der Ärzte, Zahnärzte, Psychotherapeuten, Krankenhäuser und Krankenkassen in Deutschland. Er bestimmt in Form von Richtlinien den Leistungskatalog der Gesetzlichen Krankenversicherung (GKV) für mehr als 70 Millionen Versicherte und legt damit fest, welche Leistungen der medizinischen Versorgung von der GKV erstattet werden.
Health Technology Assessment (HTA)	Systematische und mit hohen methodischen Standards durchgeführte umfassende Bewertung medizinischer Technologien, Prozeduren und Hilfsmittel sowie auch der organisatorischen Strukturen und des gesellschaftspolitischen Umfeldes, in denen sie eingesetzt werden.

© 2015 AQUA-Institut GmbH

Begriff	Beschreibung
ICD-GM	engl.: International Classification of Diseases – German Modification. Auf deutsche Verhältnisse angepasste Version der ICD, die vom Deutschen Institut für Medizinische Dokumentation und Information (DIMDI) erstellt wird.
IK-Nummer	Institutionskennzeichen. Wird gemäß §293 SGB V bei der Datenübermittlung zwischen den gesetzlichen Krankenkassen und den Leistungserbringern als eindeutige Identifizierung verwendet.
Indexleistung	Medizinische Leistung, die den Startpunkt für ein QS-Verfahren (die QS-Auslösung) genau definiert (im Regelfall OPS-Kodes).
Indexpatient	Patient, der eine Indexleistung erhalten hat und in einem QS-Verfahren für einen definierten Zeitraum dokumentationspflichtig ist.
Indikator	„Anzeiger“. Auch: Qualitätsindikator. Quantitatives Maß zur Bewertung der Erreichung eines Qualitätsziels einer bestimmten gesundheitlichen Versorgungsleistung.
Ergebnis~	Bildet Behandlungsergebnisse ab bzw. bewertet diese. Er zeigt, wie oft ein gewünschtes oder unerwünschtes Behandlungsergebnis festgestellt wird.
Prozess~	Bildet Versorgungsprozesse ab bzw. bewertet diese. Er erfasst für definierte Behandlungsschritte (als Teile des Versorgungsprozesses), ob oder in welcher Häufigkeit ein definierter Prozess durchgeführt wurde.
Struktur~	Bildet Versorgungsstrukturen ab bzw. bewertet diese. Er bezieht sich auf sachliche Voraussetzungen der Leistungserbringung, die Einfluss auf das Behandlungsergebnis haben. Dazu können bauliche Gegebenheiten genauso wie die apparative und personelle Ausstattung gehören.
Indikatorenregister	Durch das AQUA-Institut erstellte Liste von recherchierten und neu entwickelten Indikatoren eines (neuen) QS-Verfahrens. Die Indikatoren des Registers werden einem strukturierten Auswahl- und Bewertungsprozesses (RAM-Prozess) zugeführt, an dessen Ende ein abgestimmtes Indikatorenset steht.
Indikatorenset	Gesamtheit der im Anschluss an den strukturierten Auswahl- und Bewertungsprozess konsentierten Indikatoren des Indikatorenregisters.
Inhaltsvalidität	engl.: Content Validity. Teilaspekt der Validität, der die Vollständigkeit der Inhalte eines zu messenden Konstrukts bezeichnet.
Instrument	siehe: Erhebungsinstrument In einem weitergehenden Verständnis beschreibt ein Instrument im Bereich der Qualitätsförderung eine Methode bzw. ein Werkzeug, das mit dem Ziel der Verbesserung der Versorgungsqualität eingesetzt wird.
Interessenkonflikt	auch: Conflict of Interest. Bezeichnet das Risiko, dass ein professionelles Urteilsvermögen oder Handeln, welches sich auf ein primäres Interesse bezieht, durch ein sekundäres Interesse unangemessen beeinflusst wird.
Inzidenz	statistisch: siehe Inzidenzrate. Epidemiologisch: Maß zur Charakterisierung des Krankheitsgeschehens in einer bestimmten Population.
Inzidenzrate	In einem bestimmten Zeitraum neu aufgetretene Anzahl an Krankheitsfällen in einer definierten Population (Zähler: Anzahl der neu Erkrankten; Nenner: Betrachtete Zeitspanne x Anzahl der betrachteten Individuen).
Kennzahl	siehe: Indikator

Begriff	Beschreibung
Klumpenstichprobe	auch: Clusterstichprobe. Besondere Form der Zufallsauswahl. Anders als in der einfachen Zufallsauswahl werden hier nicht einzelne Merkmalsträger ausgewählt, sondern ganze Gruppen („Klumpen“).
Konfidenzintervall	Intervall um einen berechneten Ergebniswert, z.B. um das Gesamtergebnis eines Indikators. Vereinfacht ausgedrückt gibt das Konfidenzintervall den Bereich an, in dem der tatsächliche Wert eines Qualitätsindikators unter Berücksichtigung aller zufälligen Ereignisse (z.B. Dokumentationsfehler) mit einer bestimmten, zuvor festgelegten Wahrscheinlichkeit liegt.
Konkordanz	auch (bisher): Korrektheit. Übereinstimmung der Daten mit einer Referenzquelle. Im Kontext der Qualitätssicherung in der Regel die Übereinstimmung von gelieferten Qualitätsdaten mit der Krankenakte.
Konstruktvalidität	Teilaspekt der Validität, mit der die Messeigenschaften eines Konstrukts bezeichnet werden; Konstruktvalidität liegt vor, wenn Messungen das erfassen, was sie erfassen sollen.
Kostenträger	Personen und Institutionen, die die Kosten für medizinische Versorgungsleistungen tragen. Im Rahmen der gesetzlich verankerten Qualitätssicherung sind dies die gesetzlichen Krankenkassen.
Landesauswertungen	Auswertungen von Qualitätsindikatoren nach bundeseinheitlichen Vorgaben für ein Bundesland oder einen Teil eines Bundeslandes. Anhand von Landesauswertungen können Leistungserbringer (Krankenhäuser, Arztpraxen) ihr eigenes Ergebnis mit den Ergebnissen anderer Leistungserbringer anonym vergleichen.
Lebensqualität	engl.: Quality of Life. Abseits unterschiedlicher philosophischer Grundbestimmungen umschreibt Lebensqualität ein „ganzheitliches“ Konzept der menschlichen Befindlichkeit, das über die physische Gesundheit hinaus auch subjektive, soziale, materielle und immaterielle Aspekte beinhaltet.
Leistungsbereich	hier: Leistungsbereich der gesetzlichen Qualitätssicherung (auch: QS-Verfahren). Medizinischer Themen- oder Behandlungsbereich, der im Rahmen der bundesweiten Qualitätssicherung dokumentationspflichtig ist.
Leistungserbringer	Personen und Einrichtungen, die medizinische Versorgungsleistungen erbringen bzw. bereitstellen. Der Begriff wird im SGB V auch für Ärzte und ärztliche Einrichtungen sowie für zugelassene Krankenhäuser gem. §108 SGB V genutzt.
Leitlinie, medizinische	Systematisch entwickelte, aber unverbindliche Orientierungs- bzw. Entscheidungshilfe v.a. für Ärzte und andere medizinische Leistungserbringer bezüglich einer angemessenen medizinischen Versorgung. Die Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) hat mittlerweile ein weithin anerkanntes 3-stufiges Klassifikationssystem (S1 bis S3) zur Bewertung des Entwicklungsstandes und damit der Güte von Leitlinien erstellt.
Medizin, evidenzbasierte	siehe: Evidenzbasierte Medizin
MeSH-Terms	Medical Subject Headings (MeSH). Polyhierarchischer Thesaurus zur indexierten Sacherschließung von Büchern und Zeitschriftenartikeln in Medizin und Biowissenschaften.
Missing Values	„Fehlende Werte“, z.B. fehlende Antworten und nicht auswertbare Antworten bei der Auswertung eines Fragebogens.

© 2015 AQUA-Institut GmbH

Begriff	Beschreibung
Morbidität	Häufigkeit einer Krankheit innerhalb einer bestimmten Bevölkerungsgruppe zu einem bestimmten Zeitpunkt oder in einem definierten Zeitraum. Die Morbidität kann meist nur geschätzt werden. Grundlagen solcher Schätzungen sind die Inzidenz(rate) und die Prävalenz(ratio).
Mortalität	Sterblichkeit(srate). Die Anzahl der Todesfälle bezogen auf die Gesamtbevölkerung oder eine demografisch definierte Gruppe; meist auf einen Zeitraum bezogen.
Multiple logistische Regression	Statistisches Verfahren der Risikoadjustierung zur Analyse des Einflusses verschiedener Größen (z.B. Alter, Geschlecht oder Begleiterkrankungen) auf eine binäre, d.h. nur eine von zwei Möglichkeiten zulassende Zielvariable (z.B. „Patient verstorben“: ja/nein). Jeder Patient einer Einrichtung wird nur mit Patienten der gleichen Ausprägung der Einflussgrößen (z.B. gleiche Geschlechts- und Altersgruppe, gleiche Begleiterkrankungen) verglichen.
NICU-Projekt	engl.: Neonatal Intensive Care Unit. Entwicklung einer aussagefähigen Internetdarstellung durch das AQUA-Institut im Auftrag des G-BA zur Auswertung und laienverständlichen Aufbereitung der Ergebnisqualität von Perinatalzentren der Level 1 und 2 in Deutschland, die sich an der Versorgung von Früh- und Neugeborenen mit sehr niedrigem Geburtsgewicht beteiligen.
Observed to Expected Ratio (O/E)	Das Verhältnis aus der beobachteten (O = observed) und der erwarteten (E = expected) Rate ist ein für die Risikoadjustierung wichtiger Wert. Der Wert O ist die Rate (Quotient) aus den tatsächlich beobachteten Ereignissen („roh“, d.h. ohne Risikoadjustierung) und der Grundgesamtheit der Fälle im betreffenden Erfassungsjahr. Der Wert für die erwartete Rate E ergibt sich als Verhältnis der erwarteten Ereignisse und der Grundgesamtheit der Fälle des Erfassungsjahres. Zur Berechnung des Risikoprofils werden Regressionsmodelle herangezogen, die auf Daten des Vorjahres zurückgreifen. Ein O/E-Wert von 1,20 bedeutet, dass die beobachtete Rate an Ereignissen um 20 % größer ist als die erwartete Rate. Umgekehrt bedeutet ein O/E-Wert von 0,90, dass die beobachtete Rate an Ereignissen um 10 % kleiner ist als die erwartete. Der O/E-Wert ist im Regelfall dimensionslos. Um die Entwicklung über verschiedene Jahre hinweg abzubilden, wird die Differenz aus dem O- und E-Wert für das jeweilige Jahr dargestellt. Wenn die beobachtete Rate (O) an Ereignissen der erwarteten Rate (E) entspricht, dann liegt die Differenz (O – E) aus beobachteter Rate (O) und erwarteter Rate (E) bei Null.
Operationen- und Prozedurenschlüssel (OPS)	Kodierungssystem für medizinische Leistungen bzw. verbindlicher Abrechnungsschlüssel zwischen medizinischen Leistungserbringern und Kostenträgern (Gesetzliche Krankenkassen).
Outcome	Ergebnis einer medizinischen Behandlung.
Panelverfahren	siehe: RAM-Prozess
Patient Reported Outcomes (PRO)	Sammelbegriff für die Einschätzung und Bewertung von medizinischen Behandlungserfolgen aus der Patientenperspektive. PRO werden direkt beim Patienten ermittelt (z.B. durch Fragebogen, Interview).
Patientenfragebogen	Zentrales Instrument zur Ermittlung der Patientenperspektive.
generischer ~	Allgemeiner bzw. krankheitsübergreifend angelegter Fragebogen für Patienten.
spezifischer ~	Spezieller bzw. krankheitsspezifisch angelegter Fragebogen für Patienten.

Begriff	Beschreibung
Patientenorientierung	Qualitätsdimension, in deren Zentrum die Wahrnehmung der spezifischen Interessen und Bedürfnisse erkrankter Menschen und ihres Einbezugs in das Geschehen rund um seine Krankheit steht.
Patientenperspektive	Wahrnehmung der medizinisch-pflegerischen Versorgung und der Behandlungsergebnisse, wie sie individuell durch Patienten berichtet wird oder bei Patienten erfragt werden kann.
Patientensicherheit	Abwesenheit/Vermeidung unerwünschter Ereignisse im Zusammenhang mit einer medizinisch-pflegerischen Behandlung.
Patientenvertreter	hier: An der gesetzlichen Qualitätssicherung teilnehmende Personen, die von den maßgeblichen Organisationen für die Wahrnehmung der Interessen der Patientinnen und Patienten und der Selbsthilfe chronisch kranker und behinderter Menschen auf Bundesebene im Sinne des §140f SGB V benannt werden.
Perzentile	„Hundertstelwerte“. Statistische Aus- und Bewertungsgröße: Perzentile zerlegen eine statistische Verteilung (von Fällen) in 100 gleich große Teile (1 %-Segmente). Das 95. Perzentil besagt bspw., dass 95 % aller Fälle unterhalb bzw. 5 % oberhalb des Wertes liegen.
Plausibilitätsprüfung	Statistisches Verfahren, mit dem die Dokumentationsdaten auf erlaubte und/oder fehlende Werte, Widerspruchsfreiheit, Werteverteilung und bekannte Korrelationen geprüft werden.
Poisson-Regression	Statistische Methode, die eingesetzt wird, um den Einfluss verschiedener Größen (z.B. Alter, Geschlecht oder Begleiterkrankungen) auf eine Zählvariable (z.B. eine, zwei oder drei Infektionen) zu ermitteln. Ähnlich zu den multiplen logistischen Regressionen werden Poisson-Regressionen im Kontext der Qualitätssicherung dazu eingesetzt, um risikoadjustierte Analysen durchzuführen.
Prävalenz	eigentlich: Prävalenzratio. Der Anteil Erkrankter an der Gesamtzahl der betrachteten Individuen einer definierten Population zu einem bestimmten Zeitpunkt (Zähler: Anzahl der zum Untersuchungszeitpunkt Kranken; Nenner: Anzahl aller untersuchten Individuen).
Pretest, kognitiver	Einer eigentlichen (Patienten-)Befragung vorgeschaltete Befragung mit deutlich kleinerer Stichprobe, bei der die Fragen und das Feldmaterial auf Verständlichkeit und Handhabbarkeit überprüft werden. In der Regel erfolgt der Test durch eine persönliche oder telefonische Befragung von ausgewählten Patienten.
Pseudonymisierung	Ersetzen des Namens und anderer Identifikationsmerkmale durch ein Kennzeichen zu dem Zweck, die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren. (§3 Abs. 6a BDSG).
Qualität	Bezogen auf die Gesundheitsversorgung: Grad, in dem versorgungsrelevante Ergebnisse, Prozesse und Strukturen bestimmte, definierte Anforderungen erfüllen.
Qualitätsdimension	Bestimmter Aspekt innerhalb der Gesundheitsversorgung, auf den sich eine definierte Qualität bezieht, z.B. Patientensicherheit.
Qualitätsindex	Fasst mehrere Qualitätsindikatoren in einer Maßzahl zusammen.
Qualitätsindikator	siehe: Indikator

Begriff	Beschreibung
Qualitätssicherung	Sammelbegriff für unterschiedliche Ansätze und Maßnahmen zur Sicherstellung festgelegter Qualitätsanforderungen bzw. zur Erreichung bestimmter Qualitätsziele. Hier: Gesetzliche Qualitätssicherung im Gesundheitswesen nach §§135–139 SGB V.
einrichtungsübergreifende~	Verfahren, die zur Beurteilung der Qualität eines Leistungserbringers auf Vergleich mit anderen Leistungserbringern aufbauen (§137a SGB V; §1 Qesü-RL).
externe stationäre~	Einrichtungsübergreifende Qualitätssicherung für medizinisch-pflegerische Leistungen, die ausschließlich im stationären Sektor erbracht werden.
interne~	Maßnahmen, die einrichtungsintern durchgeführt werden, um die eigene Leistung nach definierten Kriterien zu überprüfen, Verbesserungspotenziale aufzudecken und Maßnahmen zur Qualitätsverbesserung einleiten zu können.
sektorgleiche~	Verfahren, die ein Thema betreffen, bei dem die Erbringung der gleichen medizinischen Leistung in unterschiedlichen Sektoren erfolgt (§1 Qesü-RL).
sektorenübergreifende~	Verfahren, die nicht nur auf einen spezifischen Sektor bezogen Anwendung finden, sondern sektorale Grenzen in den jeweiligen Anforderungen unbeachtet lassen (§137a SGB V; §1 Qesü-RL). Erfasst sind damit insbesondere sektorenüberschreitende, sektorgleiche und sektorenüberschreitende Follow-up-Verfahren.
sektorenüberschreitende~	Verfahren, die ein Thema betreffen, bei dem mindestens zwei Sektoren am Behandlungsergebnis maßgeblich Anteil haben (§1 Qesü-RL).
sektorspezifische~	Verfahren, in denen die Qualität ausschließlich in einem Sektor betrachtet wird. Dieses kann den vertragsärztlichen, den vertragszahnärztlichen oder den stationären Sektor betreffen. Dazu zählen z.B. alle Verfahren der externen stationären Qualitätssicherung und alle Verfahren der Qualitätssicherung im ambulanten Bereich.
Qualitätsziel	Definition konkreter Anforderungen an den Inhalt, die Ausführung oder das Ergebnis einer medizinisch-pflegerischen Leistung. Anhand der Indikatorwerte bzw. unter Anlegung der Referenzbereiche lässt sich ermitteln, ob und in welchem Umfang Einrichtungen des Gesundheitswesens die definierten Qualitätsziele erreichen.
QS-Auslösung	Initiierung einer Dokumentationspflicht zu Zwecken der Qualitätssicherung (QS-Dokumentation). Bei einer Erhebung vorhandener Daten (z.B. Sozialdaten bei den Krankenkassen) analog das Kriterium, dass die Lieferung eines bestimmten Datensatzes auslöst.
QS-Daten	Sammelbegriff für alle Daten, die im Zuge eines QS-Verfahrens erhoben und ausgewertet werden.
QS-Dokumentation	Gesonderte Erhebungen der Leistungserbringer zu Diagnose- und Behandlungsdaten der Patienten für die Qualitätssicherung.
QS-Filter	Der QS-Filter ist eine Software, die auf Grundlage festgelegter Kriterien die für die Qualitätssicherung zu dokumentierenden Patienten „filtert“. Die Kriterien hierzu werden in einer Spezifikation definiert.
~-Software	Implementierung der Spezifikation für den QS-Filter.
~-Spezifikation	Enthält alle Regeln für die Auslösung einer Dokumentation und zur Prüfung der Vollzähligkeit.

Begriff	Beschreibung
QS-Verfahren	siehe: Leistungsbereich
RAM-Panel	Expertengruppe, die in Anlehnung an die RAND Appropriateness Method (RAM) des internationalen Forschungsnetzwerks „Research and Developement“ (RAND) Indikatoren auswählt und bewertet.
RAM-Prozess	Strukturiertes Auswahl- und Bewertungsverfahren von Indikatoren in Anlehnung an die RAND Appropriateness Method (RAM) des internationalen Forschungsnetzwerks „Research and Developement“ (RAND).
Rechenregeln	Hier: Regeln zur Berechnung von Indikatoren.
Referenzbereich	Gibt an, ob das Ergebnis eines Indikators rechnerisch auffällig oder unauffällig ist. Ergebnisse außerhalb des Referenzbereichs ziehen üblicherweise Qualitätssicherungsmaßnahmen nach sich. Derzeit werden Ziel- und Toleranzbereiche unterschieden.
Regelbetrieb	auch: Routinebetrieb oder Echtbetrieb. Verpflichtende und Flächendeckende Umsetzung eines QS-Verfahrens.
Reliabilität	Maß für die Zuverlässigkeit einer Messmethode bzw. für die mit dieser Methode gewonnenen Daten, wobei die Messung (Datenerhebung und Auswertung) reproduzierbar sein muss.
Review	siehe: Übersichtsarbeit, systematische
Risikoadjustierung	Methode zur Ermittlung und Gewichtung wesentlicher Einflussfaktoren (individueller Risiken) auf die Ausprägung einer Messgröße. Eine Risikoadjustierung ist vor allem bei Qualitätsindikatoren erforderlich, die sich auf Behandlungsergebnisse und einen Einrichtungsvergleich beziehen. Um hier einen fairen Vergleich zu erhalten, dürfen nur in gleichem Umfang erkrankte Patienten in Einrichtungen mit einer vergleichbaren Klientel miteinander verglichen werden.
Routinedaten	hier: Daten, die wesentlich zur Abwicklung von Geschäfts- und Verwaltungsabläufen erhoben werden (z.B. Abrechungsdaten, personenbezogene administrative Daten). Abseits des uneinheitlichen Sprachgebrauchs stehen die Sozialdaten der Gesetzlichen Krankenkassen (auch: GKV-Routinedaten) im Vordergrund des Interesses, da sie gem. §299 Abs. 1a SGB V zu Zwecken der Qualitätssicherung verwendet werden dürfen. Diese beinhalten insbesondere die abrechnungsrelevanten Daten für ambulante und stationäre Versorgungsleistungen (§§295 und 301 SGB V), für Arznei-, Heil- und Hilfsmittel (§§300 und 302 SGB V) sowie die Versichertenstammdaten (§284 SGB V). Das AQUA-Institut verwendet den Begriff derzeit für Sozialdaten bei den Krankenkassen, die Morbi-RSA-Daten (§303 a–e SGB V) und die Abrechnungsdaten nach §21 KHEntgG (Abrechnungsdaten des stationären Sektors). Siehe auch: Datenarten.
Scoping-Workshop	Dem eigentlichen Entwicklungsprozess von Indikatoren für ein QS-Verfahren vorgeschaltetes Expertentreffen.
Sentinel-Event	Sehr seltene und schwerwiegende medizinische Ereignisse (z.B. Todesfälle während eines Routineeingriffs). In der Qualitätssicherung zeigt ein Sentinel-Event-Indikator Vorkommnisse an, denen in jedem Einzelfall nachgegangen wird.
Sektor	Institutionell, d.h. durch unterschiedliche Abrechnungsmodalitäten getrennte Bereiche der medizinisch-therapeutischen Versorgung im deutschen Gesundheitswesen (z.B. ambulant/stationär).

Begriff	Beschreibung
Shared-Decision-Making	auch: Partizipative Entscheidungsfindung. Konzept zur Einbeziehung der Patienten in die Entscheidung über seine medizinische Behandlung.
Sollstatistik	Aufstellung der im Erfassungsjahr zu dokumentierenden Leistungen, die vom Krankenhaus durch Konformitätserklärung schriftlich bestätigt wird. Sie gibt die zu erwartende Anzahl von Fällen in den einzelnen Leistungsbereichen (Soll) an und bildet zusammen mit der Anzahl der tatsächlich durch die Leistungserbringer gelieferten Daten (Ist) die Grundlage der Vollzähligkeitsprüfung.
Sozialdaten	Einzelangaben über die persönlichen und sachlichen Verhältnisse (personenbezogene Daten), die von den sozialrechtlichen Leistungsträgern zur Erfüllung ihrer gesetzlichen Aufgaben gesammelt und gespeichert werden.
Spannweite	auch: Variationsbreite. Maß für die Streuung einer Messwertreihe; Differenz zwischen dem größten und dem kleinsten Wert einer Messwertreihe.
Spezifikation	Datensatzbeschreibung. Festlegung, welche Daten für die Qualitätssicherung erhoben bzw. übermittelt werden müssen, welche Prüfalgorithmen zur Anwendung kommen (z.B. für Plausibilitätsprüfungen) und wie die QS-Auslösung operationalisiert ist. Im Rahmen der Neuentwicklung von QS-Verfahren ist die Spezifikation als das Ergebnis der informationstechnischen Aufbereitung zu betrachten.
Soll-Ist-Vergleich	siehe: Vollzähligkeitsprüfung.
Stellungnahmeverfahren	▪ Für die Berichte über die Leistungen im Rahmen der Beauftragung der Entwicklung von Indikatoren, Instrumenten und der notwendigen Dokumentation sowie einiger anderer zu erbringende Leistungen ist ein schriftliches Stellungnahmeverfahren mit den dazu berechtigten Institutionen nach §137a Abs. 3 SGB V durchzuführen. Die Stellungnahmen zu den Entwicklungsberichten werden im AQUA-Institut ausgewertet und die Ergebnisse fließen in den Abschlussbericht an den G-BA ein. ▪ Nach §17 Abs. 2 Qesü-RL definierte Maßnahme gegenüber rechnerisch auffälligen Leistungserbringern.
Strata	Schichten oder Gruppen (z.B. Patientengruppen mit bestimmten gemeinsamen Merkmalen wie Alter, Geschlecht usw.).
Strukturierter Dialog	Instrument der Qualitätsförderung. Strukturiertes Verfahren, das Einrichtungen bei der kontinuierlichen Verbesserung von Prozessen und Qualität unterstützt. Auslöser des Strukturierten Dialogs sind rechnerische Auffälligkeiten im Ergebnis eines Qualitätsindikators im Rahmen der statistischen Basisprüfung der übermittelten QS-Daten (QSKH-RL). Abweichend von der Definition in der QSKH-RL bezeichnet der Begriff „Strukturierter Dialog“ in der Qesü-RL nur den Prozessschritt, der dem schriftlichen Teil der Durchführung von QS-Maßnahmen folgt.
Systempflege	Routinemäßige und kontinuierliche Evaluation und Anpassung der Qualitätsindikatoren, der Softwarespezifikation usw.
Test-Retest	Testverfahren zur Messung der Reliabilität. Hierbei wird das Messinstrument (bspw. ein Fragebogen) bei den gleichen Untersuchungspersonen wiederholt eingesetzt (nach einem definierten zeitlichen Abstand, z.B. nach zwei Wochen).

© 2015 AQUA-Institut GmbH

Begriff	Beschreibung
Toleranzbereich	Für einige Indikatoren lässt sich kein Referenzbereich auf wissenschaftlicher Basis festlegen, der eine erreichbar gute Qualität beschreibt. Hier wird der Referenzbereich so festgelegt, dass er besonders auffällige Ergebnisse abgrenzt. Dies kann sowohl über einen festen Wert als auch über ein Perzentil erfolgen (Perzentil-Referenzbereich)
Tracer-Verfahren	Hier: Exemplarisch ausgewählte Operationen oder Eingriffe, an der Entwicklungen, Stärken und Schwächen spezifischer Versorgungsmaßnahmen untersucht werden. Die Untersuchungsergebnisse können (und sollen) ggf. auf andere Operationen und QS-Verfahren übertragen werden.
Übersichtsarbeit, systematische	auch engl.: Systematic Review. Studie, die sämtliches in der Literatur verfügbare Wissen zu einem bestimmten Thema bündelt und bewertet.
Validität	Übereinstimmung von Aussagen, Daten, Messwerten usw. mit der „Realität“.
Verfahren	siehe: Leistungsbereich
direkte~	auch: Bundesbezogene Verfahren. Medizinische Leistungsbereiche, bei denen die vergleichende Qualitätssicherung aufgrund geringer absoluter Fallzahlen oder der kleinen Anzahl durchführender Einrichtungen direkt auf Bundesebene durchgeführt wird. Dies betrifft derzeit vor allem einige herzchirurgische Eingriffe und weite Teile der Transplantationsmedizin.
indirekte~	auch: Länderbezogene Verfahren. Medizinische Leistungsbereiche, bei denen die Qualitätssicherung indirekt, d.h. über die Länder (Landesgeschäftsstellen oder beauftragte Annahmestellen), abgewickelt wird.
Versorgungskette	Gesamtheit der Einrichtungen und Organisationen, die die Patienten von der Früherkennung bis zur Nachsorge mit ihren Produkten und Dienstleistungen versorgen.
Versorgungspfad	Visualisiert den idealtypischen Weg von definierten Patientengruppen mit seinen entscheidenden diagnostischen und therapeutischen Leistungen in der zeitlichen Abfolge.
Versorgungsziel	Auf Ebene des Gesundheitssystems, d.h. wesentlich gesundheitspolitisch gesetzte Vorgaben bezüglich der Versorgung einer Gesellschaft, einer Population oder einer Region mit bestimmten (medizinischen) Leistungen.
Vertrauensbereich	auch: Konfidenzintervall. Der Vertrauensbereich beschreibt ein Intervall um einen berechneten Ergebniswert, z.B. um das Gesamtergebnis eines Indikators. Vereinfacht ausgedrückt gibt der Vertrauensbereich den Bereich an, in dem der tatsächliche Wert eines Qualitätsindikators unter Berücksichtigung aller zufälligen Ereignisse (z.B. Dokumentationsfehler) mit einer bestimmten, zuvor festgelegten Wahrscheinlichkeit liegt.
Vertrauensstelle	Institution, die im Rahmen der sektorenübergreifenden Qualitätssicherung erhobene patientenidentifizierende Daten pseudonymisiert. Näheres regeln die Qesü-RL (§11), QSKH-RL (Anl. 2, §4) sowie themenspezifische Bestimmungen.
Vollerhebung	Erfassung aller Elemente einer Grundgesamtheit.
Vollständigkeit	Erfassung aller zu einem einzelnen Behandlungsfall erforderlichen Angaben (Daten).
Vollzähligkeit	Erfassung aller dokumentationspflichtigen Behandlungsfälle zu einer bestimmten Versorgungsleistung.

Begriff	Beschreibung
Vollzähligkeitsprüfung	Abgleich der laut QS-Filter-Software zu erwartenden Anzahl von Fällen in einem Leistungsbereich (Soll) mit der Anzahl der tatsächlich durch die Leistungserbringer gelieferten Daten (Ist).
Zielbereich	Definiert durch einen festen Referenzbereich für erreichbar gute Qualität auf Basis wissenschaftlicher Untersuchungen.

Literatur

Abdallah, MS; Wang, K; Magnuson, EA; Spertus, JA; Farkouh, ME; Fuster, V; et al. (2013). Quality of life after PCI vs CABG among patients with diabetes and multivessel coronary artery disease: A randomized clinical trial. JAMA 310(15): 1581-1590.

AHCPR (1992). Clinical Practice Guidelines, No. 1. Acute Pain Management: Operative or Medical Procedures and Trauma. Rockville: Acute Pain Management Guideline Panel. Agency for Health Care Policy and Research.

AHRQ (2011). Quality Indicator Measure Development, Implementation, Maintenance, and Retirement. Rockville: Agency for Healthcare Research and Quality.

AHRQ (2010). AHRQ Quality Indicators. Rockville: Agency for Healthcare Research and Quality.

Alexandrescu, R; Jen, MH; Bottle, A; Jarman, B; Aylin, P (2011). Logistic Versus Hierarchical Modeling: An Analysis of a Statewide Inpatient Sample. J Am Coll Surg 213(3): 392-401.

Anastasi, A; Urbina, S (1997). Psychological testing. 7. Auflage. New York: Prentice Hall.

Andersen, T; Mooney, G (1990). The Challenges of Medical Practice Variations. London: McMillan Press.

AQUA (2014a). Bericht zur Datenvalidierung 2013. Erfassungsjahr 2012. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2014b). Allgemeine Spezifikation für die Nutzung der Sozialdaten bei den Krankenkassen – Technische Dokumentation (Stand: 31. März 2014). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2014c). Entwicklung einer Patientenbefragung im Rahmen des Qualitätssicherungsverfahrens Perkutane Koronarintervention (PCI) und Koronarangiographie. Vorbericht (Stand: 23. Juli 2014). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2014d). Perkutane Koronarintervention (PCI) und Koronarangiographie. Koordination des Gesamtverfahrens (Stand: 31. März 2014). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2014e). Qualitätsreport 2013. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2014f). Bericht zum Strukturierten Dialog 2013. Erfassungsjahr 2012 (Stand: 25. August 2014). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2014g). Technische Dokumentation für Leistungserbringer zur Basisspezifikation. QS-Filter, QS-Dokumentation, Soll- und Risikostatistik. Erfassungsjahr 2015 (Stand: 3. November 2014, V03). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2014h). Arthroskopie am Kniegelenk. Abschlussbericht (Stand: 14. Juli 2014). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2014i). Bericht zur Prüfung und Bewertung der Indikatoren der externen stationären Qualitätssicherung hinsichtlich ihrer Eignung für die einrichtungsbezogene öffentliche Berichterstattung (Stand: 21. März 2014). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2013a). Allgemeine Spezifikation für die Nutzung der Sozialdaten bei den Krankenkassen – Projektabschlussbericht. AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2013b). Bericht zur Datenvalidierung 2012. Erfassungsjahr 2011 (Stand: 15. Mai 2013). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2013c). DEK – Pflege: Dekubitusprophylaxe. Qualitätsindikatoren. Bundesauswertung zum Erfassungsjahr 2013. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2013d). Bericht zur Prüfung und Bewertung der Indikatoren der externen stationären Qualitätssicherung hinsichtlich ihrer Eignung für die einrichtungsbezogene öffentliche Berichterstattung (Stand: 18. März 2013). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2012a). Länderauswertung mit Geodarstellung für Indikatoren mit besonderem Handlungsbedarf (Stand: 25. Januar 2012). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2012b). Bericht zur Datenvalidierung 2011. Erfassungsjahr 2010 (Stand: 12. Juni 2012). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2012c). Bericht zur „Aufbereitung einer aussagefähigen Internetdarstellung der Daten des Anhangs der Anlage 1 der Vereinbarung über Maßnahmen zur Versorgung von Früh- und Neugeborenen" (Stand vom 29. Februar 2012, noch nicht zur Veröffentlichung freigegeben). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2012d). Entwicklung eines generischen Patientenfragebogens für die sektorenübergreifende Qualitätssicherung: Übersicht über die Entwicklungsschritte und Darstellung des Befragungsinstruments. Kurzbericht an den Gemeinsamen Bundesausschuss, 2012 (unveröffentlicht). AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen; Abteilung Allgemeinmedizin und Versorgungsforschung des Universitätsklinikums Heidelberg.

AQUA (2011a). Bericht zur Schnellprüfung und Bewertung der Indikatoren der externen stationären Qualitätssicherung hinsichtlich ihrer Eignung für die einrichtungsbezogene öffentliche Berichterstattung [Überarbeitete, ergänzte und korrigierte Version vom 15.05.2011]. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2011b). Bericht zur Prüfung der Nutzung von Routinedaten im Leistungsbereich Dekubitus. Sonderbericht (Stand: 11. November 2011). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2011c). Bericht zur Datenvalidierung 2010. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2011d). Kolorektales Karzinom. Abschlussbericht (Stand: 5. November 2011). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2011e). Konzept zur Sicherung der Dokumentationsqualität vor dem Hintergrund des Datenvalidierungsverfahrens in der sektorenübergreifenden Qualitätssicherung (Stand: 14. April 2011). Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

AQUA (2010). Qualitätsreport 2009. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen.

Arah, OA; Westert, GP; Hurst, J; Klazinga, NS (2006). A conceptual framework for the OECD Health Care Quality Indicators Project. Int J Qual Health Care 18(Suppl 1): 5-13.

Arts, DGT; de Keizer, NF; Scheffer, GJ (2002). Defining and Improving Data Quality in Medical Registries: A Literature Review, Case Study, and Generic Framework. J Am Med Inform Assoc 9(6): 600-611.

Ash, AS; Shwartz, M; Peköz, EA (2003). Comparing outcomes across providers. In: Risk Adjustment for Measuring Health Care Outcomes. 3. Auflage. Iezzoni, LI (Ed.). Chicago: Health Administration Press: 297-333.

Aylin, P; Bottle, A; Majeed, A (2007). Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models. BMJ 334(7602): 1044.

ÄZQ (2009). Qualitätsindikatoren – Manual für Autoren. Programm für Nationale VersorgungsLeitlinien von BÄK, KBV und AWMF. Berlin: Ärztliches Zentrum für Qualität in der Medizin.

Badura, B (1999). Evaluation und Qualitätsberichterstattung im Gesundheitswesen – Was soll bewertet werden und mit welchen Maßstäben? In: Evaluation im Gesundheitswesen. Ansätze und Ergebnisse. Badura, B; Siegrist, J (Eds.). Weinheim-München: Juventa: 15-42.

Badura, B; Strodtholz, P (2003). Qualitätsforschung und Evaluation im Gesundheitswesen. In: Das Public Health Buch - Gesundheit und Gesundheitswesen. Schwartz, FW; Badura, B; Leidl, R; Raspe, H; Siegrist, J; Walter, U (Eds.). München and Jena: Urban & Fischer. 2., völlig neu bearbeitete und erweiterte Auflage: 714-24.

BÄK (2013). Texte und Materialien zur Fort- und Weiterbildung: Curriculum „Ärztliches Peer Review“. 2. Aufl. Berlin: Bundesärztekammer.

BÄK/KBV (1997). Beurteilungskriterien für Leitlinien in der medizinischen Versorgung (Bundesärztekammer und Kassenärztliche Bundesvereinigung). Dtsch Arztebl 137: 1085-1086.

Balain, B; Ennis, O; Kanes, G; Singhal, R; Roberts, SN; Rees, D; et al. (2009). Response shift in self-reported functional scores after knee microfracture for full thickness cartilage lesions. Osteoarthritis Cartilage 17(8): 1009-1013.

Bardach, NS; Chien, AT; Dudley, RA (2010). Small numbers limit the use of the inpatient pediatric quality indicators for hospital comparison. Acad Pediatr 10(4): 266-273.

Batini, C; Scannapieca, M (2006). Data Quality: Concepts, Methodologies and Techniques. Berlin: Springer.

Baumgarten, E (2007). Aktuelles aus der Reha-Qualitätssicherung: Peer Review-Verfahren ausgewertet - bessere Reha-Qualität, aber deutliche Unterschiede zwischen Reha-Einrichtungen. RVaktuell 54: 152-154.

Behrens, J; Langer, G (2006). Evidence-based Nursing and Caring. Bern: Verlag Hans Huber.

Beske, F; Hallauer, JF; Bechtel, H (2005). Das Gesundheitswesen in Deutschland. Struktur - Leistungen - Weiterentwicklung. Köln: Deutscher Ärzte-Verlag.

Bitzer, EM; Grobe, TG; Dörning, H; Schwartz, FW (2008). One-year mortality after surgical and non-surgical approaches to coronary revascularisation – results based on administrative data of a German health insurance. GMS Med Inform Biom Epidemiol 4(2): Doc05.

Bitzer, EM; Petrucci, M; Lorenz, C; Hussein, R; Dorning, H; Trojan, A; et al. (2011). A comparison of conventional and retrospective measures of change in symptoms after elective surgery. Health Qual Life Outcomes 9(23).

Black, C; Roos, NP (1998). Administrative Data. Baby or Bathwater? Med Care 36(1): 3-5.

Bloom, B (2005). Effects of continuing medical education on improving physician clinical care and patient health: A review of systematic reviews. Int J Technol Assess Health Care 21(3): 380-385.

Blümle, A; Meerpohl, JJ; Wolff, R; Antes, G (2009). Evidenzbasierte Medizin und systematische Übersichtsarbeiten. Die Rolle der Cochrane Collaboration. MKG-Chirurg 2(2): 86-92.

Bokhour, BG; Pugh, MJ; Rao, JK; Avetisyan, R; Berlowitz, DR; Kazis, LE (2009). Improving Methods for Measuring Quality of Care. A Patient-Centered Approach in Chronic Disease. Med Care Res Rev 66(2): 147-166.

Bölt, U; Graf, T (2012). 20 Jahre Krankenhausstatistik. Wirtsch Stat (Februar): 112-138.

Bortz, J; Döring, N (2005). Forschungsmethoden und Evaluation. Heidelberg: Springer Medizin Verlag.

Bottle, A; Aylin, P (2009). Application of AHRQ patient safety indicators to English hospital data. Qual Saf Health Care 18(4): 303-308.

BQS (2007). Qualität sichtbar machen. BQS-Qualitätsreport 2006. Düsseldorf: Bundesgeschäftsstelle Qualitätssicherung.

Bratzler, DW; Nsa, W; Houck, PM (2007). Performance measures for pneumonia: are they valuable, and are process measures adequate? Curr Opin Infect Dis 20(2): 182-189.

Broge, B (2008). EPA - Europäisches Praxisassessment. In: Praxishygiene und Qualitätsmanagement. Zinn, G-C; Tabori, E; Weidenfeller, P (Eds.). Friedberg: Verlag für medizinische Praxis: 217-228.

Broge, B; Heller, G; Kaufmann-Kolle, P; Kazmaier, T; Pauletzki, J; Willms, G; et al. (2014). Stand und Perspektiven der Qualitätssicherung. In: Qualitätsreport 2013. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen: 238-244.

Bührlen, B (2010). Innovation im Gesundheitswesen: Die Rolle von HTA bei der Einführung neuer Technologien. Z Evid Fortbild Qual Gesundhwes 104(10): 703-8.

Busse, R; Drösler, S; Glaeske, G; Greiner, W; Schäfer, T; Schrappe, M (2007). Wissenschaftliches Gutachten für die Auswahl von 50 bis 80 Krankheiten zur Berücksichtigung im morbiditätsorientierten Risikostrukturausgleich. Wissenschaftlicher Beirat zur Weiterentwicklung des Risikostruturausgleichs.

Campbell, S; Reeves, D; Kontopantelis, E; Middleton, E; Sibbald, B; Roland, M (2007). Quality of Primary Care in England with the Introduction of Pay for Performance. N Engl J Med 357(2): 181-190.

Campbell, SM; Braspenning, J; Hutchinson, A; Marshall, MN (2003). Improving the quality of health care. Research methods used in developing and applying quality indicators in primary care. BMJ 326(7393): 816-819.

Campbell, SM; Reeves, D; Kontopantelis, E; Sibbald, B; Roland, M (2009). Effects of Pay for Performance on the Quality of Primary Care in England. N Engl J Med 361(4): 368-78.

Campbell, SM; Roland, MO; Buetow, SA (2000). Defining quality of care. Soc Sci Med 51(11): 1611-1625.

Capitulo, KL (1998). Improving patient satisfaction through focus group interviews. Nurs Leadersh Forum 3(4): 124-9.

Cathey, J; Al Hajeri, AA; Fedorowicz, Z (2006). A comparison of handsearching versus EMBASE searching of the “Annals of Saudi Medicine” to identify reports of randomized controlled trials. Ann Saudi Med 26(1): 49-51.

Chassin, MR (2002). Achieving And Sustaining Improved Quality: Lessons From New York State And Cardiac Surgery. Health Aff 21(4): 40-51.

Cheng, EM; Crandall, CJ; Bever, CT, Jr.; Giesser, B; Haselkorn, JK; Hays, RD; et al. (2010). Quality indicators for multiple sclerosis. Mult Scler 16(8): 970-80.

Chertow, GM; Soroko, SH; Paganini, EP; Cho, KC; Himmelfarb, J; Ikizler, TA; et al. (2006). Mortality after acute renal failure: Models for prognostic stratification and risk adjustment. Kidney Int 70(6): 1120-6.

Chin, R; Lee, BY (2008). Section V. Analysis of Results. In: Principles and Practice of Clinical Trial Medicine. Chin, R; Yoonsik, R; Lee, BY (Eds.). London: Academic Press: 302-323.

Christiansen, CL; Morris, CN (1997). Improving the Statistical Approach to Health Care Provider Profiling. Ann Intern Med 127(8.2): 764-768.

CIHI (2010). Hospital Standardized Mortality Ratio (HSMR). Canadian Institute for Health Information. http://www.cihi.ca/cihi-ext-portal/internet/en/tabbedcontent/health+system+performance/quality+of+care+and+outcomes/hsmr/cihi022025 (Zugriff am 13.11.2014).

Clark, DE; Hannan, EL; Wu, C (2010). Predicting Risk-Adjusted Mortality for Trauma Patients: Logistic versus Multilevel Logistic Models. J Am Coll Surg 211(2): 224-31.

Cleary, PD; Edgman-Levitan, S; Walker, JD; Gerteis, M; Delbanco, TL (1993). Using patient reports to improve medical care: a preliminary report from 10 hospitals. Qual Manag Health Care 2(1): 31-8.

Cochrane Collaboration (2014). Glossary of Cochrane terms. http://www.cochrane.org/glossary (Zugriff am 13.11.2014).

Cohen, DJ; Van Hout, B; Serruys, PW; Mohr, FW; Macaya, C; Den Heijer, P; et al. (2011). Quality of life after PCI with drug-eluting stents or coronary-artery bypass surgery. N Engl J Med 364(11): 1016-1026.

Cohen, J; Cohen, P; Aiken, LS; West, SG (2003). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. 3. Aufl. Mahwah: Lawrence Erlbaum Associates.

Cohen, MA; Lee, HL (1985). The Determinants of Spatial Distribution of Hospital Utilization in a Region. Med Care 23(1): 27-38.

Commonwealth Fund, The (2004). First Report and Recommendations of the Commonwealth Fund's International Working Group on Quality Indicators. A Report to Health Ministers of Australia, Canada, New Zealand, the United Kingdom, and the United States. New York: The Commonwealth Fund.

Contencin, P; Falcoff, H; Doumenc, M (2006). Review of performance assessment and improvement in ambulatory medical care. Health Policy 77(1): 64-75.

Credé, WB; Hierholzer, WJ, Jr. (1990). Surveillance for Quality Assessment: III. The Critical Assessment of Quality Indicators. Infect Control Hosp Epidemiol 11(4): 197-201.

Damman, OC; Stubbe, JH; Hendriks, M; Arah, OA; Spreeuwenberg, P; Delnoij, DM; et al. (2009). Using Multilevel Modeling to Assess Case-Mix Adjusters in Consumer Experience Surveys in Health Care. Med Care 47(4): 496-503.

de Cruppé, W; Geraedts, M (2011a). Wie wählen Patienten ein Krankenhaus für elektive operative Eingriffe? Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 54(8): 951-957.

de Cruppé, W; von dem Knesebeck, O; Gerstenberger, E; Link, C; Marceau, L; Siegrist, J; et al. (2011b). Hausärztliche Entscheidungen bei Symptomen für Diabetes mellitus Typ 2. Dtsch Med Wochenschr 136(8): 359-64.

de Koning, J (2007). Development and validation of a measurement instrument for appraising indicator quality: appraisal of indicators through research and evaluation (AIRE) instrument. Kongress Medizin und Gesellschaft 2007, Augsburg 17.–21.09.2007. Düsseldorf: German Medical Science GMS Publishing House.

de Koning, J; Smulders, AW; Klazinga, NS (2007). Appraisal of Indicators through Research and Evaluation (AIRE). Versie 2.0. Amsterdam: Academisch Medisch Centrum.

DesHarnais, SI; Forthman, MT; Homa-Lowry, JM; Wooster, LD (2000). Risk-adjusted clinical quality indicators: indices for measuring and monitoring rates of mortality, complications, and readmissions. Qual Manag Health Care 9(1): 14-22.

Diekmann, A (2006). Empirische Sozialforschung. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.

Diener, MK; Wolff, RF; von Elm, E; Rahbari, NN; Mavergames, C; Knaebel, H-P; et al. (2009). Can decision making in general surgery be based on evidence? An empirical study of Cochrane Reviews. Surgery 146(3): 444-461.

Dimick, JB; Osborne, NH; Hall, BL; Birkmeyer, JD (2010a). Risk Adjustment for Comparing Hospital Quality with Surgery: How Many Variables Are Needed? J Surg Res 158(2): 310.

Dimick, JB; Staiger, DO; Birkmeyer, JD (2010b). Ranking Hospitals on Surgical Mortality: The Importance of Reliability Adjustment. Health Serv Res 45(6p1): 1614-1629.

Dimick, JB; Staiger, DO; Osborne, NH; Nicholas, LH; Birkmeyer, JD (2012). Composite Measures for Rating Hospital Quality with Major Surgery. Health Serv Res 47(5): 1861-1879.

Dimick, JB; Welch, HG; Birkmeyer, JD (2004). Surgical mortality as an indicator of hospital quality: the problem with small sample size. JAMA 292(7): 847-851.

DIN EN ISO (2005). Qualitätsmanagementsysteme – Grundlagen und Begriffe (ISO 9000:2005). Berlin: Beuth Verlag.

Ding, YY (2009). Risk adjustment: towards achieving meaningful comparison of health outcomes in the real world. Ann Acad Med Singapore 38(6): 552-7.

DNEbM (o.J. (2011)). Interessenkonfliktregulierung: Internationale Entwicklungen und offene Fragen. (Diskussionspapier des IKR). Deutsches Netzwerk evidenzbasierte Medizin e. V.

Donabedian, A (1966). Evaluating the quality of medical care. Milbank Mem Fund Q 44(3 Suppl): 166-206.

Dr. Foster (2013). The Dr Foster Hospital Guide 2013. London: Dr Foster Intelligence.

Dreier, M; Borutta, B; Stahmeyer, J; Krauth, C; Walter, U (2010). Vergleich von Bewertungsinstrumenten für die Studienqualität von Primär- und Sekundärstudien zur Verwendung für HTA-Berichte im deutschsprachigen Raum (Schriftenreihe Health Technology Assessment, Bd. 102). Köln: Deutsches Institut für Medizinische Dokumentation und Information (DIMDI).

Drösler, SE; Cools, A; Köpfer, T; Stausberg, J (2007). Eignen sich Qualitätsindikatoren aus Routinedaten zur Qualitätsmessung im Krankenhaus? Erste Ergebnisse mit den amerikanischen Indikatoren zur Patientensicherheit in Deutschland? Z ärztl Fortbild Qual Gesundhwes 101(1): 35-42.

Droste, S (2008). Systematische Gewinnung von Informationen zu ethischen Aspekten in HTA-Berichten zu medizinischen Technologien bzw. Interventionen. Z Evid Fortbild Qual Gesundhwes 102(5): 329-341.

Droste, S; Dintsios, CM; Gerber, A (2010). Kommentar zu: Rauprich O, Nolte M, Vollmann J (2010). Systematische Literaturrecherchen in den Datenbanken PubMed und BELIT. Ein Werkstattbericht (Ethik Med 22(1):59-67). Ethik Med 22(4): 355-358.

Eccles, M; Clapp, Z; Grimshaw, J; Adams, PC; Higgins, B; Purves, I; et al. (1996). North of England evidence based guidelines development project: methods of guideline development. BMJ 312(7033): 760-762.

Eckardt, J (2008). Erste Erfahrungen mit dem praktischen Einsatz von Qualitätsindikatoren. http://www.qmr-kongress.de/download/01_Dienstag_Qualimanagement/04_Eckardt_Quali/inhalt.pdf (Zugriff am 13.11.2014).

Eichenhofer, W; Wenner, U (Eds.) (2013). Sozialgesetzbuch V. Gesetzliche Krankenversicherung. Kommentar. Köln (Wolters & Kluwer): Luchterhand.

Elashoff, JD; Lemeshow, S (2005). Sample Size Determination in Epidemological Studies. In: Handbook of epidemiology. Ahrens, W; Pigeot, I (Eds.). Berlin/Heidelberg: Springer: 559-594.

Epstein, RM; Franks, P; Fiscella, K; Shields, CG; Meldrum, SC; Kravitz, RL; et al. (2005). Measuring patient-centered communication in patient-physician consultations: theoretical and practical issues. Soc Sci Med 61(7): 1516-1528.

Erler, A (2007). Garbage in – Garbage out? Validität von Abrechnungsdiagnosen in hausärztlichen Praxen als Voraussetzung für die Einführung eines diagnosebasierten risikoadjustierten Vergütungssystems (Diss.). Berlin: Medizinische Fakultät der Charité Universitätsmedizin.

Farin, E; Dudeck, A; Meffert, C; Glattacker, M; Jäckel, W; Beckmann, U; et al. (2009). Direkte und indirekte Veränderungsmessung in der ambulanten Rehabilitation muskuloskeletaler und kardiologischer Erkrankungen. Phys Med Rehab Kuror 19: 149-161.

Fayers, P; Machin, D (2000). Quality of Life: Assessment, Analysis, and Interpretation. Chichester: Wiley.

Fink, A; Kosecoff, J; Chassin, M; Brook, RH (1984). Consensus methods: characteristics and guidelines for use. Am J Public Health 74(9): 979-983.

Fitch, K; Bernstein, SJ; Aguilar, MD; Burnand, B; et al. (2001). The RAND/UCLA Appropriateness Method User's Manual. Santa Monica: RAND.

Freise, D (2003). Teilnahme und Methodik bei Patientenbefragungen (Diss.). Köln: Universität Köln.

Friedemann, J; Schubert, H; Schwappach, D (2009). Zur Verständlichkeit der Qualitätsberichte deutscher Krankenhäuser: Systematische Auswertung und Handlungsbedarf. Gesundheitswesen 71: 3-9.

Friedrich, A (2011). Krankenhaushygiene in den Niederlanden und Deutschland. So nah und doch so fern. Vortrag auf der Qualitätssicherungskonferenz des G-BA am 28.11.2011 in Berlin. http://www.g-ba.de/downloads/17-98-3120/4_Prof.%20Friedrich.pdf (Zugriff am 13.11.2014).

Fung, CH; Lim, YW; Mattke, S; Damberg, C; Shekelle, PG (2008). Systematic review: the evidence that publishing patient care performance data improves quality of care. Ann Intern Med 148(2): 111-23.

Gail, M; Bénichou, J (2000). Encyclopedia of Epidemiological Methods. New York: John Wiley & Sons.

Galea, S (2007). Macrosocial Determinants of Population Health. New York: Springer Science+Business Media, LLC.

Ganser, A (2011). Interessenkonflikte in der medizinischen Forschung und Vorschläge für deren Minimierung. In: Interessenkonflikte in der Medizin. Hintergründe und Lösungsmöglichkeiten. Lieb, K; Klemperer, D; Ludwig, W-D (Eds.). Berlin/Heidelberg: Springer.

Gatsonis, CA; Epstein, AM; Newhouse, JP; Normand, SL; McNeil, BJ (1995). Variations in the utilization of coronary angiography for elderly patients with an acute myocardial infarction. An analysis using hierarchical logistic regression. Med Care 33(6): 625-42.

Geraedts, M; Auras, S; Hermeling; de Cruppé, W (2010). Abschlussbericht zum Forschungsauftrag zur Verbesserung der gesetzlichen Qualitätsberichte auf der Basis einer Krankenhaus-, Patienten- und Einweiserbefragung. Revidierte Fassung auf der Basis der Diskussionen in der AG Qualitätsbericht des UA Qualitätssicherung des G-BA. Witten.

Geraedts, M; Auras, S; Hermeling, P; de Cruppé, W (2009). Public Reporting - Formen und Effekte öffentlicher Berichterstattung. Dtsch Med Wochenschr 134: 232-233.

Geraedts, M; Hermeling, P; de Cruppé, W (2011). Communicating quality of care information to physicians: A study of eight presentation formats. Patient Educ Couns. DOI: 10.1016/j.pec.2011.11.005.

Gerlach, FM (2001). Qualitätsförderung in Klinik und Praxis. Eine Chance für die Medizin. Stuttgart: Thieme.

Gerste, B; Gutschmidt, S (2006). Datenqualität von Diagnosen aus dem ambulanten Bereich. Gesundheits- und Sozialpolitik 60(3-4): 29-43.

Geyer, S (2003). Forschungsmethoden in den Gesundheitswissenschaften. Eine Einführung in die empirischen Grundlagen. Weinheim und München: Juventa.

Ghali, WA; Rothwell, DM; Quan, H; Brant, R; Tu, JV (2000). A Canadian comparison of data sources for coronary artery bypass surgery outcome “report cards”. Am Heart J 140(3): 402-408.

Giersiepen, K; Pohlabeln, H; Egidi, G; Pigeot, I (2007). Die ICD-Kodierqualität für Diagnosen in der ambulanten Versorgung. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 50(8): 1028-1038.

Gigerenzer, G (2009). Making sense of health statistics. Bull World Health Organ 87: 567.

Gigerenzer, G; Gaissmaier, W; Kurz-Milcke, E; Schwartz, L; Woloshin, S (2007). Helping doctors and patients to make sense of health statistics. Psychol Sci Public Interest 8(2): 53-96.

Gillis, JZ; Panopalis, P; Schmajuk, G; Ramsey-Goldman, R; Yazdany, J (2011). Systematic review of the literature informing the systemic lupus erythematosus indicators project: reproductive health care quality indicators. Arthritis Care Res (Hoboken) 63(1): 17-30.

Goldstein, H; Spiegelhalter, DJ (1996). League tables and their limitations: Statistical issues in comparisons of institutional performance. J R Stat Soc Ser A Stat Soc 159(3): 385-443.

Grande, G; Romppel, M (2005). Subjektive Qualitätskonzepte von Patienten in der Rehabilitation. In: Auf dem Weg zur integrierten Versorgung. Badura, B; Iseringhausen, O (Eds.). Göttingen: Huber: 215-230.

Greenland, S (2000). Principles of multilevel modelling. Int J Epidemiol 29(1): 158-167.

Groene, O (2006). Vorschläge der WHO zur umfassenden Leistungsbewertung von Krankenhäusern. Gesundh ökon Qual manag 11(4): 226-233.

Grol, R (1993). Development of guidelines for general practice care. Br J Gen Pract 43(369): 146-151.

Grol, R; Berwick, DM; Wensing, M (2008). On the trail of quality and safety in health care. BMJ 336(7635): 74-76.

Grol, R; Dutzenberg, M; Brinkmann, H (2005a). Quality Management in Primary Care. European Practice Assessment. Gütersloh: Bertelsmann Stiftung.

Grol, R; Grimshaw, J (2003). From best evidence to best practice: effective implementation of change in patients' care. Lancet 362(9391): 1225-1230.

Grol, R; Wensing, M; Eccles, M (2005b). Improving patient care. The implementation of change in clinical practice. London: Elsevier.

Gyrd-Hansen, D; Olsen, KR; Sorensen, TH (2012). Socio-demographic patient profiles and hospital efficiency: does patient mix affect a hospital's ability to perform? Health Policy 104(2): 136-45.

Haller, G; Stoelwinder, J; Myles, PS; McNeil, J (2009). Quality and Safety Indicators in Anesthesia: A Systematic Review. Anesthesiology 110(5): 1158-1175.

Hannan, EL; Wu, C; DeLong, ER; Raudenbush, SW; et al. (2005). Predicting risk-adjusted mortality for CABG surgery: logistic versus hierarchical logistic models. Med Care 43(7): 726-735.

Harold, E; Means, W (2004). XML in a Nutshell. 3. Auflage. Beijing u.a.: O'Reilly Media.

Hasford, J; Staib, AH (1994). Arzneimittelprüfungen und Good Clinical Practice. München: MMV Medizin Verlag.

Hearnshaw, HM; Harker, RM; Cheater, FM; Baker, RH; Grimshaw, GM (2001). Expert consensus on the desirable characteristics of review criteria for improvement of health care quality. Qual Health Care 10(3): 173-178.

Hebel, JR; Kessler, II; Mabuchi, K; McCarter, RJ (1982). Assessment of hospital performance by use of death rates. A recent case history. JAMA 248(23): 3131-5.

Heller, G (2010). Qualitätssicherung mit Routinedaten - Aktueller Stand und Weiterentwicklung. In: Krankenhaus-Report 2010. Schwerpunkt: Krankenhausversorgung in der Krise? Klauber, J; Geraedts, M; Friedrich, J (Eds.). Stuttgart: Schattauer: 239-254.

Heller, G (2009). Administrative Data from Germanys Statutory Health Insurances for Social, Economic and Medical Research. Working Paper No. 122 of the Council for Social and Economic Data. German Council for Social and Economic Data (RatSWD).

Heller, G (2008). Zur Messung und Darstellung von medizinischer Ergebnisqualität mit administrativen Routinedaten in Deutschland. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 51(10): 1173-1182.

Heller, G (2006). Sind risikoadjustierte Analysen mit administrativen Routinedaten möglich? In: Kursbuch Versorgungsforschung. Hey, M; Maschewsky-Schneider, U (Eds.). München: Medizinisch Wissenschaftliche Verlagsgesellschaft: 253-257.

Heller, G; Babitsch, B; Gunster, C; Möckel, M (2008). Sterblichkeitsrisiko von Frauen und Männern nach Myokardinfarkt. Dtsch Arztebl 105(15): 279-285.

Heller, G; Broge, B; Szecsenyi, J (2014). Nutzung von Sekundärdaten in der gesetzlichen Qualitätssicherung. In: Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven. 2., vollständig überarbeitete und aktualisierte Auflage. Swart, E; Ihle, P; Gothe, H; Matusiewicz (Eds.). Bern: Hans Huber: 460-473.

Heller, G; Jeschke, E (2011). Einrichtungsübergreifende Qualitätssicherung der Gallenblasenentfernung auf der Basis von Routinedaten. In: Krankenhaus-Report 2012; Schwerpunkt: Qualität durch Wettbewerb. Klauber, J; Geraedts, M; Friedrich, JW, J. (Eds.). Stuttgart: Schattauer.

Heller, G; Schnell, R (2007). Hospital mortality risk adjustment using claims data. JAMA 297(18): 1983-1984.

Heller, G; Swart, E; Mansky, T (2004). Qualitätsanalysen mit Routinedaten im stationären Sektor. Ansatz und erste Analysen aus dem Gemeinschaftsprojekt „Qualitätssicherung mit Routinedaten“. In: Krankenhaus-Report 2004. Klauber, J; Robra, BP; Schellschmidt, H (Eds.). Stuttgart: Schattauer: 271-288.

Hermens, RP; Ouwens, MM; Vonk-Okhuijsen, SY; van der Wel, Y; Tjan-Heijnen, VC; van den Broek, LD; et al. (2006). Development of quality indicators for diagnosis and treatment of patients with non-small cell lung cancer: a first step toward implementing a multidisciplinary, evidence-based guideline. Lung Cancer 54(1): 117-124.

Hibbard, JH (2008). What can we say about the impact of public reporting? Inconsistent execution yields variable results. Ann Intern Med 148(2): 160-161.

Hibbard, JH; Stockard, J; Tusler, M (2005). It Isn't Just about Choice: The Potential of a Public Performance Report to Affect the Public Image of Hospitals. Med Care Res Rev 62(3): 358-371.

Hibbard, JH; Stockard, J; Tusler, M (2003). Does publicizing hospital performance stimulate quality improvement efforts? Health Aff 22(2): 84-94.

Higgins, JPT; Green, S; (eds) (2011). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0: The Cochrane Collaboration, Available from www.cochrane-handbook.org. .

Hofhuis, H; van den Ende, CH; De Bakker, DH (2006). Effects of visitation among allied health professionals. Int J Qual Health Care 18(6): 397-402.

Hopewell, S; Clarke, M; Lefebvre, C; Scherer, R (2007). Handsearching versus electronic searching to identify reports of randomized trials. Cochrane Database Syst Rev 2007(2): MR000001.

Horenkamp-Sonntag, D; Linder, R; Wenzel, F; Gerste, B; Ihle, P (2014). Prüfung der Datenqualität und Validität von GKV-Routinedaten. In: Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven. 2., vollständig überarbeitete und aktualisierte Auflage. Swart, E; Ihle, P; Gothe, H; Matusiewicz (Eds.). Bern: Hans Huber: 314-330.

Hosmer, DW; Lemeshow, S (2000). Applied Logistic Regression. New York: John Wiley & Sons.

Hox, JJ; Maas, CJM (2001). The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. Struct Equ Modeling 8(2): 157-174.

Hussey, PSM, Soeren; Morse, Lindsey; Ridgely, M. Susan; RAND Health (2007). Evaluation of the Use of AHRQ and Other Quality Indicators.

ICH (1999). International Conference on Harmonisation E9 Expert Working Group: Harmonised Tripartite Guideline. Statistical principles for clinical trials. Stat Med 18(15): 1905-1942.

Iezzoni, LI (2003). Risk Adjustment for Measuring Health Care Outcomes. Chicago: Health Administration Press.

Iezzoni, LI (1997). Assessing quality using administrative data. Ann Intern Med 127(8.2): 666-674.

IOM; Committee to Design a Strategy for Quality Review and Assurance in Medicare; Division of Health Care Services; Institute of Medicine (1990). Health, Health Care, and Quality of Care. In: Medicare: A Strategy for Quality Assurance. Volume I. Lohr, KN (Ed.). Washington: National Academy Press: 19-44.

IQWiG (2013). Allgemeine Methoden. Version 4.1 vom 28.11.2013. Köln: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen.

Ivers, N; Jamtvedt, G; Flottorp, S; Young, JM; Odgaard-Jensen, J; French, SD; et al. (2012). Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 6(CD000259).

Jäckel, WH (2009). Definition und Typen von Qualitätsindikatoren. In: Qualitätsindikatoren: Manual für Autoren. Ärztliches Zentrum für Qualität in der Medizin (Ed.). Neukirchen: Make A Book: 2-4.

Jamtvedt, JG; Young, JM; Kristoffersen, DT; O'Brian, MA; Oxman, AD (2006). Audit and feedback: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2006(2): CD000259.

Jarman, B; Pieter, D; van der Veen, AA; Kool, RB; Aylin, P; Bottle, A; et al. (2010). The hospital standardised mortality ratio: a powerful tool for Dutch hospitals to assess their quality of care? Qual Saf Health Care 19(1): 9-13.

JCAHO (1989). Characteristics of clinical indicators. Joint Commission on Accreditation of Healthcare Organisations. QRB Qual Rev Bull 15(11): 330-339.

Jensen, HI; Ammentorp, J; Kofoed, PE (2010). User satisfaction is influenced by the interval between a health care service and the assessment of the service. Soc Sci Med 70(12): 1882-7.

Jha, AK; Li, Z; Orav, EJ; Epstein, AM (2005). Care in U.S. hospitals - the Hospital Quality Alliance program. N Engl J Med 353(3): 265-274.

Joos, S; Miksch, A; Szecsenyi, J; Wieseler, B; Grouven, U; Kaiser, T; et al. (2008). Montelukast as add-on therapy to inhaled corticosteroids in the treatment of mild to moderate asthma: a systematic review. Thorax 63(5): 453-462.

Kanwal, F; Barkun, A; Gralnek, IM; Asch, SM; Kuipers, EJ; Bardou, M; et al. (2010). Measuring quality of care in patients with nonvariceal upper gastrointestinal hemorrhage: development of an explicit quality indicator set. Am J Gastroenterol 105(8): 1710-8.

Katz, MH (2011). Multivariable Analysis: A Practical Guide for Clinicians and Public Health Researchers. Cambridge: Cambridge University Press.

Kay, M (2008). XSLT 2.0 and XPath 2.0 Programmer's Reference. Indianapolis: Wiley Publishing.

KBV (2014). Das QEP-Zertifizierungsverfahren. Kassenärztliche Bundesvereinigung. http://www.kbv.de/html/898.php (letzte Aktualisierung am 2014, Zugriff am 13.11.2014).

Kelly, A; Thompson, JP; Tuttle, D; Benesch, C; Holloway, RG (2008). Public reporting of quality data for stroke: is it measuring quality? Stroke 39(12): 3367-3371.

Kessner, DM; Kalk, CE; Singer, J (1973). Assessing health quality – the case for tracers. N Engl J Med 288(4): 189-94.

Ketelaar, N; Faber, MJ; Flottorp, S; Rygh, LH; Deane, KHO; Eccles, MP (2011). Public release of performance data in changing the behaviour of healthcare consumers, professionals or organisations. Cochrane Database Syst Rev 2011(11): CD004538.

Kilian, R; Gonnerman, C; Seiler, R; Voss, P; Neumann, T; Zunhammer, A; et al. (2009). Die Entwicklung eines Fragebogens zur Beurteilung integrierter Gesundheitsversorgungsprogramme durch Patienten (BiGPAT). Gesundheitswesen 71(8-9): 460-468.

Klemperer, D (1996). Qualität in der Medizin. Der patientenzentrierte Qualitätsbegriff und seine Implikationen. Dr. med. Mabuse (Heft Januar/Februar): 22-27.

Kohlmann, T; Raspe, H (1998a). Zur Messung patientennaher Erfolgskriterien in der medizinischen Rehabilitation: Wie gut stimmen „indirekte" und „direkte" Methoden der Veränderungsmessung überein? Rehabilitation (Stuttg) 37 Suppl 1: S30-7.

Kohlmann, T; Raspe, H (1998b). Zur Messung patientennaher Erfolgskriterien in der medizinischen Rehabilitation: Wie gut stimmen „indirekte" und „direkte" Methoden der Veränderungsmessung überein? Rehabilitation (Stuttg) 37 (Suppl 1): 30-37.

König, T; Barnewold, L; Heller, G (2014). Risikoadjustierung und Fallzahl-Prävalenz-Problem. In: Qualitätsreport 2013. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen: 215-221.

Kopp, I; Geraedts, M; Jäckel, WH; Altenhofen, L; Thomeczek, C; Ollenschläger, G (2007). Nationale VersorgungsLeitlinien – Evaluation durch Qualitätsindikatoren. Med Klin 102(8): 678-682.

Köppen, M; Ruppel, J; Kazmaier, T (2014). Strukturierter Dialog. In: Qualitätsreport 2013. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen: 196–204.

Körner, T; Schmidt, J (2006). Qualitätsberichte – ein neues Informationsangebot über Methoden, Definitionen und Datenqualität der Bundesstatistiken. Wirtsch Stat 2: 109-117.

Kosinski, A; Raspe, H (1998). Patientenzufriedenheit nach einer stationären Rehabilitation : Unterschiedliche Ergebnisse bei unterschiedlichen Befragungszeitpunkten? Gesundheitswesen 60(2): 75-79.

Kötter, T; Blozik, E; Scherer, M (2012). Methods for the guideline-based development of quality indicators – a systematic review. Implement Sci 7: 21.

Kötter, T; Schaefer, F; Blozik, E; Scherer, M (2011). Die Entwicklung von Qualitätsindikatoren – Hintergrund, Methoden und Probleme. Z Evid Fortbild Qual Gesundhwes 105(1): 7-12.

Kristensen, S; Mainz, J; Bartels, P (2009). Selection of indicators for continuous monitoring of patient safety: recommendations of the project 'safety improvement for patients in Europe'. Int J Qual Health Care 21(3): 169-75.

Kunz, R; Khan, K; Kleijnen, J; Antes, G (2009). Systematische Übersichten und Meta-Analysen: Einführung in Instrumente der evidenzbasierten Medizin für Ärzte, klinische Forscher und Experten im Gesundheitswesen. Bern: Hans Huber Verlag.

Lack, N; Gerhardinger, U (2009). Qualitätsvergleich mit Funnelplots - Plädoyer für eine einheitliche Methodik. Z Evid Fortbild Qual Gesundhwes 103(8): 536-41.

Lee, E; Forthofer, E (2006). Analyzing complex survey data. Thousand Oaks/London/New Delhi: SAGE Publications.

Lester, HE; Hannon, KL; Campbell, SM (2011). Identifying unintended consequences of quality indicators: a qualitative study. BMJ Qual Saf 20(12): 1057-61.

Lieb, K; Klemperer, D; Koch, K; Baethge, C; Ollenschläger, G; Ludwig, W-D (2011). Interessenkonflikte in der Medizin: Mit Transparenz Vertrauen stärken. Dtsch Arztebl 108(6): A 256-260.

Lingard, EA; Wright, EA; Sledge, CB; Kinemax Outcomes, G (2001). Pitfalls of using patient recall to derive preoperative status in outcome studies of total knee arthroplasty. J Bone Joint Surg Am 83-A(8): 1149-56.

Lo, B; Field, M (2009). Conflict of Interest in Medical Research, Education, and Practice. Washington, D.C.: Institute of Medicine of the National Academies, The National Academies Press.

Loer, S (2009). Qualitätsmanagement durch kollegiale Visitationen - Erfahrungen aus den Niederlanden. Vortrag, 5. Hamburger Landesergebniskonferenz: Unveröffentlichtes Manuskript.

Lohr, S (2010). Sampling: Design and Analysis. Boston, MA: Brooks/Cole.

Lombarts, MJ; Klazinga, NS (2003). Supporting Dutch medical specialists with the implementation of visitatie recommendations: a descriptive evaluation of a 2-year project. Int J Qual Health Care 15(2): 119-129.

Luft, HS; Hunt, SS (1986). Evaluating individual hospital quality through outcome statistics. JAMA 255(20): 2780-4.

Maas, CJM; Hox, JJ (2005). Sufficient sample sizes for multilevel modeling. Methodology 1(3): 86-92.

MacLean, CH; Louie, R; Shekelle, PG; Roth, CP; Saliba, D; Higashi, T; et al. (2006). Comparison of administrative data and medical records to measure the quality of medical care provided to vulnerable older patients. Med Care 44(2): 141-148.

Mainz, J (2003). Defining and classifying clinical indicators for quality improvement. Int J Qual Health Care 15(6): 523-530.

Mancuso, CA; Charlson, ME (1995). Does recollection error threaten the validity of cross-sectional studies of effectiveness? Med Care 33(4 Suppl): AS77-88.

Marshall, MN; Campbell, S; Hacker, J; Roland, M (2002). Quality indicators for general practice. A practical guide for health professionals and managers. London: Royal Society of Medicine Press.

Marshall, MN; Roland, MO; Campbell, S; Kirk, S; Reeves, D; Brook, RH; et al. (2003). Measuring general practice. A demonstration project to develop and test a set of primary care clinical quality indicators. London: The Nuffield Trust.

Maydanchik, A (2007). Data Quality Assessment. Bradley Beach, NJ: Technics Publications, LLC.

Maynard, H; Chalmers, J (Eds.) (1997). Non-random Reflections on Health Services Research. BMJ Publishing Group.

McColl, A; Roderick, P; Gabbay, J; Smith, H; Moore, M (1998). Performance indicators for primary care groups: an evidence based approach. BMJ 317(7169): 1354-1360.

McDonald, R; Harrison, S; Checkland, K; Campbell, SM; Roland, M (2007). Impact of financial incentives on clinical autonomy and internal motivation in primary care: ethnographic study. BMJ 334(7608): 1357.

McGlynn, EA (2003). An evidence-based national quality measurement and reporting system. Med Care 41(1, Suppl): I8-15.

McGory, ML; Kao, KK; Shekelle, PG; Rubenstein, LZ; Leonardi, MJ; Parikh, JA; et al. (2009). Developing quality indicators for elderly surgical patients. Ann Surg 250(2): 338-47.

Meffert, C; Mittag, O (2010). Ergebnismessung in der medizinischen Rehabilitation: Vergleich unterschiedlicher Methoden der Outcome-Messung anhand von Rehabilitationswissenschaftlichen Datensätzen und kognitiven Interviews über Veränderungsitems. Freiburg: Universitätsklinikum Freiburg.

Meixner, K (2006). Weiterentwicklung und Validierung eines Verfahrens zur Visitation von Rehabilitationseinrichtungen. Rehabilitation (Stuttg) 45: 152-160.

Meyer, T; Richter, S; Raspe, H (2013). Agreement between pre-post measures of change and transition ratings as well as then-tests. BMC Med Res Methodol 13: 52.

Michaels, RK; Makary, MA; Dahab, Y; Frassica, FJ; Heitmiller, E; Rowen, LC; et al. (2007). Achieving the National Quality Forum's "Never Events". Ann Surg 245(4): 526-532.

Middel, B; Goudriaan, H; de Greef, M; Stewart, R; van Sonderen, E; Bouma, J; et al. (2006). Recall bias did not affect perceived magnitude of change in health-related functional status. J Clin Epidemiol 59(5): 503-511.

Mihrshahi, S; Brand, C; Ibrahim, JE; Evans, S; Jolley, D; Cameron, P (2010). Validity of the indicator 'death in low-mortality diagnosis-related groups' for measuring patient safety and healthcare quality in hospitals. Intern Med J 40(4): 250-257.

Mor, V (2005). Improving the Quality of Long-Term Care with Better Information. Milbank Q 83(3): 333-364.

Morgan, D (2009). Focus groups as qualitative research. 2. Aufl. Thousand Oaks: SAGE Publications.

Moscucci, M; Eagle, KA; Share, D; Smith, D; De Franco, AC; O'Donnell, M; et al. (2005). Public reporting and case selection for percutaneous coronary interventions: an analysis from two large multicenter percutaneous coronary intervention databases. J Am Coll Cardiol 45(11): 1759-1765.

Mourad, SM; Hermens, RP; Nelen, WL; Braat, DD; Grol, RP; Kremer, JA (2007). Guideline-based development of quality indicators for subfertility care. Hum Reprod 22(10): 2665-2672.

Mühlhauser, I; Müller, H (2009). Patientenrelevante Endpunkte und patient-reported outcomes in klinischer Forschung und medizinischer Praxis. In: Informiert und selbstbestimmt. Der mündige Bürger als mündiger Patient (Beiträge zum Gesundheitsmanagement, Bd. 24). Klusen, N; Fließgarten, A; Nebling, T (Eds.). Baden-Baden: Nomos: 34-65.

Mukerji, N; Jenkins, A; Nicholson, C; Mitchell, P (2012). Unplanned reoperation rates in pediatric neurosurgery: a single center experience and proposed use as a quality indicator. J Neurosurg Pediatr 9(6): 665-669.

Murphy, G; Scheu, C (2010). Warum ist Transparenz so wichtig? In: Jahrbuch Qualitätsmedizin 2010. Kuhlen, R; Rink, O; Zacher, J (Eds.). Berlin: Medizinisch wissenschaftliche Verlagsgesellschaft: 33-38.

Murphy, MK; Black, NA; Lamping, DL; McKee, CM; Sanderson, CF; Askham, J; et al. (1998). Consensus development methods, and their use in clinical guideline development. Health Technol Assess 2(3): 1-88.

Murray, C (1994). Quantifying the burden of disease: the technical basis for disability-adjusted life years. Bull World Health Organ 73(3): 429-445.

Nadzam, DM; Turpin, R; Hanold, LS; White, RE (1993). Data-driven performance improvement in health care: the Joint Commission's Indicator Measurement System (IMSystem). Jt Comm J Qual Improv 19(11): 492-500.

Nasser, M; Al, HA (2006). A comparison of handsearching versus embase searching of the archives of Iranian medicine to identify reports of randomized controlled trials. Arch Iran Med 9(3): 192-195.

NHS (2007). DoCDat Data Collection Manual. London: The Information Centre for Health and Social Care, National Health Service.

Nieuwkerk, PT (2006). Highly active antiretroviral therapy for HIV-1 infection: patients' quality of life and treatment adherence. Faculty of Medicine. Amsterdam: University of Amsterdam.

Noest, S; Ludt, S; Klingenberg, A; Glassen, K; Heiss, F; Ose, D; et al. (2014). Involving patients in detecting quality gaps in a fragmented healthcare system: development of a questionnaire for Patients' Experiences Across Health Care Sectors (PEACS). Int J Qual Health Care 26(3): 240-9.

Nonnemacher, M; Weiland, D; Stausberg, J (2007). Datenqualität in der medizinischen Forschung: Leitlinie zum adaptiven Management von Datenqualität in Kohortenstudien und Registern. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft.

O'Brian, MA; Rogers, S; Jamtvedt, JG; Oxman, AD; Odgaard-Jensen, J; Kristoffersen, DT; et al. (2007). Educational outreach visits: Effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2007(4): CD000409.

O'Brien, RM (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. Qual Quant 41(5): 673-690.

Ollenschläger, G; Helou, A; Lorenz, W (2000). Kritische Bewertung von Leitlinien. In: Lehrbuch Evidenzbasierte Medizin in Klinik und Praxis. Kunz, R; Ollenschläger, G; Raspe, H; Jonitz, G; Kolkmann, F (Eds.). Köln: Deutscher Ärzte-Verlag.

Ollenschläger, G; Thomeczek, C; Bungart, B; Lampert, U; Oesingmann, U; Kolkmann, F; et al. (1999). Leitlinien – Bedeutung, Verbindlichkeit und Qualität: Welche Problematik liegt in der Entwicklung und Einführung von Leitlinien? In: Krankenhausmanagement im Spannungsfeld zwischen Qualität und Kosten – mit einer Evidenz-basierten Medizin (EBM) und Leitlinien zu einer verbesserten Patientenversorgung? Osnabrücker Studien Band 16. Wirtschafts- und sozialwissenschaftlicher Arbeitskreis Osnabrück (Ed.). Osnabrück: Fachhochschule Osnabrück: 71-92.

Olsen, JE (2003). Data quality: the accuracy dimension. San Francisco: Morgan Kaufman Publishers.

Ose, D; Grande, G; Badura, B; Greiner, W (2008). Patienteninformation zur Bewertung von Gesundheitseinrichtungen. Möglichkeiten und Perspektiven zur Entwicklung eines patientenorientierten Informationsportals vor dem Hintergrund aktueller Initiativen zur Qualitätssicherung im Gesundheitssystem. Präv Gesundheitsf 3(3): 152-162.

Ouwens, MM; Hermens, RR; Termeer, RA; Vonk-Okhuijsen, SY; Tjan-Heijnen, VC; Verhagen, AF; et al. (2007). Quality of integrated care for patients with nonsmall cell lung cancer: variations and determinants of care. Cancer 110(8): 1782-1790.

Pearson, RW (2010). Statistical Persuasion: How to Collect, Analyze, and Present Data ... Accurately, Honestly, and Persuasively. Tousand Oaks, CA: Sage.

Pellise, F; Vidal, X; Hernandez, A; Cedraschi, C; Bago, J; Villanueva, C (2005). Reliability of retrospective clinical data to evaluate the effectiveness of lumbar fusion in chronic low back pain. Spine (Phila Pa 1976) 30(3): 365-8.

Pfaff, H; Glaeske, G; Neugebauer, E; Schrappe, M (2009). Memorandum III: Methoden für die Versorgungsforschung (Teil I). Gesundheitswesen 71: 505-510.

Pfaff, H; Schrappe, M (2011). Evidence-based Health Care. In: Lehrbuch Versorgungsforschung. Systematik - Methodik - Anwendung. Paff, H; Neugebauer, EAM; Glaeske, G; Schrappe, M (Eds.). Stuttgart: Schattauer: 33-39.

Pfandzelter, R; Walter, A; Wegscheider, K (2009). Stichproben in der ambulanten Qualitätssicherung am Beispiel der Koloskopie. Z Evid Fortbild Qual Gesundhwes 103(3): 159-164.

Phibbs, CS; Mark, DH; Luft, HS; Peltzman-Rennie, DJ; Garnick, DW; Lichtenberg, E; et al. (1993). Choice of Hospital for Delivery: A Comparison of High-Risk and Low-Risk Women. Health Serv Res 28(2): 201-222.

Pine, M; Jordan, HS; Elixhauser, A; Fry, DE; Hoaglin, DC; Jones, B; et al. (2007). Enhancement of claims data to improve risk adjustment of hospital mortality. JAMA 297(1): 71-76.

Poloniecki, J; Valencia, O; Littlejohns, P (1998). Cumulative risk adjusted mortality chart for detecting changes in death rate: observational study of heart surgery. BMJ 316(7146): 1697-1700.

Prüfer, P; Rexroth, M (2005). Kognitive Interviews. Mannheim.

Prüfer, P; Rexroth, M (2000). Zwei-Phasen-Pretesting. Mannheim: GESIS - Leibniz-Institut für Sozialwissenschaften.

Rabash, J; et al. (2001). A User's Guide to MLwiN. Version 2.1c. University of London: Centre for Multilevel Modelling. Institute of Education.

Rammstedt, B (2004). Zur Bestimmung der Güte von von Multi-Item-Skalen: Eine Einführung. Mannheim.

Raudenbush, SW; Bryk, AS (2002). Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks, CA: SAGE Publications.

Rauprich, O; Nolte, M; Vollmann, J (2010). Systematische Literaturrecherche in den Datenbanken PubMed und BELIT. Ein Werkstattbericht. Ethik Med 22(1): 59-67.

Rawlings, JO; Dickey, DA; Pantula, SG (2001). Applied Regression Analysis: A Research Tool. New York: Springer.

Reeves, D; Campbell, SM; Adams, J; Shekelle, PG; Kontopantelis, E; Roland, MO (2007). Combining multiple indicators of clinical quality: an evaluation of different analytic approaches. Med Care 45(6): 489-96.

Reiter, A; Fischer, B; Kötting, J; Geraedts, M; Jäckel, WH; Döbler, K (2008). QUALIFY: Ein Instrument zur Bewertung von Qualitätsindikatoren. Z ärztl Fortbild Qual Gesundhwes 101(10): 683-688.

Robinson, S; Brodie, M (1997). Understanding the quality challenge for health consumers: the Kaiser/AHCPR Survey. Jt Comm J Qual Improv 23(5): 239-244.

Rodriguez, HP; von Glahn, T; Chang, H; Rogers, WH; Safran, DG (2007). Patient samples for measuring primary care physician performance: who should be included? Med Care 45(10): 989-996.

Rogers, CA; Reeves, BC; Caputo, M; Ganesh, JS; Bonser, RS; Angelini, GD (2004). Control chart methods for monitoring cardiac surgical performance and their interpretation. J Thorac Cardiovasc Surg 128(6): 811-9.

Rollins, G (2004). Most hospitals lack enough procedure volume to spot surgical quality problems. Rep Med Guidel Outcomes Res 15(18): 7-9.

Romano, PS; Mull, HJ; Rivard, PE; Zhao, S; Henderson, WG; Loveland, S; et al. (2009). Validity of selected AHRQ patient safety indicators based on VA National Surgical Quality Improvement Program data. Health Serv Res 44(1): 182-204.

Romano, PS; Zach, A; Luft, HS; Rainwater, J; Remy, LL; Campa, D (1995). The California Hospital Outcomes Project: using administrative data to compare hospital performance. Jt Comm J Qual Improv 21(12): 668-682.

Roos, EM (2012). The 2012 User's Guide to: Knee injury and Osteoarthritis Outcome Score KOOS.

Roos, EM; Lohmander, LS (2003). The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes 1: 64.

Rose, D; Evans, J; Sweeney, A; Wykes, T (2011). A model for developing outcome measures from the perspectives of mental health service users. Int Rev Psychiatry 23(1): 41-6.

Rothman, KJ; Greenland, S; Lash, TJ (2008). Modern Epidemiology (Third Edition). Philadelphia: Lippincott, Williams & Wilkins.

Rubin, HR; Pronovost, P; Diette, GB (2001). The advantages and disadvantages of process-based measures of health care quality. Int J Qual Health Care 13(6): 469-474.

Ruprecht, T (2001). Patientenbefragungen (Verfahrensvorschlag). In: Handbuch für Netzberater. Controlling, Evaluation, Audit, Berichterstattung. Köln: Kassenärztliche Bundesvereinigung: 91-110.

Sackett, D; Rosenberg, W; Gray, J; Haynes, R; Richardson, W (1997). Was ist Evidenz-basierte Medizin und was nicht? Munch Med Wochenschr 139(44): 644-5.

Sankar, A; Johnson, SR; Beattie, WS; Tait, G; Wijeysundera, DN (2014). Reliability of the American Society of Anesthesiologists physical status scale in clinical practice. Br J Anaesth 113(3): 424-432.

Saur, P; Junker, U; Gaus, P; Haeske-Seeberg, H; Blöchle, C; Neugebauer, E (2008). Implementierung eines standardisierten perioperativen Schmerzmanagementkonzepts in drei Krankenhäusern eines Klinikverbundes. Schmerz 22(1): 34-42.

Schäfer, E; Birkner, N; Grothaus, FJ; Dombrowski, M; Handstein, S; Lorenz-Wangard, A; et al. (2010). BQS-Patientinnenbefragung für Frauen mit primärem Mammakarzinom. Senologie 7: 245-248.

Schiavo-Campo, S (1999). Strengthening “Performance” in Public Expenditure Management. Asian Rev Public Admin 11(2): 23-44.

Schmidt, J; Steffanowski, A; Nübling, R; Lichtenberg, S; Wittmann, W (2004). EQUA-Studie. Erfassung der Ergebnisqualität stationärer psychosomatischer Rehabilitationsbehandlungen. Vergleich unterschiedlicher Evaluationsstrategien und Entwicklung neuer Messinstrumentarien. Abschlussbericht. Privatinstitut für Evaluation und Qualitätssicherung im Gesundheits- und Sozialwesen mbH, Karlsruhe/Universität Mannheim, Lehrstuhl Psychologie II.

Schneeweiss, S; Sangha, O (2001). Leistungsvergleiche in der Medizin: Bedarf, Anforderung und Wege zur Akzeptanz. Dtsch Med Wochenschr 126(33): 918-924.

Schneider, A; Broge, B; Szecsenyi, J (2003). Müssen wir messen, um (noch) besser werden zu können? Die Bedeutung von Qualitätsindikatoren in strukturierten Behandlungsprogrammen und Qualitätsmanagement. Z Allg Med 79(11): 547-552.

Schnell, R; Hill, PB; Esser, E (2013). Methoden der empirischen Sozialforschung. 10., überarbeitete Auflage. München: Oldenbourg Wissenschaftsverlag.

Schrappe, M (2015). Qualität 2030. Die umfassende Strategie für das Gesundheitswesen. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft.

Schrappe, M (2011). Qualitätswettbewerb. In: Lehrbuch Versorgungsforschung. Systematik - Methodik - Anwendung. Paff, H; Neugebauer, E; Glaeske, G; Schrappe, M (Eds.). Stuttgart: Schattauer: 23-207.

Schubert, J; Kaiser, A; Kazmaier, T (2014). Datenvalidierung. In: Qualitätsreport 2013. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen: 205-211.

Schulz, M; Mack, B; Renn, O (2012). Fokusgruppen in der empirischen Sozialwissenschaft: von der Konzeption bis zur Auswertung. Berlin: Springer VS.

Schumacher, M; Schulgen-Kristiansen, G (2008). Methodik klinischer Studien: Methodische Grundlagen der Planung, Durchführung und Auswertung. Berlin/Heidelberg: Springer.

Schwartz, WF; Klein-Lange, M (2003). Konzepte und Strukturen der Krankenversorgung. In: Das Public Health Buch - Gesundheit und Gesundheitswesen. 2., völlig neu bearbeitete und erweiterte Aufl. Schwartz, FW; Badura, B; Busse, R; Leidl, R; Raspe, H; Siegrist, J; et al. (Eds.). München, Jena: Urban & Fischer: 270-93.

Sebastian, J (2009). The Art of Xsd - SQL Server XML Schemas. Cambridge: Red Gate Books.

Seber, GAF; Lee, AJ (2003). Linear Regression Analysis. Hoboken: Wiley-Interscience.

Semlitsch, T; Jeitler, K; Kopp, IB; Siebenhofer, A (2014). Entwicklung einer praktikablen Mini-Checkliste zur Bewertung der methodischen Leitlinienqualität. Z Evid Fortbild Qual Gesundhwes 108(5-6): 299-312.

Sens, B; Fischer, B; Bastek, A; Eckardt, J; Kaczmarek, D; Paschen, U; et al. (2007). Begriffe und Konzepte des Qualitätsmanagements. 3. Aufl. GMS Med Inform Biom Epidemiol 3(1): Doc05.

Shahian, DM; Normand, SL; Torchiana, DF; Lewis, SM; Pastore, JO; Kuntz, RE; et al. (2001). Cardiac surgery report cards: comprehensive review and statistical critique. Ann Thorac Surg 72(6): 2155-2168.

Shahian, DM; Silverstein, T; Lovett, AF; Wolf, RE; Normand, SL (2007). Comparison of clinical and administrative data sources for hospital coronary artery bypass graft surgery report cards. Circulation 115(12): 1518-1527.

Simansiki, C; Lefering, R; Paffrath, T; Riess, P; Yülcel, N; Maegele, M (2006). Die Qualität der postoperativen Schmerztherapie beeinflusst die Krankenhauswahl. Schmerz 20(4): 327-333.

Singer, JD (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. J Educ Behav Stat 24(4): 323-355.

Singer, S; Gotze, H; Mobius, C; Witzigmann, H; Kortmann, RD; Lehmann, A; et al. (2009). Quality of care and emotional support from the inpatient cancer patient's perspective. Langenbecks Arch Surg 394(4): 723-731.

Singh, S (2003). Advanced sampling theory with applications: How Michael 'selected' Amy. Dordrecht: Kluwer Academic Publishers.

Slobogean, GP; Verma, A; Giustini, D; Slobogean, BL; Mulpuri, K (2009). MEDLINE, EMBASE, and Cochrane index most primary studies but not abstracts included in orthopedic meta-analyses. J Clin Epidemiol 62(12): 1261-1267.

Smith, P (1995). On the unintended consequences of publishing performance data in the public sector. Int J Publ Admin 18(2&3): 277-310.

Snijders, T (1999). Multilevel Analysis. An Introduction to Basic and Adanced Multilevel Modeling. London: Sage Publications.

Socialstyrelsen (2010). Bilaga 2. Täckningsgradsjämförelser mellan hälsodataregistren vid Socialstyrelsen och Nationella Kvalitetsregister. Stockholm: Sveriges Kommuner och Landsting.

SooHoo, NF; Lieberman, JR; Farng, E; Park, S; Jain, S; Ko, CY (2011). Development of quality of care indicators for patients undergoing total hip or total knee replacement. BMJ Qual Saf 20(2): 153-157.

Spiegelhalter, DJ; Thomas, A; Best, NG; Lund, D (2003). WinBUGS Version 1.4 Users Manual. http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf (Zugriff am 15.11.2012).

Staiger, DO; Dimick, JB; Baser, O; Fan, Z; Birkmeyer, JD (2009). Empirically derived composite measures of surgical performance. Med Care 47(2): 226-233.

Stausberg, J; Nonnemacher, M; Weiland, D; Antony, G; Neuhäuser, M (2006). Management of Data Quality - Development of a Computer-Mediated Guideline. In: Ubiquity: Technologies for Better Health in Aging Societies. Proceedings of MIE2006. Hasman, A; Haux, R; Van der Lei, J (Eds.). Amsterdam: IOS Press: 477-482.

Steen, M; Ammann, C; Seyderhelm, A; Vey, S (2014). Pflege bestehender Leistungsbereiche. In: Qualitätsreport 2013. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen: 192-194.

Sullivan, LM; Dukes, KA; Losina, E (1999). Tutorial in biostatistics. An introduction to hierarchical linear modelling. Stat Med 18(7): 855-888.

Supekar, K; Marwadi, A; Lee, Y; Medhi, D (2002). Fuzzy Rule-Based Framework for Medical Record Validation. In: Intelligent Data Engineering and Automated Learning – IDEAL 2002. Lecture Notes in Computer Science. Yin, H; Allinson, N; Freeman, R; Keane, J; Hubbard, S (Eds.). Berlin/Heidelberg: Springer. 2412: 1-27.

SVR-E (2005). Die Chancen nutzen - Reformen mutig voranbringen. Jahresgutachten 2005/06. Wiesbaden: Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung.

SVR-G (2007). Kooperation und Verantwortung - Voraussetzungen einer zielorientierten Gesundheitsversorgung. Gutachten 2007. Bonn: Sachverständigenrat zur Begutachtung der Entwicklung im Gesundheitswesen.

SVR-KAG (2003). Finanzierung, Nutzerorientierung und Qualität. Gutachten 2003. Bonn: Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen.

Swart, E; Heller, G (2007). Nutzung und Bedeutung von (GKV-)Routinedaten für die Versorgungsforschung. In: Medizinsoziologische Versorgungsforschung. Theoretische Ansätze, Methoden, Instrumente und empirische Befunde. Janßen, C; Borgetto, B; Heller, G (Eds.). Weinheim: Juventa: 93-111.

Swart, E; Ihle, P; Gothe, H; Matusiewicz (2014). Routinedaten im Gesundheitswesen. Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven. 2., vollständig überarbeitete und aktualisierte Aufl. Bern: Hans Huber.

Szecsenyi, J; Broge, B; Stock, J (2009). QiSA. Qualitätsindikatorensystem für die ambulante Versorgung. Band A. Berlin: KomPart-Verlag.

Szecsenyi, J; Campbell, S; Broge, B; Laux, G; Willms, S; Wensing, M; et al. (2011). Effectiveness of a quality-improvement program in improving management of primary care practices. CMAJ 183(18): E 1326-33.

Tarling, R (2009). Statistical modelling for social researchers: Principles and practice. Abingdon: Routledge.

Tekkis, PP; McCulloch, P; Steger, AC; Benjamin, IS; Poloniecki, JD (2003). Mortality control charts for comparing performance of surgical units: validation study using hospital mortality data. BMJ 326(7393): 786-788.

Terwee, CB; Bot, SD; de Boer, MR; van der Windt, DA; Knol, DL; Dekker, J; et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 60(1): 34-42.

Thomas, T; Konheiser, S; Heller, G (2014). Transparente Darstellung der Versorgungsqualität von Frühgeborenen online. In: Qualitätsreport 2013. Göttingen: AQUA – Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen: 233-235.

Thornton, A; Lee, P (2000). Publication bias in meta-analysis: its causes and consequences. J Clin Epidemiol 53(2): 207-216.

Trojan, A; Satzinger, W (2001). Bilanz: Was ist zu beachten, damit Patientenbefragungen die Patientenversorgung verbessern helfen? In: Patientenbefragungen in Krankenhäusern. Konzepte, Methoden, Erfahrungen. St. Augustin: Asgard-Verlag: 377-387.

van den Hombergh, P; Grol, R; van den Hoogen, HJ; van den Bosch, WJ (1998). Assessment of management in general practice: validation of a practice visit method. Br J Gen Pract 48(436): 1743-1750.

van der Lei, J (1991). Use and abuse of computer-stored medical records. Methods Inf Med 30(2): 79-80.

Verbraucherzentrale NRW (2014). Qualitätsberichte der Krankenhäuser – Nutzen und Mängel für die Krankenhaussuche von Patienten. Düsseldorf: Verbraucherzentrale Nordrhein-Westfalen. Gruppe Gesundheits- und Pflegemarkt.

Vespermann, A; Borcherding, E (2009). Systematische Literaturrecherche für HTA-Berichte: Qualitätsgesichertes Vorgehen beim DIMDI. Z Evid Fortbild Qual Gesundhwes 103(6): 316-318.

Vonhoegen, H (2009). Einstieg in XML. Grundlagen, Praxis, Referenz. Bonn: Galileo Press.

W3C (2008). XML Encryption WG. http://www.w3.org/Encryption/2001/ (letzte Aktualisierung am 02.06.2008, Zugriff am 13.11.2014).

Weisberg, S (2005). Applied Linear Regression. Hoboken, NJ: Wiley Interscience.

Welke, KF; Karamlou, T; Ungerleider, RM; Diggs, BS (2010). Mortality rate is not a valid indicator of quality differences between pediatric cardiac surgical programs. The Annals of thoracic surgery 89(1): 139-144; discussion 145-146.

Wennberg, J (1990). On the Need for Outcomes Research and the Prospects for Evaluative Clinical Science. In: The Challenges of Medical Practice Variations. Andersen, T; Mooney, G (Eds.). London: McMillan Press: 158-73.

Wennberg, J; Gittelsohn, A (1973). Small Area Variations in Health Care Delivery. Science 182(4117): 1102-8.

Wensing, M; Broge, B; Kaufmann-Kolle, P; Andres, E; Szecsenyi, J (2004). Quality circles to improve prescribing patterns in primary medical care: what is their actual impact? J Eval Clin Pract 10(3): 457-466.

Wensing, M; Broge, B; Riens, B; Kaufmann-Kolle, P; Akkermans, R; Grol, R; et al. (2009). Quality circles to improve prescribing of primary care physicians. Three comparative studies. Pharmacoepidemiol Drug Saf 18(9): 763-769.

Wensing, M; van der Eijk, M; Koetsenruijter, J; Bloem, BR; Munneke, M; Faber, M (2011). Connectedness of healthcare professionals involved in the treatment of patients with Parkinson's disease: a social networks study. Implement Sci 6(67): 1-8.

Wensing, M; van Lieshout, J; Jung, HP; Hermsen, J; Rosemann, T (2008). The Patients Assessment Chronic Illness Care (PACIC) questionnaire in The Netherlands: a validation study in rural general practice. BMC Health Serv Res 8(182).

WIdO; et al. (2007). Qualitätssicherung der stationären Versorgung mit Routinedaten (QSR) - Abschlussbericht. Bonn: AOK-Bundesverband, Forschungs- und Entwicklungsinstitut für das Sozial- und Gesundheitswesen Sachsen-Anhalt, HELIOS Kliniken, Wissenschaftliches Institut der AOK.

Williams, SC; Schmaltz, SP; Morton, DJ; Koss, RG; Loeb, JM (2005). Quality of care in U.S. hospitals as reflected by standardized measures, 2002-2004. N Engl J Med 353(3): 255-264.

Willms, G; Bramesfeld, A; Pottkämper, K; Broge, B; Szecsenyi, J (2013). Aktuelle Herausforderungen der externen Qualitätssicherung im deutschen Gesundheitswesen. Z Evid Fortbild Qual Gesundhwes 107(8): 523-7.

Wilson, C (1987). Hospital-Wide Quality Assurance: Models for Implementation and Development. London: W.B. Saunders.

Winkler, WE (2004). Methods for evaluating and creating data quality. Information Systems 29(7): 531-550.

Wollersheim, H; Hermens, R; Hulscher, M; Braspenning, J; Ouwens, M; Schouten, J; et al. (2007). Clinical indicators: development and applications. Neth J Med 65(1): 15-22.