

Aus dem Institut für Gesundheitsökonomie und Klinische Epidemiologie  
der Universität zu Köln

Direktor (beurlaubt):

Universitätsprofessor Dr. med. Dr. Sc. (Harvard) K. W. Lauterbach MdB

Kommissarische Direktorin: Frau Privatdozentin Dr. med. S. A. K. Stock

# Wirksamkeit von Qualitätssicherung und Qualitätsmanagement in der stationären Versorgung

Evaluierte Verfahren und Ergebnisse

Inaugural-Dissertation zur Erlangung der Doktorwürde  
der Hohen Medizinischen Fakultät  
der Universität zu Köln

vorgelegt von

Carmen Khan

aus Berlin

promoviert am 06. August 2014

Dekan: Universitätsprofessor Dr. med. Dr. h.c. Th. Krieg  
1. Berichterstatter: Professor Dr. med. Dr. rer. nat. G. Ollenschläger  
2. Berichterstatter: Universitätsprofessor Dr. phil. H. Pfaff

#### Erklärung

Ich erkläre hiermit, dass ich die vorliegende Dissertationsschrift ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskriptes habe ich keine Unterstützungsleistungen erhalten.

Weitere Personen waren an der geistigen Herstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die Hilfe einer Promotionsberaterin/eines Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertationsschrift stehen.

Die Dissertationsschrift wurde von mir bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.



Köln, den 06.03.2014

Die Dissertation erfolgte in Kooperation mit dem Ärztlichen Zentrum für Qualität in der Medizin (gemeinsame Einrichtung der Bundesärztekammer und der Kassenärztlichen Bundesvereinigung) unter Leitung von Professor Dr. rer. nat. Dr. med. Günter Ollenschläger. Sämtliche Tätigkeiten bei der Erstellung der vorliegenden Arbeit wie Durchführung der Literaturrecherche, qualitative Bewertung und Klassifikation der Studien, Auswertung der Ergebnisse sowie Erstellung der Grafiken wurden von mir selbst durchgeführt.

## Danksagung

Ich danke ganz besonders Herrn Professor Ollenschläger für seine beharrliche Förderung und stets wohlwollende Unterstützung der Arbeit. Seine vielen konstruktiven und anspornenden Vorschläge waren für die Realisierung des Projekts von unschätzbarem Wert.

Des Weiteren möchte ich meiner Familie, meinen Freunden und den Kolleginnen im ÄZQ für die sehr verlässliche und kraftspendende Wegbegleitung danken.

# Inhaltsverzeichnis

Abkürzungsverzeichnis .....	VII
1 Einleitung .....	1
1.1 Begriffe .....	3
1.2 Entwicklung von Qualitätsmanagement und Qualitätssicherung in der stationären Versorgung.....	10
1.3 Problemstellung .....	14
1.4 Fragestellung und Zielsetzung .....	18
2 Methodik.....	20
2.1 Studienauswahl und -bewertung.....	23
2.1.1 Studienauswahl.....	23
2.1.2 Bewertung und Evidenzdarstellung.....	25
3 Ergebnisse.....	27
3.1 Deutschland .....	31
3.1.1 Fächerübergreifend.....	31
3.1.1.1 Externe vergleichende Qualitätssicherung nach §137 SGB V .....	32
3.1.1.2 Disease Management Programme im stationären Sektor .....	36
3.1.1.3 Qualitätssicherungsverfahren bei der Helios Kliniken GmbH.....	37
3.1.1.4 Die Initiative Qualitätsmedizin (IQM) .....	39
3.1.2 Operative Medizin.....	42
3.1.2.1 Schmerztherapie .....	42
3.1.2.2 Unfallchirurgie .....	44
3.1.3 Konservative Medizin .....	46
3.1.3.1 Krankenhaushygiene .....	46
3.1.3.2 Nephrologie.....	46
3.1.3.3 Rehabilitation .....	47
3.1.4 Onkologie .....	51
3.1.4.1 Gynäkologische Onkologie .....	51
3.1.4.2 Pädiatrische Onkologie.....	52
3.2 International .....	57
3.2.1 Länderübergreifend .....	57
3.2.2 Frankreich.....	58
3.2.3 Spanien .....	59
3.2.4 USA.....	59
3.2.5 China .....	61
3.3 Effekte von Instrumenten für Qualitätsmanagement/Qualitätssicherung .....	63

3.3.1	Leitlinien .....	63
3.3.2	Klinische Behandlungspfade .....	66
3.3.3	Patientensicherheit .....	68
3.3.4	Zertifizierung.....	69
3.3.5	Sonstiges.....	71
4	Diskussion .....	74
4.1	Kritische Analyse der Studien .....	74
4.2	Ausblick.....	83
5	Zusammenfassung .....	86
6	Literaturverzeichnis .....	89
7	Anhang .....	103
7.1	Suchstrategie .....	103
7.2	Evidenztabelle – Studiendarstellung Deutschland .....	106
7.3	Evidenztabelle – Studiendarstellung international .....	117
7.4	Kriterien für die Klassifikation der Studienbewertung .....	125
7.5	Checkliste zur Studienbewertung.....	126
7.6	Dokumentation Volltextsichtung für die primäre Forschungsfrage .....	128
7.7	Abbildungsverzeichnis .....	140
7.8	Tabellenverzeichnis.....	140
8	Lebenslauf .....	141

## Abkürzungsverzeichnis

AHRQ	Agency for Healthcare Research and Quality
AWMF	Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften
ÄZQ	Ärztliches Zentrum für Qualität in der Medizin
BÄK	Bundesärztekammer
BMG	Bundesministerium für Gesundheit
BQS	Bundesgeschäftsstelle Qualitätssicherung
DMP	Disease Management Programm
DRV	Deutsche Rentenversicherung
G-BA	Gemeinsamer Bundesausschuss
GKV	Gesetzliche Krankenversicherung
IQM	Initiative Qualitätsmedizin
KBV	Kassenärztliche Bundesvereinigung
KfH	Kuratorium für Dialyse und Nierentransplantation
OECD	Organisation for Economic Co-operation and Development
QiN	Qualität in der Nephrologie
QM	Qualitätsmanagement
QS	Qualitätssicherung
QUIPS	Qualitätsverbesserung in der postoperativen Schmerztherapie
RCT	Randomised Controlled Trial (Randomisiert kontrollierte Studie)
SGB V	Sozialgesetzbuch (Fünftes Buch)
SOP	Standard Operating Procedure (Standardisierte Arbeitsanweisungen)
TQM	Total Quality Management (Umfassendes Qualitätsmanagement)

# 1 Einleitung

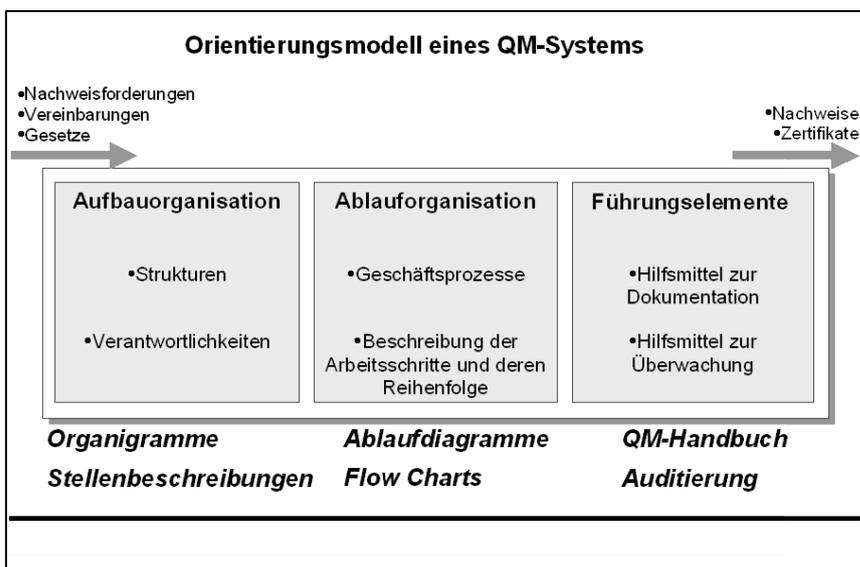
„Qualität der Patientenversorgung“ gehört seit Jahren zu den am häufigsten bemühten Schlagworten in der Gesundheitspolitik und Versorgungsforschung. Begründet wird dies unter anderem mit den zunehmend komplexer werdenden Versorgungsprozessen, der Ökonomisierung des Gesundheitswesens und mit dem berechtigten Wunsch der Öffentlichkeit nach Transparenz in den Dienstleistungssektoren [132; 149; 175].

Vor diesem Hintergrund wurde in Deutschland eine Fülle von Methoden, Maßnahmen und Organisationen etabliert, die auf Erhalt bzw. Verbesserung der medizinischen Versorgungsqualität und Patientensicherheit zielen. Angesichts des Aufwands und der Kosten wurde immer wieder gefordert, etablierte „Qualitäts-Programme“ auf ihre Praxistauglichkeit, Wirksamkeit und Kostenangemessenheit hin zu überprüfen [93; 134; 183].

**Ziel dieser Arbeit war es, mithilfe einer systematischen Literaturliteraturanalyse existierende Qualitätsmanagement/Qualitätssicherungs-Verfahren kritisch hinsichtlich ihrer Wirksamkeit zu beleuchten.**

Die Untersuchung analysiert medizinische Qualitätsmanagement(QM)-Programme, das heißt komplexe Maßnahmen zur Qualitätssicherung/Qualitätsverbesserung (siehe Abbildung 1), welche in verschiedenen Bereichen einer oder mehrerer Kliniken zur Anwendung kamen und einer Evaluation hinsichtlich des Nutzens für eine verbesserte medizinische Patientenversorgung unterzogen wurden. Diese Programme bestehen üblicherweise aus einer Kombination von verschiedenen Methoden und Instrumenten, die zur Optimierung der Versorgungsqualität beitragen sollen.

**Abbildung 1 Grundelemente komplexer Qualitätsmanagementprogramme [11]**



Außerdem wurden ausgewählte, **evaluierte Einzelkomponenten** von Qualitätssicherungsprogrammen untersucht. Aus der Fülle der gebräuchlichen Instrumente (siehe Tabelle 1) konnten durch die systematische Literaturrecherche nur Studien zur Wirksamkeit von Leitlinien, klinischen Behandlungspfaden, Projekten der Patientensicherheit und Zertifizierungen identifiziert werden.

**Tabelle 1: Beispiele für QM/QS-Instrumente und Verfahren modifiziert nach [93]**

<b>Ziele</b>	<b>Beispiele für QM/QS-Instrumente / Verfahren</b>
<b>Nachhaltige Qualitätsverbesserung</b>	<ul style="list-style-type: none"> <li>• Umfassendes Qualitätsmanagement</li> <li>• Risikomanagement</li> </ul>
<b>Beschreibung der Versorgungsqualität, Analyse von Qualitätsproblemen</b>	<ul style="list-style-type: none"> <li>• Instrumente zur Selbst- und Fremdbeurteilung, z. B. Checklisten, Arztbriefe, Beschwerdewesen</li> <li>• Hygienekontrollen, Sicherheitskontrollen, Qualitätskontrollen, Ringversuche</li> <li>• Dokumentation von Primär- oder Sekundärdaten mit qualitätsfördernder Intention, z. B. problemorientierte Krankenakten, Erfassung von unerwünschten Arzneimittelwirkungen, Register, Datenerhebung anhand von Qualitätsindikatoren (z. B. extern-vergleichende Verfahren), Krankheitspässe für Patienten und Patientenbücher</li> <li>• Patienteninterviews, Mitarbeiterbefragungen</li> </ul>
<b>Maßstäbe zur Beurteilung der Qualität</b>	<ul style="list-style-type: none"> <li>• Versorgungsziele, Leitlinien, Pflegestandards, Gesetze, Richtlinien, Rahmenvereinbarungen</li> <li>• Vergleich mit anderen Einrichtungen oder Regionen, mit „Best Practice“-Modellen (Benchmarking)</li> <li>• Qualitätsindikatoren</li> </ul>
<b>Feed-back-Methoden</b>	<ul style="list-style-type: none"> <li>• Rückmeldung interner Datenerhebungen und Qualitätsbefunde, z. B. in Teambesprechungen</li> <li>• Rückmeldung extern-vergleichender Datenerhebungen, z. B. Verordnungsspiegel von Arzneimitteln</li> <li>• Peer-Reviews, z. B. Obduktionen, medizinische Audits, Zweitmeinungen</li> </ul>
<b>Identifizierung/ Verbreitung/ Umsetzung von Lösungsansätzen („Best Practice“)</b>	<ul style="list-style-type: none"> <li>• Fortbildung, Qualitätszirkel, Qualitätshandbücher, Unterstützung von Entscheidungen und/oder der Koordination von Handlungsabläufen, z. B. Leitlinien, Pflegestandards, Behandlungspfade, Patienteninformationen, Reminder, Expertensysteme</li> <li>• Supervision, Arztvisiten, Konsile, interdisziplinäre Visiten, Fallkonferenzen</li> <li>• Externer Peer-Review, z. B. Praxisbesuche (outreach-visit), Konsultationen, Audit</li> <li>• Qualitätsmessungen, Leistungsberichte, Qualitätsberichte</li> </ul>
<b>Darstellung von QM/QS-Verfahren und ihrer Ergebnisse</b>	<ul style="list-style-type: none"> <li>• Selbst- und/oder Fremdbeurteilung des QM/QS-Systems, von Versorgungsstrukturen und -prozessen, ggf. von Patientenzufriedenheit und gesundheitlichen Ergebnissen, z. B. Zertifizierung, Qualitätspreise, Qualitätsberichte</li> </ul>

## 1.1 Begriffe

Es existiert eine Vielzahl an Begrifflichkeiten im Bereich der Qualitätsförderung wie Qualität, Qualitätsmanagement, Qualitätssicherung und Qualitätsindikatoren. Diese werden zum Teil sehr heterogen, zum Teil auch synonym in ihrer Bedeutung verwendet. Auf Grund der bestehenden Begriffsverwirrung sollten diese definiert und deren Inhalte und Zusammenhänge erklärt werden [75]. Nachfolgend werden daher die wichtigsten Begriffe dieser Thematik erläutert.

### Qualität

Die Qualität der Gesundheitsversorgung umfasst mehrere Komponenten. 1990 beschrieb Donabedian, ein bekannter Mitbegründer der Qualitätsforschung im Gesundheitswesen und international führend in der Definition von Qualitätsbegriffen, sieben Aspekte, nach denen die Qualität der Gesundheitsversorgung beurteilt werden kann [54]. Diese sind zum einen die potentielle Wirksamkeit einer Behandlung hinsichtlich der Verbesserung des Gesundheitszustands des Patienten unter optimalen Versorgungsbedingungen (efficacy), zum anderen die tatsächlich erzielte Wirksamkeit unter realen Bedingungen (effectiveness). Weiterhin werden der Aspekt der Effizienz durch das Erzielen einer größtmöglichen Gesundheitsverbesserung unter geringstmöglichen Kosten (efficiency) und die optimale Berücksichtigung von Kosten und Nutzen (optimality) angeführt. Als weitere relevante Aspekte von Qualität gelten die Angemessenheit der Versorgung in Übereinstimmung mit Patientenpräferenzen (acceptability), die Festlegung einzelner Qualitätskriterien in Konformität mit den Interessen der Gemeinschaft (legitimacy) und die gerechte Ressourcenallokation im Gesundheitssystem (equity) [54].

Diese Qualitätsaspekte sind weiterhin relevant, jedoch für die konkrete Erhebung und Beurteilung der Qualität nur zum Teil anwendbar. Zur greifbareren Beurteilung von Qualität werden drei Qualitätsdimensionen unterschieden, die für Messmethoden geeignet sind.

Diese sind

- **Strukturqualität** (materielle, personelle und organisatorische Ausstattung des Arbeitsplatzes)
- **Prozessqualität** (Aktivitäten zur Diagnosestellung und Behandlung)
- **Ergebnisqualität** (erzielte Verbesserung des Gesundheitszustands oder der Lebensqualität des Patienten),

die in gegenseitiger Wechselwirkung miteinander stehen [52; 53].

Die Ergebnisqualität kann auch unbeeinflusst von den anderen Qualitätsdimensionen bleiben. So ist es beispielsweise auch bei geringer Qualität eines Behandlungsprozesses möglich, eine gute Ergeb-

nisqualität auf Grund der hohen Anzahl an selbstlimitierenden Spontanverläufen von Krankheiten zu erzielen [199]. Auf der anderen Seite ist auch bei optimaler Struktur- und Prozessqualität - beispielsweise bei fehlender Adhärenz des Patienten - eine schlechte Ergebnisqualität nicht auszuschließen.

In der sogenannten DIN EN ISO 9000 ff Normenreihe wurden Definitionen für Begriffe und Anforderungen des Qualitätsmanagements in mehreren Überarbeitungen branchenunabhängig formuliert und international anerkannt [11; 79]. Die Normenreihe DIN EN ISO 9000 wurde 1987 eingeführt. Qualität wird laut DIN EN ISO 9000:2005 definiert als „Grad, in dem ein Satz inhärenter Merkmale Anforderungen erfüllt“ [79].

Konkret übertragen auf die Strukturen im Krankenhaus würden nach dieser Definition die „inhärenten Merkmale“ z.B. in den einzelnen Komponenten eines Behandlungsprozesses wie Indikation, Durchführung und erzieltm Ergebnis einer Therapie bestehen. Die Verwendung der Bezeichnung umfasst den Ausprägungsgrad der angestrebten Qualitätskriterien, den man messen kann. Das bedeutet, dass der Begriff „Qualität“ per se entsprechend dieser Definition und seines lateinischen Ursprungs (qualitas = Beschaffenheit oder Eigenschaft) wertneutral ist und die Bereiche hoch- bis minderwertiger Qualität umschließt.

Eine etwas weniger abstrakte und technokratische Definition als die sehr gängige Formulierung nach der Norm DIN EN ISO 9000 ist die des Institute of Medicine<sup>1</sup> in den USA, die im internationalen Raum zitiert wird. **Nach dieser ist Qualität „der Grad, in dem medizinische Leistungen die Wahrscheinlichkeit für jeden einzelnen oder für ganze Populationen erhöhen, angestrebte Gesundheitsziele zu erreichen, die mit dem aktuellen medizinischen Wissensstand übereinstimmen“ [100].**

Die Kriterien, die den Begriff „Qualität“ füllen, können in Abhängigkeit des Blickwinkels der Betrachtung sehr heterogen sein. So stehen sich die unterschiedlichen Positionen der Gesundheitspolitik, der Gesundheitsökonomie, der Fachkreise und sonstiger Partikularinteressen gegenüber [27].

### Qualitätsmanagement

**Qualitätsmanagement** wird - entsprechend der Formulierung nach der Norm DIN EN ISO 9000:2005 „als aufeinander abgestimmte Tätigkeiten zum Leiten und Lenken einer Organisation definiert, die darauf abzielen, die Qualität der produzierten Produkte oder der angebotenen Dienstleistung zu verbessern [11; 79]. Konkret handelt es sich dabei um alle Maßnahmen, durch die die Qualitätspolitik, Ziele und Verantwortlichkeiten in einem Betrieb (z.B. Klinik oder Arztpraxis) festgelegt sowie diese durch Mittel wie Qualitätsplanung, Qualitätslenkung, Qualitätssicherung und Qualitätsverbesserung verwirklicht werden (siehe Abbildung 2) [11].

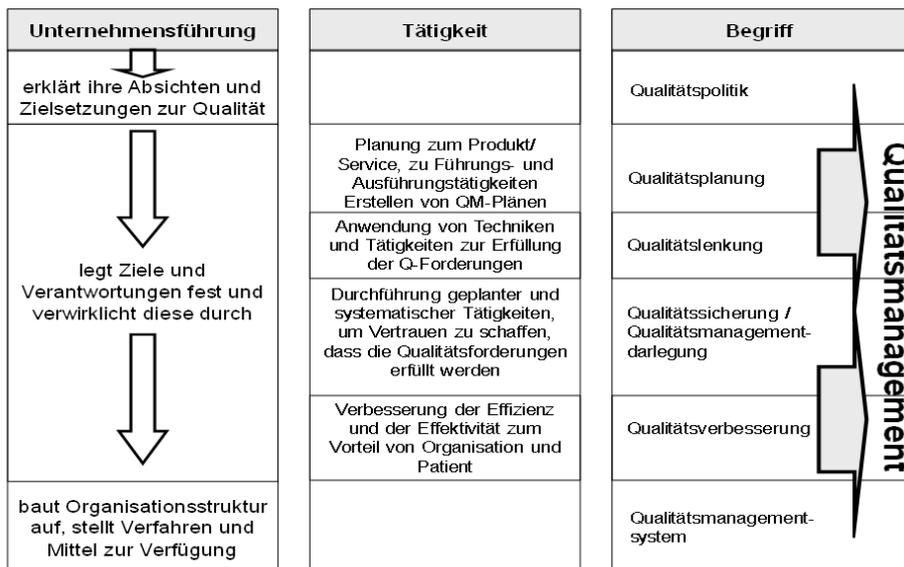
---

<sup>1</sup> Das Institute of Medicine ist eine unabhängige, gemeinnützige Organisation in den USA, die Entscheidungsträger der Gesellschaft bei Fragen zur Gesundheitsversorgung berät [98].

Es handelt sich um einen mehrdimensionalen Ansatz, der den Fokus auf ein strategisches Vorgehen im Sinne eines übergreifenden Managementkonzepts mit Lösungs- und Zielorientierung legt [59].

Qualitätsmanagement (QM) gilt demnach als Oberbegriff für alle qualitätsbezogenen Aktivitäten. Dieser Begriffsinhalt wurde früher ausschließlich mit „Qualitätssicherung“ benannt. Nachfolgend wurde dieser als Unterbezeichnung in das QM integriert [80].

**Abbildung 2 Zusammenhang der Begriffe des Qualitätsmanagements [11]**



Im QM sollte das kontinuierliche Bestreben zur Verbesserung der Qualität des eigenen Handelns ein fester Bestandteil der täglichen Praxis sein, für den alle Mitarbeiter selbst Verantwortung übernehmen. In diesem Sinn ist QM auch eine partizipative Führungsstrategie, die eine Stärkung von Verantwortung, Einfluss und Entscheidungskompetenz aller Beteiligten anstrebt [75]. QM ist die Aufgabe der obersten Führungsebene eines Krankenhauses oder einer Abteilung. Dabei sollte das QM des Krankenhauses von der Führungsebene ausgehend alle Mitarbeiter einbeziehen und überzeugen [187].

Im stationären Sektor übernimmt häufig die Krankenhausleitung die Steuerung der Prozesse einschließlich der Festlegung der jeweiligen Prioritäten. Die Krankenhausleitung ist folglich verantwortlich für die operative Umsetzung des QM. Im Mittelpunkt der Prozessoptimierung steht die Patientenorientierung. Krankenhausträger, Krankenhausleitung und alle übrigen Mitarbeiter sind aufgefordert, sich im Bewusstsein ihrer Verantwortung für die Qualität ihrer Leistungen, für die Beachtung der Qualitätssicherung und für die Realisierung der Regelungen zum Qualitätsmanagement einzusetzen [71]. Besondere Bedeutung hat in diesem Zusammenhang die berufsgruppen-,

hierarchie- und fachübergreifende Zusammenarbeit sowie die stetige interne, systematische Bewertung des erreichten Standes der Qualitätssicherungsanstrengungen [71].

Die grundsätzlichen Anforderungen an ein QM lassen sich in folgenden Aspekten zusammenfassen:

- Patientenorientierung
- Verantwortung und Führung
- Wirtschaftlichkeit
- Prozessorientierung
- Mitarbeiterorientierung und -beteiligung
- Zielorientierung und Flexibilität
- Fehlervermeidung und Umgang mit Fehlern
- kontinuierlicher Verbesserungsprozess [71].

Als wesentliches Element des Qualitätsmanagements gilt die Festlegung von Qualitätszielen, um die Qualität messbar und analysierbar machen zu können. Diese vordefinierten Qualitätsziele sollen regelmäßig hinsichtlich der Zielerreichung bzw. notwendiger Verbesserungsmaßnahmen analysiert werden. Ein bewährtes Konzept stellt der sogenannte PDCA-Zyklus nach Deming dar [11; 181], der mit P (Plan) die Planung, mit D (Do) die Ausführung, mit C (Check) die Überprüfung und Evaluation sowie mit A (Act) die erneute Umsetzung der optimierten Prozesse umfasst [45]. Der PDCA-Zyklus lässt sich auf alle Bereiche des QM, der Qualitätssicherung sowie auf alle Instrumente wie Leitlinien, Behandlungspfade, Qualitätsindikatoren, Patientensicherheitsprojekte etc. anwenden (siehe Abbildung 3).

### Qualitätssicherung

Nach DIN EN ISO 9000:2005 versteht man unter **Qualitätssicherung** (QS) den Teil des QM, der auf das Erzeugen von Vertrauen ausgerichtet ist, dass Qualitätsanforderungen erfüllt werden („Qualitätssicherung = Qualitäts**z**sicherung“). Hierunter sind Aktivitäten zu verstehen, die bei Kunden und Partnern im Gesundheitswesen Vertrauen schaffen, dass eine Organisation alle festgelegten, üblicherweise vorausgesetzten und verpflichtenden Erfordernisse und Erwartungen erfüllt. In diesem Sinne kann Qualitätssicherung z.B. in Form von Qualitätsdarlegungen bzw. Qualitätsnachweisen erfolgen (z.B. Qualitätsberichte, Zertifizierungen, Qualitätspreise) [11; 79].

Dieser Definition steht die nach wie vor übliche Verwendung des Begriffs „Qualitätssicherung“ im Sozialgesetzbuch bzw. im Fachjargon des Gesundheitswesens gegenüber [11]. Unter medizinischen QS-Maßnahmen im weiteren Sinne versteht man häufig alle Maßnahmen, durch welche die Qualität der gesundheitlichen Versorgung positiv beeinflusst werden kann [27].

In den internationalen Publikationen zu QS- oder QM-Verfahren ist diese Unterscheidung nicht in der geschilderten Art wiederzufinden. Die im Rahmen von QS-Maßnahmen beschriebenen Inhalte entsprechen häufig der deutschen Definition (siehe oben) eines QM und umgekehrt.

Für eine systematische Literaturanalyse kommt erschwerend hinzu, dass die Begriffe „*QS- oder QM-Programm*“ häufig gleichgesetzt werden mit dem sogenannten „umfassenden Qualitätsmanagement“ (UQM) oder „Total Quality Management“ (TQM), welches in allen Bereichen des Krankenhauses oder eines Klinikverbunds eingesetzt wird. Die Etikette QM/QS-Programm oder TQM für Krankenhäuser haben wenig Aussage bezüglich der Inhalte, die sich in der Auswahl der Bereiche für die Qualitätsbetrachtungen, in der Art und dem Ausmaß der Umsetzung sehr zwischen den Institutionen unterscheiden können [136].

Der Begriff der externen QS wird vielfach beschränkt auf die Bedeutung als Verfahren zur systematischen Methode der Messung und Erhebung von Ergebnissen von qualitätssichernden oder -fördernden Maßnahmen. In § 137 des Fünften Buch des Sozialgesetzbuches (SGB V) wird unter QS ebenfalls hauptsächlich die Qualitätsdatenerhebung und das Feedback an die Kliniken durch die „vergleichende externe QS“ subsumiert. Da die Begriffe im Bereich der *Programme* zur QS und zum QM häufig noch weniger abgrenzbar sind, fließend ineinander übergehen und sehr uneinheitlich verwendet werden, werden diese in der vorliegenden Arbeit gemeinsam betrachtet.

Innerhalb der QS werden Verfahren der *internen* und *externen* QS unterschieden. Unter externer Qualitätssicherung wird häufig die Messung und Sicherstellung der Erfüllung von Standards verstanden, die Qualität charakterisieren und die Transparenz für die Entscheidungsfindung des Patienten erzeugen soll [149]. Dies sind Qualitätsbeurteilungen anhand externer Maßstäbe durch externe Institutionen, die selbst nicht an der Versorgung beteiligt sind wie z.B. die gesetzliche externe QS (siehe Kapitel 1.2 und Kapitel 3.1.1.1). Die interne QS hingegen stellt ein Verfahren der institutionsinternen Qualitätsbeurteilung dar, welches von den an der Versorgung beteiligten Personen selbst und anhand eigener Daten durchgeführt wird (z.B. Peer-Review) [75].

Ein QS-Programm benötigt die Integration in ein QM, um positive Effekte für die Gesundheitsversorgung ausüben zu können [69]. Der Gemeinsame Bundesausschuss (G-BA) definiert Qualitätssicherungsmaßnahmen ebenfalls als integralen Bestandteil des Qualitätsmanagements. Damit sollen die externen Qualitätssicherungsmaßnahmen Erkenntnisse und Unterstützung für ein systematisches einrichtungsinternes Qualitätsmanagement liefern [71]. In § 135a des SGB V ist die Verpflichtung der Leistungserbringer zur Sicherung und Weiterentwicklung der Qualität der von ihnen erbrachten Leistungen dargelegt [29]. Die Leistungen müssen dem jeweiligen Stand der wissenschaftlichen Erkenntnis entsprechen und in der fachlich gebotenen Qualität erbracht werden. Die

zur Finanzierung durch die gesetzlichen Krankenversicherungen zugelassenen Krankenhäuser (§108 SGB V) sind nach Maßgabe des § 137 SGB V verpflichtet, sich an einrichtungsübergreifenden Maßnahmen der Qualitätssicherung zu beteiligen und einrichtungsintern ein Qualitätsmanagement einzuführen und weiterzuentwickeln. Das explizit genannte Ziel dieser einrichtungsübergreifenden Maßnahmen der Qualitätssicherung ist insbesondere, die Ergebnisqualität in der Patientenversorgung zu verbessern [29].

Orientierend am Nutzen für den Patienten verfolgen Maßnahmen zur Qualitätssicherung und zur Weiterentwicklung der Qualität von Krankenhausleistungen mehrere Ziele. Die wichtigsten Zielsetzungen bestehen darin, durch Erkenntnisse über Qualitätsdefizite Leistungsbereiche systematisch zu identifizieren, für die Qualitätsverbesserungen erforderlich sind, und Unterstützung zur systematischen, berufsgruppenübergreifenden, einrichtungsinternen Qualitätssicherung (internes Qualitätsmanagement) zu geben [73]. Darüber hinaus wird intendiert, Vergleichbarkeit von Behandlungsergebnissen (insbesondere durch Qualitätsindikatoren – siehe unten) herzustellen und durch diese Erkenntnisse die Qualität von Krankenhausleistungen in der Indikationsstellung, in den strukturellen Voraussetzungen zur Erbringung der Leistungen und im Erzielen guter Ergebnisse zu sichern [73].

#### Instrumente für Qualitätsmanagement/Qualitätssicherung

Als **Qualitätsindikatoren** bezeichnet man Messgrößen, anhand derer man quantitativ den Ausprägungsgrad der Qualität messen kann [5; 74; 146; 198]. Qualitätsindikatoren ermöglichen die Unterscheidung zwischen guter und schlechter Qualität von Strukturen, Prozessen und/oder Ergebnissen der Versorgung. Sie sind Hilfsgrößen, die die Qualität einer Einheit durch Zahlen bzw. Zahlenverhältnisse indirekt abbilden [198]. Häufig ist der Indikatorwert ein Prozentwert. Der Zähler ist die Anzahl der Patienten, für die das Qualitätsziel je nach Zielsetzung des Qualitätsindikators erreicht bzw. nicht erreicht wurde. Der Nenner ist die Gesamtheit aller Patienten, die sich einer entsprechenden Behandlung unterzogen haben [9]. Für die meisten Qualitätsindikatoren wurden Referenzbereiche definiert, die eine qualitativ gute Versorgung abbilden sollen [5; 74; 146; 198].

**Leitlinien** gehören zu den wichtigsten Instrumenten des Qualitätsmanagements. Die Integration von Leitlinien in QM-Programme ist eine der effektivsten Leitlinien-Implementierungsmaßnahmen [11; 62].

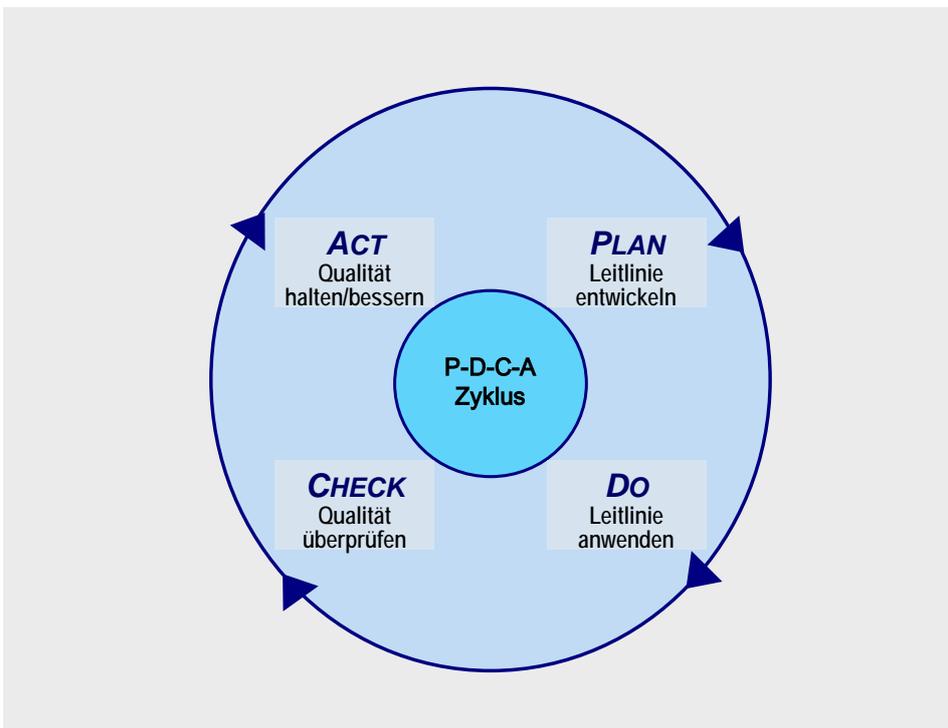
Leitlinien sind definiert als systematisch entwickelte Entscheidungshilfen für Leistungserbringer und Patienten über die angemessene Vorgehensweise bei speziellen Gesundheitsproblemen. Sie sind Orientierungshilfen im Sinne von „Handlungs- und Entscheidungskorridoren“, von denen in begründeten Fällen abgewichen werden kann oder sogar muss [25; 62].

Sie sollten folgende Kriterien erfüllen:

- Bearbeitungsprozess durch eine multidisziplinäre Expertenkommission, die die wichtigsten Adressaten repräsentiert
- Berücksichtigung von Patientenvertretern und Patientenpräferenzen
- klarer und transparenter Prozess zur Minimierung von Verzerrungen und Interessenkonflikten
- transparente Verknüpfung und Begründung der Evidenzbewertung und Stärke des Empfehlungsgrads
- Revision der Leitlinie bei relevanter neuer Evidenz [12; 101].

Unabhängig von der Art der QM/QS-Verfahren sind Leitlinien als integrale Bestandteile dieses QM/QS-Systems in einer Gesundheitseinrichtung hilfreich. Die systematische Vorgehensweise bei der Entwicklung, Disseminierung, Implementierung und Evaluation von Leitlinien entspricht den bereits dargelegten Elementen qualitativ hochwertiger Arbeit in Form von Planen, Ausführen, Prüfen und Handeln (Plan-Do-Check-Act-Zyklus (siehe Abbildung 3)) [11; 62].

**Abbildung 3 Integration von Leitlinien in Qualitätsmanagement-Modelle (Beispiel: Deming-PDCA-Zyklus) [11]**



Zusätzlich zu Leitlinien existieren weitere Instrumente für QS- oder QM-Programme wie z.B. klinische Behandlungspfade, Projekte zur Patientensicherheit und Zertifizierungen. Auf die Evaluation dieser qualitätssichernden Instrumente wird im Kapitel 3.3 näher eingegangen. Nachfolgend werden die Verfahren lediglich kurz mit ihren Definitionen erklärt.

**Klinische Behandlungspfade** sind strukturierte Instrumente, die den Weg des Patienten mit einer bestimmten Erkrankung von der Aufnahme bis zur Entlassung vorgeben. Sie stellen eine bestimmte Abfolge und Terminierung von festgelegten diagnostischen und therapeutischen Versorgungsschritten mit festen Verantwortlichkeiten aller involvierten Fachdisziplinen wie Ärzte, Pflege, Physio-, Ergotherapeuten und Sozialarbeiter dar [32; 133].

**Patientensicherheit** wird als „Abwesenheit von unerwünschten Ereignissen“ definiert [2; 99]. Patientensicherheit zu gewährleisten, setzt die Etablierung von QM-Systemen und Prozessen voraus, die die Wahrscheinlichkeit von Fehlern minimieren und die Wahrscheinlichkeit, diese abfangen zu können, vergrößern sollen [99].

Es werden im Bereich der Patientensicherheit mehrere Begrifflichkeiten unterschieden [2; 11; 40; 197]:

- **Unerwünschtes Ereignis** (*engl.: adverse event*): Ein schädliches Vorkommnis, das eher auf der Behandlung als auf der Erkrankung beruht. Es kann vermeidbar oder unvermeidbar sein.
- **Kritisches Ereignis** (*engl.: critical incident*): Ein Ereignis, das zu einem unerwünschten Ereignis führen könnte oder dessen Wahrscheinlichkeit deutlich erhöht.
- **Fehler** (*engl.: error*): Eine Handlung oder ein Unterlassen bei dem eine Abweichung vom Plan, ein falscher Plan oder kein Plan vorliegt. Ob daraus ein Schaden entsteht, ist für die Definition des Fehlers irrelevant.

Unter **Zertifizierung** versteht man ein Verfahren, in dem ein (unparteiischer) Dritter schriftlich bestätigt, dass ein Produkt, Prozess oder eine Dienstleistung mit festgelegten Anforderungen konform ist. Die Begriffe Zertifizierung und Akkreditierung werden international uneinheitlich verwendet [11]. Akkreditierung wird im internationalen Sprachgebrauch häufig als Synonym für Zertifizierung gebraucht [170]. Zur Zertifizierung eines QM gehören z.B. Systeme wie die European Foundation for Quality Management (EFQM) und die Kooperation für Transparenz und Qualität im Krankenhaus (KTQ) [187], auf die im Kapitel 3.3.4 noch näher eingegangen werden wird.

## 1.2 Entwicklung von Qualitätsmanagement und Qualitätssicherung in der stationären Versorgung

Die Entwicklung des QM wurde durch mehrere Aspekte in den letzten 20 Jahren gefördert:

1. wachsendes medizinisches Wissen in vielen Bereichen der klinischen Versorgung, für welches eine systematische Datenaufbereitung mit Integration ins QM hilfreich ist
2. zunehmend komplexer werdende medizinische Prozesse

3. knapper werdende finanzielle Mittel und zunehmender ökonomischer Druck auf die Gesundheitsdienstleister, welcher durch Nachweis von Qualitätsdaten abgemildert werden kann
4. sich verstärkender Wettbewerb im Gesundheitsmarkt
5. die sich verändernde Haltung der Patienten gegenüber dem Gesundheitssystem durch vermehrten Informationsbedarf bei medizinischen Entscheidungen, vermehrte Dienstleistungserwartungen und vermehrter Nachfrage nach Transparenz durch Qualitätsdaten für den Vergleich von Einrichtungen
6. kommerzielle und lobbyistische Ziele [132; 149; 175].

Externe Qualitätssicherungsmaßnahmen in der stationären Versorgung begannen - auf Initiative einiger Landesärztekammern - bereits viele Jahre vor der Gesundheitsreform, die 1989 durch das „Gesetz zur Strukturreform im Gesundheitswesen“ eingeleitet wurde [27; 195]. In der Gesundheitsreform 1989 wurde im SGB V die Beteiligung der Krankenhäuser sowie die Vorsorge- oder Rehabilitationseinrichtungen an Maßnahmen zur Qualitätssicherung festgeschrieben (§137 SGB V) [36; 75; 195]. Bevor die gesetzliche externe QS in den 1990er-Jahren verbindlich vom Gesetzgeber beschlossen wurde [31; 195], haben in den 1970er-Jahren bereits Ärzte in Bayern die Peri- und Neonatalerhebung als freiwilliges Instrument kontinuierlicher Qualitätsverbesserung initiiert, die sich im Verlauf auf die anderen Bundesländer ausdehnte [59; 195]. Die Peri- und Neonatalerhebung entstand aus einer Initiative von Frauenärzten im Bemühen, Ergebnisse über die Umsetzung wissenschaftlicher Erkenntnisse zu objektivieren und vergleichbar zu machen, und stellte damit erste Schritte der Versorgungsforschung dar [158].

**In einer gemeinsamen Bestandsaufnahme der Bundesärztekammer (BÄK) und der Kassenärztlichen Bundesvereinigung (KBV) über die Aktivitäten der Spitzenorganisationen der ärztlichen Selbstverwaltung wurde ein Überblick ausgewählter QS-Aktivitäten bis 1995 zusammengestellt, welcher in Tabelle 2 dargelegt ist [27].**

**Tabelle 2 Historischer Überblick ausgewählter Qualitätssicherungsaktivitäten in und bei den Ärztekammern [27]**

1967	- Fachnormenausschuss der BÄK und des deutschen Normenausschusses
1970	- Arbeitskreis Richtlinien der BÄK zur Durchführung der statistischen Qualitätskontrolle und von Ringversuchen im Bereich der Heilkunde
1975	- Mit Münchner Perinatalstudie Beginn überregionaler QS-Programme in Kooperation mit Ärztekammern - Einrichtung von Schlichtungsstellen bei Landesärztekammern
seit 1978	- Ausschüsse für Qualitätssicherung bei Landesärztekammern - QS-Projektgeschäftsstellen bei Landesärztekammern

1979	- Beschluss des 82. Deutschen Ärztetages zur Einrichtung eines Ausschusses der Bundesärztekammer „Rationalisierung und Qualitätssicherung der ärztlichen Berufsausübung“
1980	- Leitlinien des 83. Deutschen Ärztetages zur Qualitätssicherung - Einrichtung von Arbeitskreisen des Ausschusses Qualitätssicherung der BÄK
1983	- Arbeitsgemeinschaft „Münchner Perinatalgespräch“ - Beschluss des 86. Deutschen Ärztetags zur Förderung der Qualitätssicherung - Ständige Konferenz Qualitätssicherung der Landesärztekammern
1984	- Positionspapier der Bundesärztekammer zur Qualitätssicherung
1986	- Qualitätssicherung in „Blauen Papier“ des 89. Deutschen Ärztetages - Kooperationsvereinbarung zwischen BÄK und Deutsche Krankenhausgesellschaft zur Qualitätssicherung - mehrseitige Verträge mit Ärztekammern zur Qualitätssicherung beginnend mit Baden-Württemberg und Nordrhein
1988	- Verankerung der Qualitätssicherung als ausdrückliche Verpflichtung in der ärztlichen Berufsordnung - Verankerung der Fortbildungspflicht in der Berufsordnung
1989	- Sachstandsbericht der Bundesärztekammer zur Qualitätssicherung
1990	- Gründung des gemeinsamen Ausschusses QS der BÄK und der KBV
1991	- Gründung der Bundesarbeitsgemeinschaft Qualitätssicherung Herzchirurgie - Gründung der Arbeitsgemeinschaft Chirurgie bei der Bundesärztekammer
1992	- Sicherung der Qualität ärztlicher Berufsausübung als Ziel der Weiterbildung in der (Muster-) Weiterbildungsordnung - Beschluss des Deutschen Ärztetages zur Einrichtung eines Kuratoriums für Qualitätssicherung - Gründung der Arbeitsgemeinschaft „Zentraler Erfahrungsaustausch der ärztlichen Stellen nach § 16 RöV“ bei BÄK und KBV
1993	- Gründung der „Arbeitsgemeinschaft zur Förderung der Qualitätssicherung in der Medizin“ (BÄK, KBV, Deutsche Krankenhausgesellschaft, Gesetzliche Krankenversicherung)
1994	- Acht Thesen des 97. Deutschen Ärztetages zur Qualitätssicherung
1995	- Gründung der Ärztlichen Zentralstelle Qualitätssicherung – jetzt Ärztliches Zentrum für Qualität in der Medizin (ÄZQ)
1996	- Publikation des Curriculums Qualitätssicherung - Ärztliches Qualitätsmanagement durch BÄK und KBV

Nachdem 1996 erstmals das Konzept und die Inhalte einer strukturierten Zusatzweiterbildung für das Curriculum „Ärztliches Qualitätsmanagement“ von BÄK, KBV und der Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) zustimmend zur Kenntnis genommen wurden, starteten im Oktober 1996 die Bundesländer Baden-Württemberg und Niedersachsen zeitnah mit der Umsetzung eines dreistufigen Fortbildungskonzeptes, welches Ärzte im Krankenhaus für die neue Rolle als Qualitätsmanager qualifizieren sollte [26; 166]. Dieses Curriculum wird bis heute bei den Landesärztekammern fortgeführt und weiterentwickelt.

Im Jahr 2000 wurden die bestehenden Verpflichtungen zu Maßnahmen der Qualitätssicherung in den Krankenhäusern durch das Gesundheitsreformgesetz der Gesetzlichen Krankenversicherung (GKV) um weitere qualitätssichernde, bzw. -fördernde Regelungen erweitert [36; 75; 195]. So wurde festgelegt, dass sich alle an der Patientenversorgung beteiligten Gesundheitseinrichtungen zur Teilnahme an Maßnahmen der *einrichtungsübergreifenden* externen QS verpflichten [139; 195].

Die externe QS in Krankenhäusern war bis 2004 in der Hand des Bundeskuratoriums bestehend aus GKV-Spitzenverbänden und der Deutschen Krankenhausgesellschaft unter Beteiligung des Verbands der privaten Krankenversicherungen, der Bundesärztekammer sowie der Berufsorganisationen der Krankenpflegeberufe [21]. Nachfolgend übernahm diese der Gemeinsame Bundesausschuss (G-BA) [139; 195], der hierzu eine „Richtlinie über Maßnahmen der Qualitätssicherung in Krankenhäusern“ beschlossen hat [18; 73]. Seit 2005 sind die Krankenhäuser darüber hinaus gesetzlich verpflichtet, strukturierte Qualitätsberichte zu veröffentlichen [72]. Für Krankenhäuser eröffnen sie die Möglichkeit, Leistungen und Qualität darzustellen. Qualitätsberichte sollen der Information und Transparenz auf dem Gebiet der Krankenhausbehandlung dienen und stellen damit zugleich für Krankenhäuser ein mögliches Wettbewerbsinstrument dar [72].

Um für Qualitätssicherungsverfahren in der ambulanten und der stationären Versorgung einen einheitlichen Rahmen vorzugeben, etablierte der G-BA im Jahr 2010 die Richtlinie zur **einrichtungs- und sektorenübergreifenden Qualitätssicherung**. Dadurch sollten longitudinale Betrachtungen im Follow-up-Verfahren ermöglicht werden, bei denen die Ergebnisqualität einer in einem Sektor erbrachten Leistung in einem anderen Sektor gemessen werden kann [4; 195]. Die Umsetzung gestaltete sich allerdings schwieriger als zuvor angenommen wurde. Es zeigte sich, dass die Abbildung von Behandlungsverläufen über Zeiträume, Einrichtungen und Sektoren hinweg auf datentechnische und rechtliche Probleme sowie auf Schwierigkeiten bei der Zuordnung von Verantwortlichkeiten gestoßen ist [195].

### 1.3 Problemstellung

QM/QS ist ein kontroverses Thema. Befürworter nennen die verstärkte Transparenz, die informierte Entscheidungsfindung des Patienten und die Kosteneinsparungen als Vorteile der Verfahren für die Patientenversorgung und Patientensicherheit [149]. Auf der anderen Seite werden Gegenargumente genannt wie die Entstehung eines unvorteilhaften Wettbewerbs zwischen den Kliniken, eine für die Versorgung nachteilige Selektion eines bestimmten Patientenkontingents, eine eingeschränkte ärztliche Behandlungsfreiheit und die Hemmung von Innovationen [149].

In den letzten Jahren ist eine Vielzahl an verschiedenen Verfahren und Maßnahmen für qualitätssichernde oder verbessernde Strategien in Deutschland entwickelt worden. Auch im internationalen Bereich ist die gleiche Entwicklung erkennbar [84]. Verschiedene Akteure des Gesundheitswesens aus der klinischen Versorgung, der Gesundheitspolitik und der Wissenschaft vertraten die Auffassung, dass die Qualität weiterhin durch vielerlei Maßnahmen verbessert werden musste, starteten Programminitiativen mit unterschiedlichen Ausgangsperspektiven und führten Verfahren wie Zertifizierungen, TQM, externe QS mit Messung der Qualität, Risikomanagement, Patientensicherheit und Disease Management Programme (DMP) ein. Alle Initiativen setzen sich das gleiche Ziel, die Versorgung der Patienten zu verbessern [84].

Die Einführung von neuen Medikamenten folgt einem strengen Prozess mit Prüfung und Evaluation von Teilschritten [28; 63]. Auch im Bereich der QS/QM-Verfahren sind solche Ansätze vor und nach Einführung solcher Maßnahmen notwendig [84; 93; 154].

So finden sich Publikationen zur Evaluation von QS/QM-Maßnahmen mit Darstellung von positiven Effekten in Folge dieser qualitätssichernden Verfahren sowohl im ambulanten als auch im stationären Sektor.

In einer Auswertung wurde beispielsweise dargelegt, dass durch die qualitätssichernden Maßnahmen der gesetzlichen Rentenversicherung in kardiologischen Rehabilitationseinrichtungen hinsichtlich der Qualität und der Laufzeit der Entlassungsberichte im Verlauf von 2001 bis 2004 deutliche Verbesserungen der Ergebnisse der Prozessqualität erzielt werden konnten [13]. Im Rahmen einer weiteren wissenschaftlichen Evaluation der externen Qualitätssicherung der Deutschen Rentenversicherung (DRV) wurde 2008 mit dem Ziel der Weiterentwicklung des Visitations-Verfahrens das Projekt „Visit II“ durch den Arbeitsbereich Rehabilitationswissenschaften der Universität Würzburg in Abstimmung mit der DRV durchgeführt. Dieses wissenschaftliche Evaluationsprojekt bestand in einer schriftlichen Befragung und in Telefoninterviews mit den visitierten Kliniken mit der Bitte um Rückmeldungen der Visitoren zu dem Verfahren, die im Ganzen ein sehr positives Meinungsbild gegenüber den Visitationen widerspiegeln [128]. Die Ergebnisse der Auswertung der Studie wurden für die

Weiterentwicklung des Visitationsverfahrens berücksichtigt und eine Projektgruppe zur Erarbeitung von einheitlichen Vorgaben für ein rentenversicherungsweit einzusetzendes Visitationsverfahren wurde daraufhin von der DRV beauftragt.

Eine 2009 anhand von Veränderungen in Qualitätskennzahlen einzelner Krankenhäuser durchgeführte Wirksamkeitsanalyse der gesetzlichen externen QS (siehe Kapitel 3.1.1.1) ergab ebenfalls insgesamt einen Hinweis für eine Effektivität dieser Maßnahmen [115].

Auch im ambulanten Bereich findet sich eine Reihe von Publikationen zur Wirksamkeit von QM-Programmen. So finden sich in Publikationen aus Deutschland zum DMP Diabetes mellitus Typ 2 oder aus dem internationalen Raum zu Disease Management Programmen zu chronisch obstruktiven Lungenerkrankungen deutliche Hinweise für eine Wirksamkeit hinsichtlich der Reduktion von Mortalität und weiteren klinischen Endpunkten sowie Endpunkten der Prozessqualität [1; 113; 124].

Darüber hinaus wurden international in einem europaweiten und gesundheitssystemübergreifenden Forschungsprojekt zu qualitätsverbessernden Strategien wie TQM, Zertifizierung, Patientensicherheitsprojekten, Leitlinien und Qualitätsindikator-Messungen in acht partizipierenden EU-Ländern positive Zusammenhänge zu verbesserten Prozessindikatoren berichtet [179].

Die dargestellte Effektivität der QS/QM-Programme wurde jedoch vielfach kritisiert, da durch das zum Teil limitierte methodische Vorgehen der Analysen und die zugrunde liegende Datenvalidität ein Wirksamkeitsbeleg für eine verbesserte Patientenversorgung zunehmend hinterfragt wurden [4; 77; 118; 121; 130; 139; 157].

Ein weiterer großer Kritikpunkt im Rahmen der QM/QS-Verfahren in Krankenhäusern stellt die auf den stationären Aufenthalt begrenzte Messung der Qualität dar, sodass ein ausreichend langes Follow-up für eine adäquate Beurteilung der Qualität fehlen würde [4]. Die sektorenübergreifende Qualitätssicherung wurde als mögliche Lösung gesehen. Im Rahmen der vielen konzeptionellen Herausforderungen der sektorenübergreifende Qualitätssicherung wurde zunehmend auch über das Verhältnis von Aufwand zur Wirksamkeit diskutiert [4].

Durch die gesetzliche Verpflichtung, ein internes Qualitätsmanagement in den Krankenhäusern einzuführen, sind die Krankenhäuser gezwungen, das meist geeignete unter den Qualitätsmanagementmodellen für sich auszuwählen. Da vom Gesetzgeber nur wenig konkrete Vorgaben existieren, stellt dies für die meisten Krankenhäuser vor dem Hintergrund der unsicheren Datenlage zur Wirksamkeit eine Herausforderung dar [130]. Die Auswahl bleibt dem Krankenhaus überlassen. Der Gesetzgeber und die Länder müssen lediglich die Rahmenbedingungen schaffen [130].

Donabedian unterschied mehrere Ebenen der Einflussmöglichkeiten auf die Effektivität von QS-Maßnahmen. Er unterteilte zum einen in Rahmenbedingungen der Organisationskultur mit Art der Führung und Trägerschaft und zum anderen in operationale Faktoren, die die Effektivität der Maßnahmen fördern. Er sah Effektivität nur unter der Voraussetzung als möglich an, dass die vorhandene Versorgung hinsichtlich ihrer Qualität überhaupt verbesserungswürdig, die Verbesserung realisierbar und kongruent mit den eigenen beruflichen Idealvorstellungen sei [55].

Es wurde bereits im Jahr 1996 in einer Gemeinsamen Bestandsaufnahme der Bundesärztekammer und der Kassenärztlichen Bundesvereinigung zur Qualitätssicherung und den Grundlagen einer bedarfsgerechten Gesundheitsversorgung im Konsens mit der Ärzteschaft gefordert, dass die Angemessenheit von QS-Programmen konsequent immer wieder auf ihre Effektivität und Effizienz hin überprüft, evaluiert und gegebenenfalls modifiziert werden muss [27].

Auch im internationalen Raum wurde zeitgleich diese Forderung laut, da auch hier viele Qualitätsinitiativen geschaffen wurden, jedoch keine hinreichenden Beweise bezüglich der Wirksamkeit und nur sehr wenige kontrollierte Studien vorlagen [55].

Heute fast 20 Jahre später werden vielfach die gleichen Forderungen gestellt, die Wirksamkeit von QM/QS-Maßnahmen in den bereits etablierten Programmen zu evaluieren und die Evidenzbasis von QM/QS damit zu stärken. Zahlreiche Aspekte der Wirksamkeit und des konkreten Nutzens der vorgesehenen Maßnahmen sind immer noch weitgehend ungeklärt und es ist fraglich, ob Qualitätsprobleme mit diesen Verfahren tatsächlich adressiert und verringert werden können [4].

Auch der Gemeinsame Bundesausschuss legte in seiner Vereinbarung über die grundsätzlichen Anforderungen an ein einrichtungsinternes Qualitätsmanagement für Krankenhäuser eine Evaluation des QM fest. So heißt es § 4, Absatz 2 und 3: *„Maßnahmen des Qualitätsmanagements müssen, wie alle anderen Krankenhausleistungen auch, der Forderung nach Effektivität und Wirtschaftlichkeit genügen. (...) Kosten und Nutzen des Qualitätsmanagements sollten im Einzelfall evaluiert, anhand der eigenen Ausgangsbedingungen ermittelt und bewertet werden. (...) Der Gemeinsame Bundesausschuss strebt an, die möglichen Auswirkungen dieser Vereinbarungen zu einem späteren Zeitpunkt zu bewerten.“* [71].

Auf Grund des enormen personellen und finanziellen Aufwands und der weiten Verbreitung der QM/QS-Programme stellt sich weiterhin die Frage nach den vielfach geforderten Nachweisen zur Effektivität und Effizienz dieser Verfahren [27; 55; 71; 84; 93; 154].

Für das externe QS-Programm der Gesetzlichen Krankenversicherung in der stationären medizinischen Rehabilitation wurden beispielsweise bereits mehr als 1,5 Millionen Euro aufgewendet [77].

Die geschätzten Bürokratie-Kosten für die sektorenübergreifende Qualitätssicherung im Bereich der Katarakt-OP beliefen sich auf bis zu 7,6 Millionen Euro [4].

In einem sehr umfangreichen Gutachten des Kompetenz-Centrum Qualitätssicherung/Qualitätsmanagement (KCQ) des GKV-Spitzenverbandes und der Gemeinschaft der Medizinischen Dienste der Krankenversicherung wurde unter anderem nach systematischen Untersuchungen zu Effekten der Einführung von internem Qualitätsmanagement in bundesdeutschen Kliniken gesucht. Entsprechende aussagekräftige Studien konnten von den Autoren nicht eruiert werden [121].

In einer Literaturanalyse von Glattacker et al. zeigte sich ebenfalls, dass Maßnahmen zur Qualitätsmessung und Qualitätsverbesserung in allen Bereichen des Gesundheitsversorgungssystems etabliert sind, bislang jedoch nur eine limitierte Evidenz bezüglich der Effektivität und Effizienz vorliegt. Als Ursache dafür sieht er zum einen die Komplexität der zu analysierenden Intervention, die mangelnde methodische Güte der Studien und einen „publication gap“. Dieser beinhaltet, dass die Ergebnisse vieler Qualitätssicherungsprojekte lediglich informell mitgeteilt, jedoch nicht systematisch publiziert wurden [77].

Um die limitierten Ressourcen im Gesundheitssystem effizient und zielgerichtet einsetzen zu können, nimmt die Bedeutung der Sicherung und Förderung von Qualität kontinuierlich zu. Die Frage nach der Effektivität und Effizienz der QM/QS ist umso mehr von Bedeutung, da in Zukunft möglicherweise die Verteilung der Ressourcen an Qualität der Leistungen gekoppelt sein könnte, wie es z.B. im internationalen Raum bereits im Rahmen von Pay-for-Performance<sup>2</sup> Ansätzen gezeigt wurde [130; 189]. Auch in Deutschland bestehen vielfältige Diskussionen zur Einführung von Zuschlägen bei erbrachter hochwertiger Qualität bzw. Abschlägen bei unterdurchschnittlicher Qualität. Qualität könnte in Zukunft ein Kriterium für die Krankenhausplanung sein, sodass die Verteilung der Gelder in nicht unerheblichem Maß von Qualitätsfragen bestimmt sein könnte.

Deutschland weist im internationalen Vergleich anhand von OECD<sup>3</sup>-Gesundheitsdaten von 12 Ländern aus dem Jahr 2010 die höchste Rate an Krankenhausbetten mit 5.7 pro 1000 Bürgern mit der nahezu längsten durchschnittlichen Verweildauer der Behandlung bei 7.6 Tagen in Akutkrankenhäusern auf [174]. Diese Zahlen verdeutlichen den hohen Stellenwert, den die Krankenhausbehandlung in Deutschland einnimmt, und verstärken die Relevanz des Themas der Effektivität von QM/QS in deutschen Krankenhäusern.

---

<sup>2</sup> Pay-for-Performance: leistungsorientierte Vergütung

<sup>3</sup> Organisation for Economic Co-operation and Development (OECD) ist eine internationale Organisation mit 34 Mitgliedstaaten für wirtschaftliche Zusammenarbeit und Entwicklung [135].

Die Untersuchung der Wirksamkeit von QM/QS-Maßnahmen ist eine komplexe Herausforderung und umfasst die Berücksichtigung von einer Reihe an Prozessschritten wie die Einführung, Implementierung, Verhaltensänderung und schließlich die Verbesserung der Versorgung [55].

#### 1.4 Fragestellung und Zielsetzung

In der heutigen Zeit wird für alle neu eingeführten medikamentösen Maßnahmen in der Behandlung von Patienten ein wissenschaftlich fundierter Wirksamkeitsnachweis geltend gemacht [28; 63]. Wie im Kapitel 1.3 dargestellt wurde, wird seit langem auch für die QS/QM-Verfahren diese Evaluation bezüglich eines Wirksamkeitsnachweises berechtigterweise postuliert, um Nutzen und Schaden für den Patienten, die Gesundheitsdienstleister und für das gesamte Gesundheitssystem ermessen zu können.

Obwohl Publikationen und Studien mit dargestellten positiven Effekte der einzelnen Verfahren und Programme zu QS/QM in der stationären Versorgung in Deutschland existieren, mehren sich Hinweise, dass bislang die Wirksamkeit im Sinne einer verbesserten und nachhaltigen Patientenversorgung trotz allem weder hinreichend wissenschaftlich untersucht noch durch hochwertige, aussagekräftige Evidenz belegt ist [4; 77; 118; 121; 130; 139; 157].

Es stellt sich also die Frage, ob all diese Bemühungen für QM/QS, die mit so viel Engagement und großem Aufwand vorangetrieben werden, tatsächlich für den Patienten nutzbringend sind und welche konkrete Evidenzgrundlage diese besitzen. Es gilt die Frage zu erörtern, ob und in wie weit die vielfach dargestellten positiven Effekte in den Publikationen auch einer detaillierten Analyse unter methodischen Gesichtspunkten standhalten und einen kausalen Zusammenhang zwischen den Verfahren und einer verbesserten Patientenversorgung nachweisen können. Des Weiteren wäre es hilfreich zu wissen, ob gegebenenfalls auch Studien mit Hinweisen auf negative Effekte der Verfahren existieren.

Darüber hinaus drängt sich die Frage auf, ob der bisherige Fokus, der auf die weitere wissenschaftliche Ausgestaltung oder Neukonzeption von QM/QS-Verfahren gesetzt wird, bei der aktuellen Evidenzlage weiterhin zielführend ist oder ob die vorhandene Studienlage vielmehr Hinweise auf eine andere einzuschlagende Richtung wie der vermehrten Evaluationsforschung und der verstärkten Förderung und Rückbesinnung auf das ärztliche Ethos gibt.

Zur Beantwortung dieser Fragestellungen erfolgte eine systematische Literaturrecherche zur Identifizierung von Publikationen, die die Evaluation von QS/QM-Programmen hinsichtlich ihrer Wirk-

samkeit im nationalen und internationalen Raum thematisieren. Diese Studien wurden einer kritischen Analyse bezüglich ihrer methodischen Güte der internen und externen Validität nach Kriterien der Evidenzbasierten Medizin<sup>4</sup> unterzogen, zusammengefasst und im Vergleich zu Erfahrungen aus dem Ausland ausgewertet und diskutiert.

Ziel dieser Arbeit war es, auf Grund des großen personellen und finanziellen Aufwands selbst kleinerer QM/QS-Programme, der zunehmenden Ressourcenknappheit und der gesundheitspolitischen Bedeutung dieser Thematik die existierenden QM/QS-Programme national und international kritisch hinsichtlich ihrer öffentlich dargelegten Wirksamkeit zu beleuchten. Mit der vorliegenden Arbeit sollte eine Bestandsaufnahme zum aktuellen wissenschaftlichen Stand der Wirksamkeit dieser Programme sowie daneben von einzelnen qualitätssichernden Instrumenten im stationären Sektor geschaffen werden, um gegebenenfalls als Hilfestellung für die Weiterentwicklung der Evaluationsforschung für Qualitätsinitiativen dienen zu können.

---

<sup>4</sup> Die Evidenzbasierte Medizin wurde 1996 von dem Internisten und Nephrologen David Sackett et al. definiert als die Integration der besten verfügbaren externen Evidenz (wissenschaftliche Belege) mit der persönlichen klinischen Erfahrung und den Patientenwünschen für Entscheidungen über die medizinische Versorgung einzelner Patienten [156].

## 2 Methodik

Zur Beantwortung der Forschungsfrage wurde eine systematische Übersichtsarbeit (Review) erstellt. Diese dient dem Ziel, die gesamte Evidenz zusammenzutragen, die zu den definierten Auswahlkriterien passt. Die systematisch angewandten Methoden sollen das Risiko eines sogenannten Bias reduzieren, um aus den Ergebnissen verlässliche Schlussfolgerungen ziehen und Entscheidungen treffen zu können (Antman 1992, Oxman 1993 zitiert nach [37]). Ein Bias (systematischer Fehler) ist die Tendenz der Studienergebnisse, systematisch von den „wahren“ Ergebnissen abzuweichen (Verzerrungsrisiko). Bias führt entweder zu einer Über- oder Unterschätzung der wahren Wirkung einer Maßnahme oder Exposition. Die Ursachen dafür liegen vor allem im Design und der Durchführung der Studie [114].

Zu den Hauptmerkmalen einer solchen Übersichtsarbeit gehören:

- eine systematische Literaturrecherche in den am besten geeigneten Literaturdatenbanken (siehe unten) [50]
- klar definierte Ziele mit a priori festgelegten Ein- und Ausschlusskriterien für die Auswahl der Studien
- eine reproduzierbare Methodik
- die Erhebung des Verzerrungsrisikos (Bias) der eingeschlossenen Einzelstudien zur Einschätzung der Validität der Ergebnisse
- eine systematische Darstellung der Charakteristik und Ergebnisse der eingeschlossenen Einzelstudien mit Zusammenfassung [37].

### Systematische Literaturrecherche

Die systematische Recherche ist der notwendige erste Schritt zur Identifikation der bestverfügbaren Evidenz [50].

Eine systematische Suche sollte folgende Schritte umfassen:

- Auswahl geeigneter Recherchequellen
- Festlegung des Suchvokabulars und Entwicklung der Strategie
- Durchführung der Suche in den ausgewählten Quellen
- Sichtung der Ergebnisse und Anpassung der Recherchestrategie
- erneute Durchführung der Suche
- Sichtung der Treffer auf Relevanz und
- Dokumentation der Recherche [50].

Die in der beschriebenen Weise durchgeführte systematische Literaturrecherche [37; 50] wurde in den Datenbanken Medline<sup>5</sup> via <http://www.pubmed.org/>, in der Cochrane Library via <http://www.thecochranelibrary.com> sowie zusätzlich über weitere Quellen der Handsuche [37] durchgeführt. Diese Quellen umfassten Referenzen aus den Literaturverzeichnissen relevanter Publikationen, „related citations“ der Medline-Treffer und Bücher. Um die sogenannte „Graue Literatur“ erfassen zu können, erfolgte darüber hinaus eine Suche auf ausgewählten Internetseiten von themenrelevanten, nationalen und internationalen Institutionen, die übergeordnet für QS/QM-Programme zuständig sind, und über die Suchmaschine Google Scholar. „Graue Literatur“ umfasst Berichte und Gutachten, die nicht in formal publizierten Quellen wie Zeitschriften und Büchern veröffentlicht werden [50].

Recherchestrategie und -vokabular richteten sich nach den Möglichkeiten der jeweiligen Recherchequelle und wurden entsprechend jeweils angepasst. Für Deutschland wurde für die Suche in Medline und Google Scholar keine zeitliche Limitation vorgenommen; Medline enthält Publikationen zurückreichend bis 1966. Für den internationalen Raum wurde der Suchzeitraum in der Datenbank Medline auf die Zeit von Dezember 2011 bis Dezember 2013 eingeschränkt. In der Cochrane Database of Systematic Reviews wurde die gemeinsame Recherche für Deutschland und für den internationalen Raum auf die letzten 10 Jahre begrenzt.

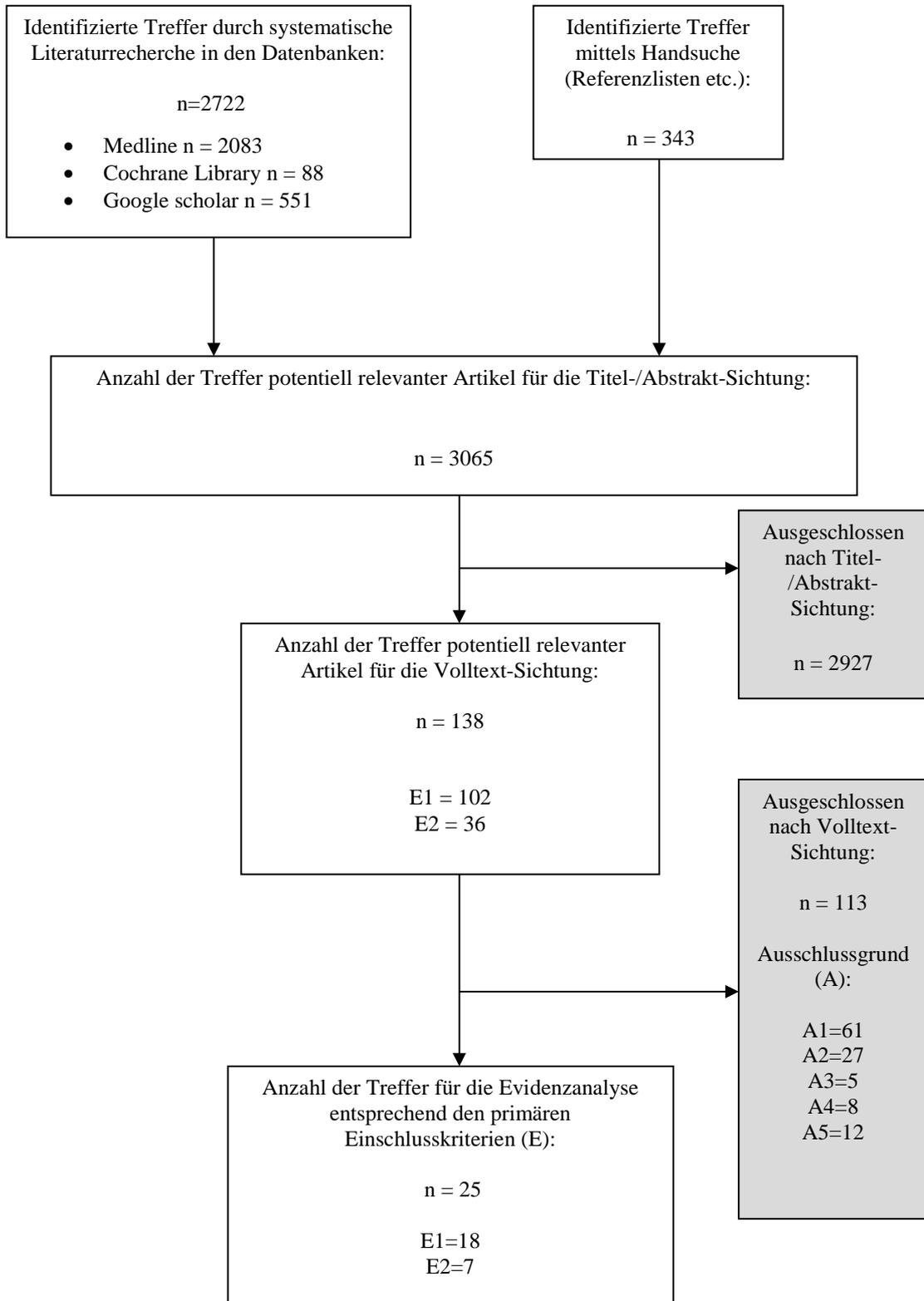
Die gesamte detaillierte Recherchestrategie mit den genutzten Datenbanken, Suchbegriffen, Suchzeiträumen und Treffern ist im Anhang 1 dargelegt.

Die Studien wurden in weiteren Schritten zunächst auf Titel-/Abstraktebene, nachfolgend auf Volltextebene gesichtet und nach im Kapitel 2.2 beschriebenen, a priori festgelegten Kriterien ausgewählt. Eine grafische Zusammenfassung der einzelnen Stufen der Literatursuche mit Anzahl der Treffer und Dokumentation der Ausschlussgründe findet sich im nachfolgenden Flussdiagramm entsprechend dem PRISMA –Statement [126].

---

<sup>5</sup> MEDLINE (Medical Literature Analysis and Retrieval System Online) ist eine bibliografische Datenbank der National Library of Medicine in den USA.

**Abbildung 4 Flussdiagramm zum Prozess der Studienauswahl<sup>6</sup>**  
 (nach PRISMA-Statement [126])



<sup>6</sup> Darstellung der Ein- und Ausschlusskriterien siehe Tabelle 3

## 2.1 Studienauswahl und -bewertung

### 2.1.1 Studienauswahl

Ziel der Recherche war das Identifizieren von Evaluationsstudien, die die Wirksamkeit von QS/QM-Programmen oder –verfahren auf die Qualität der Patientenversorgung im Krankenhaussektor untersuchten.

#### Studientyp

Da in Anbetracht der Thematik eher ein niedriger Evidenzgrad der Publikationen mit wenigen kontrollierten Studien zu erwarten war, wurden auch Studien von methodisch geringerer Studienqualität in die Analyse eingeschlossen, wenn sie bestimmte Voraussetzungen der Vergleichsgruppe erfüllten (siehe unten).

Es wurden als Primärstudien randomisiert-kontrollierte Studien (RCT), nicht-randomisierte kontrollierte Studien (Kohortenstudien, kontrollierte Vorher-Nachher-Studien mit Kontrollarm ohne Intervention, Querschnittsstudien), Kohortenstudien ohne Kontrollarm (Verlaufsbeobachtung zu Beginn und im Verlauf einer Intervention, unkontrollierte Vorher-Nachher-Studie und Zeitreihenuntersuchungen) oder als aggregierte Evidenz systematische Übersichtsarbeiten aus genannten Studientypen eingeschlossen. Im Bewusstsein, dass aus unkontrollierten Studien auf Grund des hohen Verzerrungsrisikos nur eine limitierte sachgerechte Aussage zur Beantwortung der Forschungsfrage abgeleitet werden kann, wurden diese ihrer jeweiligen Studienqualität entsprechend in der Zusammenfassung der Evidenzanalyse berücksichtigt. Die Berücksichtigung niedriger Evidenzstufen hatte auch das Ziel, eine umfassende Erfassung und Darstellung der existierenden gegebenenfalls minderwertigen Studienlage zu erhalten.

Bei kontrollierten Vorher-Nachher-Studien werden in der Gruppe mit Intervention als auch in der Gruppe ohne Intervention sowohl vor sowie nach Einführung der Intervention die festgelegten Endpunkte erhoben. Bei unkontrollierten Vorher-Nachher-Studien werden in einer Gruppe vor und nach Intervention Daten erhoben mit Vergleich nur innerhalb dieser Gruppe. Ein klar definierter Zeitpunkt der Intervention war dabei gefordert [39].

#### Population

Es wurden nur Publikationen aus dem stationären Sektor oder klinikähnlichen Einrichtungen berücksichtigt; Publikationen aus der ambulanten Versorgung wie z.B. die nicht-stationären DMP-Auswertungen wurden daher ausgeschlossen.

### Intervention und Kontrolle

Eingeschlossen wurden als Intervention alle übergeordneten QS/QM-Programme, die mindestens mehr als 2 Abteilungen in einem Krankenhaus oder mehrere Kliniken (Klinikverbünde) umfassten. Als Vergleich galten Gruppen ohne Intervention für den gesamten Beobachtungsverlauf (kontrollierte Studien) oder Vergleichsgruppen, deren Daten vor oder direkt zu Beginn der Einführung der Intervention erhoben wurden, wenn noch keine Wirkung zu erwarten war, und deren Ergebnisse im zeitlichen Verlauf oder mit den Daten nach Einführung der Intervention verglichen wurden (unkontrollierte Vorher-Nachher-Studien).

### Endpunkte

Als Endpunkte wurden alle Parameter berücksichtigt, die eine Wirksamkeit auf die Ergebnis-, Prozess- oder Strukturqualität in der Patientenversorgung in den Krankenhäusern widerspiegeln. Gesundheitsökonomische Auswirkungen waren nicht Ziel der Untersuchung und wurden nicht in die Auswertung der Studien miteinbezogen.

Für die Studienausswahl wurden die in Tabelle 3 dargelegten Ein- und Ausschlusskriterien zu Grunde gelegt.

**Tabelle 3 Ein- und Ausschlusskriterien der Studienselektion**

Ausschlussgründe	
A1	Keine Ergebnisse einer Evaluation des QS/QM-Programms
A2	Ergebnisse vorhanden, aber <ul style="list-style-type: none"><li>- keine Evaluationsergebnisse eines <i>übergeordneten</i> Programms für ein Krankenhaus (in &gt; 2 Abteilungen) oder mehrere Krankenhäuser/Verbund, nur singuläre QS-/QM-Intervention in einer Abteilung oder</li><li>- keine Evaluation des ganzen Programms, sondern nur von Einzelfaktoren, z.B. der Datenvalidität/-qualität, Mindestmengen oder</li><li>- keine Evaluation zur <i>Wirksamkeit</i> (z.B. nur Darstellung der Ergebnisse der Qualitätsindikatoren ohne Verlauf oder Vergleich)</li></ul>
A3	Ergebnisse vorhanden, aber <ul style="list-style-type: none"><li>- kein <i>QS/QM-Programm</i> (z.B. <i>Behandlungs-Programm</i> mit Evaluation von Behandlungsinterventionen)</li></ul>
A4	nicht stationär/kein Krankenhaus bzw. nicht klinisch (z.B. technische Labormessungen, Geräte, begrenzt auf Labor oder QS von Biomarkern)
A5	Methodisch: Anderes Studiendesign als unter E1 genannt (z.B. fehlende Vergleichsgruppe <i>zu Beginn</i> der Intervention oder fehlende Kontrollgruppe) oder keine Studie (Tagungsbericht, Poster, Abstrakt, Umfragen, Editorial)

Einschlussgründe	
E1	RCT, nicht-randomisierte kontrollierte Studie (Kohortenstudien, kontrollierte Vorher-Nachher-Studien mit Kontrollarm ohne Intervention, Querschnittsstudien), Kohortenstudien ohne Kontrollarm (Verlaufsbeobachtung zu Beginn und im Verlauf einer Intervention, unkontrollierte Vorher-Nachher-Studie und Zeitreihenuntersuchungen) oder systematische Übersichtsarbeiten aus genannten Studientypen mit Intervention QS/QM-Programm mit Evaluationsergebnissen zur Wirksamkeit – <i>Deutschland</i>
E2	RCT, nicht-randomisierte kontrollierte Studie (Kohortenstudien, kontrollierte Vorher-Nachher-Studien mit Kontrollarm ohne Intervention, Querschnittsstudien), Kohortenstudien ohne Kontrollarm (Verlaufsbeobachtung zu Beginn und im Verlauf einer Intervention, unkontrollierte Vorher-Nachher-Studie und Zeitreihenuntersuchungen) oder systematische Übersichtsarbeiten aus genannten Studientypen mit Intervention QS/QM-Programm mit Evaluationsergebnissen zur Wirksamkeit – <i>international</i>
E3	Allgemeine Hintergrundinformationen
E4	Instrumente der QS/QM (Leitlinien, Behandlungspfade, Patientensicherheit, Zertifizierung etc.) mit Evaluationsergebnissen zur Wirksamkeit

Neben dem eigentlichen Ziel der Auswahl von Studien zur primären Fragestellung wurden darüber hinaus Publikationen zu qualitätssichernden Instrumenten wie Leitlinien, Behandlungspfade, Patientensicherheitsprogramme, Zertifizierungen etc., die QS/QM-Programme unterstützen, aus der Treffermenge der systematischen Recherche selektiert. Hierbei wurde ebenfalls der Fokus auf Evaluationsstudien zur Wirksamkeit gesetzt. Da die zugrundeliegende Recherchestrategie nicht auf diese Studienthematik ausgerichtet war, ist diese Auswahl als selektive Literaturrecherche und nicht mit dem Anspruch der Vollständigkeit anzusehen.

### 2.1.2 Bewertung und Evidenzdarstellung

Vor dem Hintergrund der Fragestellung der vorliegenden Arbeit waren als bestverfügbare Evidenz hauptsächlich Untersuchungen in Form von Vorher-Nachher-Studien sowie zu einem weitaus geringeren Anteil in Form von kontrollierten randomisierten, nicht-randomisierten und Kohortenstudien zu erwarten. Es existieren nur wenige Checklisten mit konkreten Kriterien zur Bewertung der Studienqualität für das erstgenannte Studiendesign. Die häufig angewandten Checklisten von wissenschaftlichen Institutionen der Evidenzbasierten Medizin wie dem Scottish Intercollegiate Guidelines Network oder dem Oxford Centre for Evidence-based Medicine [137; 165] sind nur bedingt für diese Art von Studien anwendbar. Daher wurden die Checklisten der Cochrane Effective Practice and Organisation of Care (EPoC) Group [38; 39], einer Subgruppe der Cochrane Collaboration

adaptiert, die auf diesen Studientyp konkreter eingehen. Die Effective Practice and Organisation of Care Group erstellt systematische Reviews mit der Frage nach Effektivität von übergeordneten, regulierenden Interventionen zur Verbesserung der Gesundheitsversorgung wie z.B. Programmen aus QS/QM. Diese Checklisten wurden in Kombination mit den Checklisten des Cochrane Handbook for Systematic Reviews of Interventions [144; 145] herangezogen, um zwei weitere Kriterien zu Angaben von Interessenkonflikten<sup>7</sup> und statistischen Maßzahlen erweitert und in modifizierter Form angewandt. Die verwendeten modifizierten Kriterien für die Studienbewertung sind im Anhang 5 einsehbar. Da Vorher-Nachher-Studien bedingt durch das Studiendesign ein hohes Maß an Verzerrungsrisiko mit sich bringen, sollten bestimmte Qualitätskriterien eingehalten werden, damit eine adäquate Deutung der Daten zulässig ist. Die Bewertung von RCT oder systematischen Übersichtsarbeiten erfolgte nach den Checklisten des Scottish Intercollegiate Guidelines Network [165]. Die methodische Auswahl der einzuschließenden Studien und die Zuordnung des Studiendesigns erfolgten ebenfalls angelehnt an die Definitionen und Empfehlungen [38; 39] der Effective Practice and Organisation of Care Group der Cochrane Collaboration.

Nach Auswahl der Studien wurden die relevanten Daten der den Einschlusskriterien entsprechenden Publikationen nach einem systematischen Schema in Tabellenform dargestellt, um eine einheitliche und strukturierte Berücksichtigung der Studieninhalte für eine faire kritische Bewertung zu gewährleisten. Diese Darstellung umfasste unter anderem neben einer Beschreibung des Studiendesigns und wesentlichen Aspekten der Studienmethodik auch Angaben zur Population, zu den QS/QM-Verfahren als Intervention und zu den Ergebnissen der gemessenen Endpunkte.

Die Gesamtbewertung der Studien wurde nach dem Grad des Verzerrungsrisikos in vier Stufen von geringem bis sehr hohem Verzerrungsrisiko klassifiziert. Dieses Vorgehen ist im Ansatz angelehnt an die Methodik der Arbeitsgruppe „Grading of Recommendations Assessment Development and Evaluation“ (GRADE), die sich in der Studienbewertung vor allem am Verzerrungsrisiko und nachrangig am Studiendesign orientiert [86; 87]. Die zugrunde gelegten Kriterien für die Zuordnung der Kategorien sind in Anhang 4 dargestellt.

Auf Grund der hohen Heterogenität der Studien und der limitierten Angabe an statistischen Maßzahlen wurde davon Abstand genommen, die Studien meta-analytisch zusammenzufassen. Es erfolgte eine qualitative Synthese der verfügbaren Studiendaten.

---

<sup>7</sup> Interessenkonflikte sind definiert als Situationen, die ein Risiko dafür schaffen, dass professionelles Urteilsvermögen oder Handeln, welches sich auf ein primäres Interesse bezieht, durch ein sekundäres Interesse unangemessen beeinflusst wird [97; 117].

### 3 Ergebnisse

In der Anzahl der 3.065 Treffer potentiell relevanter Artikel für die nachfolgende Titel-/Abtrakt-Sichtung wurden nur wenige Publikationen (n=138) identifiziert, die Evaluationen zur Wirksamkeit thematisierten. Es fanden sich im nächsten Schritt der Volltext-Sichtung nur 25 Studien, die die Wirksamkeit zumindest mit einer geringen Validität der Ergebnisse durch Wahl einer Vergleichsgruppe oder Kontrollgruppe evaluierten.

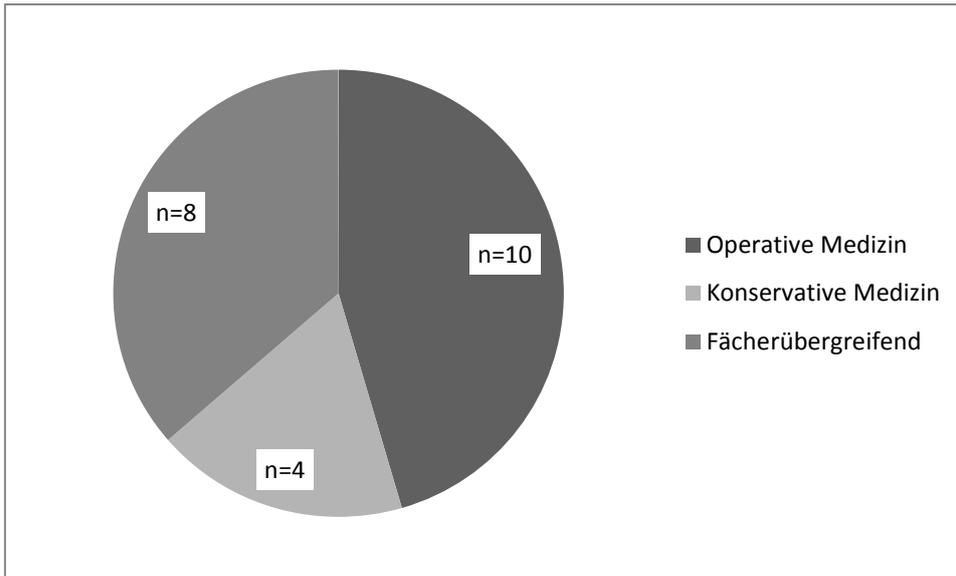
Unter den 25 identifizierten Publikationen, die den a priori festgelegten Einschlusskriterien (siehe Kapitel 2.1.1) entsprachen, waren 22 Primärstudien und drei systematische Reviews. Von den 22 Primär-Publikationen analysierten 17 die deutschen und 5 die internationalen QM-/QS-Verfahren. Ein Review stammte aus Deutschland, zwei stammten aus dem internationalen Raum. Die Reviews folgten zum Teil anderen Ein- und Ausschlusskriterien hinsichtlich der berücksichtigten Publikationen als denen der vorliegenden Arbeit, waren aber auf Grund der partiellen Überschneidungen aussagekräftig.

Den Einschlusskriterien entsprechende Studien zu den größten in Deutschland existierenden QM/QS-Programmen der gesetzlichen externen QS nach §137 SGB V und zu den QS-Programmen aus dem Rehabilitationsbereich der Gesetzlichen Krankenversicherung und der Deutschen Rentenversicherung waren nicht zu identifizieren. Auf Grund der Relevanz dieser drei großen Programme wurden diese dennoch in der vorliegenden Arbeit thematisiert und themenspezifische sonstige Publikationen zur Evaluation wiedergegeben.

#### Beschreibung der Studienmerkmale

Von den 22 den Einschlusskriterien entsprechenden Primär-Publikationen war der größere Anteil mit 10 Studien aus dem Bereich der operativen Medizin, acht Studien stammten aus der konservativen Medizin und bei vier Studien handelte es sich um Publikationen zu fächerübergreifenden Verfahren (siehe Abbildung 5). Bei den meisten Studien (13 von 22 Primärstudien) bestand die Datengrundlage für die Endpunktmessungen aus zusätzlich erhobenen Daten, bei sechs von 22 Studien aus Routinedaten und bei drei der 22 Studien aus Registerdaten. Die Größe der Populationen für die Datenanalyse variierte von minimal 224 bis maximal ca. 100.000 Fälle mit einem annähernden Mittelwert von 20.044 Datensätzen. Die Analysezeiträume dauerten von minimal sechs Monaten bis zu maximal neun Jahren, im Mittelwert 3,7 Jahre.

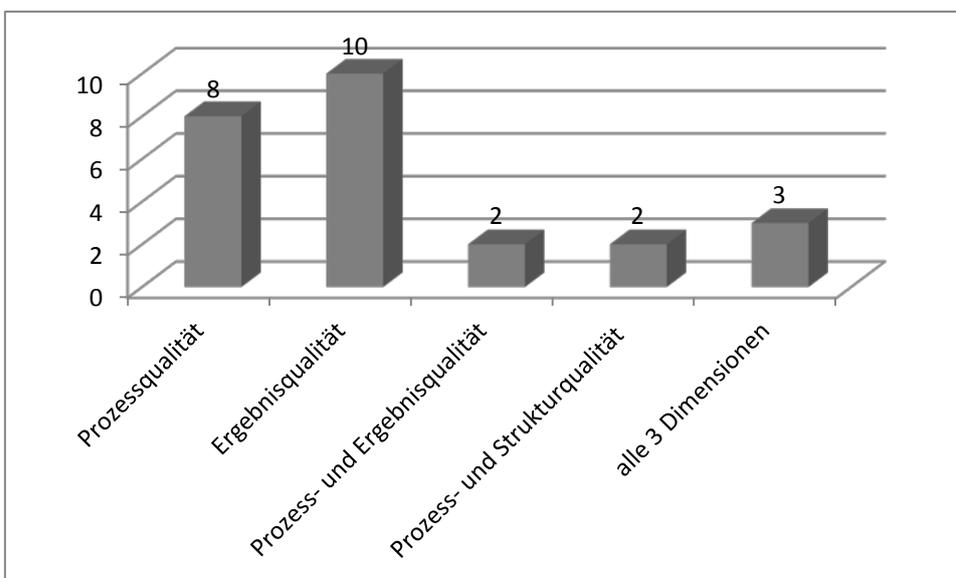
**Abbildung 5 Anteil der Primär-Publikationen aus den einzelnen Fachgebieten**



Zusammenfassende Ergebnisse der Studienbewertung unter methodischen Gesichtspunkten

In acht der 25 Studien [22; 23; 33; 69; 171; 175; 176; 194] wurden lediglich Surrogatparameter der Prozessqualität als möglicher Hinweis auf die zu erwartende Ergebnisqualität berücksichtigt; in zehn der 25 Studien [41; 70; 85; 111; 116; 123; 129; 147; 148; 153] wurden ausschließlich Parameter der Ergebnisqualität untersucht. In den übrigen Studien wurden mehrere Dimensionen als Endpunkte abgedeckt (siehe Abbildung 6).

**Abbildung 6 Verteilung der in den Studien gemessenen Qualitätsdimensionen**



Insgesamt waren die Ergebnisse der Studien in den meisten Fällen einem hohen Verzerrungsrisiko ausgesetzt (siehe Abbildung 7). Dieses Verzerrungsrisiko war beispielsweise durch eine retrospek-

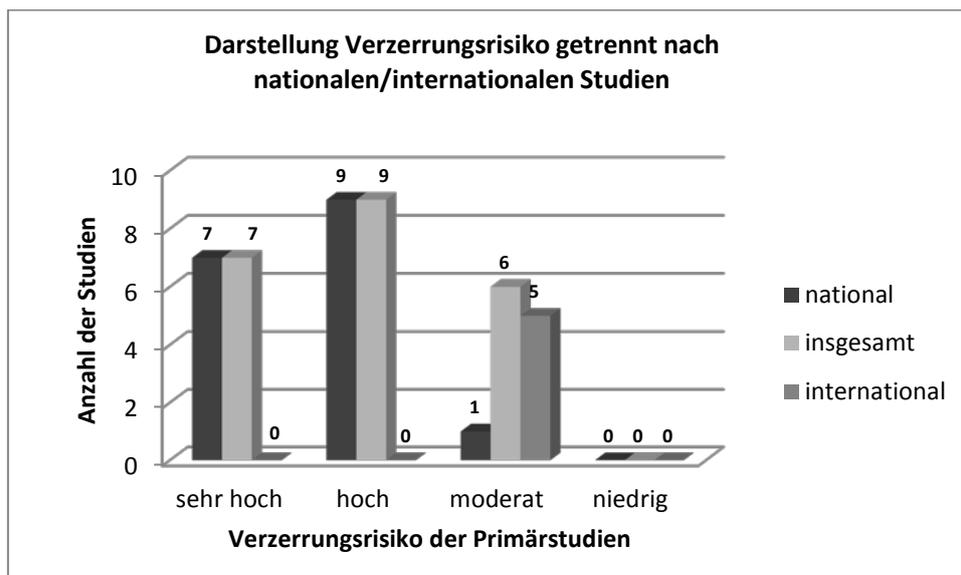
tive Datenerhebung im Vorher-Nachher-Vergleich mit fehlender oder nicht ausreichend transparent dargestellter Vergleichbarkeit der gegenübergestellten Gruppe bezüglich weiterer potentiell relevanter Ko-Faktoren wie Alter, Komorbidität oder Krankheitsstadien oder durch insgesamt geringe Angaben zum methodischen Vorgehen der Datenerhebung und Auswertung bedingt.

Bei mehr als zwei Drittel der 22 Primärstudien (n=16) war das Verzerrungsrisiko hoch bis sehr hoch [14; 22; 23; 51; 111; 116; 123; 129; 147; 148; 151; 153; 175; 176; 194; 199] und nur bei 6 der 22 Primärstudien [33; 41; 69; 70; 85; 171] war das Verzerrungsrisiko moderat. Von diesen sechs Studien waren fünf aus dem internationalen Raum und nur eine Studie [70] aus Deutschland. Es fand sich keine Studie mit einem geringen Verzerrungsrisiko und entsprechend hoher Studienqualität.

Bei 14 von 22 Primärstudien handelte es sich um ein unkontrolliertes Vorher-Nachher-Studiendesign mit einer Vergleichsgruppe vor der Intervention oder zu Beginn der Intervention [22; 23; 41; 69; 85; 111; 123; 129; 147; 148; 151; 153; 175; 194], aber ohne Kontrollgruppe ohne Intervention mit Beobachtung im selben kompletten Zeitverlauf. 5 der 22 Primärstudien [33; 70; 116; 171; 199] wiesen einen kontrollierten Ansatz der Verlaufsbeobachtungen auf, darunter war ein RCT [33]. Dieser hatte ebenfalls methodische Schwächen, sodass dieser gleichfalls ein moderates Verzerrungsrisiko trug. 11 Studien (50%) wurden als prospektiv beschrieben.

Angaben zu Interessenkonflikten, die das Urteilsvermögen der Autoren beeinträchtigen und damit zu Verzerrungen (Bias) und Fehleinschätzungen hinsichtlich der Ergebnisse führen können [117], fanden sich nur bei 14 der 25 Publikationen.

**Abbildung 7 Verzerrungsrisiko der Primärstudien**



### Ergebnisse der Wirksamkeit von QM/QS

Die Ergebnisse innerhalb aller 25 Studien waren ohne Berücksichtigung der Studienqualität insgesamt sehr widersprüchlich. In 13 der 25 Studien [14; 42; 51; 70; 85; 116; 121; 123; 151; 170; 175; 176; 199] wurden inkonsistente Ergebnisse gezeigt, d.h. innerhalb derselben Studie wurden für einen Teil der Endpunkte signifikante Ergebnisse, für einen anderen Teil nicht-signifikante Ergebnisse zur Wirksamkeit gemessen. Der Anteil an Ergebnis- und Prozessqualität war innerhalb dieser inkonsistenten Endpunkte nahezu gleichverteilt. Einen positiven Wirksamkeitsnachweis mit signifikanten Werten der Endpunkte stellten zehn von den 25 Studien dar [22; 23; 33; 69; 111; 129; 147; 148; 153; 194]; kein Effekt wurde bei zwei Studien beschrieben [41; 171].

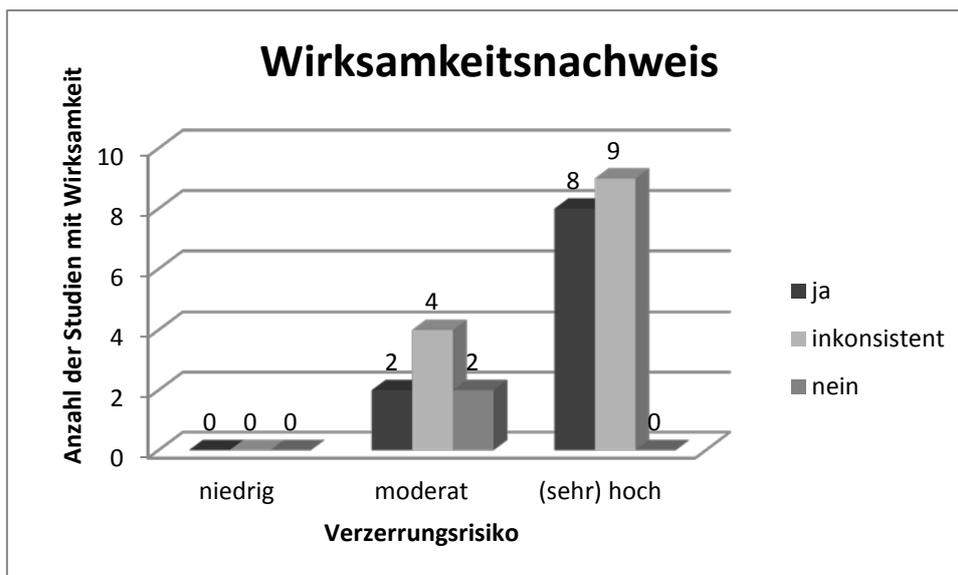
Unter Beachtung des Verzerrungsrisikos der Ergebnisse aller Publikationen (inklusive systematischer Reviews) ergab sich ein weniger positives Bild. Acht von zehn Studien mit positivem Wirksamkeitsnachweis stammten aus der Kategorie mit hohem bis sehr hohem Verzerrungsrisiko [22; 23; 111; 129; 147; 148; 153; 194]. Die Studien mit inkonsistenten Ergebnissen waren anteilmäßig zu nahezu gleichen Teilen in den einzelnen Verzerrungskategorien vorzufinden (vier von acht mit moderatem Verzerrungsrisiko [42; 70; 85; 170], bzw. neun von 17 mit hohem bis sehr hohem Verzerrungsrisiko [14; 51; 116; 121; 123; 151; 175; 176; 199]). Die Studien von methodisch höherer Güte mit moderatem Verzerrungsrisiko (acht von 25) [33; 41; 42; 69; 70; 85; 170; 171] zeigten jedoch überwiegend (sechs von acht Studien dieser Studienkategorie) inkonsistente oder fehlende Belege einer Wirksamkeit bei nicht-signifikanten Werten der Studien-Ergebnisse [41; 42; 70; 85; 170; 171].

Nur in zwei Studien mit moderatem Verzerrungsrisiko konnte ein signifikanter Wirksamkeitsnachweis in Bezug auf die Prozessqualität erbracht werden [33; 69]. Ein negativer Effekt im Sinne einer schädlichen Wirkung der Interventionen konnte in keiner Studie abgebildet werden.

Bei den systematischen Übersichtsarbeiten wurde die zusammengefasste Gesamtbewertung des Verzerrungsrisikos der Einzelstudien in der Auswertung berücksichtigt und nicht die allgemeine Qualität und das Verzerrungsrisiko der methodischen Durchführung der Übersichtsarbeit.

Die Abbildung 8 veranschaulicht die Ergebnisse der Studien unter Berücksichtigung des Verzerrungsrisikos der Studien in einer differenzierten Darstellung.

Abbildung 8 Wirksamkeitsnachweis aller Studien



In der Gesamtschau aller beschriebenen Ergebnisse ließ sich weder ein eindeutig positiver Effekt noch ein negativer Effekt der QM/QS-Programme feststellen. Sowohl die Studien mit hohem Verzerrungsrisiko als auch die methodisch etwas besseren Studien mit moderatem Verzerrungsrisiko zeigten überwiegend inkonsistente Ergebnisse mit partiell positiven oder nicht nachweisbaren Effekten. Damit ist weder eine Wirksamkeit noch eine fehlende Wirksamkeit belegt.

In Tabelle 4 (Seite 54 ff) werden die evaluierten QS/QM-Programme aus Deutschland in einer Übersicht zusammenfassend dargestellt.

### 3.1 Deutschland

In den folgenden Kapiteln werden die evaluierten QS/QM-Programme aus Deutschland mit ihren dazugehörigen Evaluationsstudien beschrieben. Eine detaillierte Beschreibung der Einzelstudien mit Darstellung relevanter Daten in Form von Evidenztabelle ist im Anhang 2 einsehbar.

#### 3.1.1 Fächerübergreifend

In einem umfassenden Gutachten des Kompetenz-Centrums Qualitätssicherung/Qualitätsmanagement (KCQ) wurde im Auftrag des GKV Spitzenverbands eine Aufarbeitung des internen QM in der medizinischen Versorgung in Bezug auf Umsetzung und Nutzenfrage im stationären und ambulanten Sektor vorgenommen. In einer systematischen Recherche mit einem Recherchezeitraum bis zum 1.11.2011 wurden 27 Volltexte als themenrelevante Treffer für den nationalen und internationalen Raum identifiziert [121].

Es fanden sich in der Gesamtauswertung der Daten überwiegend unkontrollierte Vorher-Nachher-Studien von niedriger Studienqualität mit lediglich zwei Vorher-Nachher-Studien mit Kontroll-

gruppe. Innerhalb dieser Treffer fanden sich keine aussagekräftigen Studien zu den Effekten der Einführung von einem internen QM in bundesdeutschen Kliniken [121].

Die Studienanalyse ist für den Gutachtenteil, der den Krankenhaussektor analysiert, insgesamt methodisch gut durchgeführt worden und das Verzerrungsrisiko der Gesamtergebnisse ist gering. Es wurde keine methodische Studieneinschränkung festgelegt, die Einzelstudien wurden mit ihren Studienlimitationen ausführlich beschrieben, Effektstärken oder Signifikanzwerte wurden jedoch nicht angegeben.

#### 3.1.1.1 Externe vergleichende Qualitätssicherung nach §137 SGB V

Die externe QS nach §137 SGB V ist ein seit 2001 etabliertes Verfahren in der stationären Versorgung. Vom 1.1.2001 bis zum 31.12.2009 war die Bundesgeschäftsstelle Qualitätssicherung (BQS) für Qualität und Patientensicherheit [20] auf Bundesebene verantwortlich für den Aufbau und die Umsetzung der datengestützten externen Qualitätssicherung für alle deutschen Krankenhäuser. Dies erfolgte zunächst im Auftrag des Bundeskuratoriums Qualitätssicherung und seit 2004 im Auftrag des Gemeinsamen Bundesausschusses (G-BA), der hierzu eine „Richtlinie über Maßnahmen der Qualitätssicherung in Krankenhäusern“ beschlossen hat [18; 73; 195].

Seit dem Jahr 2010 hat die externe QS das AQUA-Institut [8] für angewandte Qualitätsförderung und Forschung im Gesundheitswesen übernommen, nachdem es im Rahmen einer europaweiten Ausschreibung den Zuschlag als fachlich unabhängige Institution erhalten hat [31; 159].

Als Grundlage für die Qualitätsmessung im Rahmen der externen QS dienen Qualitätsindikatoren [159]. Die Aufgabe des AQUA-Instituts beinhaltet die Auswertung von Qualitätsindikatoren, die von den regionalen Landesgeschäftsstellen für Qualitätssicherung in den einzelnen Bundesländern unterstützt wird [31]. Soweit erforderlich, erlässt der G-BA Vergütungsabschläge für Leistungserbringer, die ihre Verpflichtungen zur Qualitätssicherung nicht einhalten [29]. Diese Krankenhäuser müssen pro fehlendem Dokumentationsfall finanzielle Verluste durch verminderte Kostenerstattung der Krankenkassen in Kauf nehmen [31].

In der Richtlinie des Gemeinsamen Bundesausschusses über Maßnahmen der Qualitätssicherung in Krankenhäusern werden die Details zu den Maßnahmen der externen stationären Qualitätssicherung, zu den Zuständigkeiten und zur Finanzierung festgelegt [73]. Zur Sicherung der Qualität von Krankenhausleistungen sind definierte Leistungsbereiche von den Krankenhäusern verbindlich zu dokumentieren (§4 Absatz 1 der Richtlinie zur Qualitätssicherung in Krankenhäusern) [73].

Die gezielte Rückkopplung statistisch auffälliger Ergebnisse an die Krankenhäuser findet in dem sogenannten „Strukturierten Dialog“ statt [17]. Nach § 11 der Richtlinie des Gemeinsamen Bundesausschusses über Maßnahmen der Qualitätssicherung in Krankenhäusern [73] wird ein Strukturierter Dialog eingeleitet, wenn rechnerische Auffälligkeiten in den Datenvalidierungsverfahren oder durch Abweichung bei einem Qualitätsindikator vom festgelegten Referenzbereich erkennbar

sind [7]. Im Strukturierten Dialog sollen Hinweise zur Optimierung der Dokumentation gegeben werden bzw. Schwachstellen im internen QM identifiziert werden, um strukturelle Prozesse durch kritische Auseinandersetzung mit den Ergebnissen zu verbessern. Kritisiert wird allerdings die Validität der Daten, da die Dokumentation der Daten ausschließlich durch die Einrichtungen stattfindet [193]. Durch umfangreiche Bemühungen des AQUA-Instituts bezüglich der Datenvalidierungsverfahren konnten diese jedoch in den letzten Jahren verbessert werden [10]. Darüber hinaus wurde auch die Validität einzelner Qualitätsindikatoren hinterfragt [193].

Für die externe stationäre Qualitätssicherung des Erfassungsjahres 2012 wurden für bestimmte, vom G-BA festgelegte Leistungsbereiche über vier Millionen Datensätze zu 464 Qualitätsindikatoren über 30 Leistungsbereiche von bundesweit 1.658 Krankenhäusern ausgewertet [10].

In der geschilderten Darstellung der gesetzlichen externen QS werden der immense Aufwand und die Komplexität des Verfahrens deutlich. Umso erstaunlicher ist die Tatsache, dass keine Evaluationsstudie zur Wirksamkeit hinsichtlich einer verbesserten Patientenversorgung mit den Minimalanforderungen einer Vergleichsgruppe entsprechend den Einschlussbedingungen der vorliegenden Arbeit in der Literaturrecherche identifiziert werden konnte.

Diese Beobachtung machten auch die Autoren einer Studie zur externen QS am Universitätsklinikum Carl Gustav Clarus in Dresden, die feststellten, dass der Nutzen dieser flächendeckenden, umfangreichen Dokumentation der externen QS bisher noch nicht systematisch in Bezug auf eine verbesserte Patientenversorgung evaluiert wurde. Sie führten eine systematische Evaluation in Bezug auf Aufwand und Nutzen dieses Verfahrens in den Jahren 2003 bis 2011 an der eigenen Klinik durch. Eine Evaluation der Wirksamkeit in Bezug auf eine Verbesserung der Behandlungsqualität wurde auch hier nicht vorgenommen. Eine Vergleichsgruppe ohne Intervention oder zu Beginn des Verfahrens war gleichmaßen nicht vorhanden [139].

Es wurden 45.639 Datensätze in Form von ausgefüllten Qualitätssicherungsbögen in die Analyse eingeschlossen, von denen in 167 Fällen statistische Auffälligkeiten festgestellt wurden und zu denen Stellungnahmen verfasst werden mussten. Im Ergebnis der Auswertung des Strukturierten Dialogs wurden nur in neun Fällen qualitative Auffälligkeiten festgestellt, in keinem Fall handelte es sich um tatsächliche Qualitätsdefizite. Dieser Detektion von neun qualitativ auffälligen Qualitätsindikatoren stand eine reine Dokumentationsdauer von bis zu 247 Tagen pro Jahr mit zusätzlich nicht messbarer Zeit für weitere Aufwendungen wie Stellungnahmen entgegen. Die Autoren hinterfragten kritisch, ob der Strukturierte Dialog in seiner jetzigen Form zu einer Verbesserung der Behandlung führen kann und der Aufwand in Relation zur Aussagefähigkeit der Daten gerechtfertigt ist. Da die externe Qualitätssicherung mit einer zusätzlichen Erhebung von Daten verbunden ist, wurde empfohlen, die begonnene Verwendung von zusätzlichen Datenquellen wie Register- und

Abrechnungsdaten weiter fortzuführen. Die Autoren kommen zu dem Schluss, dass auf Grundlage der Daten der externen QS der erhoffte Nachweis eines Nutzens für die verbesserte Behandlungsqualität nur sehr eingeschränkt möglich ist und forderten daher eine transparente Evaluationsforschung als integralen Bestandteil der Weiterentwicklung der externen QS [139].

Das Verfahren der externen QS ist Gegenstand zahlreicher Diskussionen hinsichtlich des hohen Aufwands und der Validität der Daten. Exemplarisch wurden in einer Studie die Daten von sechs Qualitätsindikatoren auf der Datenbasis der externen QS 2007 mit Qualitätsindikatoren auf Basis von Routinedaten der DRG<sup>8</sup>-Statistik verglichen, welche signifikante Unterschiede zeigten [120].

Von 2003 bis 2005 wurde in einer Untersuchung in Nordrhein Westfalen anhand dreier Parameter zur hüftgelenknahen Fraktur geprüft, ob sich die Umsetzung einer leitlinienkonformen Versorgung durch die erhobenen Daten der externen QS messen lassen kann [162]. Dafür wurden 48.831 Fälle ausgewertet; orientiert hatte man sich an den damals existierenden themenspezifischen unfallchirurgischen Leitlinien [162]. Bei den Kriterien präoperative Verweildauer, Thrombose- und Antibiotikaprophylaxe zeigte sich eine zum Teil hohe Übereinstimmung mit den Empfehlungen der Leitlinie (64%, 99% und 94%). Die risikoadjustierten Ergebnisse für die Letalität und thrombembolischen Komplikationen zeigten im Zusammenhang mit der präoperativen Verweildauer (<24 versus >48h) keinen signifikanten Unterschied, was durch die limitierten Auswertungszeitfenster des stationären Aufenthalts und die geringe Fallzahl dieser Komplikationen bedingt sein könnte [162]. Es lässt sich durch diese Untersuchung ableiten, dass die Daten der externen QS grundsätzlich geeignet wären, die Implementierung der Leitlinie zu messen. Weitere inhaltliche Aussagen zur Wirksamkeit bzw. zur konkreten Implementierung in diesem untersuchten Bereich sind jedoch nicht möglich, da zum einen keine Kontrollgruppe vorhanden war und zum anderen keine Daten zu Infektionen oder zur Länge der stationären Nachbeobachtung beschrieben wurden. Damit ist auch die Angemessenheit des Beobachtungszeitraums nicht beurteilbar.

In einem ähnlichen unkontrollierten Ansatz einer Verlaufsauswertung untersuchte man ebenfalls an dem Beispiel der Qualitätsindikatoren der präoperativen Verweildauer bei der hüftgelenknahen Femurfraktur und zusätzlich an dem Qualitätsindikator des postoperativen Präparatröntgen beim Mammakarzinom die Umsetzung der damaligen Leitlinie [112], um in der Funktion des Monitorings der Datensätze Erkenntnisse für die angewandte Versorgungsforschung gewinnen zu können [161]. Im Zeitverlauf von 2003 bis 2006 wurden 189.756 Datensätze zur Mamma-Chirurgie und 331.087 Datensätze zur hüftgelenknahen Femurfraktur aus 1.000 bis 1.300 Krankenhäusern analysiert. Beim postoperativen Präparatröntgen wurde ein Zielwert von mindestens 95% angestrebt.

---

<sup>8</sup> DRG=Diagnosis Related Groups (diagnosebezogene Fallgruppen)

2003 wurde bei 36% der dokumentierten Fälle und bis zum Jahre 2006 bei 84% der Fälle eine leitlinienkonforme Behandlung gesehen. Bei der präoperativen Verweildauer ergaben sich keine relevanten Veränderungen bei einem Zielwerte von <15% für eine OP nach 48 h und einer Gesamtrate von 18.5% im Jahr 2003 und einer Rate von 16% im Jahr 2006 [161]. Auch hier zeigt sich die grundsätzliche Möglichkeit, die Daten der externen QS zur Messbarkeit der Implementierung von Handlungsempfehlungen der Leitlinie anhand von Prozessqualitätsindikatoren nutzen zu können. Eine Kausalität für die Verbesserung der Versorgung durch die externe QS lässt sich anhand der Daten allerdings nicht ableiten.

Ein Feedback-System in der Art des Strukturierten Dialogs wurde nach Auswertung mehrerer Studien in einem internationalen systematischen Review von de Vos et al. aus dem Jahr 2009 als effektiver im Vergleich zu alleinigem Feedback ohne Dialog bewertet [44]. In diesem Review wurden Studien zu Feedback in Form von Qualitätsberichten in Kombination mit damit einhergehenden Implementierungshilfen für Qualitätsverbesserungen mit Feedback in Form von alleiniger Mitteilung der Ergebnisse für die interne Evaluation verglichen [44]. Die Implementierungshilfen bestanden aus Unterstützung durch externe Teams, mit welchen gemeinsam die Ergebnisse der Qualitätserhebung diskutiert, analysiert und anschließend Optimierungskonzepte erstellt wurden. Die kombinierten Maßnahmen waren signifikant wirksamer im Hinblick auf die Ergebnisse der erhobenen Qualitätsindikatoren [44].

Eine 2009 anhand von Veränderungen in Qualitätskennzahlen ohne Kontrollgruppe durchgeführte Wirksamkeitsanalyse externer Qualitätssicherungsmaßnahmen ergab insgesamt einen Hinweis für eine Effektivität externer Qualitätssicherung in Bezug auf verschiedene Indikatoren vor allem aus dem Bereich der Koronarangiografie und der gynäkologischen Operationen in Krankenhäusern in Bayern [115]. Vorjahresvergleiche zeigten in der Mehrheit der Indikatoren deutliche Verbesserungen. Auf Bundesebene waren nahezu alle Veränderungen aufgrund der hohen Fallzahl signifikant. Ursächlich hierfür wurden neben den Aktivitäten der Landesgeschäftsstellen sowie der Eigeninitiative von Krankenhäusern der zunehmende Formalismus des Strukturierten Dialogs sowie vor allem der Publikationsdruck durch die zweijährlich erscheinenden Qualitätsberichte von den Autoren angesehen [115]. In dieser Wirksamkeitsanalyse wurde aus der statistischen Datenauswertung von den Autoren geschlussfolgert, dass der direkte Einfluss der Landesgeschäftsstellen im Dialog mit den Krankenhäusern verglichen mit der Initiative der Krankenhäuser selbst nicht überbewertet werden sollte [115].

Auf Grund der fehlenden Kontrollgruppe ohne Intervention sind auch diese Studienableitungen mit einem sehr hohen Verzerrungsrisiko versehen und nur sehr eingeschränkt zu verwenden.

**Zusammenfassend ist festzustellen, dass einige Studien zu verschiedenen Aspekten der externen QS existieren, jedoch keine einzige Evaluationsstudie zur Wirksamkeit auf eine verbes-**

**serte Patientenversorgung mit den Minimalanforderungen einer Vergleichsgruppe ohne oder zu Beginn der Intervention in der Literaturrecherche identifiziert werden konnte.**

### 3.1.1.2 Disease Management Programme im stationären Sektor

Im Jahre 2002 wurden die gesetzlichen Grundlagen für die Einführung von sogenannten „Strukturierten Behandlungsprogrammen bei chronischen Krankheiten“ oder „Disease Management Programmen“ (DMP) geschaffen [172]. Die DMP sollen den Behandlungsablauf und die Qualität der medizinischen Versorgung chronisch Kranker gezielt verbessern. Über die DMP sollen insbesondere eine systematische Kooperation der Leistungserbringer, eine dem aktuellen Stand der Wissenschaft entsprechende Versorgung und eine aktive Mitarbeit der Patienten an der Behandlung erreicht werden. Für die Entwicklung der DMP ist seit 2004 der Gemeinsame Bundesausschuss und für die Zulassung das Bundesversicherungsamt zuständig. Für die Versicherten ist die Teilnahme am Programm freiwillig [14; 51].

Die im Jahr 2002 eingeführten DMP müssen nach gesetzlichen Vorgaben evaluiert werden [29]. Die Krankenkassen oder ihre Verbände haben nach § 137f Absatz 4 SGB V eine externe Evaluation der strukturierten Behandlungsprogramme durchzuführen. Im Original heißt es : *„Die Krankenkassen oder ihre Verbände haben nach den Richtlinien des Gemeinsamen Bundesausschusses nach Absatz 2 eine externe Evaluation (...) durch einen vom Bundesversicherungsamt im Benehmen mit der Krankenkasse oder dem Verband auf deren Kosten bestellten unabhängigen Sachverständigen auf der Grundlage allgemein anerkannter wissenschaftlicher Standards zu veranlassen, die zu veröffentlichen ist“* [172]. Nach dieser Vorgabe soll die Evaluation medizinische Parameter, Ausgabebereiche der Leistungserbringer sowie eine Erhebung zur Lebensqualität umfassen. Ein Vergleich mit einer Kontrollgruppe aus der Regelversorgung ist allerdings nicht vorgesehen, lediglich die Analyse im zeitlichen Verlauf [56].

Eine Studie des Deutschen Krankenhausinstituts untersuchte 2006 die Auswirkungen des Programms im Krankenhausbereich. Diese waren am ehesten beim DMP Brustkrebs zu erwarten, weil hier im Gegensatz zu den anderen DMP-Themen ein Großteil der Behandlung im Krankenhaus erfolgt. Die Deutsche Krankenhausgesellschaft beauftragte das Deutsche Krankenhausinstitut, eine erste Bestandsaufnahme zur Umsetzung der DMP Brustkrebs im Krankenhaus vorzunehmen. In zwei Repräsentativbefragungen aus den Jahren 2004/2005 wurden Strukturen und Entwicklungen in DMP-Krankenhäusern mit Nicht-DMP-Krankenhäusern verglichen [14; 51].

In dem Gutachten wurden strukturelle Unterschiede zwischen DMP- und Nicht-DMP-Krankenhäusern, Entwicklung der Fallzahlen in den beiden Gruppen, Art und Ausmaß der Kooperation zwischen DMP-Krankenhäusern und DMP-Vertragsärzten sowie Änderungen im Versor-

gungsgeschehen in den DMP-Häusern dargestellt. Ein direkter Vergleich zwischen DMP und Nicht-DMP-Krankenhaus wurde lediglich zu den ersten beiden Themen der Strukturqualität durchgeführt. Die Auswirkungen des DMP auf die Versorgungsqualität wurden nicht vergleichend analysiert; die Auswertung erfolgte lediglich innerhalb der DMP-Häuser. Die Autoren führten an, dass es nicht das Ziel war, Auswirkungen auf die Versorgungsqualität differenziert zu evaluieren [14; 51].

Die qualitativen Auswirkungen des DMP Brustkrebs auf Versorgungsprozesse wurden bei der Befragung neben anderen Faktoren durch Veränderungen bei diagnostischen oder therapeutischen Standards, bei operativen Verfahren und bei der Verwendung von evidenzbasierten Leitlinie erhoben. Bei einer Mehrheit von jeweils knapp zwei Drittel der Häuser hatte das DMP Brustkrebs gar keine Änderungen der Versorgung bewirkt. Bei weiteren 15% ließen sich nur wenige Auswirkungen bei den genannten Parametern feststellen. Nur jeweils ca. 10-15% nannten mittelmäßige bis sehr große Veränderungen. Die geringen Änderungen könnten unter anderem darin begründet gewesen sein, dass für die Teilnahme am DMP ohnehin eine Reihe von Anforderungen an die Struktur- und Prozessqualität erfüllt gewesen sein müssen [14; 51].

Die Studie hat bei hohem Verzerrungsrisiko erhebliche Limitationen, um Aussagen zur Wirksamkeit des DMP ableiten zu können. Zum einen handelt es sich vom Studiendesign her lediglich um eine Querschnittsstudie mit einer Kontrollgruppe ohne DMP. Diese Kontrollgruppe wurde jedoch nur bezüglich struktureller Parameter und nicht in Bezug auf die Patientenversorgung miteinander verglichen. Angaben zur Datenauswertung, zur Vergleichsgruppe oder zu den Interessenkonflikten der Autoren sind nicht vorhanden. Eine Verlaufskontrolle aus longitudinalen Daten ist ebenfalls nicht durchgeführt worden.

### 3.1.1.3 Qualitätssicherungsverfahren bei der Helios Kliniken GmbH

Die Helios Kliniken GmbH besitzt nach Angaben des Internetauftritts vom Jahr 2014 72 Akut- und Rehabilitationskliniken [92].

Seit 2000 betreibt Helios ein QM-System für seine Akutkliniken. Dieses beinhaltet die Messung von Qualitätsindikatoren auf Basis von administrativen Daten, ein monatliches Feedback an die Häuser, eine öffentliche Berichterstattung der Daten, halbjährliche Qualitätskonferenzen und eine Analyse zur Verbesserung durch ein Peer-Review-Verfahren [129].

Im Jahr 2000 hat die Helios Klinikette die Krankenhaussterblichkeit bezogen auf die Hauptdiagnose aus den Abrechnungsdaten für die Krankenkassen ausgewertet und diese bei verschiedenen Hauptdiagnosen des Jahres 1999 mit den zu diesem Zeitpunkt publizierten Daten des Statistischen Bundesamtes verglichen. Für die Dokumentation der Abrechnungsdaten wurde die Internationale

Klassifikation von Krankheiten (ICD-9-Schlüssel) verwendet. Die signifikant höhere Krankenhaussterblichkeit bei den Krankheitsbildern Myokardinfarkt, Pneumonie und Herzinsuffizienz und die sich daraus ergebenden Diskussionen führten zur Entwicklung des Peer-Review-Verfahrens zum Zweck der Qualitätssicherung [111].

Beim Peer-Review-Verfahren wurden von einer unabhängigen ärztlichen Arbeitsgruppe mindestens zwei externe Chefärzte des Konzerns (Peers) ausgewählt, die eine in Bezug auf die Mortalitätsdaten auffällige Klinik besuchten und eine retrospektive Aktenanalyse und -bewertung vornahmen. Zuvor wurden diese Akten von den an der Behandlung beteiligten Ärzten ebenfalls bezüglich Verbesserungspotential analysiert. Das Ziel war es, mögliche Probleme zu identifizieren und Verbesserungen durch z.B. verstärkte Leitlinien-Adhärenz, verbesserte Arbeits- oder Kooperationsprozesse anzustoßen [111; 129]. Nachfolgend wurden die Verbesserungen kontrolliert und gegebenenfalls der Prozess wiederholt. Als finanzieller Anreiz für das Erlangen der Qualitätsziele wurden Bonuszahlungen ausgeführt [129].

In einer Auswertung des Verlaufs von 2000 bis 2010 lag die Krankenhaussterblichkeit in den 16 Konzern-Kliniken der Krankheitsbilder Herzinfarkt, Herzinsuffizienz und Pneumonie im Jahr 2000 bei 12,9%, 17,4% und 13,1%. Bis zum Jahr 2010 sank sie auf Werte von 9,1% für Herzinfarkt, 6,4% für Herzinsuffizienz und 8,4% für Pneumonie mit Unterschreiten der bundesweiten Krankenhaussterblichkeit des Statistischen Bundesamtes [111].

Als limitierend ist für diese Auswertung anzusehen, dass keine Maßzahlen zur Signifikanz der Ergebnisse oder zur statistischen Berechnung angegeben wurden. Die Beeinflussung der Ergebnisse in Bezug auf andere Krankheitsbilder bleibt offen. Ein monokausaler Zusammenhang ist nicht belegt; auch die Autoren der Studie sehen die Abnahme der Krankenhaussterblichkeit über die 10 Jahre als multifaktoriell bedingt an. Alleine die Einführung eines Benchmarkings innerhalb der Klinikette und die Einführung anderer verpflichtender QS-Maßnahmen wie z.B. die gesetzliche externe QS können diesen Entwicklungsprozess ebenfalls positiv beeinflusst haben [31]. Ein Wirksamkeitsnachweis ist demnach bei hohem Verzerrungsrisiko der publizierten Daten nicht vorhanden.

In einer weiteren retrospektiven Vorher-Nachher-Studie wurde der Effekt dieses QM/QS-Systems ebenfalls anhand administrativer Daten der Helios Kliniken in Bezug auf die innerklinische Mortalität für Pneumonie, Schlaganfall und Herzinsuffizienz im Vergleich zur bundesweiten Mortalität untersucht. Die Gruppe wurde ein Jahr vor im Vergleich bis zu drei Jahre nach Einführung der Intervention beobachtet. Weiterhin wurden die Subgruppen der Kliniken mit initial suboptimalen Werten, die unter dem bundesweiten Durchschnitt lagen, und die Subgruppe der Häuser mit initial bereits guten Werten gesondert betrachtet. Der Studienzeitraum umfasste die Jahre 2003 bis 2012.

Der Beginn der Beobachtung war abhängig vom Zeitpunkt des Eintritts einer Klinik in den Konzern [129].

Im Ergebnis zeigte die Subgruppe der Kliniken mit initial schlechten Werten eine signifikante Verbesserung der standardisierten Mortalitätsrate<sup>9</sup> bis unter den Bundesdurchschnitt. Bei der Subgruppe der Häuser mit initial bereits guten Werten war kein signifikanter Unterschied zu verzeichnen [129].

Merkmale wie die Komplexität des Falls waren in der Gruppe nach Einführung des QM bei den Krankheitsbildern Pneumonie und Herzinfarkt signifikant höher, die Dauer des stationären Aufenthalts bei den Krankheitsbildern Herzinsuffizienz und Pneumonie mit zum Teil erhöhten Verlegungsraten signifikant verkürzt, sodass durch die ungleichen Gruppen eine Beeinflussung der Ergebnisse mit erhöhtem Verzerrungsrisiko in Betracht gezogen werden muss. Die Autoren erhielten dauerhafte Fördermittel von Helios [129].

Als Limitierung der Ergebnisse fällt neben den Interessenkonflikten auf, dass keine Angabe bestand, ob die Subgruppen a priori definiert wurden, eine Auswertung der Gesamtgruppe wurde nicht dargestellt. Bei insgesamt hohem Verzerrungsrisiko sind die Ergebnisse sehr eingeschränkt verwertbar.

Ogleich die Wirksamkeit des Verfahrens nicht belegt ist, gab es die Möglichkeit, andere Faktoren für Verbesserungsmöglichkeiten in den Behandlungsprozessen zu identifizieren, wie z.B. die Etablierung von konzerninternen Behandlungsempfehlungen und -standards auf Basis von Leitlinien der Fachgesellschaften für die analysierten Krankheitsbilder, eine entsprechende Schulung von Ärzten und die Einführung von Morbiditäts- und Mortalitätskonferenzen. Damit sind andere Bausteine des QM im Klinikverbund initiiert worden [111].

#### 3.1.1.4 Die Initiative Qualitätsmedizin (IQM)

Die Helios Kliniken GmbH, die Johanniter GmbH Gemeinnützige Gesellschaft für soziale Einrichtungen, die Medizinische Hochschule Hannover (MHH), die Malteser Trägergesellschaft gGmbH, die SRH Kliniken GmbH, das Universitätsklinikum Carl Gustav Carus Dresden an der TU Dresden, die Vereinigung Berufsgenossenschaftlicher Kliniken und das Universitätsspital Basel gründeten im Jahre 2008 den gemeinnützigen Verein Initiative Qualitätsmedizin (IQM) [96].

Bis 2013 waren 230 Kliniken in Deutschland, Österreich und der Schweiz Mitglied bei IQM [60]. 3,5 Millionen stationäre Behandlungsfälle wurden abgebildet, was einem Anteil von 17 % aller Fälle in Deutschland entspricht [148]. Der Anteil freigemeinnütziger Träger betrug 11,6%, interna-

---

<sup>9</sup> Die standardisierte Mortalitätsrate gibt das Verhältnis von eingetretenen zu den nach Alter und Geschlecht adjustierten, erwarteten Todesfällen an.

tionale Träger (Österreich und Schweiz) 14,3%, Universitäten 20%, private Träger 22,9% und die kommunalen Häuser 31,1% [148].

Die Methodik von IQM basiert auf Erfahrungen der Helios-Kliniken-Gruppe (siehe Kapitel 3.1.1.3.) mit der Verwendung von Routine-Daten zur Qualitätsmessung in Kombination mit einem Peer-Review-Verfahren [147].

Diese Initiative hat das selbsterklärte Ziel, ein QS-Verfahren in die Öffentlichkeit zu tragen, das mit wenig zusätzlichem Dokumentationsaufwand transparent Ergebnisqualität analysiert und Ärzte als Qualitätsverantwortliche in den Mittelpunkt stellt [60]. Ein weiteres Ziel des IQM Peer-Review-Verfahrens ist es, Optimierungen durch einen gegenseitigen Lernprozess und eine offene Fehlerkultur zu erreichen [148].

Das IQM-Verfahren hat die Grundsätze übernommen, dass die Indikatormessungen aus Routinedaten erfolgen, Transparenz durch Veröffentlichung der Ergebnisse hergestellt wird und ein aktives Qualitätsmanagement durch eine kollegiale Diskussion im Peer-Review-Verfahren auf Augenhöhe durchgeführt wird [60; 147; 148].

In der Weiterentwicklung des Verfahrens besteht das Peer-Team aus einem Teamleiter und drei aktiv tätigen Chefarzten. Nach Aktenanalyse der Peers und der Kollegen vor Ort erfolgt die gemeinsame kollegiale Diskussion aller Fälle. In dieser Diskussion werden gemeinsam konkrete Maßnahmen erarbeitet, die eine Optimierung erwarten lassen, welche zum Abschluss anhand der wichtigsten Punkte der Klinikleitung vorgestellt werden, die neben dem Chefarzt für die Umsetzung der benannten Maßnahmen als Kontrollinstanz mitverantwortlich ist. Der Erwartungswert ist gleichfalls die alters- und geschlechtsspezifisch adjustierte Sterblichkeit der Klinik, die sich aus den Bundesdurchschnittswerten pro Klinik errechnen lässt [148].

Mit Berücksichtigung der Pilotprojekte 2009 wurden inzwischen 136 Peer-Review-Verfahren bis 2012 durchgeführt [60]. Das identifizierte Optimierungspotenzial lag 2011 wie im Vorjahr bei 64%. Die Sterblichkeitsrate von drei ausgewählt dargestellten Krankheitsbildern sank nach erfolgreichem Peer-Review-Verfahren von 2009 bis 2011. So sank die Todesfallrate bei Herzinfarkt von 12,2 % vor dem Review in 2009 auf 5,3% in 2011, bei Herzinsuffizienz von 19,8% auf 11,7% in 2011 und beim Schlaganfall von 15,8% in 2008 auf 9,1% in 2011 [3; 148].

Durch die Analyse der Daten zu Hirninfarkt, Herzinfarkt und Herzinsuffizienz wurden übergreifende Optimierungspotenziale unter anderem in den Themenbereichen der leitlinienorientierten, zeitgerechten Diagnostik und Therapie, der Schnittstelle zwischen den verschiedenen Kranken-

haussektoren und der konsequenten Patientenbehandlung bei Entscheidungen zur Therapiebegrenzung dargestellt [60].

Die BÄK und die Landesärztekammern haben diese Initiative ab 2010 begleitet und die Ergebnisse einer Auswertung bereits für eine Optimierung des Verfahrens berücksichtigt. So wurden z.B. zu verschiedenen Themen Checklisten entwickelt, um die Kliniken in der Organisation zu unterstützen, und Schulungen der Peers mit Fokus auf das Kommunikationstraining (Moderation und Gesprächsführung) durchgeführt. Des Weiteren hat diese Begleitung der BÄK zu einer gemeinsamen Erstellung des Curriculums „Ärztliches Peer-Review-Verfahren“ in Kooperation mit anderen Institutionen unter Federführung der BÄK geführt [147; 148].

Es ist eine unabhängige Studie geplant, die die Auswirkungen dieses Peer-Review-Verfahrens genau untersuchen soll [147; 148]. Bisher liegen noch keine Ergebnisse vor, sodass eine abschließende Beurteilung der Wirksamkeit anhand der vorliegenden Publikationen nicht möglich ist.

Positiv stellen sich der geringe Aufwand und die fehlende zusätzliche Erhebung von Daten dar. Anhand der vorliegenden publizierten Ergebnisse lässt sich jedoch nicht ableiten, welchen Anteil das Peer-Review-Verfahren des IQM konkret an den Qualitätsverbesserungen hat. Man sieht lediglich Verbesserungen zu ausgewählten Erkrankungen in einzelnen Kliniken ohne Analyse weiterer konkurrierender Ko-Faktoren wie z.B. andere QS-Maßnahmen, die in dem gleichen Zeitraum eingeführt wurden. Diese werden nicht in den Publikationen [147; 148] diskutiert.

Es werden Mortalitätsdaten im zeitlichen Verlauf im Vergleich zu Daten vor Einführung des Peer-Review-Verfahren angegeben. Bei den Ergebnissen finden sich kaum Angaben zur statistischen Maßzahlen wie Signifikanzniveaus. Weiterhin ist anzumerken, dass bei der Risikoadjustierung lediglich demografische Faktoren wie Alter und Geschlecht berücksichtigt werden; weitere Adjustierungsvariablen wie z.B. Komorbiditäten bleiben unberücksichtigt und schränken dadurch einen Vergleich zwischen den Gruppen ein.

Lediglich anhand von Mortalitätsdaten die Qualität einer Klinik zu messen, ist kritisch zu hinterfragen, da es Hinweise gibt, dass ein Zusammenhang der Mortalitätsrate mit der Versorgungsqualität nicht hinreichend verlässlich ist [140]. Die Verwendung von Routinedaten führt ebenfalls oftmals zu kontroversen Diskussionen, da diese primär zu Abrechnungszwecken und nicht zur Qualitätssicherung erhoben werden, was einen begrenzten Informationsgehalt durch fehlende qualitätsrelevante klinische Details und gegebenenfalls eine eingeschränkte Datenvalidität durch die Abhängigkeit der Güte und Systematik des Kodiervorgangs bedingen kann [188].

Potentielle Interessenkonflikte sind in der Tatsache zu sehen, dass fast alle Publikationen zu Ergebnissen dieser Initiative von leitenden Helios-Mitarbeitern oder von Autoren mit dauerhafter Förderung durch Helios verfasst wurden.

### 3.1.2 Operative Medizin

#### 3.1.2.1 Schmerztherapie

##### Initiative Schmerzfremde Klinik mit Zertifikat Qualitätsmanagement Akutschmerztherapie

Die durch die Chirurgische Arbeitsgemeinschaft Akutschmerz [49] gegründete Initiative „Schmerzfremde Klinik“ hat zusammen mit dem TÜV Rheinland das Zertifikat „Qualitätsmanagement Akutschmerztherapie“ [186] mit dem Ziel der Verbesserung des Umgangs mit akuten Schmerzen in deutschen Kliniken entwickelt. Diese beinhaltet im Wesentlichen die Implementierung der Leitlinie „Behandlung akuter perioperativer und posttraumatischer Schmerzen“ [47] durch Anpassung an die jeweilige Klinik mit Entwicklung von Standards und Behandlungspfaden. Voraussetzung ist die Bildung eines Qualitätszirkels, der sich aus Mitgliedern aller an der Schmerztherapie beteiligten Berufsgruppen zusammensetzt, und Schulungen durch Workshops und Kurse. Zum Stand 2011 waren 48 Krankenhäuser seit dem Beginn im Jahr 2006 zertifiziert worden, davon 83 % aus operativen Abteilungen [16; 116].

Diese Initiative wurde in einer Studie im Vergleich mit dem nachfolgend aufgeführten Projekt evaluiert [116; 143].

##### Qualitätsverbesserung in der postoperativen Schmerztherapie (QUIPS)

Das Projekt QUIPS [143] wird unter federführender Leitung der Deutschen Gesellschaft für Anästhesiologie und Intensivmedizin (DGAI) und dem Berufsverband der Deutschen Anästhesisten (BDA) in mehr als 130 deutschen Kliniken als Routineinstrument der Qualitätssicherung eingesetzt. Es beinhaltet eine Qualitätsindikatoren-Datenerhebung und Analyse von Patientenangaben zur Schmerztherapie mit einem webbasierten automatisierten Feedback für ein internes und externes Benchmarking sowie eine kontinuierliche Verlaufsbeobachtung in den teilnehmenden Kliniken. Das Ziel ist, die Qualität der postoperativen Schmerztherapie durch eine standardisierte Erhebung von Ergebnisparametern, ihre Analyse und anschließendes Feedback unter Einbeziehung von Benchmarkstrategien zu verbessern. Außerdem sollen Ressourcen gezielt eingesetzt und der Erfolg von Optimierungsmaßnahmen evaluiert werden [116; 122; 143].

Das Projekt hat 2006 begonnen und wurde durch das Bundesministerium für Gesundheit (BMG) bis 2007 gefördert, anschließend übernahmen die Deutsche Gesellschaft für Anästhesiologie und Intensivmedizin (DGAI) und der Berufsverbandes Deutscher Anästhesisten (BDA) die Schirmherrschaft. Die Deutsche Gesellschaft für Chirurgie (DGCH), der Berufsverband Deutscher Chirurgen (BDC) sowie die Österreichische Gesellschaft für Anästhesiologie, Reanimation und Intensivmedi-

zin (ÖGARI) sind Kooperationspartner von QUIPS. Das QUIPS-Projektteam setzt sich aus Mitgliedern aller an der Schmerztherapie beteiligten Berufsgruppen zusammen [116; 122; 143].

In einer Vorher-Nachher-Studie mit z.T. prospektiver, zum Teil retrospektiver standardisierter Erhebung mittels eines validierten Fragebogens sollten drei Gruppen vor und nach Einführung dieser Maßnahme verglichen werden. Eine der Gruppen war ursprünglich als Kontrollgruppe geplant. Diese Gruppen waren Kliniken mit dem Zertifikat „Qualitätsmanagement Akutschmerztherapie“ [186], Kliniken mit Teilnahme an dem Benchmarking-Projekt zur Qualitätsverbesserung in der postoperativen Schmerztherapie (QUIPS) und eine Kontrollgruppe ohne Beteiligung an beiden Qualitätsmaßnahmen [116].

In der QUIPS-Gruppe waren Kliniken der Traumatologie, Orthopädie, Allgemein Chirurgie vertreten; in der Gruppe mit dem Zertifikat „Qualitätsmanagement Akutschmerztherapie“ waren Kliniken der Kardio-, Viszeral-, Allgemein-, plastischen Chirurgie und Unfallchirurgie, Orthopädie, Gynäkologie, Urologie beteiligt. Erhoben wurde die patientenseitige Bewertung bezüglich 1) postoperativer Schmerzen, 2) Nebenwirkungen, 3) schmerzbedingter Einschränkungen und 4) Patientenzufriedenheit (Skala 1-15) mit der erhaltenen Schmerztherapie. Es zeigte sich eine signifikante Verbesserung beider Verfahren in allen vier Indikatorbereichen mit stärkerer Verlaufsverbesserung in der Gruppe des Zertifikats „Qualitätsmanagement Akutschmerztherapie“ [116].

Ob diese Ergebnisse kausal auf die Qualitätsinitiativen zurückzuführen sind, lässt sich nicht beurteilen, da sich mehrere Limitationen ergaben. Diese sind zum einen in der fehlenden adäquaten Kontrollgruppe zu sehen, die wegen einer zu geringen Fallzahl weder mit den anderen Gruppen noch in der Nachbeobachtung ausgewertet wurde; zum anderen ist eine weitere Limitation in der fehlenden Darstellung der Populationscharakteristik der Kliniken und Patienten (Alter, Komorbidität, Allgemeinzustand, Art der Operation, andere Qualitätsmaßnahmen etc.) zu sehen, die eine Beurteilbarkeit der Vergleichbarkeit der Gruppen und möglicher Confounder<sup>10</sup> erschwert.

In einer weiteren Studie im Design eines unkontrollierten Vorher-Nachher-Vergleichs zu Beginn und im Verlauf des Projekts wurde QUIPS als Pilotprojekt evaluiert. Zwischen Oktober 2003 bis September 2006 nahmen 30 operative Abteilungen von sechs Krankenhäusern teil. Zu Beginn erfolgte die Förderung durch das BMG, im Verlauf durch Fachgesellschaften der Chirurgie und Anästhesie. Es wurde die maximale Schmerzintensität postoperativ mit der Numerischen Rating-Skala (NRS) gemessen, die die Schmerzintensität auf einer Skala von eins bis zehn beurteilt; dies erfolgte mit drei Messpunkten zu Beginn (Baseline), in der Mitte und am Ende des Studienzeit-

---

<sup>10</sup> Confounder sind alle wesentlichen Faktoren, die nicht direkt Gegenstand der Untersuchung sind, aber Auswirkungen auf den gemessenen Endpunkt haben können [114].

raums. Es ließ sich nur nach den ersten drei Monaten in fünf von sechs Kliniken eine signifikante Verbesserung feststellen, im Verlauf war bis zum Ende des Beobachtungszeitraums nur in zwei Kliniken weiterhin ein positiver Effekt, in zwei Kliniken ein nicht-signifikanter Effekt und in zwei Kliniken eine signifikante Verschlechterung zu verzeichnen. Die Studie hat diverse Limitationen durch das Studiendesign und durch die fehlende Berücksichtigung von Confoundern, die den schwachen gemessenen Effekt noch unwahrscheinlicher werden lassen [123].

### 3.1.2.2 Unfallchirurgie

#### TraumaRegister

Das Traumaregister [185] wurde als QS-Programm 1993 gegründet, um multizentrisch anonyme Behandlungsdaten prospektiv zu sammeln und die Qualität der klinischen und präklinischen Traumapatienten analysieren und kontrollieren zu können. Die Daten betreffen die komplette Behandlung vom Unfallort bis hin zur Entlassung. Dabei soll vor allem die Prozessanalyse den teilnehmenden Krankenhäusern helfen, Schwachpunkte zu identifizieren und gegebenenfalls verbesserte Behandlungskonzepte aufzustellen. Die teilnehmenden Kliniken erhalten seit 1999 für diesen Zweck jährliche individuelle Feedback-Berichte in anonymisierter Form mit Daten zu diagnostischen und therapeutischen Maßzahlen und Daten zu Mortalität, Krankenhausaufenthalt und Zustand bei Entlassung. Zusätzlich werden Auswertungen von Qualitätsindikatoren vorgenommen. Vergleiche werden mit dem Vorjahr desselben Krankenhauses und mit den gesamten Registerdaten getätigt. Zum Vergleich zwischen den einzelnen Krankenhäusern wird nach Schwere der Traumatisierung mittels eines Prognose-Scores adjustiert (Revised Injury Severity Score (RISC-Score)) [153].

Eine retrospektive Studie untersuchte die Überlebensraten vor und nach Einführung des Feedback-Berichts in einem Zeitraum von 1999 bis 2005. Mit dem Revised Injury Severity Score wurde die Prognose der Verletzten in Form der erwarteten Mortalität mit der tatsächlich eingetretenen standardisierten Mortalitätsrate auf einer Zeitachse nach Einführung des Feedback-Berichts verglichen. 1999 zeigte sich, dass der Anteil der tatsächlichen Mortalität 22,8% versus der prognostisch erwarteten 22,0% betrug. Bei stetiger Verbesserung im Lauf der Jahre wurde 2005 erstmalig ein statistisch signifikanter Effekt erreicht. In 2005 war der Anteil 18,7% versus 22,1%. Der Zeitraum bis zur Notfall-OP (45 Minuten versus 39 Minuten) oder bis zur ersten Bildgebung (25 Minuten versus 15 Minuten) hatte sich im Verlauf ebenfalls verbessert [153].

Als Limitation der Ableitung von Aussagen zur Wirksamkeit ist zu nennen, dass kaum Angaben zur statistischen Grundlage der Berechnung und nur wenig Angaben zur Adjustierung gegeben werden. Zusatzangaben von weiteren relevanten Variablen der Population werden nicht näher dargestellt. Von wem die Auswertung vorgenommen wurde, wurde nicht beschrieben.

Die Kausalität der positiven Ergebnisse ist nicht nur in der Intervention, sondern eher in mehreren Faktoren zu sehen. Ein Selektionsbias ist möglich, da die Teilnahme der Kliniken auf freiwilliger Basis und keine Kontrolle auf Vollständigkeit der Daten erfolgten, sodass die Möglichkeit besteht, dass schlechte Ergebnisse nicht berichtet wurden [153]. Weitere Faktoren von QS-Maßnahmen, die im Beobachtungszeitraum eingeführt wurden (Leitlinien, gesetzliche externe QS), werden nicht diskutiert.

In einer weiteren Untersuchung erfolgte der Vergleich zu Beginn der Einführung am 1.1.1999 bis zum 31.12.2001 [151]. Daten von 6.936 Patienten des Registers wurden ausgewertet und Veränderungen in den Kliniken per Fragebogen erhoben. In 85 % der Krankenhäuser wurden seit Einführung des TraumaRegisters qualitätsverbessernde Maßnahmen als Folge der Rückmeldung aus dem Datenregister eingeführt, insbesondere im Rahmen der Schockraumversorgung.

Als Ergebnis zeigte sich eine signifikante Verbesserung im Sinne einer Beschleunigung diagnostischer Maßnahmen, der Einleitung therapeutischer Maßnahmen und der Beatmungsdauer. In Bezug auf den Endpunkt der tatsächlichen versus der erwarteten Letalität war im gesamten Beobachtungszeitraum die tatsächliche Letalität niedriger; im Verlauf war jedoch kein signifikanter Unterschied zu sehen [151].

Die Studie hatte ebenfalls ein sehr hohes Verzerrungsrisiko durch das Studiendesign und die fehlenden Angaben zu den Vergleichsgruppen.

### TraumaNetzwerk

Darüber hinaus existiert seit 2008 das TraumaNetzwerk der Deutschen Gesellschaft für Unfallchirurgie (DGU) [184]. Das Projekt TraumaNetzwerk DGU dient der Optimierung der Prozess- und Strukturqualität durch eine zertifizierte Vernetzung von Traumazentren über Netzwerke geprüfter Krankenhäuser einer Region, die regelhaft an der Versorgung Schwerverletzter teilnehmen. Ziel des TraumaNetzwerks ist die Förderung der optimalen Versorgung Schwerverletzter [152]. Die daran teilnehmenden Einrichtungen übernehmen entsprechend ihrer Ausstattung und Struktur unterschiedliche Aufgaben in diesem Netzwerk als überregionale, regionale oder lokale Traumazentren verteilt über Deutschland [184]. In diesem Projekt wurden von 2008 bis Mai 2013 insgesamt 39 Traumanetzwerke zertifiziert. Die Finanzierung erfolgt über Audit- und Teilnahmegebühren am TraumaRegister. Rehabilitation ist ein Bestandteil des TraumaNetzwerks [152].

Kernpunkt ist ein QM auf Grundlage des „Weißbuchs Schwerverletztenversorgung der DGU“ [46]. Das Weißbuch der Schwerverletztenversorgung wurde 2006 von der Deutschen Gesellschaft für Unfallchirurgie mit Empfehlungen zur Struktur, Organisation und Ausstattung der Schwerverletztenversorgung in Deutschland publiziert. Auf Grund von veränderten Empfehlungen der interdisziplinären Leitlinie zur Schwerverletztenversorgung der DGU hat diese das Weißbuch überarbeitet und 2012 eine zweite Auflage herausgegeben [152; 184].

**Publikationen zur Evaluation der Effektivität des TraumaNetzwerks sind in kontrollierten Studien bislang nicht identifizierbar [152].**

### 3.1.3 Konservative Medizin

#### 3.1.3.1 Krankenhaushygiene

In einer prospektiv kontrollierten Studie, die vom BMG gefördert wurde, erfolgte eine Auswertung über einen Zeitraum von 26 Wochen in vier Krankenhäuser mit einem QM zur Reduktion von nosokomialen Infektionen und in vier Kontrollkliniken ohne QM [70]. Es wurden die Daten zu Beginn der Einführung des QM und im Verlauf bei insgesamt 11.295 Patientenfällen verglichen. Das QM umfasste die Einrichtung eines Qualitätszirkels mit daraus abgeleiteter Entwicklung von QS-Instrumenten und der Erhebung von nosokomialen Ereignissen.

In der Auswertung der ersten Beobachtungsperiode ergab sich ein signifikanter Unterschied in der Ereignisrate an nosokomialen Infektionen zugunsten der QM-Kliniken, im weiteren Verlauf der zweiten Beobachtungsperiode war dieser Effekt jedoch bei grenzwertig nicht-signifikantem Ergebnis nicht mehr sichtbar [70]. Da Patienten in der Kontrollklinik höhere Risikomerkmale aufwiesen und damit eine höhere Ereignisrate zu erwarten wäre, ist das Ergebnis vermehrt aussagekräftig. Limitierend für die Studie waren die zum Teil eingeschränkte Vergleichbarkeit der Gruppen und die Wahl der Vergleichsgruppe zu Beginn der Intervention anstelle von einem Zeitpunkt *vor* der Intervention.

#### 3.1.3.2 Nephrologie

Qualität in der Nephrologie (QiN) [142] ist ein QM-System, welches 1999 mit dem Ziel der Verbesserung und Sicherung der Qualität von Dialyseeinrichtungen gegründet wurde. QiN wurde von zwei Instituten der Universität Köln (Klinik für Nephrologie und Institut für Gesundheitsökonomie und Klinische Epidemiologie) in Kooperation mit dem Kuratorium für Dialyse und Nierentransplantation e.V. (KfH) entwickelt, welches ein Behandlungsnetzwerk von Dialysezentren in Deutschland darstellt [175; 176]. QiN ist ein externes QM-System, welches eine Datenerhebung in einer zentralen Datenbank für terminale Nierenerkrankungen und einen vierteljährlichen strukturierten, einrichtungsspezifischen Bericht in pseudonymisierter Form bezüglich der Ergebnisse der Qualitätsindikatoren und Daten zum Benchmark umfasst [175; 176].

In einem Vorher-Nachher-Vergleich von Datenerhebungen zum Zeitpunkt der Einführung des QM-Programms 1999 und nach einer Beobachtungszeit bis 2003 zeigte sich eine Verbesserung von einzelnen Surrogatparametern - wie Daten zur von KfH definierten Prozessqualität (Dialyseeffizienz, Blutdruck, einzelne Laborparameter). Weitere Endpunkte (z.B. Calcium/Phosphat-Haushalt, Ernährung, Entzündungsparameter) blieben ohne Veränderung [175]. Daten zur Mortalität, Le-

bensqualität und Morbidität (Häufigkeit und Dauer von Krankenhausaufenthalten etc.) sollten zwar gemessen werden, wurden jedoch nicht dargestellt.

In einer weiterführenden Analyse sollte untersucht werden, ob die Verbesserungen dem QiN Programm oder einem allgemeinen Trend zuzuschreiben sind [176]. Dafür wurde ein Vergleich der Beurteilungsparameter für die Dialyseeffizienz zwischen Einrichtungen mit früher (zu Beginn im Jahr 1999) versus später (im Jahr 2002) Programmteilnahme vorgenommen. Es zeigten sich marginal schlechtere Ergebnisse der Mittelwerte der „späten“ Gruppe, aber eine signifikante Verbesserung im Verlauf in beiden Gruppen [176].

Als wesentliche Limitation sind das Fehlen einer adäquaten Kontrollgruppe und die fehlenden Angaben zu Begleitvariablen der Gruppen zu sehen. Die Einführung beispielsweise neuer Leitlinien oder Veränderungen bei Medikamenten wurde in der Auswertung ebenfalls nicht berücksichtigt. Die Kausalität ist nicht bewiesen und in mehreren Faktoren zu sehen. Darüber hinaus werden keine Angaben zur Ergebnisqualität, nur zu Surrogatparameter der Prozessqualität getätigt, die zwar zum Teil statistisch signifikante Werte aufzeigen, deren klinische Relevanz jedoch diskussionswürdig ist. Publierte Daten zur Mortalität, Lebensqualität und Morbidität wurden derzeit nicht identifiziert.

### 3.1.3.3 Rehabilitation

Die beiden größten Leistungsträger der medizinischen Rehabilitation sind in Deutschland die Deutsche Rentenversicherung (DRV) und die Gesetzliche Krankenversicherung (GKV). Je nach federführendem Träger sind die Einrichtungen verpflichtet, sich an einem der beiden Programme zu beteiligen [64]. Bei beiden Programmen erfolgte die Entwicklung der eingesetzten Instrumente und Verfahren durch wissenschaftliche Institute. Die Finanzierung der Durchführung und Auswertung werden in beiden Fällen von dem jeweiligen Rehabilitations-Träger übernommen. Beim QS-Programm der GKV [76] ist dies erst seit 2007 der Fall, davor mussten die Einrichtungen die Kosten selbst tragen. Das DRV-Programm [48] wird in der Routine durch Mitarbeiter der DRV organisiert und durchgeführt. Die GKV beauftragt externe wissenschaftliche Institute von Universitätsklinikum sowie zur unabhängigen Datenauswertung das BQS Institut für Qualität und Patientensicherheit [19]. Dabei stehen die analytischen Messungen und risikoadjustierten Einrichtungsvergleiche von Qualitätsindikatoren im Mittelpunkt [64]. Die Strukturqualität wird bei beiden Trägern mittels Strukturhebungsbögen erfasst und durch vor-Ort Begehungen (Visitationen) überprüft. Die Ergebnisqualität wird in beiden Programmen primär über die Angaben des Patienten in Fragebögen gemessen, um die Patientenperspektive miteinzubeziehen. Dies wurde sehr kontrovers diskutiert, da generell die Befähigung des Patienten in Frage gestellt wurde, sich zur Qualität der rehabilitativen Behandlung, die als vorwiegend ärztlich-medizinisch definiert wurde, zu äußern [105].

2004 wurde zwischen den Spitzenverbänden der Krankenkassen und den für die Wahrnehmung der Interessen der Vorsorge- und Rehabilitationseinrichtungen maßgeblichen Spitzenorganisationen zur Qualitätssicherung und zum Qualitätsmanagement in der stationären Vorsorge und Rehabilitation nach §137 d Absatz 1 und 1a SGB V eine Vereinbarung getroffen [190]. Diese hatte das Ziel, dass Rehabilitationsleistungen mit Hilfe von qualitätssichernden Maßnahmen ausreichend, zweckmäßig, wirtschaftlich, zielgerichtet und nach dem allgemein anerkannten Stand der medizinischen Erkenntnisse entsprechend erbracht werden. Des Weiteren sollte die Qualität der Leistungserbringung transparent dargestellt, bewertet, verglichen und damit einer Qualitätsverbesserung der Strukturen, Prozesse und Ergebnisse zugeführt werden [190].

Grundlage dafür war ein externes Qualitätssicherungsverfahren mit folgenden Instrumenten:

- Strukturerhebungsbogen mit definierten Mindestanforderungen zur Messung der Strukturqualität
- Peer-Review-Verfahren zur Messung der Prozessqualität
- Fragebogen für den Patienten und Arzt/Therapeut zur Messung der Ergebnisqualität
- Fragebogen für den Patienten zur Messung der Patientenzufriedenheit
- Visitationen der Einrichtung durch externe Begutachter für eine validierte Gesamtbeurteilung [182; 190].

Durch eine gemeinsame Erklärung der Träger der gesetzlichen Kranken-, Renten- und Unfallversicherung aus dem Jahr 1999 wurde eine Vereinheitlichung der QS-Programme der Krankenkassen und der Rentenversicherung verbindlich festgelegt [104]. Diese ist nur bis zu einem gewissen Grad umgesetzt worden, in einigen Bereichen gibt es weiterhin Divergenzen [64].

**Publizierte Evaluationen für einen Beleg zur Wirksamkeit für eine verbesserte Patientenversorgung waren für die QS-Programme der DRV und der GKV nicht identifizierbar.**

#### Qualitätssicherung bei der Rehabilitation der Deutschen Rentenversicherung (DRV)

1994 führte die gesetzliche Rentenversicherung auf Initiative der Bundesversicherungsanstalt für Angestellte (BfA) ein umfassendes Qualitätssicherungsprogramm für die stationäre medizinische Rehabilitation ein, was seit 1997 routinemäßig eingesetzt und kontinuierlich weiterentwickelt wird [13; 78; 104]. Dieses wird von den Initiatoren als „Reha-QS“ der Rentenversicherung bezeichnet und beteiligt rund 950 Kliniken. Um zielgerichtet qualitätssichernde und -verbessernde Maßnahmen einleiten zu können, erhalten die Verantwortlichen in den Rehabilitationseinrichtungen und beim Rentenversicherungsträger eine detaillierte Berichterstattung der Ergebnisse zur Qualitätssicherung [13; 64].

Erklärte Ziele sind

- eine am Patienten orientierte Qualitätsverbesserung der medizinischen Rehabilitation
- erhöhte Transparenz und Ergebnisorientierung
- Erschließung von Leistungsreserven
- Förderung von einem internen QM
- Sicherstellung von bundesweit vergleichbaren Qualitätsstandards in allen Rehabilitations-Kliniken [78].

Im Peer-Review-Verfahren werden die Rehabilitationsleistungen anhand von Entlassungsberichten und Patiententherapieplänen mittels einer Checkliste qualitätsrelevanter Prozessmerkmale bewertet [13; 78; 105; 106].

Eine weitere Methode der externen Qualitätssicherung ist das Visitations-Verfahren der durch Zufall ausgewählten Kliniken, welches durch den DRV zum einen zur Sicherstellung der Einhaltung definierter Qualitätsstandards, zum anderen zum Austausch im persönlichen Gespräch und zur Beratung der Einrichtung erfolgt. Die Visitationen werden durch ein Visitationsteam aus einem Verwaltungsvisitor und einem ärztlichen Experten der DRV in einem eintägigen Besuch mit Feedback- und Beratungsgespräch und Klinikrundgang durchgeführt. Die Begehung soll einen direkten Eindruck von den strukturellen Bedingungen, Abläufen und der Organisation vermitteln [128].

Im Rahmen der bereits erwähnten wissenschaftlichen Evaluation mit dem Ziel der Weiterentwicklung des Verfahrens wurde 2008 das Projekt „Visit II“ durch den Arbeitsbereich Rehabilitationswissenschaften der Universität Würzburg in Abstimmung mit dem DRV durchgeführt. Durch schriftliche Befragungen und Telefoninterviews wurden von den visitierten Kliniken und den Visitoren Rückmeldungen zu dem Verfahren erbeten [128]. Auf Grund der fehlenden Darstellung der Effektivität des Verfahrens auf die Verbesserung der Qualität der Versorgung und der fehlenden Kontrollgruppe bleibt die Aussage der Studie sehr eingeschränkt.

Die Ergebnisqualität und die Patientenzufriedenheit werden über Patientenbefragung gemessen [78; 105]. Als wesentlicher Unterschied zur „QS-Reha“ der GKV ist die Erfassung des sozialmedizinischen Verlaufs nach abgeschlossener Rehabilitation zu nennen. Dies dient der Darstellung über einen mehrjährigen Zeitraum, wieviele der Rehabilitanden im Erwerbsleben bleiben oder durch Berentung oder Tod ausscheiden [106].

Ein weiterer von dem GKV Verfahren abweichender Bestandteil der „Reha-QS“ der DRV ist das Rehabilitations-Leitlinienprogramm als langfristige Strategie der Qualitätsverbesserung [24; 78;

105; 106]. Die gesetzliche Rentenversicherung fördert seit 1998 zunächst unter dem Namen „BfA-Reha-Leitlinienprogramm" die Projekte zur Erstellung von Leitlinien für die Rehabilitation [24]. Zur Sicherung der Prozessqualität werden die sogenannten „Therapie-Standards", die zuvor Rehabilitations-Leitlinien genannt wurden, eingesetzt [64; 105]. Diese Therapie-Standards wurden von wissenschaftlichen Instituten im Auftrag der DRV erstellt [78]. Im Rahmen der Qualitätssicherung wird die Erfüllung dieser Standards mit einer einrichtungsinternen Auswertung der dokumentierten therapeutischen Leistungen überprüft [64; 106].

Zunächst wurde im Entwicklungsprozess der Leitlinien in einer Literaturrecherche nach Verfahren und Therapien gesucht, die wissenschaftlich belegt sind. In einem zweiten Schritt wurde in einem Soll-Ist-Abgleich der Bedarf zur Erstellung von spezifischen Rehabilitations-Leitlinienthemen analysiert. Dafür wurden die im Entlassungsbrief kodierten therapeutischen Leistungen, die mit Qualitätsmerkmalen verknüpft sind, auf Konformität zur wissenschaftlichen Literatur analysiert. Nach Erstellung der Therapie-Standards in einem multiprofessionellem Gremium wurden diese zunächst im Rahmen von Pilotprojekten in ausgewählten Einrichtungen eingeführt und auf Praxistauglichkeit geprüft, bevor sie flächendeckend in den von der DRV belegten Kliniken eingesetzt wurden [24].

**Eine Evaluation zur Effektivität des Leitlinien-Programms konnte unter den gesichteten Publikationen nicht identifiziert werden.**

#### Qualitätssicherung bei der Rehabilitation der Gesetzlichen Krankenversicherung (GKV)

Für den Bereich der GKV ist die QS in der Rehabilitation seit 1999 gesetzlich verankert [78]. Nachdem in den Jahren 2000 bis 2004 sukzessive Konzeptions- und Erprobungsprojekte für die indikationsspezifischen Abläufe und Verfahrensweisen der QS in den oben genannten Indikationen der medizinischen Rehabilitation durchlaufen wurden, ist das sogenannte „QS-Reha"-Verfahren seit dem Jahr 2005 für alle Indikationen in der Routinephase [65; 104]. Ab 2007 haben sich stationäre Rehabilitationseinrichtungen ferner an Zertifizierungsvereinbarungen durch ein internes QM zu beteiligen (§20 SGB IX) [65; 78].

Das QS-Verfahren der GKV in der stationären medizinischen Rehabilitation [76] ist für alle Rehabilitationseinrichtungen der GKV verpflichtend. Es ist ein einrichtungsübergreifendes und vergleichendes QS-Verfahren, welches in regelmäßigen Zyklen die Qualitätsdimensionen der Struktur-, Prozess- und Ergebnisqualität und die Patientenzufriedenheit messen soll sowie klinikvergleichende Ergebnisrückmeldungen vorsieht [104]. Diese Klinikvergleiche basieren auf einem Risikoadjustierungsmodell, bei dem patientenrelevante Confounder wie Komorbidität, Alter und Geschlecht kontrolliert werden [78]. Träger des Verfahrens ist der GKV-Spitzenverband. Auch hier werden Visitationen bei einer Zufallsstichprobe der Einrichtungen zur Validierung der Klinikangaben durchgeführt. Die Messung der Prozessqualität erfolgte mittels Peer-Review-Verfahren, welches

seit 2011 nicht mehr bei der GKV durchgeführt wird. Zur Messung der Patientenzufriedenheit werden per Fragebogen Angaben eingeholt. Auch zur Messung der Ergebnisqualität werden die Daten über Patientenangaben erhoben [65; 78; 104].

**Zusammenfassend ist festzustellen, dass eine systematische Evaluation im Sinne einer Wirksamkeitsanalyse der Gesamtprogramme für die stationäre Rehabilitation bislang nicht vorliegt [64].** Auch ist bislang ungeklärt, ob die Einführung oder Aufrechterhaltung eines internen QM bewirken kann, dass Rehabilitanden medizinisch besser behandelt werden oder zufriedener mit der Behandlung sind [64].

### 3.1.4 Onkologie

#### 3.1.4.1 Gynäkologische Onkologie

2003 wurde ein Qualitätssicherungsprogramm durch Bildung von zertifizierten, multidisziplinären Brustzentren und einem Benchmarking-Programm gegründet, welches von der Deutschen Krebsgesellschaft und der Deutschen Gesellschaft für Senologie initiiert und deren Datenanalyse von der externen Organisation Westdeutsches Brustzentrum (WBC) durchgeführt wurde [22]. Das Benchmarking-Programm umfasste eine nationale Netzbildung der Zentren mit einem Ranking auf Basis von Qualitätsindikatoren mit dem Ziel, dass die Zentren die Vorgehensweisen der Besten modifizieren und adaptieren. Die Zentren erhielten halbjährlich Feedback-Berichte ihrer eigenen Leistung und der Leistung im Vergleich mit den anderen Teilnehmern in anonymer Form [22; 23; 194].

Teilnehmende Zentren kamen aus Deutschland, Schweiz, Österreich und dem deutschsprachigen Italien. Im Verlauf der Jahre war eine ansteigende Teilnahme von 59 auf 210 Zentren bzw. 5.994 auf 34.678 Brustkrebsfälle zu verzeichnen. Das Ziel des ganzen Programms war die Sicherstellung, dass die Versorgung der Patientinnen mit Brustkrebs leitliniengerecht auf Basis von Qualitätssicherungsmaßnahmen erfolgt [22; 23; 194].

Da Langzeitergebnisse wie Auftreten eines lokalen Rezidivs, Metastasen und Langzeitüberlebensdaten erst mit einer Latenz von mindestens fünf bis zehn Jahren zu erwarten sind, wurde auf kurz- bis mittelfristige Surrogatparameter mittels Messung der Prozessqualität in Form von der Therapie und Diagnostik in Konkordanz mit Leitlinien zurückgegriffen. Die Messung beinhaltete 173 Parameter zusammengefasst in 11 ratenbasierten<sup>11</sup> Qualitätsindikatoren-Sets (Qualitätsindikatoren zum Großteil auf Basis von Leitlinien). Die Qualitätsindikatoren umfassten die präoperativen und operativen Aspekte der brusterhaltenden Therapie, die Erhebung des Hormonrezeptorstatus und der ent-

---

<sup>11</sup> Indikatoren basierend auf einer Rate mit einem definierten Zähler und Nenner (Zähler als Anzahl der Patienten mit dem zu messenden Qualitätsmerkmal, Nenner als Gesamtheit aller Patienten der Zielpopulation, auf die sich der Indikator bezieht) [5; 9]

sprechenden endokrinen Therapie, die neoadjuvante und adjuvante Chemotherapie in Form von Standard- und Studienregimen sowie Strahlentherapie nach brusterhaltender Therapie oder Mastektomie. Die Einführung von Ergebnisqualitätsparametern wie Komplikationen nach Entlassung, krankheitsfreies Überleben, Mortalität, Rezidivraten war zu einem späteren Zeitpunkt geplant [22; 23; 194].

Ergebnisse aller Qualitätsindikatoren verbesserten sich im Verlauf des Beobachtungszeitraums von Beginn des Programms 2003 bis 2010 [194]. Auch in der Subgruppenanalyse zu den Qualitätsindikatoren der Radiotherapie war der gleiche Effekt zu sehen [23].

Da die Teilnahme an dem Programm freiwillig war, ist ein Selektionsbias nicht auszuschließen. Die Studie wurde als eine prospektive, multizentrische Interventionsstudie mit drei Phasen beschrieben; de facto handelt es sich um eine Verlaufsbeobachtung mit Vorher-Nachher-Vergleich. Die Kausalität der Verbesserung der Qualitätsindikatoren ist bei Studienlimitationen in Form von wenig konkreten Angaben zur Art der Auswertung, fehlenden Angaben zu den Begleitvariablen als mögliche Störgrößen (Alter, Komorbidität) bzw. gegebenenfalls eingesetzten Adjustierungsverfahren nicht belegt.

Es werden mehrere andere Faktoren für die Verbesserung der Ergebnisse zum Teil auch von den Autoren der Studie diskutiert [23]. So könnte auch die verbesserte Dokumentation im zeitlichen Verlauf und die zu Beginn des Programms noch nicht vorhandene interdisziplinäre Leitlinie zum Mammakarzinom [112] zu einer Verbesserung der Qualitäts-Kennzahlen geführt haben. Diese wurde in der 1. Auflage erst 2004 publiziert. Des Weiteren wird von den Autoren eingeräumt, dass die geringere Anzahl an Teilnehmern zu Beginn des Programms Zweifel bezüglich der Repräsentativität der anfänglichen Daten aufkommen lassen [23].

#### 3.1.4.2 Pädiatrische Onkologie

Das QM-Programm Schmerztherapie in der pädiatrischen Onkologie (STOP-Programm) [177] wurde 1997 auf Initiative der QM-Gruppe der Deutschen Gesellschaft für pädiatrische Hämatologie und Onkologie ins Leben gerufen. Das Programm umfasste mehrere zeitliche Abschnitte, welche von einer mehrphasigen Evaluationsstudie von 1997 bis 2000 begleitet wurde. Diese sollte sowohl den Vergleich von Abteilungen mit und ohne QM-Programm als auch den Verlauf innerhalb der QM-Gruppe auswerten. Das STOP-Programm beinhaltete Fortbildungsmaßnahmen, Audits und strukturierte halbjährliche Feedbacks. Das Ziel des Programms war die Verbesserung der Schmerzkontrolle in der pädiatrischen Onkologie mit Evaluation der Maßnahmen [199].

Die prospektiv angelegte Studie zeigte eine signifikante Reduktion der Häufigkeit von starken Schmerzen innerhalb der QM-Gruppe im Verlauf. Im Vergleich QM-Gruppe versus non-QM-Gruppe war eine vermehrte Anwendung von empfohlenen Instrumenten zur Schmerzerhebung und -dokumentation der Ärzte zugunsten der QM-Gruppe zu verzeichnen. Sowohl im Verbrauch von Opioiden und Non-Opioiden innerhalb der QM-Gruppe im Verlauf als auch im Vergleich der QM-Gruppe versus non-QM-Gruppe in der Selbsteinschätzung von Patienten und Eltern bezüglich der Schmerzkontrolle stellte sich jedoch kein signifikanter Unterschied dar. In der QM-Gruppe wurde sogar signifikant mehr Unzufriedenheit mit der Schmerzkontrolle bei schmerzhaften Prozeduren als in der Non-QM-Gruppe angegeben [199].

Die Vergleichbarkeit der Gruppen QM versus non-QM ist auf Grund der fehlenden detaillierten Angaben der Population der non-QM-Abteilungen nur sehr eingeschränkt möglich, sodass sich mögliche Confounder der externen Beurteilung entziehen. Der Vergleich der Endpunkte Schmerzkontrolle und Schmerzmittelverbrauch erfolgte lediglich innerhalb der QM-Gruppe im zeitlichen Verlauf ohne Kontrollgruppe, sodass die Wirksamkeit des STOP-Programms auch in Anbetracht der dargestellten Ergebnisse nicht belegt ist.

**Tabelle 4 Übersicht der evaluierten Verfahren in Deutschland**

<sup>1</sup> Struktur-, Prozess-, Ergebnisqualität nach [52; 53]

<sup>2</sup> ja:↑ , nein:↓ inkonsistent:↔

<sup>3</sup> niedrig, moderat, hoch, sehr hoch (siehe Kapitel 2.1.2 Bewertung und Evidenzdarstellung, Anhang 2 ausführliche Studiendarstellung, Anhang 4 und Anhang 5 Kriterien Verzerrungsrisiko)

QS/QM-Verfahren	Beginn (-Ende)	Ziel	Ergebnisse der Evaluierung - Qualitätsdimension <sup>1</sup> - positiver Effekt <sup>2</sup> - Verzerrungsrisiko der Ergebnisse <sup>3</sup>	Referenzen
Peer-Review – Helios Kliniken GmbH (Pilotprojekt)	2000-2008	Optimierung der Behandlungsqualität i.S. eines kontinuierlichen Verbesserungsprozesses	- Ergebnisqualität - positiver Effekt: ↑ - Verzerrungsrisiko: hoch	[31; 111; 129]
Gemeinnütziger Verein Initiative Qualitätsmedizin (IQM) –  aus Peer-Review-Verfahren der Helios-Kliniken entsprungen	2008	Analyse der Ergebnisqualität, Indikatormessungen aus Routinedaten, Transparenz durch Veröffentlichung der Ergebnisse, aktives Qualitätsmanagement durch eine kollegiale Diskussion auf Augenhöhe	- Ergebnisqualität - positiver Effekt: ↑ - Verzerrungsrisiko: hoch	[3; 60; 147; 148]
TraumaRegister und TraumaNetzwerk der Deutschen Gesellschaft für Unfallchirurgie QM-System	1999	Prozessanalyse der teilnehmenden Krankenhäuser für Identifizierung von Schwachpunkten und ggf. Erstellung von verbesserten Behandlungskonzepten	- Ergebnisqualität und Prozessqualität - positiver Effekt: ↑ - Verzerrungsrisiko: sehr hoch	[151-153]
Schmerztherapie in der pädiatrischen Onkologie (STOP)-QM-Programm  Initiator: QM-Gruppe der Deutschen Gesellschaft für pädiatrische Hämatologie und Onkologie	1997-2000	Verbesserung der Schmerzkontrolle in der pädiatrischen Onkologie und Evaluation	- Ergebnisqualität und Prozessqualität - positiver Effekt: ↔ - Verzerrungsrisiko: hoch	[199]

QS/QM-Verfahren	Beginn (-Ende)	Ziel	Ergebnisse der Evaluierung - Qualitätsdimension <sup>1</sup> - positiver Effekt <sup>2</sup> - Verzerrungsrisiko der Ergebnisse <sup>3</sup>	Referenzen
<p>QiN – Qualität in der Nephrologie externes QM-System</p> <p>Träger: Klinik für Nephrologie und Institut für Gesundheitsökonomie und Klinische Epidemiologie der Universität zu Köln in Kooperation mit dem Kuratorium für Dialyse und Nierentransplantation e.V.</p>	1999	Verbesserung und Sicherung der Qualität der KfH Dialyseeinrichtungen	<ul style="list-style-type: none"> <li>- Prozessqualität</li> <li>- positiver Effekt: ↔</li> <li>- Verzerrungsrisiko: sehr hoch</li> </ul>	[175; 176]
<p>Qualitätssicherungsprogramm durch Zertifizierung und Benchmarking von Brustzentren</p> <p>Träger: Deutsche Krebsgesellschaft und Deutsche Gesellschaft für Senologie – externe Organisation durch Westdeutsches Brustzentrum (WBC)</p>	2002	Sicherstellung leitliniengerechter Versorgung auf Basis von Qualitätssicherungsmaßnahmen	<ul style="list-style-type: none"> <li>- Prozessqualität</li> <li>- positiver Effekt: ↑</li> <li>- Verzerrungsrisiko: sehr hoch</li> </ul>	[22; 23; 194]
<p>Disease Management Programm Brustkrebs</p> <p>Gesetzlich verankert auf Grundlage von §137 SGB V</p>	2002	Verbesserung des Behandlungsablaufs und der Qualität der medizinischen Versorgung chronisch Kranker durch eine dem aktuellen Stand der Wissenschaft entsprechenden Versorgung	<ul style="list-style-type: none"> <li>- Prozess- und Strukturqualität</li> <li>- positiver Effekt: ↔</li> <li>- Verzerrungsrisiko: hoch</li> </ul>	[14; 51]
<p>„Reha-QS“ Qualitätssicherungsprogramm in der Rehabilitation</p> <p>Träger: DRV</p>	1994	Qualitätsverbesserung der medizinischen Rehabilitation mit Patientenorientierung, erhöhter Transparenz und Ergebnisorientierung	keine aussagekräftigen Evaluationsergebnisse identifizierbar	
<p>„QS-Reha“ Qualitätssicherungsprogramme in der Rehabilitation</p>	1999	Verbesserung der Struktur-, Prozess- und Ergebnisqualität und der Patientenzufriedenheit	keine aussagekräftigen Evaluationsergebnisse identifizierbar	

QS/QM-Verfahren	Beginn (-Ende)	Ziel	Ergebnisse der Evaluierung - Qualitätsdimension <sup>1</sup> - positiver Effekt <sup>2</sup> - Verzerrungsrisiko der Ergebnisse <sup>3</sup>	Referenzen
Träger: GKV				
Externe QS nach §137 SGB V	2001	Qualitätsverbesserungen durch Erkenntnisse über Qualitätsdefizite, Unterstützung beim einrichtungsinternen Qualitätsmanagement, Sicherstellung der angemessenen Indikationsstellung, Strukturqualität und Ergebnisqualität	keine aussagekräftigen Evaluationsergebnisse identifizierbar	
Initiative Schmerzfreie Klinik mit Zertifikat „QM Akutschmerztherapie“ (durch TÜV Rheinland vergeben)  Träger: Chirurgische Arbeitsgemeinschaft Akutschmerz zusammen mit TÜV Rheinland	2006	Verbesserung des Umgangs mit akuten Schmerzen in deutschen Kliniken	- Ergebnisqualität - positiver Effekt: ↑ - Verzerrungsrisiko: hoch	[16; 116]
Qualitätsverbesserung in der postoperativen Schmerztherapie (QUIPS)  Träger: BMG bis 2007, ab 2007 Deutsche Gesellschaft für Anästhesiologie und Intensivmedizin (DGAI), Berufsverband Deutscher Anästhesisten (BDA)	2006	Verbesserungsmaßnahmen in der postoperativen Schmerztherapie, gezielter Einsatz von Ressourcen und Evaluierung von Optimierungsmaßnahmen	- Ergebnisqualität - positiver Effekt: ↑ - Verzerrungsrisiko: hoch	[116; 123; 143]
QM in der Krankenhaushygiene	2001	Vermeidung bzw. Reduktion von nosokomialen Ereignissen	- Ergebnisqualität - positiver Effekt: ↓ - Verzerrungsrisiko: moderat	[70]

## 3.2 International

Die detaillierte Beschreibung der Einzelstudien mit Darstellung relevanter Daten in Form von Evidenztabellen ist im Anhang 3 dargestellt.

Eine allgemeine Übersicht über die internationalen Evaluationsstudien findet sich in Tabelle 5 (Seite 62).

Im internationalen Bereich haben die Joint Commission on Accreditation of Healthcare Organizations und andere Organisationen in den 70er Jahren die ersten formalen QS-Programme initiiert [149]. Seitdem sind - wie in Deutschland - in vielen Ländern die Krankenhäuser durch nationale Regelungen zur QS verpflichtet worden, z.B. in Großbritannien seit 1990, in Frankreich seit 1991, in Schweden und Niederlanden seit 1993 und in Norwegen seit 1994 [170].

### 3.2.1 Länderübergreifend

In einem systematischen Review zur Effektivität qualitätssichernder Verfahren wurden die eingesetzten Maßnahmen in 18 Ländern anhand von Publikationen begutachtet [170]. Es wurden Länder mit bereits etablierten Qualitätssicherungsmaßnahmen einbezogen wie Australien, USA, Kanada, die vier Länder Skandinaviens, Niederlande, Belgien, Großbritannien, Frankreich, Schweiz, Österreich, Italien, Portugal, Spanien, Japan und Neuseeland. Insgesamt ergaben sich aus der Analyse und Recherche der Datenlage eine nur limitierte und uneinheitliche Evidenz für die Effektivität qualitätssichernder Verfahren. Belege für positive Auswirkungen auf Endpunkte der Versorgungsqualität für Patienten konnten nicht gefunden werden. Letztendlich fehlt in den Gesundheitssystemen das Wissen, welche Ansätze für welche Form der Verbesserung in welchem Kontext angemessen sind [170]. Limitierend ist ein moderates Verzerrungsrisiko dieser Übersichtsarbeit, da die Autoren keine festen Ein- und Ausschlusskriterien formulierten und wenig konkrete Darstellungen der Details, des Studiendesigns und der Qualität der Einzelstudien gaben. Eine Differenzierung zwischen ambulantem und stationärem Setting wurde nicht vorgenommen.

Conry et al. erstellten 2012 einen systematischen Review zur Untersuchung der Effektivität von Interventionen zur Verbesserung der Versorgungsqualität in einzelnen Abteilungen oder für viele Krankenhäuser in Form von Programmen mit einem Recherchezeitraum von 2000 bis 2010. Es wurden Studien mit kleineren qualitätsfördernden Maßnahmen wie Zusammenlegung von Stationen, Befragungen von einzelnen Patientengruppen oder Workshops bis hin zu übergeordneten Programmen mit Erstellung von Behandlungspfaden, Schulungen und Feedback in mehreren Kliniken in den Review eingeschlossen. Von 120 potentiellen Volltexten wurden 20 Studien als relevant

eingeschätzt und für die Auswertung berücksichtigt. Es wurden keine Angaben gemacht, für welche Länder gesucht wurde oder wo die Studien erfolgten.

Nach eingeschränkter Betrachtung der Studien, die übergeordnete QS/QM-Programme thematisierten und damit den Einschlusskriterien der vorliegenden Arbeit entsprachen, zeigten sich zum Teil positive Auswirkungen mit Verbesserungen auf unterschiedliche Endpunkte wie Patientenzufriedenheit und Prozessparameter. Die Einzelstudien waren jedoch überwiegend Vorher-Nachher-Vergleiche und wiesen nur moderate bis niedrige Qualität auf. Der einzige RCT mit hoher Studienqualität zeigte keine eindeutige Wirksamkeit der Intervention, die eine Schulung des Personals und die Förderung der Aufnahme einer Begleitperson bei der Entbindung in zehn Krankenhäusern hinsichtlich vermehrter Aufnahmen beinhaltete [42].

In dem bereits in Kapitel 3.1.1 dieser Arbeit beschriebenen Gutachten des Kompetenz-Centrum Qualitätssicherung/Qualitätsmanagement (KCQ) im Auftrag des GKV Spitzenverbands zur Aufarbeitung des internen QM in der medizinischen Versorgung wurde auch nach internationalen Evaluationsstudien gesucht. In der Zusammenfassung der 49 relevanten Publikationen zeigte sich eine sehr heterogene Datenlage mit überwiegend Vorher-Nachher-Untersuchungen aus unterschiedlichsten Gesundheitssystemen, basierend auf unterschiedlichen QM-Ansätzen mit verschiedenen Zielparametern. Die meisten Studien waren von limitierter methodischer Güte und daher mehrheitlich nicht zur Ableitung von kausalen Zusammenhängen geeignet. Die Studien mit höherwertigem Design konnten keine konsistenten Ergebnisse zum Patientennutzen ableiten, es war fast durchgehend - trotz verbesserter Prozessqualität - keine Optimierung patientenrelevanter Endpunkte festzustellen [121].

### 3.2.2 Frankreich

Eine prospektive, unkontrollierte Vorher-Nachher-Studie in Frankreich untersuchte die Effektivität eines internen QM-Programms, welches die Erhebung von Qualitätskriterien, Feedback, verstärkte Leitlinienanwendung, die Einführung von standardisierten Arbeitsanweisungen (Standard Operating Procedures (SOP)) und Schulungen beinhaltete. Es wurde das Ziel verfolgt, in der Notaufnahme eines Traumazentrums der Orthopädie, Unfallchirurgie und plastischen Chirurgie das Bewusstsein beim Personal für präventive und verbessernde qualitätsfördernden Maßnahmen zu schaffen. Es wurde das Auftreten von Diagnose-, Behandlungs- oder Organisationsfehler, die zu einem kritischen, unerwünschten Ereignis führen könnten, anhand von jeweils rund 2000 Krankenakten vor und nach Einführung des QM im Jahr 1993 ausgewertet. Es zeigte sich eine signifikante Verbesserung der drei gemeinsam betrachteten Kategorien von 204 Ereignissen (9,9%) vor der Einführung auf 71 Ereignissen (5,9%) nach der Einführung des QS-Programms [69].

Auf Grund der fehlenden Kontrollgruppe sind die Ergebnisse nur mit Einschränkungen verwertbar.

### 3.2.3 Spanien

In einer randomisiert, kontrollierten Studie in zehn allgemeinmedizinischen Gesundheitszentren wurden drei Gruppen vor und sechs Monate nach der Intervention miteinander verglichen [33]. Eine Gruppe hatte nach einwöchiger QM-Fortbildung ein eigenes, selbstorganisiertes internes QM mit dem Ziel der Verbesserung der Behandlung von unkomplizierten, respiratorischen Infekten entsprechend existierender Leitlinien-Empfehlungen aufgebaut. Die zweite Gruppe erhielt über eine externe QS Feedbackberichte zu ihren Ergebnissen und die dritte Gruppe war ohne weitere QS/QM-Maßnahmen. Alle drei Gruppen wussten nichts über die Ziele der Studie [33].

Es zeigte sich eine signifikante Verbesserung der Leitlinien-Adhärenz für alle festgelegten Kriterien in der Behandlung von unkomplizierten respiratorischen Infekten in der Gruppe mit dem internen QM. In den anderen beiden Gruppen war dieser Effekt weniger ausgeprägt oder gar nicht zu sehen [33].

Limitierend sind die sehr kleine Fallzahl in der Gruppe mit QM, die sehr geringen Angaben zu den konkreten QM-Maßnahmen und zur Studienpopulation der Patienten und Ärzte, was die Beurteilung der Repräsentativität und der Vergleichbarkeit der Gruppen sehr einschränkt. Ferner werden eingeschränkte Angaben zur Art der Datenauswertung oder zum Randomisierungsverfahren gegeben.

### 3.2.4 USA

In einer Studie wurde - im Auftrag der Agency for Healthcare Research and Quality (AHRQ) - über drei Jahre der Einfluss von vier QM-Programmen in fünf Bundesstaaten in den USA auf Krankenhäuser mit freiwilliger QM-Anwendung (n = ca. 152) versus Krankenhäuser ohne QM (n = ca. 195) verglichen. Medicare [34], eine der staatlichen Krankenversicherungen in den USA, beauftragte externe Institutionen mit der Durchführung der QM-Programme [171]. Diese beinhalteten verschiedene Methoden wie Schulungsmaterialien, Datenerhebung, Feedbackberichte zu Qualitätsindikator-Ergebnissen, Hilfe bei der Erstellung von Behandlungspfaden und SOPs. Die Erhebung erfolgte retrospektiv auf Basis von Krankenakten mit den Endpunkten der Prozessqualität mittels 15 Qualitätsindikatoren (basierend auf Leitlinien) zu fünf Krankheiten: Herzinfarkt, Pneumonie, Vorhofflimmern, Herzinsuffizienz, Schlaganfall. Als Studiendesign wurde eine kontrollierte Vor-Nachher-Studie mit initialer Erhebung 1998 vor Beginn der Einführung und einer Follow-up-Erhebung im Jahr 2000 bis 2001. Die teilnehmenden Bundesstaaten waren Utah, Nevada, Washington, Maryland und New York.

Es zeigte sich kein signifikanter Unterschied in der Verbesserung bei 14 von 15 Qualitätsindikatoren zwischen den Gruppen. Einziger Unterschied war bei dem Qualitätsindikator „Screening für Pneumokokken-Impfung“ zugunsten der teilnehmenden Kliniken zu sehen. Beide Gruppen wiesen

bei 11 der 15 Qualitätsindikatoren zu Herzinfarkt, Pneumonie, Vorhofflimmern, Herzinsuffizienz und Schlaganfall Verbesserungen auf [171].

Ein moderates Verzerrungsrisiko ist in dem retrospektiven Studiendesign der Vorher-Nachher-Studie zu sehen, was die Aussagen der Studie abschwächt.

#### National Surgical Quality Improvement Program (NSQIP)

Das National Surgical Quality Improvement Program (NSQIP) [6] ist ein chirurgisches Register des American College of Surgeons (ACS) mit dem Ziel, risikoadjustierte Ergebnisse von chirurgischen Interventionen zu messen und für die jeweilige Klinik und im Vergleich mit anderen Kliniken zu analysieren. Das National Surgical Quality Improvement Program wurde ursprünglich in den 90er Jahren im Auftrag der Regierung zur Qualitätsverbesserung in den 133 Krankenhäusern der US-Kriegsveteranen gegründet, da in diesen damals eine hohe Rate an Mortalität und chirurgischen Komplikationen beobachtet wurden. 2004 folgte die Übernahme und Neu-Gründung durch das American College of Surgeons. 350 Krankenhäuser nehmen in den USA und im Nahen Osten an dem Programm teil [68].

Im Mai 2008 wurde eine Unter-Initiative des QS-Programms der „National Surgical Quality Improvement Program (NSQIP)“ mit Zusammenschluss von zehn Krankenhäusern in Tennessee gegründet. Guillaumonegui et al. haben in einer unkontrollierten Vorher-Nachher-Studie mit Vergleich der Datenerhebung zu Beginn und Verlauf die Auswirkungen des Programms auf diese Subgruppe untersucht. Dafür wurden die Registerdaten mit risikoadjustierten Prozess- und Ergebnisparametern aus dem Zusammenschluss der zehn Krankenhäuser in Tennessee erhoben, analysiert und anschließend als halbjährliches Feedback an die Kliniken zurückgespiegelt [85]. Im Vergleich zwischen Anfang 2009 und dem Verlauf bis Ende 2010 zeigte sich in den rund 30.000 ausgewerteten chirurgischen Fällen in den Daten von 20 Kategorien zu postoperativen Komplikationen eine signifikante Verbesserung in fünf Kategorien als primäre Endpunkte (oberflächliche Wundinfektion, Beatmung länger als 48 Stunden, Prothesen-/Transplantatversagen, akutes Nierenversagen, Wundheilungsstörung). Außerdem stellte sich eine Verschlechterung in drei Kategorien (tiefe Venenthrombose/Thrombophlebitis Pneumonie, Harnwegsinfekte) und kein signifikanter Unterschied bezüglich des sekundären Endpunkts der 30-Tages-Mortalität dar. Die Autoren erwogen mehrere Ursachen für die Ergebnisse wie z.B. das Durchführen von Peer-Reviews oder anderen QS-Maßnahmen in Folge des Feedbacks durch die Registerdaten. Des Weiteren wird ein möglicher Hawthorne-Effekt<sup>12</sup> durch verstärkte Bemühungen der Studienteilnehmer in Folge einer Selbstreflexion unter der Studien-Beobachtung von den Autoren in Betracht gezogen [85]. Es erfolgte eine umfassende Berücksichtigung von möglichen Confoundern (Risikofaktoren, soziodemographische Daten etc.)

---

<sup>12</sup> Wirkung allein durch die Tatsache, dass Teilnehmer an Beobachtungsstudien vermehrt Aufmerksamkeit erfahren [114]

mit transparenter Darstellung der vergleichbaren zwei Gruppen. Leider wurden kaum Angaben zu den Maßnahmen, die die Kliniken als Konsequenz aus den Daten zogen, dargestellt.

In einer weiteren Studie zu NSQIP wurden die aus dem Register resultierenden QS-Maßnahmen in einer chirurgischen Abteilung einer Mayo-Klinik dargestellt [35]. In dieser Studie wurden nach Analyse der Register-Daten mögliche kausale Faktoren für postoperative Wundinfektionen an die Kliniken zurückgespiegelt, die die Entwicklung mehrerer qualitätsverbessernder Maßnahmen auslösten. Diese basierten auf einer Literaturanalyse durch ein multidisziplinäres Team. Die Autoren analysierten in einer Vorher-Nachher-Studie den Zeitraum ein Jahr (2009 und 2010) vor Einführung dieser Maßnahmen in der chirurgischen Abteilung und ein Jahr danach im Jahr 2011. Als Ergebnis war eine signifikante Reduktion der postoperativen Wundinfektionen von 9,8% auf 4,0% bei 531 Patienten vor und 198 Patienten nach Einführung der Maßnahmen unter gegebener Vergleichbarkeit der Gruppen zu verzeichnen [35]. Auf Grund des Studiendesign ohne Kontrollgruppe ist ein kausaler Zusammenhang nicht sicher nachweisbar, aber unter Berücksichtigung der Effektstärke und der methodisch guten Auswertung mit Kontrolle der potentiellen Confounder sehr in Betracht zu ziehen.

### 3.2.5 China

In China wurde in einer prospektiven, unkontrollierten Vorher-Nachher-Studie in einem Lehr-Krankenhaus mit 3000 Betten das Auftreten von Dekubitalgeschwüren vor und sechs Monate nach Einführung eines spezifischen und sehr umfangreichen QM zu Dekubitalgeschwüren miteinander verglichen. Das QM umfasste einen multidimensionalen Ansatz mit Erhebung, Rückmeldung, Audit, Fortbildung, Diskussion in Qualitäts-Konferenzen und in multidisziplinären Wundteams und Erstellung evidenzbasierter Handlungsanweisungen für präventive Maßnahmen. Zwischen der Gruppe mit 2.913 Patienten vorher und 3.019 Patienten nachher ergab sich insgesamt eine nicht signifikante Reduktion der Patienten mit Geschwüren von 1,8 auf 1,4 %, die auch in weiteren Subgruppen nicht signifikant waren [41].

Bei fehlender Kontrolle, jedoch bestehender Vergleichbarkeit der Gruppen sind die Ergebnisse mit Einschränkung verwertbar.

**Tabelle 5 Übersicht der evaluierten Verfahren international**

<sup>1</sup> Struktur-, Prozess-, Ergebnisqualität nach [52; 53]

<sup>2</sup> ja:↑, nein:↓, inkonsistent:↔

<sup>3</sup> niedrig, moderat, hoch, sehr hoch (siehe Kapitel 2.1.2 Bewertung und Evidenzdarstellung, Anhang 3 ausführliche Studiendarstellung, Anhang 4 und Anhang 5 Kriterien Verzerrungsrisiko)

<b>QS/QM-Verfahren</b>	<b>Beginn (-Ende)</b>	<b>Ziel</b>	<b>Ergebnisse der Evaluierung</b> - <b>Qualitätsdimension</b> <sup>1</sup> - <b>positiver Effekt</b> <sup>2</sup> - <b>Verzerrungsrisiko der Ergebnisse</b> <sup>3</sup>	<b>Referenzen</b>
National Surgical Quality Improvement Program (NSQIP) der American College of Surgeons in USA	1994	Transparente Darstellung der Ergebnisqualität für Benchmarking, Verbesserung der chirurgischen klinischen Versorgung	- Ergebnisqualität - positiver Effekt:↔ - Verzerrungsrisiko: moderat	[35; 68; 85]
QM zur Prävention von Dekubitalgeschwüren in China	2009	Reduktion der Rate an Dekubitalgeschwüren im Krankenhaus	- Ergebnisqualität - positiver Effekt: ↓ - Verzerrungsrisiko: moderat	[41]
Internes QM in Spanien, Murcia	k.A.	Verbesserung der Behandlungsqualität durch verstärkte Leitlinien-Adhärenz	- Prozessqualität - positiver Effekt:↑ - Verzerrungsrisiko: moderat	[33]
Staatliches QM-Programm in den USA in den Bundesstaaten Utah, Nevada, Washington, Maryland und New York  Träger: Medicare (staatliche Krankenversicherung in USA)	1999	Verbesserung aller Versorgungsqualitäten in Krankenhäusern	- Prozessqualität - positiver Effekt: ↓ - Verzerrungsrisiko: moderat	[171]
Internes QM in Frankreich	1993	Reduktion von Diagnose-, Behandlungs- oder Organisationsfehlern, die zu einem kritischen, unerwünschten Ereignis führen könnten	- Prozessqualität - positiver Effekt: ↑ - Verzerrungsrisiko: moderat	[69]

### 3.3 Effekte von Instrumenten für Qualitätsmanagement/Qualitätssicherung

Es existiert eine Reihe von qualitätssichernden Maßnahmen, die im Rahmen eines QM oder als singuläre Maßnahme eingesetzt werden können. Diese sind z.B. Leitlinien, Zertifizierungen, Mortalitäts- und Morbiditätskonferenzen, klinische Behandlungspfade oder Projekte zur Patientensicherheit. Auch diese Maßnahmen erfordern einen hohen Einsatz an Zeit und Aufwand der konzeptionellen Erarbeitung und Bemühungen um Implementierung. Vor diesem Hintergrund stellt sich auch für diese Instrumente der QS die Frage nach der Wirksamkeit auf die Verbesserung der Patientenversorgung.

Im Nachfolgenden werden einige wichtige Verfahren im Hinblick auf ihre Wirksamkeit und ihren Stellenwert dargestellt. Diese Studien wurden neben den Publikationen zu der primären Fragestellung aus der Treffermenge der systematischen Recherche selektiert. Das Auswahlkriterium war hierbei ebenfalls die Thematik der Evaluation zur Wirksamkeit der Verfahren auf die Patientenversorgung. Diese Auswahl ist als selektive Literaturrecherche zu betrachten und erhebt daher nicht den Anspruch auf Vollständigkeit (siehe auch Kapitel 2.1.1 Studienauswahl).

#### 3.3.1 Leitlinien

Für die objektive Beurteilung der Qualität der medizinischen Versorgung ist eine klare Vorstellung von „guter Qualität“ als „Soll-Zustand“ notwendig [109]. Diese Empfehlungen werden in hochwertigen Leitlinien nach den in Kapitel 1.1 genannten, anerkannten methodischen Kriterien entwickelt [12; 101; 109].

Die Erwartungen in evidenzbasierte Leitlinien sind hoch, da die dort geltenden Standards und Qualitätsziele in die verschiedensten Verfahren des QM/QS wie Audit, Zertifizierung, klinische Behandlungspfade, Qualitätsindikatoren etc. eingehen [170]. Leitlinien dienen als modernes Informationsmanagement für die ärztliche Tätigkeit, als Referenz für Qualitätsziele und als mögliche Grundlage für die Erstellung von innerklinischen SOPs, Algorithmen und Behandlungspfaden [133].

Handlungsleitlinien und Richtlinien haben aus diesem Grund im Rahmen der Qualitätssicherung eine bedeutende Rolle eingenommen, die auch innerhalb der Selbstverwaltung durch die Erstellung von Richtlinien und Leitlinien des wissenschaftlichen Beirats der BÄK oder von Handlungsempfehlungen der Arzneimittelkommission der Ärzteschaft (AkdÄ) zu fachübergreifenden Themen seit langem etabliert sind [27].

### Wirksamkeit von Leitlinien auf die Qualität der Versorgung in den Niederlanden

In einem systematischen Review zur Wirksamkeit von evidenzbasierten klinischen Praxis-Leitlinien auf die Qualität der Versorgung in den Niederlanden zeigte sich, dass in 17 von 19 Studien eine signifikante Verbesserung der Prozess- und Strukturqualität gezeigt werden konnte. Signifikante Effekte der patientenrelevanten Ergebnisqualität waren in sechs von neun Studien zu verzeichnen, jedoch weitaus weniger deutlich ausgeprägt. Die Effekte waren innerhalb einer Leitlinie zwischen den Empfehlungen sehr variabel, sodass die Autoren empfahlen, Implementierungsstrategien und Barrierenanalysen nicht für die Leitlinie im Ganzen, sondern gegebenenfalls für einzelne Empfehlungen in Erwägung zu ziehen [119].

Die Evidenz für die Frage, welche Implementierungs- und Disseminierungsstrategie für Leitlinien die effektivste und effizienteste ist, ist noch unzureichend vorhanden, sodass die sinnvollste Methode für die unterschiedlichen klinischen Kontexte in Abhängigkeit der Strukturen und Ressourcen variieren können und das Verhältnis von Nutzen und Aufwand im Einzelfall abgewogen werden sollte [67; 83].

Unbestritten dagegen ist, dass Leitlinien auf der einen Seite eine der wichtigsten Implementierungsmaßnahmen mit Umsetzung der Empfehlungen in die tägliche Praxis darstellen, wenn sie in ein gut funktionierendes QM/QS Programm integriert werden. Auf der anderen Seite dienen Leitlinien als Grundlage für QM/QS-Verfahren durch ihre Funktion als Quelle aufbereiteter Evidenz [57; 58; 107; 133; 161; 162].

### Leitlinien-Implementierung durch das Qualitätssicherungsprogramm der deutschen Arbeitsgemeinschaft Gynäkologische Onkologie – Thema Ovarialkarzinom

Das Qualitätssicherungsprogramm der deutschen Arbeitsgemeinschaft Gynäkologische Onkologie (AGO) Organkommission OVAR (QS-OVAR) hat 1999 ein Aktionsprogramm zur Verbesserung der Versorgung von Patienten mit Ovarialkarzinomen unter Berücksichtigung von Leitlinien gestartet [58]. Dieses umfasste die Herausgabe von Leitlinien, Projekte zur Leitlinien-Implementierung, Symposien und Untersuchungen zur Versorgungsqualität. In diesem Rahmen wurde die Versorgungsstruktur im Jahr 2001 mit der Frage untersucht, inwieweit die empfohlenen Standards der aktuellen Leitlinien in der Realität umgesetzt wurden. Alle gynäkologischen Abteilungen wurden angeschrieben. 481 (42,8%) der Kliniken beteiligten sich an der Analyse mit 476 Patientinnen. Die Autoren-Gruppe kam zu dem Ergebnis, dass nur 58% der Patienten die standardgerechte Behandlung erhielt, obwohl der Effekt dieser empfohlenen Therapiestandards sowohl beim frühen als auch beim fortgeschrittenen Karzinom mit einem signifikanten Einfluss auf das Überleben nachgewiesen wurde [58].

Ein hausinternes QS/QM-Programm mit Behandlungs-Standards auf Basis der nationalen Leitlinien (SOPs) für die Therapie von Ovarialkarzinomen im frühen Stadium FIGO<sup>13</sup> IA-IIIa mit jährlichen Qualitätsberichten und Qualitätskonferenzen wurde im Jahr 2001 in einer deutschen Klinik eingeführt. In einer retrospektiven Studie im Vorher-Nacher-Design sollte durch die Auswertung von Daten eines eigenen Tumorregisters von 1997 bis 2007 mit Ergänzung von Daten aus den Krankenakten evaluiert werden, wie mittels Einführung von internen Standards und QS die Adhärenz zu Leitlinien-Empfehlungen im Vergleich vor und nach der Einführung beeinflusst wurde [107]. 117 operierte Patienten mit einem Alters-Median von 57,5 Jahren wurden in die Studie eingeschlossen (27% vor und 83% nach Einführung des Programms). Die Leitlinien-Adhärenz mit Durchführung einer Operation nach Standard war nach Einführung des Programms signifikant höher. Die Leitlinien-Adhärenz bezüglich Chemotherapie veränderte sich nicht durch das Programm, die Abweichungen von Leitlinien-Empfehlungen beinhalteten sowohl Über- als auch Untertherapie. Die Überlebenszeitanalysen waren - wie auch bei einer der nachfolgenden Studien zum Endometrium-Karzinom [57] - zu niedrig für adäquate Analysen, zeigten jedoch einen Trend für eine verbesserte Gesamtmortalitätsrate in der Gruppe nach Einführung der Standardtherapie gemäß Leitlinien [107]. Die geringe Fallzahl und das Studiendesign sind deutlich limitierend für die Aussage der Studie.

In einer weiteren Auswertung dieses Registers für dieselbe Klinik für Behandlungsfälle mit fortgeschrittenem Ovarialkarzinom im Stadium FIGO IIB-IV von 1997 bis 2008 wurden ebenfalls die Gruppen vor und nach Einführung des QM-Programms im Jahr 2001 bezüglich der Leitlinien-Adhärenz untersucht [89]. Das QM-Programm beinhaltete die Einführung des Tumorregisters, spezielle Schulungen, routinemäßige interdisziplinäre Konsultationen und jährliche Qualitätskonferenzen. 438 Patienten wurden in der Kohorte untersucht, davon nur 57 vor 2001. Die in den Leitlinien empfohlene komplette Resektion stieg von 33% auf 62% im Verlauf an, das Gesamtüberleben besserte sich signifikant von 26 Monaten vor Einführung des Programms auf nachfolgend 45 Monate [89]. Da sich in den Gruppen der komplett resezierten Fälle kein Unterschied zwischen den Studienintervallen ergab, prognostische Merkmale wie Alter und FIGO-Stadium sich nicht günstiger im Verlauf verhielten und auch keine relevante Änderung der Chemotherapie-Regime auftraten, kann die verbesserte OP-Technik im Rahmen des QM-Programms als ein relevanter Faktor für die Verbesserung der Überlebenszeit interpretiert werden. Die exakte Darstellung von Daten der Patientencharakteristiken im Vergleich zwischen den Gruppen fehlt allerdings. Auch die geringe Fallzahl vor Einführung des Programms ist limitierend für die Interpretation der Ergebnisse.

---

<sup>13</sup> FIGO-Klassifikation: Stadieneinteilung gynäkologischer Tumoren von der Fédération Internationale de Gynécologie et d'Obstétrique (FIGO)

### Leitlinien-Implementierung durch Qualitätssicherungsprogramm der deutschen Arbeitsgemeinschaft Gynäkologische Onkologie – Thema Endometrium-Karzinom

In derselben Klinik wurde die gleiche Untersuchung wie die Studie zum Ovarialkarzinom [107] ebenfalls von 1997 bis 2007 zum Einfluss des QS/QM-Programms auf die chirurgischen Behandlungsverfahren des Endometrium-Karzinoms durchgeführt [57]. Das hausinterne QS/QM-Programm mit Behandlungs-Standards (SOPs) auf Basis der nationalen Leitlinien für chirurgische Behandlungsverfahren des Endometrium-Karzinoms mit jährlichen Qualitätsberichten und Qualitätskonferenzen wurde im Jahr 2001 eingeführt [57]. In dieser retrospektiven Studie im Vorher-Nacher-Design mit Auswertung von Daten des eigenen Tumorregisters von 1997 bis 2007 mit Ergänzung von Daten aus den Krankenakten wurde die Adhärenz zu Leitlinien-Empfehlungen im Vergleich vor und nach der Einführung evaluiert. Eingeschlossen wurden in die Auswertung alle operierten Patienten mit FIGO I-III dieser Klinik mit einer Population von 206 Patienten (66 vor und 140 nach dem Jahr 2001). Mögliche Confounder, die die Behandlungsergebnisse beeinflussen könnten - wie Alter, Allgemeinzustand, Tumorstadium oder Histologie - wurden mittels Adjustierungsverfahren berücksichtigt. Übertherapie analog der hausinternen Standards konnte signifikant nach der Implementierung der Maßnahmen reduziert werden (21% versus 7%). Im Gegensatz dazu ergaben sich bezüglich der Untertherapie keine signifikanten Unterschiede (17% versus 22%). Insgesamt erfolgte die Behandlung analog der Standards sowohl vor als auch nach Einführung ohne signifikante Verbesserung der Adhärenz mit 62% vorher und 71 % nachher [57].

Die Ergebnisse dieser explorativen Analyse bekräftigen die Relevanz von Negativ-Empfehlungen in Leitlinien und die Möglichkeit des Implementierungserfolgs von Leitlinien durch QS-Programme. Die fehlende Veränderung bezüglich der Untertherapie ist jedoch nicht als Versagen der Implementierungsstrategie zu sehen, da diese Untertherapie sehr häufig durch individuell berechnete Faktoren der einzelnen Patientenfälle begründet ist, wie auch in dieser Studie dargestellt wurde.

#### 3.3.2 Klinische Behandlungspfade

Behandlungspfade bieten als strukturierte Instrumente für den Behandlungsverlauf sowohl Mitarbeitern und als auch Patienten eine gewisse Transparenz, Handlungssicherheit und eine mögliche Optimierung von Schnittstellenproblemen [32; 133]. Die Erstellung erfolgt ähnlich wie bei Leitlinien durch ein multidisziplinäres Gremium. Grundlage für diese Pfade sind häufig Leitlinien oder andere Evidenzquellen, die an die lokalen Gegebenheiten angepasst werden. Sie stellen einen Weg dar, die Empfehlungen von nationalen Leitlinien in die klinische Praxis vor Ort zu überführen [32; 133].

### Einfluss von klinischen Behandlungspfaden in der Chirurgie

In einem systematischen Review prüften Ronellenfitsch et al. die Wirksamkeit der Implementierung von klinischen Behandlungspfaden in der Chirurgie. Von den 30 passenden Studien waren die meisten im Vorher-Nachher-Design, nur ein RCT wurde identifiziert. Die meisten Studien zeigten eine deutliche Kostenreduktion, vor allem zurückzuführen auf die Verkürzung des stationären Aufenthalts. In Bezug auf Mortalität und Morbidität waren ebenfalls deutliche Verbesserungen in den meisten Studien dargestellt worden, in dem einzigen RCT war dies allerdings nicht nachweisbar [150]. Limitierend war bei diesem Review die fehlende Angabe zur Existenz einer Kontrollgruppe der Einzelstudien, zum Verzerrungsrisiko der Studienergebnisse sowie zu Maßzahlen zur konkreten Beurteilung der dargestellten Effekte.

Eine unkontrollierte Vorher-Nachher-Studie, die jedoch entsprechende Confounder bestmöglich berücksichtigte, konnte durch die Einführung eines Behandlungspfads für Patienten mit Nierentransplantation keine signifikanten Unterschiede in Bezug auf die Ergebnisqualität wie Morbidität (Infektionen, Thrombose), Re-Operationsrate oder 30-Tages-Mortalität nachweisen. Auch die Funktionalität der Spenderniere war nach einem Jahr ohne Unterschied zwischen den Gruppen. Die einzig signifikanten Unterschiede stellten sich für nahezu alle gemessenen Parameter der Prozessqualität wie z.B. zeitgerechtes Entfernen von Kathetern dar, was auch einem Hawthorne-Effekt<sup>14</sup> geschuldet sein kann [163].

Auf Grund der zunehmenden Fluktuation von Personal in den Kliniken ist für die strukturierte Einarbeitung und Ausbildung allerdings auch ohne eindeutige Studienevidenz die Sinnhaftigkeit von Pfaden naheliegend. Vor allem bei elektiven Patientenaufnahmen im Krankenhaus mit bereits gesicherten Diagnosen ist dies eine etablierte und sinnvolle Hilfe, insbesondere für neue oder unerfahrene Mitarbeiter.

Bestätigt wurde dies in einer qualitativen Studie mit persönlichen Interviews einer chirurgischen Abteilung in Deutschland [43]. Die meisten Befragten waren der Auffassung, dass Pfade für jüngeres und neu eingestelltes ärztliches oder pflegerisches Personal relevante Informationen über standardisierte Prozesse und damit eine gewisse Absicherung und Hilfe für eigenständiges Arbeiten, vor allem bei Behandlungsentscheidungen und in der Kommunikation mit Patienten enthalten. Es wurde die Wichtigkeit eines multidisziplinären Ansatzes aus verschiedenen Hierarchieebenen bei der Entwicklung und Implementierung von klinischen Pfaden in den Interviews hervorgehoben [43].

---

<sup>14</sup> Wirkung allein durch die Tatsache, dass Teilnehmer an Beobachtungsstudien vermehrt Aufmerksamkeit erfahren [114]

### Klinische Behandlungspfade in England

Ein in England bis 2013 weit verbreiteter klinischer Behandlungspfad ist der „Liverpool Care Pathway for the Dying Patient“ (LCP), ein umfassendes Betreuungskonzept für die Betreuung und Begleitung von unheilbar kranken Patienten am Ende ihres Lebens. Dieses Betreuungskonzept findet auch in Deutschland in der Palliativmedizin Anwendung. Trotz fast 10-jähriger Umsetzung existiert bisher keine unabhängige, formale, prospektive Evaluation von diesem Behandlungspfad [127]. Die Autoren eines vom Gesundheitsministerium in England beauftragten, unabhängigen Gutachtens zu diesem Behandlungspfad schlussfolgerten, dass hinreichende Evidenz durch eine vergleichende Untersuchung dieses Pfads mit anderen Versorgungsformen am Ende des Lebens weder für den Nutzen noch für mögliche Risiken existieren. Die Auswertung von 483 Hinweisen von medizinischem Personal, Institutionen, Betroffenen und Angehörigen löste jedoch massive Kritik und kontroverse Diskussionen zum „Liverpool Care Pathway for the Dying Patient“ aus, da dieser häufig nicht adäquat angewendet wurde und damit eher zu Schaden als Nutzen führte. Es wurde empfohlen, das Konzept des Pfads einer Überarbeitung zu unterziehen [127].

An diesem Beispiel wird deutlich, wie wichtig der Kontext der Implementierung von QS-Maßnahmen mit sachgerechter Einweisung, Schulung und die in regelmäßigen Abständen nachfolgende Evaluation ist.

#### 3.3.3 Patientensicherheit

Grundvoraussetzung für Patientensicherheit ist ein funktionierendes System wie z.B. ein QM-System, in welches die Projekte der Patientensicherheit integriert werden können [99].

Im Patientenrechtegesetz vom 20.02.2013 werden erstmalig wesentliche Maßnahmen zur Patientensicherheit mit Mindeststandards zu Risikomanagement und zu Fehlermeldesystemen über grundsätzliche Anforderungen an ein einrichtungsinternes Qualitätsmanagement in Krankenhäusern festgelegt, um unerwünschte Ereignisse zu vermeiden [30].

In einem narrativen Review zu Maßnahmen zur Verbesserung der Patientensicherheit wurde nach Sichtung der Literatur zusammengefasst, dass im Medikationsprozess, bei invasiven Prozeduren und bei Patientenübergaben eine nicht unerhebliche Anzahl an sogenannten iatrogenen Krankheiten entstehen, die ihre Wurzeln in den Systemen selbst haben. Insbesondere ist dies in mangelhafter Kommunikation, Schnittstellenproblematik, fehlenden Standardisierungen von Prozessen und Nicht-Beachtung von Regeln und Checklisten zu sehen [40]. Da ein gutes QM diese Aspekte abdecken sollte und die organisatorische und strukturelle Basis dafür zur Verfügung stellen kann, ist QM eine gute Voraussetzung, dass solche Projekte zur Patientensicherheit in die Praxis umgesetzt werden können.

So wie im ganzen QM ist auch im Bereich der Patientensicherheit als wichtiger Faktor zur Etablierung einer Sicherheitskultur die Rolle der Führung zu nennen, die die Relevanz von Patientensicherheit als integralen Bestandteil der Unternehmenskultur glaubhaft vermitteln muss. Dafür ist die Einbeziehung der persönlichen Verantwortung aller einzelnen Mitarbeiter und eine geschulte Kommunikation über Fehler als Bestandteil der Aus- und Weiterbildung relevant [40].

Es gibt mehrere Instrumente zur Wahrung oder Optimierung der Patientensicherheit. Zum einen die Erfassung von allen unerwünschten oder kritischen Ereignissen, um eine Analyse und eine daraus resultierende Prozessmodifikation möglich zu machen. Dies wird in Fehlerberichtssystemen wie dem Critical Incident Reporting System (CIRS) umgesetzt, welche an vielen Krankenhäusern fest etabliert sind [2; 11; 40; 187]. Diese bieten die Möglichkeit, auf freiwilliger Basis im Inter- oder Intranet anonym Berichte von Ereignissen einzugeben, die von Experten gegebenenfalls unter Erarbeitung von Veränderungsvorschlägen analysiert werden.

#### Effektivität von Maßnahmen zur Patientensicherheit

In einer ausführlichen, systematischen Evidenzanalyse der AHRQ von 2013, die die Ergebnisse des Berichts aus dem Jahr 2001 aktualisierte, wurde die Effektivität und Umsetzung von Maßnahmen zur Patientensicherheit untersucht [167]. Von den initial 100 verschiedenen Maßnahmen zeigten zehn eine ausreichend hohe Evidenz bezüglich Wirksamkeit und Implementierbarkeit.

Diese umfassten die nachfolgend aufgelisteten Verfahren:

- Präoperative Checklisten
- Checklisten für die Prävention von ZVK-Infektionen
- Interventionen zur Reduktion von Harnwegskathetern
- Pflegerische Maßnahmen zur Reduktion von beatmungsassoziierten Pneumonien
- Händehygienemaßnahmen
- Liste zur Vermeidung von riskanten Abkürzungen
- Maßnahmen zur Reduktion von Dekubitalgeschwüren
- Maßnahmen zur Vermeidung von nosokomialen Infektionen
- Sonographie zur ZVK-Anlage
- Maßnahmen zur Verbesserung der Thromboseprophylaxe

[167].

#### 3.3.4 Zertifizierung

Die grundlegenden Inhalte und Verfahrensweisen der medizinischen Zertifizierung der Akutkrankenhäuser aus Australien, Kanada, Frankreich, Großbritannien und den USA stimmen im Kern weitgehend überein [170].

Es existieren mehrere Zertifizierungssysteme zur Darlegung eines normgerechten Qualitätsmanagements. Die bekanntesten sind das European Foundation for Quality Management (EFQM) als international in Europa verbreitetes Zertifizierungssystem, das KTQ (Kooperation für Transparenz und Qualität im Krankenhaus) für Deutschland [108], das Qualitätsmodell Krankenhaus (QMK) der AOK [196], die Din EN ISO 9000 ff – Normierung und im angloamerikanischen Raum die Joint Commission on Accreditation of Healthcare Organizations [103]. All diese QM-Konzepte dienen vor allem in erster Linie der Bewertung des Qualitätsmanagements und entsprechender Zertifizierung von Einrichtungen [11; 138]. Ihnen allen ist die Selbst- und Fremdbewertung gemeinsam. Sie variieren in der unterschiedlichen Fokussierung auf zum Teil prozess- oder ergebnisorientierte Konzepte des Bewertungsmaßstabs. Sie beurteilen nicht die Qualität selbst, lediglich das Managementsystem im Bemühen um gute Qualität [11].

Für die Zertifizierung ist der Nachweis eines internen QM-Systems mit einem prozessorientierten Ansatz, der sich an Erwartungen von Patienten, zuweisenden Kollegen und Mitarbeitern orientiert, notwendig. Das QM wird in strukturierten Prozessen wie z.B. SOPs oder klinischen Behandlungspfaden umgesetzt. Die Qualitätsanforderungen orientieren sich dabei an dem aktuellen Stand der Wissenschaft, Empfehlungen von Leitlinien und gesetzlichen Vorgaben [90; 91]. Nach Zusammenstellung eines QM-Gremiums, einer Bestandsaufnahme der „Ist-Analyse“, Erstellung eines QM-Handbuchs mit Dokumentation der QM-Struktur und der festgelegten Behandlungsstandards und Arbeitsanweisungen wird zunächst ein internes Audit durchgeführt. Anschließend erfolgt ein externes Audit von einer akkreditierten Zertifizierungsgesellschaft, die abschließend das Zertifikat mit einer Gültigkeit von 3 Jahren und jährlicher Begutachtung vergibt [90; 91].

#### Reviews zum Einfluss von Zertifizierung auf Akut-Krankenhäuser

In einer großen, kontrollierten Querschnittsstudie im Ruhrgebiet untersuchten die Autoren den Zusammenhang zwischen einer Zertifizierung von Akut-Krankenhäusern und Unterschieden in der Patientenzufriedenheit nach Entlassung [155]. Es wurden insgesamt 25 kardiologische Abteilungen (zehn nicht-zertifizierte versus 15 zertifizierte Kliniken) mit über 3.000 Patienten eingeschlossen, die mittels eines speziell für Krankenhäuser entwickelten Fragebogens untersucht wurden. Die Zertifizierung erfolgte entweder über das Verfahren der KTQ (Kooperation für Transparenz und Qualität im Krankenhaus) [108] oder über eine weitere Zertifizierungsgesellschaft mit Namen pro-Cum Cert (pCC) [141]. Beide beinhalteten für die Krankenhäuser eine externe Überprüfung der organisatorischen Prozesse und Strukturen hinsichtlich nationaler Qualitäts- und Sicherheitsanforderungen und eine interne Selbstbewertung. Es zeigten sich keine signifikanten Unterschiede in der Patientenzufriedenheit zwischen den Gruppen [155]. Mögliche Confounder wurden berücksichtigt und in der Auswertung dargestellt, die Gruppen waren vergleichbar. Eine Korrelation zwischen der

Zertifizierung von Krankenhäusern und einer verbesserten Patientenzufriedenheit ist damit nicht nachweisbar.

Zu der gleichen Schlussfolgerung kamen die Autoren eines systematischen Reviews, der die verschiedenen Einflüsse der Zertifizierung von Krankenhäusern betrachtet hat [81]. Die Auswirkungen auf die Qualitätsindikatoren zeigten in der Gesamtauswertung der Studien eine inkonsistente Studienlage mit widersprüchlichen oder nicht signifikanten Unterschieden zwischen zertifizierten und nicht zertifizierten Kliniken [81]. Limitierend bei dieser Publikation ist der Punkt, dass das Studiendesign und die Qualität der Einzelstudien nicht beschrieben wurden und kaum Detailangaben zu den Studien wie z.B. Existenz einer Kontrollgruppe angegeben wurden, sodass eine weiterführende Beurteilung des Fazits der Autoren nicht möglich ist.

Weiterhin werden diese Ergebnisse auch durch eine Cochrane-Analyse bestätigt, die die Wirksamkeit von externen Kontrollen z.B. im Rahmen von Zertifizierungen hinsichtlich verbesserter Behandlungsprozesse und einer verbesserten Patientenversorgung untersuchte. Die Autoren folgerten in Anbetracht von nur zwei identifizierten relevanten Studien mit nicht-signifikanten oder inkonsistenten Ergebnissen (ein RCT und eine unkontrollierte Vorher-Nachher-Studie), dass bei zu geringen hochwertigen Studien keine hinreichende Aussage zur Wirksamkeit getroffen werden kann [66].

### 3.3.5 Sonstiges

#### Morbiditäts- und Mortalitätskonferenzen

Morbiditäts- und Mortalitätskonferenzen sind vielerorts fest etablierte Verfahren im QM der Krankenhäuser. In diesen Konferenzen werden ausgewählte Patientenfälle mit besonders komplizierten Verläufen in einer Konferenz unter Beteiligung vieler ärztlicher Fachdisziplinen vorgestellt, um gemeinsam konstruktive Lösungen für eine Optimierung der Prozesse zu erarbeiten.

In einer Studie mit einem Erhebungszeitraum von 2004 bis 2009 wurde die Wirksamkeit von Morbiditäts- und Mortalitätskonferenzen untersucht, die in den PDCA-Zyklus integriert wurden. Die Integration sollte durch Ableitung von neuen Strategien als Ergebnis der Konferenzen stattfinden [191]. Es erfolgte eine prospektive standardisierte Dokumentation mittels Fragebogen zu allen auftretenden Anastomoseninsuffizienzen nach kolorektaler Resektion. Diese Daten wurden in Berücksichtigung von Kovariablen wie Risikofaktoren, Operations-Indikation und Operations-Technik statistisch ausgewertet und anschließend potentielle Gründe für die Anastomoseninsuffizienzen in den monatlich stattfindenden Morbiditäts- und Mortalitätskonferenzen diskutiert. Einmal jährlich wurden in einem Treffen gemeinsam neue Lösungsstrategien entwickelt.

In der Studie wurde untersucht, wie häufig vor (2004 bis 2006) und nach (2006 bis 2009) Einführung dieser aus den Morbiditäts- und Mortalitätskonferenzen entwickelten neuen Strategien Anastomoseninsuffizienzen auftraten. Die Implementierungsrate wurde kontrolliert. 637 Patienten wurden im Untersuchungszeitraum einer Kolonresektion unterzogen. Es zeigte sich eine signifikante Reduktion der Anastomoseninsuffizienzen von 18 von 317 Patienten auf 9 von 320 Patienten. Beim Vergleich der Risikofaktoren der Gruppen ergaben sich für den zweiten Zeitraum bis 2009 signifikante schlechtere Werte, die die positiven Effekte damit bekräftigen [191].

Auf Grund des Studiendesigns mit fehlender Kontrollgruppe ist jedoch kein Beweis der Wirksamkeit gegeben. Die Studie kann lediglich einen Hinweis auf einen positiven Effekt von Mortalitäts- und Morbiditätskonferenzen bieten.

### Pay-for-Performance

Auch Ansätze wie Pay-for-Performance<sup>15</sup> wurden hinsichtlich ihrer Wirksamkeit auf die Patientenversorgung untersucht. In einer großen kontrollierten Vorher-Nachher-Studie mit methodisch gutem Niveau wurden in den USA anhand von Krankenkassendaten 252 Krankenhäuser mit Teilnahme an einem Pay-for-Performance-Programm (Premier Hospital Quality Incentive Demonstration (PHQID)) und 3.363 Kontrollhäuser miteinander verglichen. Unter Berücksichtigung von mehr als sechs Millionen Patientenfällen zeigte sich weder in der 30-Tages-Mortalität wegen Herzinfarkt, Herzinsuffizienz, Koronararterien-Bypass-OP oder Pneumonie ein signifikanter Unterschied in der Verbesserung im Zeitraum zwei Jahre vor und sechs Jahre (2003 und 2009) nach Einführung des Programms [102].

Interessanterweise wurde in einer ebenfalls methodisch gut durchgeführten kontrollierten Studie, die das gleiche Programm in allen Kliniken des National Health Service (NHS) an der Nordwestküste in England einführt, eine deutliche Reduktion der risiko-adjustierten Mortalitätsrate in den teilnehmenden Häusern gemessen [180]. Diese Diskrepanz zu den Evaluationsergebnissen in den USA könnte zum einen in dem verstärkten Einsatz von zusätzlichen qualitätsfördernden Maßnahmen als Reaktion auf das Programm liegen oder zum anderen in einer höheren Bonuszahlung [180]. Deutlich wird an diesem Vergleich zwischen den Ländern, dass für die Wirksamkeit von qualitätsfördernden Maßnahmen die Art der Implementierung mit gegebenenfalls zusätzlichen qualitätssichernden Maßnahmen und die Begleitumstände eine große Rolle spielen.

Ein Gutachten im Auftrag des BMG zu Pay-for-Performance im Gesundheitswesen bestätigt diese Ergebnisse [189]. In der Zusammenfassung des Gutachtens ohne Berücksichtigung der oben genannten zwei Studien formulierten die Autoren ihr Fazit, dass es bislang noch nicht in Studien gelungen ist, zweifelsfrei die Wirksamkeit von Pay-for-Performance-Projekten nachzuweisen. Die Evidenz für die Wirksamkeit finanzieller Anreize war sowohl im stationären als auch im ambulanten

---

<sup>15</sup> Pay-for-Performance: leistungsorientierte Vergütung

ten Bereich sehr schwach und nur auf der Ebene einzelner Qualitätsindikatoren nachweisbar. In den meisten Projekten kamen vorwiegend Prozessindikatoren und kaum Ergebnisindikatoren zum Einsatz [189].

## 4 Diskussion

### 4.1 Kritische Analyse der Studien

#### Ergebnisse

Im Mittelpunkt der vorliegenden Arbeit stand eine systematische Literaturrecherche zur Beantwortung der Frage, ob und inwieweit für QM/QS-Programme eine Wirksamkeit in Bezug auf eine verbesserte Patientenversorgung nachweisbar ist und welche konkrete Evidenzgrundlage unter methodischen Aspekten die dargestellten positiven Effekte in den Publikationen besitzen. Als Konsequenz daraus stellte sich ferner die Frage, ob die weitere wissenschaftliche Ausgestaltung oder Neukonzeption von QM/QS-Verfahren bei der aktuellen Evidenzlage zielführend scheint oder ob die vorhandene Studienlage vielmehr Hinweise auf eine andere einzuschlagende Richtung gibt (wie z.B. verstärkte Evaluationsforschung).

Insgesamt wurden 3.065 potentiell relevante Arbeiten gefunden. Publikationen, die eine Evaluation zur Wirksamkeit in Hinblick auf einen direkten oder indirekten patientenrelevanten Nutzen thematisierten, fanden sich nur in sehr geringem Maße ( $n = 138$ ). Studien, die die Wirksamkeit in einem für die Frage geeigneten Studiendesign durch Wahl einer Vergleichsgruppe oder Kontrollgruppe evaluierten, waren in Anbetracht der gefundenen Publikationsfülle in noch geringerem Ausmaß vorhanden ( $n=25$ ).

Betrachtet man die Studien aus Deutschland und dem Ausland gemeinsam, waren in rund zwei Drittel der den Einschlusskriterien entsprechenden Studien (15 von 25 Studien) neben anderen Qualitätsdimensionen Parameter der Ergebnisqualität berücksichtigt worden. In acht der 25 relevanten Studien wurden lediglich Surrogatparameter der Prozessqualität als möglicher Hinweis auf die zu erwartende Ergebnisqualität untersucht. Die einzigen Studien mit akzeptabler Studienqualität (mit moderatem Verzerrungsrisiko) und signifikantem Nachweis von Wirksamkeit werteten Parameter der Prozessqualität aus [33; 69]. Diese zwei Studien stammten aus dem internationalen Raum.

Die Messung der Prozessqualität wurde zum Teil damit begründet, dass durch die Nachbeobachtungszeit von meist weniger als zehn Jahren „harte“ klinische Outcome-Variablen (wie Mortalität oder Langzeitkomplikationen) nicht ausreichend häufig auftreten, um statistisch valide Aussagen machen zu können. In einigen Studien wird in der Beschreibung des QM-Verfahrens die Erhebung von Parametern der Ergebnisqualität wie z.B. Lebensqualität angegeben, diese aber in der Auswertung nicht dargestellt [175; 176].

Insgesamt waren die Ergebnisse der Studien nur sehr eingeschränkt zu verwerten. Es fand sich keine Studie mit einem geringen Verzerrungsrisiko und entsprechend hoher Studienqualität. Bei 16

der 22 Primärstudien war das Verzerrungsrisiko hoch bis sehr hoch und nur in sechs der 22 Primärstudien war das Verzerrungsrisiko moderat. Von diesen sechs Studien waren fünf aus dem internationalen Raum und nur eine Studie [70] aus Deutschland. In dieser prospektiv, kontrollierten Studie aus dem Jahr 2002, die vom BMG gefördert wurde, erfolgte eine Auswertung über einen Zeitraum von 26 Wochen in vier Krankenhäusern mit einem QM zur Reduktion von nosokomialen Infektionen im Vergleich zu vier Kontrollkliniken ohne QM. Es ergaben sich inkonsistente Ergebnisse bei zunächst positivem Effekt in der ersten Beobachtungsperiode, jedoch im weiteren Verlauf der zweiten Beobachtungsperiode nicht mehr nachweisbarem Unterschied [70].

Bei 14 der 22 Primärstudien handelte es sich um ein unkontrolliertes Vorher-Nachher-Studiendesign; nur vier Studien wiesen einen kontrollierten Ansatz auf. Das erhöhte Verzerrungsrisiko der Studien war durch die Wahl des Studiendesigns in den meisten Fällen erklärbar, da Studien ohne Kontrollgruppe per se ein erhöhtes Verzerrungsrisiko in sich tragen und nicht-randomisierte Studien die Gefahr des Selektionsbias nur durch gute Adjustierungsverfahren partiell auffangen können. Dies wurde von keiner Studie erfüllt. Immerhin 11 der 22 Primärstudien wurden als prospektiv beschrieben. Es fand sich unter den kontrollierten Studien ein RCT [33], dieser war ebenfalls mit methodischen Schwächen und damit moderatem Verzerrungsrisiko behaftet.

Die Schwäche der unkontrollierten Vorher-Nachher-Studien liegt darin, dass die kausale Zuordnung der Wirksamkeit zur Intervention oder zu einem ohnehin vorhandenen Entwicklungstrend sehr eingeschränkt ist. So besteht die Gefahr, dass ohne Berücksichtigung des Studiendesigns die gezeigten Effekte der QS/QM-Intervention überschätzt werden [61].

Dies konnte auch in der vorliegenden Arbeit festgestellt werden. So neigten Studien mit höherem Verzerrungsrisiko zu deutlich positiveren Ergebnissen. Unter Beachtung des Verzerrungsrisikos der Ergebnisse der Einzelstudien ergab sich, dass acht von zehn Studien mit positivem Wirksamkeitsnachweis aus der Kategorie mit hohem bis sehr hohem Verzerrungsrisiko stammten [22; 23; 111; 129; 147; 148; 153; 194]. Die Studien von methodisch höherer Güte mit moderatem Verzerrungsrisiko zeigten überwiegend inkonsistente oder fehlende Belege einer Wirksamkeit.

Die methodischen Limitationen bestätigen die Ergebnisse früherer Übersichtsarbeiten zu diesem Thema [42; 77; 121; 136]. Ein Grund für die geringe Anzahl aussagekräftiger Studien liegt darin, dass die Konzeption und Realisierung einer Studie für diese Fragestellung eine große Herausforderung darstellt. Selbst bei Vorhandensein einer Kontrollgruppe, die sich nur sehr schwer finden lässt, ist eine kausale Zuordnung des Effekts zur Intervention wie bei allen Studien der vorliegenden Arbeit nur bedingt möglich. Bei allen in dieser Arbeit gefundenen Studien ist von mehreren Ursachen für die verbesserten Ergebnisse der Patientenversorgung auszugehen. Da es sich um komplexe Interventionen handelt, wäre Monokausalität selbst bei höherer Studienqualität mit Randomisie-

rung bei dieser Thematik nicht eindeutig zu erwarten. Studienanforderungen für komplexe Interventionen wären mitzubedenken. Der Aufbau eines vergleichenden Studiendesigns ohne eine randomisierte Kontrollgruppe mit z.B. Matching<sup>16</sup> oder Propensity Score<sup>17</sup> Matching stellt ein aufwändigeres Verfahren dar als die einfach zu realisierenden einarmigen Vorher-Nachher-Studien. Der hohe Einsatz von Vorher-Nachher-Studien ist darauf zurückzuführen, dass diese verhältnismäßig einfach auszuführen, dabei aber aussagekräftiger als reine Beobachtungsstudien sind [61].

Es fand sich in Deutschland keine Studie von hoher methodischer Güte (niedriges Verzerrungsrisiko), die einen positiven Effekt in einer der drei Qualitätsdimensionen Struktur-, Prozess-, Ergebnisqualität durch ein QS- oder QM-Programm belegen konnte. Zehn Studien mit eingeschränkter Qualität legen hingegen einen positiven Wirksamkeitsnachweis dar.

In der Gesamtbetrachtung lässt sich weder ein eindeutig positiver Effekt noch ein negativer Effekt der QM/QS-Programme feststellen. Sowohl die Studien mit hohem Verzerrungsrisiko als auch die methodisch etwas besseren Studien mit moderatem Verzerrungsrisiko zeigen überwiegend inkonsistente Ergebnisse mit partiell positiven oder nicht nachweisbaren Effekten. Damit ist weder eine Wirksamkeit noch die fehlende Wirksamkeit belegt. Diese Ergebnisse sind übereinstimmend mit ähnlichen Vor-Untersuchungen in der Evaluation von QM/QS [42; 77; 121].

Diese fehlenden Belege in den Studien sind bemerkenswert vor dem Hintergrund, dass bei fehlender Randomisierung bei nahezu allen Studien mit freiwilligen Initiativen ein massiver Selektionsbias anzunehmen ist, in dem nur die besonders motivierten und engagierten Kliniken an einem QM/QS-Programm teilnahmen und entsprechend ohnehin förderliche Strukturen aufzeigen konnten, die nicht der Intervention selbst zuzuschreiben waren. Damit wäre ein signifikanter Unterschied in der Verbesserung zugunsten dieser Kliniken zu erwarten gewesen.

Auffallend ist die Tatsache, dass in Studien mit Kontrollgruppen und einem signifikanten Unterschied zur Interventionsgruppe die Kontrollgruppen ebenfalls Verbesserung im zeitlichen Verlauf aufwiesen [33; 70]. Diskutiert wird als Erklärung ein möglicher sogenannter „Spillover-Effekt“. Dieser beinhaltet einen vermuteten Übertragungseffekt, sodass qualitätsfördernde Maßnahmen auch in nicht partizipierenden Kliniken ebenfalls Eingang finden und Verbesserungen auslösen. Demzufolge könnte man nicht nachweisbare Effekte zwischen den beiden sich verbessernden Gruppen und ein daraus abgeleiteter fehlender Beleg einer Wirksamkeit der QS/QM-Verfahren

---

<sup>16</sup> Matching ist eine Methode mit dem Ziel, die Verteilung von Merkmalen (Alter, Geschlecht, Risikofaktoren etc.) zwischen den Gruppen auszugleichen, in dem jedem Patient aus der Interventionsgruppe ein Patient mit ähnlichen Merkmalen in der Vergleichsgruppe zugeordnet wird [178].

<sup>17</sup> Gegenüber dem exakten Matching besteht beim Matching unter Verwendung von Propensity Score Methoden der Vorteil, dass aus beliebig vielen Variablen ein Score erstellt wird, nach dem gepaart wird. Dies erfordert keine exakte Paarbildung für jeden einzelnen Patient [56; 178].

möglicherweise im Gegenteil auch als Hinweise für eine besonders gute Wirksamkeit der Verfahren diskutieren. Dieser Effekt sollte bei Auswertungen dieser komplexen Evaluationen in der Qualitätsforschung Berücksichtigung finden [88].

In einer systematischen Übersichtsarbeit von Drabik et al. [56] konnte gezeigt werden, dass sich ähnliche Schwierigkeiten der kausalen Zuordnung der Wirksamkeit der QS-Maßnahmen auch bei der Evaluation von Disease Management Programmen im Kontrollgruppendesign ergaben. Bei diesen Auswertungen wurden Adjustierungsmethoden in sieben von acht ausgewerteten kontrollierten Evaluationsstudien angewandt. Es zeigte sich ein sehr heterogener Ansatz in der Planung und Durchführung der Studien, insbesondere in der Wahl der berücksichtigten Variablen und Methoden der Auswertung, was eine Vergleichbarkeit und Interpretation der Ergebnisse sehr einschränkte [56].

#### Darstellung der Interessenkonflikte

Angaben zu Interessenkonflikten, die das Urteilsvermögen der Autoren beeinträchtigen und damit zu Verzerrungen (Bias) und Fehleinschätzungen hinsichtlich der Ergebnisse führen können [97; 117], fanden sich nur bei 14 der 25 Publikationen. Üblicherweise sind diese Angaben vor allem bei Arzneimittelstudien gefordert, da die Gefahr einer möglichen Beeinflussung durch die Finanzierung durch die Pharmaindustrie gesehen wird [97; 117]. Bei QS/QM-Programmen sind diese Angaben ebenfalls wichtig, da auch hier Klinikbetreiber, Versicherer, privatwirtschaftliche Unternehmen oder ebenfalls die Pharmaindustrie die Programme finanzieren und damit Interesse an einer positiven Ergebnisdarstellung haben.

Die geringe Anzahl an Publikationen (14 von 25) mit Angaben zu Interessenkonflikten ist konsistent mit der deutlich optimierbaren Qualität der Studienberichterstattung [131; 173] der gefundenen Studien (siehe unten).

Da die Ergebnisse bei jeder Art von Studien durch potentielle Interessenkonflikte durch Einzel- oder Partikularinteressen verzerrt werden können, ist dieser niedrige Anteil an Interessenkonflikterklärungen für die Beurteilung der Studienergebnisse hinderlich. Da eine unabhängige Finanzierung der Evaluation der Verfahren zu fordern, jedoch häufig nicht realisierbar ist, ist diese formale Darstellung der Interessenkonflikte als Minimalanforderung solcher Evaluationsstudien anzusehen.

#### Studienberichtsqualität

Auf Grund des Studiendesigns der Vorher-Nachher-Studien ergibt sich ein hohes Risiko an Verzerrungsmöglichkeiten. Ein einfacher, rein deskriptiver Vergleich ohne Berücksichtigung von Einflussfaktoren führt zu unzulässigen Fehlschlüssen, wenn z.B. Gruppen mit überwiegend Niedrigrisiko-Patienten bessere Ergebnisse als Gruppen mit einem hohen Anteil an Hochrisiko-Patienten erzielen. Risikoadjustierung hat das Ziel, alle wesentlichen Einflussfaktoren zu erkennen und zu berücksichtigen, die direkte Auswirkungen auf den gemessenen Endpunkt haben. Kontrollen durch

Adjustierung der Confounder durch die Anwendung bestimmter statistischer Verfahren bei der Analyse können den Selektions-Bias reduzieren. Diese Verfahren wurden allerdings vielfach nicht beschrieben bzw. bei fehlender Möglichkeit der Anwendung auf Grund von zu geringer Fallzahl oder anderen Gründen auch nicht diskutierend erwähnt, sodass die Methodik vielfach intransparent blieb.

Auch die Beschreibung der Populationen mit patientenspezifischen Variablen wie Alter, Geschlecht, Schwere der Erkrankung, Risikoscores oder Komorbiditäten wurden häufig in den Studien nicht angegeben, sodass eine Vergleichbarkeit der Gruppen zur Beurteilung des Verzerrungsrisikos der Ergebnisse vielfach nicht möglich war. Weitere Einflussvariablen wie die Einführung von weiteren QS-Maßnahmen oder neuen Leitlinien, neuen medikamentösen, operativen oder diagnostischen Verfahren oder Paradigmenwechseln erfuhren nur selten in der Diskussion Berücksichtigung. Es fanden sich kaum Angaben zum Umgang mit fehlenden Daten oder anderen Details der Auswertung.

Es ist daher dringend eine bessere Darstellung der Studienberichterstattung für diese Art der Studien zu Qualitätsprojekten zu fordern. Analog zu den etablierten Anleitungen zur Berichterstattung eines RCT (CONSORT-Statement [125]), eines systematischen Reviews (PRISMA-Statement [126]) oder zu Beobachtungsstudien (STROBE-Statement [192]) existieren seit 2008 auch ähnliche Empfehlungen zum Berichten von Studien zu QS/QM-Projekten mit dem Namen SQUIRE-Statement (Standards for Quality Improvement Reporting Excellence), welches viele relevanten Punkte beinhaltet [131; 173]. Diese internationalen Bemühungen zur Professionalisierung der Evaluation der QS/QM-Projekte in den letzten Jahren sollten auch in Deutschland sowohl bei der Konzeption als auch bei der Berichterstattung der Evaluationsstudien verstärkt Berücksichtigung finden. Auch andere Autoren von systematischen Reviews zu diesem Themengebiet der evaluierenden QS/QM-Forschung beschreiben die geringe Studienberichtsqualität [82]. Die eingeschränkte Publikationsqualität mit fehlenden relevanten Angaben ist zwar kein Beweis für ein methodisch fragliches Vorgehen der Evaluation, gibt aber zumindest einen starken Hinweis darauf.

Ebenfalls ist die Tatsache bemerkenswert, dass keine zur Wirksamkeit aussagekräftige Studie zu den größten in Deutschland existierenden QS-Programmen der gesetzlichen externen QS nach §137 SGB V und zu den QS-Programmen aus dem Rehabilitationsbereich der GKV und der DRV zu identifizieren war. Themenspezifische Publikationen zur Evaluation dieser Verfahren waren meist eine Wiedergabe von Messdaten zu Qualitätsindikatoren oder anderen Variablen, jedoch ohne Vergleichsgruppe ohne Intervention, sodass selbst eine eingeschränkte Ableitung von Aussagen zur Wirksamkeit nicht möglich war.

Begründet könnte dies dadurch sein, dass durch den flächendeckenden Einsatz der gesetzlichen externen QS und der großen QS-Programme der Rehabilitation, die ebenfalls per Gesetz verpflich-

tend sind, eine Studie mit einer Vergleichsgruppe ohne Intervention nicht mehr realisierbar ist. Auf Grund der langen Laufzeit sind Vorher-Nachher-Vergleiche ebenfalls nicht mehr möglich und wurden bei der Einführung der Programme nicht bedacht. Ein weiterer Grund der nicht auffindbaren Evaluationsstudien könnte auch in der prädominanten Stellung dieser fest etablierten, von öffentlichen Geldern finanzierten Programme liegen, für die der Rechtfertigungsdruck bezüglich einer Effektivität geringer ist als bei kleineren Qualitäts-Initiativen mit nur begrenzter zeitlicher und finanzieller Projektförderung und weniger starken Lobby.

Die Publikationen z.B. in Form von umfangreichen Qualitätsreports haben wie die Grundkonzeption dieser Programme selbst vor allem den Fokus auf einen Vergleich untereinander im Sinne einer Form von Benchmarking gesetzt. Es finden lediglich Vorjahresvergleiche als Verlaufsbeschreibung statt, was eine Wirksamkeitsanalyse sehr erschwert. Diese Vergleiche sollen die Qualität verbessern, fördern jedoch auch die Konkurrenz und nicht die Kooperation zwischen den Kliniken.

Prospektive Beobachtungsstudien mit einem Vergleich der Qualitätsdaten innerhalb einer Klinik über einen längeren zeitlichen Verlauf, die alle potentiellen Confounder mit geeigneten Adjustierungsverfahren berücksichtigen, könnten zumindest explorativ in Hinblick auf die Wirksamkeit der Verfahren diskutiert werden. Diese longitudinalen Verläufe wären für die Sicherstellung und die Messung des Erfolgs oder Misserfolgs der Qualitätsbemühungen förderlich. Auch diese sind von eingeschränkter Aussagekraft und bieten keinen direkten kausalen Nachweis, der bei der üblichen jährlichen Darstellung eines Querschnitts ebenfalls nicht vorliegt, würden aber wichtige hypothesengenerierende Hinweise geben. Diese Art der Publikation war ebenfalls kaum zu finden. In drei Studien wurden longitudinale Vergleiche der Datensätze der externen QS mit der Frage nach Leitlinien-Adhärenz [161; 162] bzw. mit der Frage nach dem Verhältnis von Aufwand und Nutzen vorgenommen [139]. Diese Studien gaben Hinweise, dass diese mit viel Mühe und zunehmender methodischer Sorgfalt gesammelten Datensätze der externen QS auch alternativ in anderer sehr hilfreicher Darstellung genutzt werden könnten.

### Qualitätsindikatoren

Als Kernstück der Qualitätsmessungen werden vielfach Qualitätsindikatoren angewendet. Wenn die Entwicklung der Qualitätsindikatoren jedoch weder einem etablierten Verfahren folgt noch einer kritischen Bewertung unterzogen wird [15], ist die Basis der QS bereits einem ersten Verzerrungsschritt unterworfen, der sich dann in der weiteren Evaluation der Verfahren auswirkt [159]. Auf Grund der großen Relevanz einer korrekten Messung der Qualität als Grundvoraussetzung der Evaluation von QS/QM müssen höchste Anforderungen an die Instrumente gestellt werden, die die Versorgungsqualität messen [5; 74; 146; 198], da anderenfalls bei Nutzung von Qualitätsindikatoren von möglicherweise niedriger methodischer Güte durch Fehleinschätzung der Versorgungsrea-

lität kontraproduktive Konsequenzen mit Ressourcenverschwendung und Schaden für den Patienten resultieren könnten [159; 160].

Der G-BA beauftragte das AQUA Institut als fachunabhängige Institution mit der Entwicklung von Qualitätsindikatoren für die gesetzliche externe QS. Darüber hinaus darf jedoch auch jeder andere Verband, Institution oder Organisation Qualitätsindikatoren entwickeln und einsetzen [159]. Derzeit wird fast ausschließlich durch das AQUA Institut Validierungsarbeit für die Qualitätsindikatoren vorgenommen. Ein möglicher Interessenkonflikt ist darin zu sehen, dass das Institut die Qualitätsindikatoren nicht nur per Auftrag durch den G-BA entwickelt, sondern auch evaluiert. Es besteht eine unüberschaubare Anzahl an Qualitätsindikatoren, deren methodische Güte auf Grund der fehlenden oder mit heterogenen Bewertungsinstrumenten durchgeführten Evaluation nicht eruiertbar ist. Dies führt zu einer scheinbaren Transparenz an Qualitätsmessungen, die mitunter nicht gegeben ist [159].

Auch aus den in dieser Arbeit berücksichtigten Publikationen der QS/QM-Evaluation geht selten hervor, welche Güte und Basis die verwendeten Qualitätsindikatoren besitzen und ob eine etablierte Methodik bei der Entwicklung zur Anwendung kam.

De Vos et al. gingen 2009 in einem systematischen Review der Frage nach, in wie weit die Implementierung von Qualitätsindikatoren die Qualität in Krankenhäusern verbessern können [44]. Die am häufigsten verwendeten Implementierungsstrategien in den 21 gefundenen Studien waren Audit und Feedback. Auch hier fokussierten sich die meisten Studien auf Prozess- anstelle von Ergebnisqualität in der Effektivitätsmessung. Von den sechs Studien, die Ergebnisqualität untersuchten, zeigten nur zwei Studien signifikante Verbesserungen durch Implementierung von Qualitätsindikatoren, in vier Studien erwies sie sich als ineffektiv. In Anbetracht der Ergebnisse schlussfolgerten die Autoren, dass die Implementierung von Qualitätsindikatoren in Krankenhäusern sich am effektivsten erweist, wenn gezieltes Feedback durch Qualitätsberichte kombiniert wird mit anderen Implementierungsstrategien wie Schulungen oder einem konkreten Plan für Verbesserungsmaßnahmen [44].

Eine externe QS mit Darstellung der Qualitätsindikatoren sollte ursprünglich auch der Transparenz und Informierung der Patienten dienen. Viele der abgebildeten Prozessindikatoren sind jedoch für Patienten unbrauchbar, da sie das intendierte Surrogat für die Ergebnisqualität nicht ermessen können und die Darstellung ohnehin für Laien schwer verständlich ist. Darüber hinaus haben die meisten Patienten vor allem Interesse an Aspekten, die man kaum abbilden kann, wie eine gute Arzt-Patienten-Beziehung und ähnlich schwer messbare qualitätsrelevante Faktoren [149]. Eine verbesserte Verbraucherinformation und Transparenz sind wichtige Faktoren der patientenorientierten

Versorgung und könnten durch eine verbesserte Patientenzufriedenheit zu erhöhter Ergebnisqualität führen [27].

Prozessindikatoren als Surrogatparameter für die zu erwartende Ergebnisqualität werden ähnlich den Surrogat-Endpunkten in klinischen Studien sehr kontrovers diskutiert und als geeignetes Messinstrument hinterfragt. Auf der anderen Seite werden auch die „harten“ klinischen Parameter der Ergebnisqualität wie Mortalität kontrovers diskutiert.

In einem systematischen Review kommen die Autoren in einer Auswertung von Studien zu diesem Thema zu dem Ergebnis, dass die Datenlage hinsichtlich eines Zusammenhangs der risikoadjustierten Mortalitätsrate mit der Versorgungsqualität ebenfalls weder konsistent noch verlässlich ist [140].

#### Instrumente von Qualitätsmanagement/Qualitätssicherung

Methodisch hochwertige Leitlinien haben als Unterstützung bzw. Grundlage von QM/QS auf die Patientenversorgung hinsichtlich der Prozess-, Struktur- und Ergebnisqualität signifikant positive Effekte [119]. Wesentliche Grundvoraussetzung für diesen positiven Effekt ist die Implementierung, die durch QS/QM-Programme durch einen systematischen und strukturierten Weg vereinfacht werden kann.

In mehreren Studien zur wirksamen Implementierung gemessen an dem Maß der Leitlinien-Adhärenz in der Patientenversorgung der Kliniken konnten deutliche Hinweise gegeben werden, dass die Einführung eines QM/QS-Programms die Leitlinien-Adhärenz in der Versorgung deutlich positiv verändern kann [57; 89; 107]. Insbesondere gab es Hinweise, dass Übertherapie vermindert werden konnte, was die Relevanz von Negativ-Empfehlungen in Leitlinien unterstreicht [57].

Für eine Zertifizierung waren keine Belege für die Wirksamkeit hinsichtlich verbesserter Behandlungsprozesse oder einer verbesserten Patientenversorgung zu finden [66; 81; 155].

Klinische Behandlungspfade als strukturierte Instrumente der stationären Behandlung von Patienten sind bei elektiven Patientenaufnahmen im Krankenhaus mit bereits gesicherten Diagnosen etablierte Maßnahmen im QM der Kliniken. In einer systematischen Übersichtsarbeit und einer Einzelstudie zeigten sich jedoch keine signifikanten Unterschiede in Bezug auf die Ergebnisqualität, lediglich eine Besserung der Parameter der Prozessqualität [163] oder inkonsistente Ergebnisse [150]. Auf Grund der hochgradigen Fluktuation von Personal und dem zunehmenden Einsatz von Honorarärzten in den Kliniken ist für die strukturierte Einarbeitung und Ausbildung auch ohne Studienevidenz die Hilfe durch Behandlungspfade naheliegend [43].

Gut geführte Register bieten eine sinnvolle Möglichkeit der Erfassung von Unter- oder Überversorgung im Sinne der Versorgungsforschung und gegebenenfalls Versorgungssteuerung [170]. Auf Grund der vielfach monierten Kritik an der Validität der Routinedaten von Krankenkassen für die Erhebung von QS-Daten, der auf der anderen Seite zeitlich eingeschränkten Aussagekraft der Messdaten externer QS-Programme, die auf den stationären Aufenthalt begrenzt sind, stellt die Nutzung von Registerdaten eine sehr hilfreiche Alternative zur Erfassung von Langzeit-Verläufen der Ergebnisqualität dar. Grundvoraussetzung für solche Register sind jedoch a priori formulierte Fragestellungen und definierte Erhebungsinstrumente für die Ergebnisqualität [110].

#### Deutschland im internationalen Kontext

Die Evaluationsforschung von QM/QS-Verfahren unterscheidet sich deutlich von anderen Bereichen der Evidenzbasierten Medizin wie Therapie oder Diagnostik, da es sich meist um komplexe Interventionen handelt, die per se schwieriger zu evaluieren sind. Weiterhin kommt hinzu, dass bei vielen der existierenden Studien sowohl im Studiendesign, in der Ausführung als auch in der Studien- und Ergebnisbeschreibung auch international erhebliche Mängel festzustellen sind [168].

Da das Interesse an QS/QM sowohl wissenschaftlich, gesundheitspolitisch sowie gesundheitsökonomisch steigt, ist eine angemessene Evaluationsforschung dieser Verfahren zu fordern und weiterzuentwickeln, wie es international vielfach begonnen wurde [39; 61; 131; 164; 168; 173; 181]. Die externe QS scheint im internationalen Vergleich durch ihre flächendeckende Anwendung über alle Bundesländer hinweg einmalig. So existieren zwar ähnliche Verfahren wie z.B. in den USA über die AHRQ, diese sind dort jedoch auf bestimmte Versichertengruppen limitiert [139].

Der Aufwand, der neben Deutschland in vielen Ländern für die Implementierung, Gestaltung, Umgestaltung und Überwachung von Qualitätssicherungssystemen betrieben wird, ist sehr groß. Die Heterogenität der unterschiedlichen Verfahren von QM/QS und der verschiedenen Institutionen ließen in Kombination mit den methodischen Herausforderungen einen Evidenzbeleg für ihre Wirksamkeit auf die Versorgungsqualität auch international offen [170].

#### Limitationen der eigenen Methodik

Ein limitierender Faktor ist der anzunehmende Publikationsbias. Träger von QM-Systemen, die in QS/QM investiert haben und deren Ergebnisse nicht dem erwarteten Erfolg entsprechen, haben - wie auch in anderen Bereichen - ein geringes Interesse, negative Ergebnisse zu publizieren.

Es wurde aus methodischen Gründen und zur Reduktion eines Verzerrungsrisikos eine breite Recherchestrategie gewählt, um die Konsistenz der Ergebnisse zu erhöhen. Trotz der sensitiven Recherchestrategie und der hohen Treffermenge waren jedoch relevante Studien nicht identifizierbar, da die Verschlagwortung von QM/QS in den gängigen bibliografischen Datenbanken uneinheitlich

ist. Auf Grund der heterogenen Vielfalt der einzelnen QS/QM-Programme ist das Vokabular der Schlagworte sehr vielseitig und damit Vollständigkeit schwer zu erreichen. Die Schwierigkeit bei dieser Art der Recherchen wurden auch in einer Studie von Hempel et al. bestätigt, die verschiedene Recherchestrategien zur Suche von qualitätsfördernden Projekten miteinander verglichen haben und bei allen Strategien unzureichende Ergebnisse bezüglich Präzision und Sensitivität erzielten [94].

Auf Grund der hohen Trefferzahl bei der Suche nach QM/QS allgemein war eine einschränkende Verknüpfung der gewählten Schlagworte zu QS/QM-Programmen mit Schlagworten der Evaluation und des Krankenhauses notwendig, was möglicherweise ebenfalls zur fehlenden Identifizierung von relevanten Arbeiten beigetragen hat. Es wurde versucht, die Lücken der Recherchestrategie mittels Handsuche zu ergänzen.

Eine weitere Limitation ist die Begrenzung der Recherche auf Literaturdatenbanken wie Medline und die Cochrane Library (Cochrane Database of Systematic Reviews) ohne Suche in den Datenbanken Embase oder CCMed (Current Contents Medicine), einer Datenbank für deutschsprachige Literatur.

Darüber hinaus ist die Validität der Ergebnisse durch die Tatsache geschmälert, dass kein zweiter unabhängiger Reviewer die Sichtung und Bewertung der Literatur vorgenommen hat.

Weiterhin ist kritisch anzumerken, dass die systematische Literaturrecherche für das Thema Effektivität von QM/QS auch dadurch Schwächen aufweist, dass die Publikationen nicht alle in Datenbanken wie Medline oder Cochrane Library verfügbar sind, sondern vielfach nur in Form der sogenannten „Grauen Literatur“ mit Sondergutachten im Auftrag von Institutionen des Gesundheitssystems oder von wissenschaftlichen Instituten [94]. Die „Graue Literatur“ stellt in der systematischen Suche eine Herausforderung dar, da sie in nur eingeschränkt zugänglicher Form und schlecht auffindbar publiziert wird. Um diese Schwäche möglichst auszugleichen und die „Graue Literatur“ erfassen zu können, erfolgte eine Suche auf ausgewählten Internetseiten von themenrelevanten, nationalen und internationalen Institutionen und über die Suchmaschine Google Scholar.

## 4.2 Ausblick

In der heutigen Zeit wird für alle neu eingeführten Medikamenten zur Behandlung von Patienten ein wissenschaftlich fundierter Wirksamkeitsnachweis gefordert [28; 63]. Diese auch für die QS/QM-Verfahren geforderte Evaluation [4; 27; 55; 71; 84; 93; 154; 172] wurde bisher nur sehr rudimentär umgesetzt.

In Anbetracht der dargelegten methodischen Limitationen der wenigen identifizierbaren Evaluationsstudien sind keine hinreichenden Evidenzbelege für eine Wirksamkeit von QM/QS-Verfahren

vorhanden und bei positiven Ergebnissen allenfalls nur explorativ hypothesengenerierende Aussagen möglich. Für die dringend geforderte zukünftige gemeinsame Weiterentwicklung von Evaluationsansätzen wäre der Einbezug von Ärzten aus der klinischen Versorgung, Wissenschaftlern, Patienten, politischen Entscheidungsträger und von Personen aus dem Krankenhausmanagement hilfreich, um die Expertise aller relevanten Interessenvertreter konstruktiv miteinfließen lassen zu können [42].

Diese Weiterentwicklung der QS/QM-Forschung sollte in einer Verschiebung der derzeitigen Aufmerksamkeit liegen. Bislang lag der Fokus der Bemühungen in der externen QS auf der Erfassung von Daten, der Datenvalidierung und Optimierung der Dokumentation. Zu fordern wäre jedoch, sich eher auf hochwertige Evaluation zu konzentrieren, um adäquate Aussagen zum Nutzen für die Behandlungsqualität des Patienten ableiten zu können.

Auf Grund der unzureichenden Evidenz für die Effektivität und Effizienz der Maßnahmen sollten Evaluationselemente integrale Bestandteile jedes QS/QM-Programms werden [77]. Als adäquate Evaluation für Prozesse der Qualitätsverbesserung sollten nicht nur quantitative (Datenanalyse) sondern auch qualitative Forschungsansätze (Interviews, Umfragen) Verwendung finden [95]. Die Evaluation im QS/QM-Bereich bedarf einer ausgewogenen Methodik, angelehnt an die Methodik zur Evaluation von komplexen Interventionen [84]. Qualitätsforschung ist ein komplexes Vorhaben mit zahlreichen Variablen, die sich gegenseitig beeinflussen und einer Dynamik unterliegen [59]. Eine verstärkte QS/QM-Forschung zur Wirksamkeit der Verfahren könnte auch die Akzeptanz von solchen Programmen oder Projekten seitens der Gesundheitsdienstleister erhöhen und zu einer verstärkten Patientenorientierung im QM führen [64].

Da die Begriffe zu QM/QS und den dazugehörigen Programmen sehr uneinheitlich verwendet werden, ist eine systematische Strukturierung dieser Begrifflichkeiten eine Herausforderung. Dies legt nahe, dass diese Unübersichtlichkeit für Nutzer und Adressaten der QS/QM-Programme ohne tiefergehende Schulungen in diesem Bereich erst recht zu Konfusion führt, was für die Motivation zur Anwendung der Verfahren wenig förderlich ist. Eine konsistente Anwendung der Begriffe ist sowohl für eine transparentere Abbildung der Thematik als auch für die Weiterentwicklung der Evaluationsforschung unabdingbar, damit z.B. Ergebnisse auf einer Meta-Ebene besser zusammengefasst werden könnten.

Auch wenn keine hinreichende Evidenz nach Kriterien der Evidenzbasierten Medizin für die Wirksamkeit von QS/QM-Maßnahmen auf übergeordneter Programmebene gezeigt werden kann, trägt QS/QM viele potentiell nützliche Aspekte in sich, die in einer konsequenten Selbstkontrolle durch systematische Dokumentation der Prozesse und Ergebnisse, in einer Selbstreflexion der Arbeit im fachlichen Austausch, in einer Optimierung der Kooperation mit anderen und in der Etablierung einer positiven Fehlerkultur begründet sind [132].

Die gesetzlichen Vorgaben zu QS/QM sind relativ unspezifisch und wenig präzise formuliert. Sowohl die gesetzlich vorgegebenen als auch die freiwilligen QM/QS-Programme haben ihre Schwächen und Limitationen und stellen zum Teil konkurrierende Maßnahmen dar wie z.B. das Peer-Review-Verfahren der IQM [3; 60; 147; 148] und der Strukturierte Dialog der gesetzlichen externen Qualitätssicherung [73], die zum Teil parallel in den teilnehmenden Kliniken durchgeführt werden. Sowohl der Strukturierte Dialog als auch das Peer-Review-Verfahren umfassen Teile vom Benchmarking-Konzept in der stationären Gesundheitsversorgung und überschneiden sich in ihrer Ausführung zum Teil erheblich. Dies stellt für das in den Kliniken arbeitende ärztliche Personal eine Doppelbelastung dar.

Um die Lücken der einzelnen QS-Systeme schließen und gegenseitig ausgleichen zu können, ist eine Ko-Existenz von Verfahren für ein umfassendes Qualitätssicherungssystem sinnvoll. Diese sollte jedoch komplementär und nicht parallel betrieben werden [169].

Es wird vielfach über die Nutzung von Routinedaten zur Gewährleistung von ausreichend langem Follow-up für die sektorale oder sektorenübergreifende Qualitätssicherung als Alternative zur bisherigen und sehr aufwändigen Datenerhebung der externen QS diskutiert, um z.B. durch Sozialdaten von Krankenkassen den Erhebungsaufwand zu minimieren und die Datenqualität zu verbessern. Die gesetzliche Grundlage wurde im Jahre 2012 geschaffen, diese Daten in pseudonymisierter Form zu nutzen. Da jedoch für die grundlegende Kernfrage der Effektivität für eine verbesserte Patientenversorgung bislang keine aussagekräftige Evidenz vorliegt und auch kein hinreichendes Konzept zur Evaluation publiziert wurde, ist diese immer tiefer gehende Beschäftigung mit Details der Datenerhebung diskussionswürdig.

Da in Deutschland vielfältige Diskussionen zur Einführung von Zuschlägen bei erbrachter hochwertiger Qualität bzw. Abschlägen bei unterdurchschnittlicher Qualität bestehen und Qualität in Zukunft ein Kriterium für die Krankenhausplanung sein könnte, wäre die Verteilung der Gelder in nicht unerheblichem Maß von Qualitätsfragen bestimmt. Angesichts der fehlenden hinreichenden Evidenz und der dargestellten Kritikpunkte wäre dies zu hinterfragen.

## 5 Zusammenfassung

Die Bedeutung von Qualitätsmanagement und Qualitätssicherung hat in den letzten Jahren durch das wachsende medizinische Wissen, durch zunehmend komplexer werdende medizinische Prozesse, durch den zunehmenden ökonomischen Druck auf die Gesundheitsdienstleister und die sich verändernde Haltung der Patienten gegenüber dem Gesundheitssystem deutlich zugenommen [132; 149; 175].

Ziel dieser Arbeit war es, auf Grund des immensen personellen und finanziellen Aufwands der Qualitätsmanagement/Qualitätssicherungs-Verfahren, der zunehmenden Ressourcenknappheit und der gesundheitspolitischen Bedeutung dieser Thematik mithilfe einer systematischen Literaturanalyse die existierenden Qualitätsmanagement/Qualitätssicherungs-Programme in der stationären Versorgung national und international kritisch hinsichtlich ihrer Wirksamkeit zu beleuchten.

In der hohen Anzahl von potentiell relevanten Artikeln (n=3065) fanden sich nur wenige Publikationen, die die Evaluationen zur Wirksamkeit von Qualitätsmanagement/Qualitätssicherung thematisierten (n=138). Obwohl Publikationen mit dargestellten positiven Effekte der einzelnen Verfahren und Programme in Deutschland existieren, zeigte sich in der Gesamtauswertung aller Studien der vorliegenden Arbeit bei fehlenden konsistenten und unzureichend validen Ergebnissen bislang kein Beleg einer Wirksamkeit solch übergeordneter Qualitätsmanagement/Qualitätssicherung-Verfahren auf die Patientenversorgung im Krankenhaussektor.

In einer detaillierteren Analyse unter methodischen Gesichtspunkten der Evidenzbasierten Medizin konnte keine Studie einen kausalen Zusammenhang zwischen den Verfahren und einer verbesserten Patientenversorgung nachweisen. In der Gesamtbetrachtung lässt sich sowohl in Deutschland als auch international weder ein eindeutig positiver Effekt noch ein negativer Effekt der Qualitätsmanagement/Qualitätssicherungs-Programme feststellen. Diese Ergebnisse sind übereinstimmend mit ähnlichen Vor-Untersuchungen zu dieser Thematik [42; 77; 121].

Es lässt sich hinsichtlich der methodischen Limitationen der eingeschlossenen Studien zusammenfassen, dass nur wenige Studien ein geeignetes Studiendesign mit der Wahl einer adäquaten Vergleichsgruppe vorhalten konnten, in einem Drittel der Studien nur Surrogatparameter der Prozessqualität als möglicher Hinweis auf die zu erwartende Ergebnisqualität untersucht wurden und insgesamt die Ergebnisse der Studien bei rund zwei Drittel der Studien einem hohen bis sehr hohen Verzerrungsrisiko ausgesetzt waren. Die Studienberichterstattung der gefundenen Studien war überwiegend optimierungsbedürftig, eine bessere Darstellung für diese Art der Studien zu Qualitätsprojekten ist dringend notwendig [131; 173]. Formale Darstellungen der Interessenkonflikte, eine Beschreibung der Studienpopulation mit potentiellen oder berücksichtigten Confounder und Auswertungsverfahren sind als Minimalanforderung solcher Evaluationsstudien zu fordern. Angaben zu den Interessenkonflikten, die das Urteilsvermögen der Autoren beeinträchtigen können und

damit zu Verzerrungen hinsichtlich der Ergebnisse führen können [97; 117], fanden sich nur bei rund der Hälfte der Publikationen. Bei Qualitätsmanagement/Qualitätssicherungs-Programmen haben diese eine große Bedeutung, da Klinikbetreiber, Versicherer, privatwirtschaftliche Unternehmen oder die Pharmaindustrie die Programme finanzieren und damit unangemessenen Einfluss auf die Ergebnisse ausüben können.

Aussagekräftige Studien zur Wirksamkeitsanalyse zu den größten in Deutschland existierenden Qualitätssicherungs-Programmen der gesetzlichen externen Qualitätssicherung nach §137 SGB V und zu den Qualitätssicherungs-Programmen aus dem Rehabilitationsbereich der GKV und der DRV waren nicht identifizierbar.

Als Instrumente bzw. Grundlage sind methodisch hochwertige Leitlinien als feste Bestandteile jedes Qualitätsmanagement/Qualitätssicherung-Programm nachgewiesenermaßen sinnvoll [119]. Es gibt Hinweise, dass die wirksame Implementierung von Leitlinien in den Kliniken durch Einführung eines Qualitätsmanagement deutlich positiv unterstützt werden konnte [57; 89; 107]. Für eine Zertifizierung waren keine Belege für die Wirksamkeit hinsichtlich verbesserter Behandlungsprozesse oder einer verbesserten Patientenversorgung zu finden [66; 81; 155]. Gut geführte Register bieten unter bestimmten Voraussetzungen gute Erhebungsinstrumente für die Ergebnisqualität und eine gute Möglichkeit zur Erfassung von Unter- oder Überversorgung [110; 170].

Ein möglicher Grund für die wenigen aussagekräftigen Studien liegt darin, dass die Konzeption und Realisierung einer Studie für diese Fragestellung eine große Herausforderung darstellt. Die Evaluationsforschung von Qualitätsmanagement/Qualitätssicherungs-Verfahren unterscheidet sich deutlich von anderen Bereichen der Evidenzbasierten Medizin wie Therapie oder Diagnostik, da es sich meist um komplexe Interventionen handelt. Dafür ist es notwendig, dass mehr Aufmerksamkeit in das Verständnis von einzelnen Maßnahmen und deren Wirksamkeit gesetzt wird, um die entscheidenden Erfolgsfaktoren erkennen zu können. Die Weiterentwicklung der Qualitätsforschung sollte in einer Verschiebung der bislang starken Bemühungen und Diskussionen um die Datenerfassung und Datenvalidierung auf eine methodisch bestmögliche Evaluation zu sehen sein, um adäquate Aussagen zum tatsächlichen und spürbaren Nutzen für die Behandlungsqualität des Patienten ableiten zu können.

Auf Grund der unzureichenden Evidenz für die Effektivität und Effizienz sollten Evaluationselemente als integrale Bestandteile jedes Qualitätsmanagement/Qualitätssicherungs-Programms bereits vor der Einführung in die klinische Versorgung konzeptionell mitbedacht und eingeplant werden [77]. Als adäquate Evaluation für Prozesse der Qualitätsverbesserung sollten sowohl quantita-

tive (Datenanalyse) als auch qualitative Forschungsansätze für komplexe Interventionen Verwendung finden [84; 95].

Auch wenn keine hinreichende Evidenz nach Kriterien der Evidenzbasierten Medizin für die Wirksamkeit von Qualitätsmanagement/Qualitätssicherungs-Maßnahmen auf einer übergeordneten Programmebene gezeigt werden konnten, tragen diese auf einer tieferliegenden Ebene viele positive Aspekte für das komplexe System Krankenhaus in sich.

## 6 Literaturverzeichnis

1. Abteilung Allgemeinmedizin und Versorgungsforschung Universitätsklinikum Heidelberg. Evaluation des DMP Diabetes mellitus Typ 2 im Rahmen der ELSID-Studie. Abschlussbericht für den AOK-Bundesverband. 2012 [Letzter Zugriff am: 02.12.2013] <http://www.aok-gesundheitspartner.de/bund/dmp/evaluation/elsid/index.html>
2. Aktionsbündnis Patientensicherheit e.V. Glossar. 2014 [Letzter Zugriff am: 21.01.2014] <http://www.aps-ev.de/patientensicherheit/glossar/>
3. Albrecht M, De Meo F, Ekkernkamp A, Heber M, von Harnoncourt F, Wolfram T. Jahrbuch Qualitätsmedizin. 2011 [Letzter Zugriff am: 26.05.2013] <http://www.initiative-qualitaetsmedizin.de/downloads/downloads-publikationen>
4. Albrecht M, Loos S, Otten M. [Cross-sectoral quality assurance in ambulatory care]. Z Evid Fortbild Qual Gesundhwes 2013; 107 (8):528-33
5. Altenhofen L, Blumenstock G, Diel F, Döbler K, Geraedts M, Jäckel WH, Klakow-Franck R, Kleudgen S, Kopp I, Nothacker M, Ollenschläger G, Reiter A, Weinbrenner S, Zorn U (Hrsg). Manual Qualitätsindikatoren. Manual für Autoren. Berlin: Ärztliches Zentrum für Qualität in der Medizin (ÄZQ); 2009 (äzq Schriftenreihe; 36). <http://www.aeqz.de/mdb/edocs/pdf/schriftenreihe/schriftenreihe36.pdf>
6. American College of Surgeons. National Surgical Quality Improvement Program (NSQIP). Homepage. 2014 [Letzter Zugriff am: 14.02.2014] <http://site.acsnsqip.org/>
7. AQUA-Institut für angewandte Qualitätsförderung im Gesundheitswesen GmbH. Bericht zum Strukturierten Dialog 2010 - Erfassungsjahr 2009. 2011 [Letzter Zugriff am: 02.03.2012] [http://www.sqg.de/downloads/Themen/Strukturierter\\_Dialog/Strukturierter\\_Dialog\\_Abschlussbericht.pdf](http://www.sqg.de/downloads/Themen/Strukturierter_Dialog/Strukturierter_Dialog_Abschlussbericht.pdf)
8. AQUA-Institut für angewandte Qualitätsförderung im Gesundheitswesen GmbH. Homepage. 2014 [Letzter Zugriff am: 14.02.2014] <http://www.aqua-institut.de/de/home/index.html>
9. AQUA-Institut für angewandte Qualitätsförderung im Gesundheitswesen GmbH. Methodenpapier, Version 3.0. [Letzter Zugriff am: 04.01.2014] <http://www.aqua-institut.de/de/projekte/137a-sgb-v/methodenpapier/methodenpapier.html>
10. AQUA-Institut für angewandte Qualitätsförderung im Gesundheitswesen GmbH. Qualitätsreport 2012. 2013 [Letzter Zugriff am: 30.11.2013] <https://www.sqg.de/themen/qualitaetsreport/index.html>
11. Ärztliches Zentrum für Qualität in der Medizin (ÄZQ). Kompendium Q-M-A. Qualitätsmanagement in der ambulanten Versorgung. 3. Aufl. Gramsch E, Hoppe JD, Jonitz G, Köhler A, Ollenschläger G, and Thomeczek C (Hrsg). Köln: Dt. Ärzte-Verl.; 2008
12. Ärztliches Zentrum für Qualität in der Medizin (ÄZQ), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF). Deutsches Instrument zur methodischen Leitlinien-Bewertung (DELBI). Fassung 2005/2006 + Domäne 8. 2008 [Letzter Zugriff am: 14.02.2014] <http://www.leitlinien.de/mdb/edocs/pdf/literatur/delbi-fassung-2005-2006-domaene-8-2008.pdf>

13. Beckmann U, Klosterhuis H, Mitschele A. Qualitätsentwicklung durch Qualitätssicherung - Erfahrungen aus zehn Jahren Qualitätssicherung der Rehabilitation. DAngVers 2005; (52):431-8
14. Blum K, Offermann, DKI. DMP Brustkrebs - Krankenhäuser sehen wenig Positives. Deutsches Ärzteblatt 2006; Jg. 103 (Heft 41):A 2678-A 2680
15. Blumenstock G. [Assessing the quality of quality indicators]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2011; 54 (2):154-9
16. Bohmer AB, Poels M, Simanski C, Trojan S, Messer K, Wirtz MD, Neugebauer EA, Wappler F, Joppich R. [Quality management in acute pain therapy: results from a survey of certified hospitals]. Schmerz 2012; 26 (4):425-4
17. BQS Bundesgeschäftsstelle Qualitätssicherung gGmbH. BQS-Qualitätsreport 2002. Strukturierter Dialog - Kernstück der externen Qualitätssicherung. 2002 [Letzter Zugriff am: 23.03.2012] [www.bqs-qualitaetsreport.de/2002/grundlagen/g\\_dialog.pdf](http://www.bqs-qualitaetsreport.de/2002/grundlagen/g_dialog.pdf)
18. BQS Bundesgeschäftsstelle Qualitätssicherung gGmbH. BQS-Qualitätsreport 2008. Düsseldorf: 2009 [Letzter Zugriff am: 2012 March 04] <http://www.bqs-qualitaetsreport.de/2008>
19. BQS Institut für Qualität und Patientensicherheit. Qualitätssicherung in der medizinischen Rehabilitation (QS-Reha-Verfahren). 2014 [Letzter Zugriff am: 14.02.2014] <http://bqs-institut.de/projekte/benchmarkprojekte/qs-reha-verfahren.html>
20. BQS Institut für Qualität und Patientensicherheit. Homepage. 2014 [Letzter Zugriff am: 14.02.2014] <http://bqs-institut.de/>
21. Brenske M, Schorn K, Walger M, Schlottmann N. Der Gemeinsame Bundesausschuss - aktuelle und künftige Bedeutung für den Krankensektor. Krankenhaus 2005; (3):167-73
22. Brucker SY, Schumacher C, Sohn C, Rezai M, Bamberg M, Wallwiener D. Benchmarking the quality of breast cancer care in a nationwide voluntary system: the first five-year results (2003-2007) from Germany as a proof of concept. BMC Cancer 2008; 8 :358
23. Brucker SY, Wallwiener M, Kreienberg R, Jonat W, Beckmann MW, Bamberg M, Wallwiener D, Souchon R. Optimizing the quality of breast cancer care at certified german breast centers: a benchmarking analysis for 2003-2009 with a particular focus on the interdisciplinary specialty of radiation oncology. Strahlenther Onkol 2011; 187 (2):89-99
24. Brüggemann S. Das Reha-Leitlinienprogramm der BfA. Z ärztl Fortbild Qual Gesundh wes 2005; (99):47-50
25. Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV). Beurteilungskriterien für Leitlinien in der medizinischen Versorgung - Beschlüsse der Vorstände der Bundesärztekammer und Kassenärztlicher Bundesvereinigung, Juni 1997. Dtsch Arztebl 1997; 94 (33):A-2154-5
26. Bundesärztekammer (BÄK), Kassenärztliche Bundesvereinigung (KBV), Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) (Hrsg). Curriculum Ärztliches Qualitätsmanagement. 4. Aufl. Berlin: BÄK; 2007 (Texte und Materialien der Bundesärztekammer zur Fortbildung und Weiterbildung; Band 10). <http://www.bundesaerztekammer.de/downloads/CurrAerztlQM3.pdf>

27. Bundesärztekammer und Kassenärztliche Bundesvereinigung (Hrsg). Qualitätssicherung und Qualitätsverbesserung in der Medizin. Grundlagen einer bedarfsgerechten Gesundheitsversorgung: gemeinsame Bestandsaufnahme von BÄK und KBV über die Aktivitäten der Spitzenorganisationen der ärztlichen Selbstverwaltung auf dem Gebiet der Qualitätssicherung in der Medizin von 1955 bis 1995 / [Red.: Günter Ollenschläger]. München: Zuckschwerdt; 1996
28. Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM). Zulassungsverfahren. 2013 [Letzter Zugriff am: 15.02.2014] <http://www.bfarm.de/DE/Arzneimittel/zul/zulassungsverfahren/node.html>
29. Bundesministerium für Justiz und Verbraucherschutz. Das Fünfte Buch des Sozialgesetzbuchs - Gesetzliche Krankenversicherung (Artikel 1 des Gesetzes vom 20. Dezember 1988 (BGBl. I S. 2477, 2482), das zuletzt durch Artikel 1 des Gesetzes vom 22. Dezember 2013 (BGBl. I S. 4382) geändert worden ist. 1988 [Letzter Zugriff am: 05.01.2014] [http://www.gesetze-im-internet.de/sgb\\_5/BJNR024820988.html](http://www.gesetze-im-internet.de/sgb_5/BJNR024820988.html)
30. Bürgerliches Gesetzbuch. Gesetz zur Verbesserung der Rechte von Patienten und Patientinnen vom 20.2.2013. Bundesgesetzblatt 2013; Teil 1 (Nr. 9):277-82
31. Busse R, Nimptsch U, Mansky T. Measuring, monitoring, and managing quality in Germany's hospitals. Health Aff (Millwood ) 2009; 28 (2):w294-w304
32. Campbell H, Hotchkiss R, Bradshaw N, Porteous M. Integrated care pathways. BMJ 1998; 316 (7125):133-7
33. Canovas JJ, Hernandez PJ, Botella JJ. Effectiveness of internal quality assurance programmes in improving clinical practice and reducing costs. J Eval Clin Pract 2009; 15 (5):813-9
34. Centers for Medicare & Medicaid Services. Homepage Medicare. 2014 [Letzter Zugriff am: 14.02.2014] <http://www.medicare.gov/>
35. Cima R, Dankbar E, Lovely J, Pendlimari R, Aronhalt K, Nehring S, Hyke R, Tyndale D, Rogers J, Quast L. Colorectal surgery surgical site infection reduction program: a national surgical quality improvement program - Driven multidisciplinary single-institution experience. J Am Coll Surg 2013; 216 (1):23-33
36. Clade H. GKV-Gesundheitsreform 2000. Deutsches Ärzteblatt 2000; 97 (Heft 1-2):A17-A19
37. Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. Higgins JPT and Green S (Hrsg). 2011 [www.cochrane-handbook.org](http://www.cochrane-handbook.org)
38. Cochrane Effective Practice and Organisation of Care (EPOC) Group. EPOC Resources for review authors. Suggested risk of bias criteria for EPOC reviews. Oslo: Norwegian Knowledge Centre for the Health Services. 2012 [Letzter Zugriff am: 02.01.2014] <http://epoc.cochrane.org/epoc-author-resources>
39. Cochrane Effective Practice and Organisation of Care (EPOC) Group. EPOC Resources for review authors. What study designs should be included in an EPOC review and what should they be called? Oslo: Norwegian Knowledge Centre for the Health Services. 2012 [Letzter Zugriff am: 02.01.2014] <http://epoc.cochrane.org/epoc-author-resources>

40. Conen D. [Measures to enhance patient safety. Importance of efficiency evaluation]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2011; 54 (2):171-5
41. Cong L, Yu J, Liu Y. Implementing a continuous quality improvement program for reducing pressure prevalence in a teaching hospital in China. J Wound Ostomy Continence Nurs 2012; 39 (5):509-13
42. Conry MC, Humphries N, Morgan K, McGowan Y, Montgomery A, Vedhara K, Panagopoulou E, Mc GH. A 10 year (2000-2010) systematic review of interventions to improve quality of care in hospitals. BMC Health Serv Res 2012; 12 :275
43. De Allegri M, Schwarzbach M, Loerbroks A, Ronellenfisch U. Which factors are important for the successful development and implementation of clinical pathways? A qualitative study. BMJ Qual Saf 2011; 20 (3):203-8
44. de Vos M, Graafmans W, Kooistra M, Meijboom B, van d, V, Westert G. Using quality indicators to improve hospital care: a review of the literature. Int J Qual Health Care 2009; 21 (2):119-29
45. Deming WE (Hrsg). Out of the crisis. Massachusetts: Massachusetts Institute of Technology; 1986
46. Deutsche Gesellschaft für Unfallchirurgie e.V. (Hrsg). Weißbuch Schwerverletztenversorgung der DGU. 2. Aufl. Berlin: 2012 <http://www.dgu-online.de/qualitaet-sicherheit/schwerverletzte/weissbuch-schwerverletztenversorgung.html>
47. Deutsche Interdisziplinäre Vereinigung für Schmerztherapie (Hrsg). S3-Leitlinie „Behandlung akuter perioperativer und posttraumatischer Schmerzen“ (AWMF-Reg.-Nr. 041/001). 2008 [Letzter Zugriff am: 14.02.2014] <http://www.awmf.org>
48. Deutsche Rentenversicherung. Homepage. 2014 [Letzter Zugriff am: 14.02.2014] [www.deutsche-rentenversicherung.de](http://www.deutsche-rentenversicherung.de)
49. Deutschen Gesellschaft für Chirurgie. Internetseite Chirurgische Arbeitsgemeinschaft Akutschmerz (CAAS). 2014 [Letzter Zugriff am: 14.02.2014] <http://www.dgch.de/deutsch/sektionen-und-ags/#c430>
50. Deutsches Cochrane Zentrum (DCZ), Institut für medizinisches Wissensmanagement der AWMF (AMWF-IMWi), Ärztliches Zentrum für Qualität in der Medizin (ÄZQ). Gechter D, Nothacker M, Khan C, Mavergames C, Motschall E, Boeker M, Mucche-Borowski C, Blümle A, Lang C, Kopp I, Meerpohl J. Manual Systematische Literaturrecherche für die Erstellung von Leitlinien. 2013 [Letzter Zugriff am: 14.02.2014] <http://www.aezq.de/mdb/edocs/pdf/literatur/manual-literaturrecherche.pdf>
51. Deutsches Krankenhausinstitut (DKI). Auswirkungen der DMP Brustkrebs auf die Krankenhäuser. Blum K. and Offermann M (Hrsg). 2006
52. Donabedian A. An Exploration of Structure, Process and Outcomes as Approaches to Quality Assessment. In: Selbmann HK, Überla K (Hrsg). Quality assessment of medical care. Gerlingen: Bleicher; 1982.
53. Donabedian A. The Quality of Care. JAMA 1988; 260 (12):1743-8
54. Donabedian A. The Seven Pillars of Quality. Arch Pathol Lab Med 1990; 114 :1115-8
55. Donabedian A. The effectiveness of quality assurance. Int J Qual Health Care 1996; 8 (4):401-7

56. Drabik A, Sawicki PT, Muller D, Passon A, Stock S. [The methods within the evaluation of disease management programmes in control-group designs using the example of diabetes mellitus - a systematic literature review]. *Gesundheitswesen* 2012; 74 (8-9):496-501
57. du Bois A, Strutas D, Buhrmann C, Traut A, Ewald-Riegler N, Kommos S, Gomez R, Harter P. Impact of treatment guidelines and implementation of a quality assurance program on quality of care in endometrial cancer. *Onkologie* 2009; 32 (8-9):493-8
58. du Bois A, Rochon J, Lamparter C, Pfisterer J. Das Qualitätssicherungsprogramm der AGO Organkommission OVAR (QS-OVAR): Versorgungsstruktur und Realität in Deutschland 2001. *Zentralbl Gynakol* 2005; 127 (1):9-17
59. Eberlein-Gonska M. [What is evidence-based concerning quality management? Thoughts about an apparently simple question]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2011; 54 (2):148-53
60. Eberlein-Gonska M, Rink O. Peer Review - Fortschreiten in einem lernenden System. *Deutsches Ärzteblatt* 2013; 110 (16):A760-A766
61. Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care* 2003; 12 (1):47-52
62. Europarat. Entwicklung einer Methodik für die Ausarbeitung von Leitlinien für die optimale medizinische Praxis. Empfehlung Rec (2001)13 des Europarates am 10. Oktober 2001 und erläuterndes Memorandum. *Z Arztl Fortbild Qualitätssich* 2002; 96 Suppl 3 :5-59
63. European Medicines Agency. Marketing authorisation. 2014 [Letzter Zugriff am: 15.02.2014]  
[http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general\\_content\\_000197.jsp&mid=WC0b01ac058002251c](http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000197.jsp&mid=WC0b01ac058002251c)
64. Farin E, Jackel WH. Qualitätssicherung und Qualitätsmanagement in der medizinischen Rehabilitation. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2011; 54 (2):176-84
65. Farin E, Jackel WH, Schalaster V. [The quality assurance programme of the statutory health insurance funds in medical rehabilitation: results and further developments]. *Gesundheitswesen* 2009; 71 (3):163-74
66. Flodgren G, Pomey MP, Taber SA, Eccles MP. Effectiveness of external inspection of compliance with standards in improving healthcare organisation behaviour, healthcare professional behaviour or patient outcomes. *Cochrane Database Syst Rev* 2011; (11): CD008992
67. Francke AL, Smit MC, de Veer AJ, Mistiaen P. Factors influencing the implementation of clinical guidelines for health care professionals: a systematic meta-review. *BMC Med Inform Decis Mak* 2008; 8 :38
68. Fuchshuber PR, Greif W, Tidwell CR, Klemm MS, Frydel C, Wali A, Rosas E, Clopp MP. The power of the National Surgical Quality Improvement Program - achieving a zero pneumonia rate in general surgery patients. *Perm J* 2012; 16 (1):39-45
69. Gagneux E, Lombrail P, Vichard P. Trauma emergency unit: long-term evaluation of a quality assurance programme. *Qual Health Care* 1998; 7 (1):12-8

70. Gastmeier P, Brauer H, Forster D, Dietz E, Daschner F, Ruden H. A quality management project in 8 selected hospitals to reduce nosocomial infections: a prospective, controlled study. *Infect Control Hosp Epidemiol* 2002; 23 (2):91-7
71. Gemeinsamer Bundesausschuss. Vereinbarung des Gemeinsamen Bundesausschusses gemäß § 137 Abs.1 Satz 3 Nr.1 SGB V über die grundsätzlichen Anforderungen an ein einrichtungsinternes Qualitätsmanagement für nach § 108 SGB V zugelassene Krankenhäuser, in Kraft getreten 23.12.2005. *Bundesanzeiger* 2005; Nr. 242
72. Gemeinsamer Bundesausschuss. Qualitätsbericht der Krankenhäuser. 2011 [Letzter Zugriff am: 14.02.2014] <http://www.g-ba.de/institution/themenschwerpunkte/qualitaetsicherung/qualitaetsbericht>
73. Gemeinsamer Bundesausschuss. Richtlinie des Gemeinsamen Bundesausschusses über Maßnahmen der Qualitätssicherung in Krankenhäusern (QSKH-RL) in der Fassung vom 15. August 2006, zuletzt geändert am 20. Oktober 2011, in Kraft getreten am 1. Januar 2012. 2012 [Letzter Zugriff am: 01.02.2014] [http://www.g-ba.de/downloads/62-492-790/QSKH-RL\\_2013-06-20.pdf](http://www.g-ba.de/downloads/62-492-790/QSKH-RL_2013-06-20.pdf)
74. Geraedts M, Selbmann HK, Ollenschlaeger G. Critical appraisal of clinical performance measures in Germany. *Int J Qual Health Care* 2003; 15 (1):79-85
75. Gerlach FM (Hrsg). *Qualitätsförderung in Praxis und Klinik. Eine Chance für die Medizin.* Stuttgart: Thieme; 2001
76. GKV Qualitätssicherung Medizinische Rehabilitation. Homepage. 2014 [Letzter Zugriff am: 14.02.2014] <http://www.qs-reha.de/>
77. Glattacker M, Jackel WH. Evaluation der Qualitätssicherung - aktuelle Datenlage und Konsequenzen für die Forschung. *Gesundheitswesen* 2007; 69 (5):277-83
78. Glattacker M, Müller E, Jackel WH. Qualitätssicherung und Leitlinien in der Rehabilitation bei muskuloskelettalen Krankheiten. *Z Rheumatol* 2008; 67 (7):575-82
79. Graebig K (Hrsg). *Wörterbuch Qualitätsmanagement. Normgerechte Definitionen mit Übersetzungen. Zweisprachig Deutsch-Englisch.* Berlin: Beuth; 2006
80. Graebig K, Viethen G (Hrsg). *Qualitäts-Terminologie Gesundheitswesen, Definitionen, Erläuterungen, Fallbeispiele.* Frankfurt/Main: pmi; 1996
81. Greenfield D, Braithwaite J. Health sector accreditation research: a systematic review. *Int J Qual Health Care* 2008; 20 (3):172-83
82. Grimshaw J, McAuley LM, Bero LA, Grilli R, Oxman AD, Ramsay C, Vale L, Zwarenstein M. Systematic reviews of the effectiveness of quality improvement strategies and programmes. *Qual Saf Health Care* 2003; 12 (4):298-303
83. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, Whitty P, Eccles MP, Matowe L, Shirran L, Wensing M, Dijkstra R, Donaldson C. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004; 8 (6):iii-72
84. Grol R, Baker R, Moss F. Quality improvement research: understanding the science of change in health care. *Qual Saf Health Care* 2002; 11 (2):110-1
85. Guillaumondegui OD, Gunter OL, Hines L, Martin BJ, Gibson W, Clarke PC, Cecil WT, Cofer JB. Using the National Surgical Quality Improvement Program and the Tennessee

- Surgical Quality Collaborative to improve surgical outcomes. *J Am Coll Surg* 2012; 214 (4):709-14
86. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ. What is "quality of evidence" and why is it important to clinicians? *BMJ* 2008; 336 (7651):995-8
  87. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J,onso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y, Norris SL, Williams JW, Jr., Atkins D, Meerpohl J, Schunemann HJ. GRADE guidelines: 4. Rating the quality of evidence-study limitations (risk of bias). *J Clin Epidemiol* 2011; 64 (4):407-15
  88. Hansen LO, Herrin J, Nembhard IM, Busch S, Yuan CT, Krumholz HM, Bradley EH. National quality campaigns: who benefits? *Qual Saf Health Care* 2010; 19 (4):275-8
  89. Harter P, Muallem ZM, Buhrmann C, Lorenz D, Kaub C, Hils R, Kommos S, Heitz F, Traut A, du Bois A. Impact of a structured quality management program on surgical outcome in primary advanced ovarian cancer. *Gynecol Oncol* 2011; 121 (3):615-9
  90. Helbig M, Helbig S, Kahla-Witzsch HA, Kroll T, May A. Certifying a university ENT clinic using the ISO 9001:2000 international standard. *Int J Health Care Qual Assur* 2010; 23 (3):268-76
  91. Helbig M, Kahla-Witzsch HA, Helbig S, Knecht R, Gstottner W. Qualitätsmanagement nach DIN EN ISO 9001:2000. *HNO* 2006; 54 (12):922-8
  92. HELIOS Kliniken GmbH. Zahlen und Fakten über HELIOS. 2014 [Letzter Zugriff am: 14.02.2014] <http://www.helios-kliniken.de/ueber-helios/unternehmensportrait/zahlen-und-fakten.html>.
  93. Helou A, Schwartz FW, Ollenschläger G. Qualitätsmanagement und Qualitätssicherung in Deutschland. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2002; 45 :205-14
  94. Hempel S, Rubenstein LV, Shanman RM, Foy R, Golder S, Danz M, Shekelle PG. Identifying quality improvement intervention publications - a comparison of electronic search strategies. *Implement Sci* 2011; 6 :85
  95. Hulscher ME, Laurant MG, Grol RP. Process evaluation on quality improvement interventions. *Qual Saf Health Care* 2003; 12 (1):40-6
  96. Initiative Qualitätsmedizin. Homepage. 2014 [Letzter Zugriff am: 14.02.14] <http://www.initiative-qualitaetsmedizin.de/>
  97. Institute of Medicine . Conflict of interest in medical research, education, and practice. Lo B and Field MJ (Hrsg) Washington: National Academy Pr.; 2009
  98. Institute of Medicine. Homepage. 2014 [Letzter Zugriff am: 23.02.2014] <http://www.iom.edu/>
  99. Institute of Medicine. Committee on Quality of Health Care. To err is human: building a safer health system. Vol. 627. Kohn KL, Corrigan JM, Donaldson MS (Hrsg). Washington: National Academies Press; 1990
  100. Institute of Medicine (IOM). Medicare: a strategy for Quality assurance. (Vol. 1). Lohr KN (Hrsg). Washington DC: National Academy Press; 1990

101. Institute of Medicine (IOM). Clinical Practice Guidelines We Can Trust. Graham R, Mancher M, Miller Wolman D, Greenfield S and Steinberg E (Hrsg). Washington: National Academy Press; 2011
102. Jha AK, Joynt KE, Orav EJ, Epstein AM. The long-term effect of premier pay for performance on patient outcomes. *N Engl J Med* 2012; 366 (17):1606-15
103. Joint Commission. Internetseite. 2014 [Letzter Zugriff am: 14.02.2014] <http://www.jointcommission.org/>
104. Kawski S, Koch U. Qualitätssicherung in der medizinischen Rehabilitation in Deutschland. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2004; 47 (2):111-7
105. Klosterhuis H. Reha-Qualitätssicherung der Rentenversicherung - eine kritische Bestandsaufnahme. *RVaktuell* 2010; 57 (8):260-8
106. Klosterhuis H. [Rehabilitation guidelines as tool of quality assurance in rehabilitation]. *Z Arztl Fortbild Qualitätssich* 2005; 99 (1):41-5
107. Kommos S, Harter P, Traut A, Strutas D, Riegler N, Buhrmann C, Gomez R, du Bois A. Compliance to consensus recommendations, surgeon's experience, and introduction of a quality assurance and management program: influence on therapy of early-stage ovarian carcinoma. *Int J Gynecol Cancer* 2009; 19 (4):787-93
108. Kooperation für Transparenz und Qualität im Krankenhaus (KTQ). Homepage. 2014 [Letzter Zugriff am: 14.02.2014] <http://www.ktq.de/>
109. Kopp IB. [From clinical practice guidelines towards quality assurance]. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 2011; 54 (2):160-5
110. Kostuj T, Smektala R. Qualitätssicherung mithilfe von Routinedaten. *Unfallchirurg* 2010; 113 (12):1047-2
111. Krahwinkel W, Rink O, Liebetau M, Gunther M, Schuler E, Kuhlen R. [10 years of Peer Reviewing: improving treatment by quality indicators from administrative data]. *Dtsch Med Wochenschr* 2011; 136 (41):2083-8
112. Kreienberg R, Kopp I, Lorenz W. S3 Leitlinie Mammakarzinom der Frau: Diagnostik, Therapie und Nachsorge, 1. Aufl. (AWMF-Register-Nummer: 032 . 045OL). 2004 [nicht mehr abrufbar]
113. Kruis AL, Smidt N, Assendelft WJ, Gussekloo J, Boland MR, Rutten-van MM, Chavannes NH. Integrated disease management interventions for patients with chronic obstructive pulmonary disease. *Cochrane Database Syst Rev* 2013; 10 :CD009437
114. Kunz R, Ollenschläger G, Raspe H, Jonitz G, Donner-Banzhoff N (Hrsg). *Lehrbuch Evidenz-basierte Medizin in Klinik und Praxis*. 2. Aufl. Köln: Dt. Ärzte-Verl.; 2007
115. Lack N, Gerhardinger U. [An analysis of the effectiveness of external quality assurance programmes using changes in quality indicators of individual hospitals]. *Z Evid Fortbild Qual Gesundhwes* 2010; 104 (6):503-11
116. Lehmkuhl D, Meissner W, Neugebauer EA. [Evaluation of the "initiative pain-free clinic" for quality improvement in postoperative pain management. A prospective controlled study]. *Schmerz* 2011; 25 (5):508-15

117. Lieb K, Klemperer D, Koch K, Baethge C, Ollenschläger G, Ludwig WD. Interessenkonflikte in der Medizin: Mit Transparenz Vertrauen stärken. Ein Vorschlag zur Deklaration von Interessenkonflikten. Dtsch Arztebl 2011; 108 (6):A 256-60
118. Linder R, Ahrens S, Koppel D, Heilmann T, Verheyen F. The benefit and efficiency of the disease management program for type 2 diabetes. Dtsch Arztebl Int 2011; 108 (10):155-62
119. Lugtenberg M, Burgers JS, Westert GP. Effects of evidence-based clinical practice guidelines on quality of care: a systematic review. Qual Saf Health Care 2009; 18 (5):385-92
120. Maass C, Schleiz W, Weyermann M, Drosler SE. [Are hospital administrative data suitable for external quality assurance? Comparison of quality indicators based on separate statutory data collections (BQS) and hospital administrative data]. Dtsch Med Wochenschr 2011; 136 (9):409-14
121. Mayer ED, Boukamp KB. Internes Qualitätsmanagement in der medizinischen Versorgung. Gutachten des Kompetenz-Centrum Qualitätssicherung/Qualitätsmanagement (KCQ) des GKV-Spitzenverbandes und der Gemeinschaft der Medizinischen Dienste der Krankenversicherung. 2012 [Letzter Zugriff am: 23.12.2013] <http://www.kcqq.de/106-640.php>
122. Meissner W. [QUIPS: quality improvement in postoperative pain management]. Z Evid Fortbild Qual Gesundheitswes 2011; 105 (5):350-3
123. Meissner W, Mescha S, Rothaug J, Zwacka S, Goettermann A, Ulrich K, Schleppers A. Quality improvement in postoperative pain management: results from the QUIPS project. Dtsch Arztebl Int 2008; 105 (50):865-70
124. Miksch A, Laux G, Ose D, Joos S, Campbell S, Riens B, Szecsenyi J. Is there a survival benefit within a German primary care-based disease management program? Am J Manag Care 2010; 16 (1):49-54
125. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, Elbourne D, Egger M, Altman DG. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; (340):c869
126. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009; 6 (7):e1000097
127. Neuberger J. More care, less pathway - a review of the Liverpool Care Pathway . 2013 [Letzter Zugriff am: 30.11.2013] <https://www.gov.uk/government/publications>
128. Neuderth S, Saupe-Heide M, Bruckner U, Gross B, Wenderoth N, Vogel H. Wissenschaftliche Begleitung der Weiterentwicklung des Visitationsverfahrens für Reha-Kliniken der Deutschen Rentenversicherung Bund - Das Projekt „Visit II“. Rehabilitation 2012; 51 (3):171-80
129. Nimptsch U, Mansky T. Quality measurement combined with peer review improved German in-hospital mortality rates for four diseases. Health Aff (Millwood) 2013; 32 (9):1616-23
130. Offermanns G. Qualitätsmanagement im Krankenhaus. In: Prozess- und Ressourcensteuerung im Gesundheitssystem. Berlin Heidelberg: Springer ; 2011. S. 173-249

131. Ogrinc G, Mooney SE, Estrada C, Foster T, Goldmann D, Hall LW, Huizinga MM, Liu SK, Mills P, Neily J, Nelson W, Pronovost PJ, Provost L, Rubenstein LV, Speroff T, Splaine M, Thomson R, Tomolo AM, Watts B. The SQUIRE (Standards for Quality Improvement Reporting Excellence) guidelines for quality improvement reporting: explanation and elaboration. *Qual Saf Health Care* 2008; 17 Suppl 1 :i13-i32
132. Ollenschläger G. Qualitätssicherung - wem nützt das? *Arztebl Sachsen* 1999; (9):418-21
133. Ollenschläger G, Lelgemann M, Kopp I. Nationale VersorgungsLeitlinien - Nutzung im Qualitätsmanagement unter besonderer Berücksichtigung von Klinischen Behandlungspfaden und Regionalen Leitlinien. *Med Klin (Munich)* 2007; 102 (7):565-9
134. Ollenschläger G, Thomeczek C. Qualitätssicherung und kontinuierliche Qualitätsverbesserung. Bestandsaufnahme der ärztlichen Selbstverwaltung zur Qualitätssicherung in der Medizin 1955-1995. *Gesundh Wesen* 1996; 58 (7):360-71
135. Organisation for Economic Co-operation and Development (OECD). Homepage. 2014 [Letzter Zugriff am: 23.02.2014] <http://www.oecd.org>
136. Ovretveit J, Gustafson D. Evaluation of quality improvement programmes. *Qual Saf Health Care* 2002; 11 (3):270-5
137. Oxford Centre for Evidence-based Medicine (CEBM). EbM tools for critical appraisal. 2011 [Letzter Zugriff am: 14.02.2014] [www.cebm.net](http://www.cebm.net)
138. Paeger A. Qualitätsmanagement in der medizinischen Rehabilitation - Vergleich internationaler und nationaler Konzepte. *Z Arztl Fortbild Qualitätssich* 2002; 96 (1):17-24
139. Petzold T, Steinwitz A, Schmitt J, Eberlein-Gonska M. [Evaluation of external quality assurance in accordance with sect. 137 SGB V at the Carl Gustav Carus university hospital in Dresden]. *Z Evid Fortbild Qual Gesundhwes* 2013; 107 (8):541-7
140. Pitches DW, Mohammed MA, Lilford RJ. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. *BMC Health Serv Res* 2007; 7 :91
141. proCum Cert GmbH Zertifizierungsgesellschaft. Homepage. 2014 [Letzter Zugriff am: 14.02.2014] <http://www.procum-cert.de/>
142. Qualität in der Nephrologie (QiN). Homepage. 2013 [Letzter Zugriff am: 14.02.2014] <http://www.qin-dialyse.de/>
143. Qualitätsverbesserung in der postoperativen Schmerztherapie (QUIPS). Homepage. 2014 [Letzter Zugriff am: 26.01.2014] <http://www.quips-projekt.de>
144. Reeves BC, Deeks JJ, Higgins JPT, Wells GA. Chapter 13: Including non-randomized studies. In: Higgins JPT, Green S, Cochrane Collaboration (Hrsg). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011)*. 2011. [www.cochrane-handbook.org](http://www.cochrane-handbook.org)
145. Reeves BC, Deeks JJ, Higgins JPT, Wells GA. Supplementary material of Chapter 13: Non-randomized studies. In: Higgins JPT, Green S (Hrsg). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration. 2011 [Letzter Zugriff am: 04.01.2014] [www.cochrane-handbook.org](http://www.cochrane-handbook.org)

146. Reiter A, Fischer B, Kotting J, Geraedts M, Jackel WH, Dobler K. QUALIFY: Ein Instrument zur Bewertung von Qualitätsindikatoren. *Z Arztl Fortbild Qualitatssich* 2007; 101 (10):683-8
147. Rink O. Das IQM-Peer-Review-Verfahren. Verbesserung der Ergebnisqualität, Strategie und Ergebnisse. *Präv Gesundheitsf* 2012; 8 :22-8.
148. Rink O. [The IQM peer review procedure - results of the "initiative qualitatssicherung"]. *Z Evid Fortbild Qual Gesundhwes* 2012; 106 (8):560-5
149. Riskin L, Campagna JA. Quality assessment by external bodies: intended and unintended impact on healthcare delivery. *Curr Opin Anaesthesiol* 2009; 22 (2):237-41
150. Ronellenfitsch U, Rossner E, Jakob J, Post S, Hohenberger P, Schwarzbach M. Clinical Pathways in surgery: should we introduce them into clinical routine? A review article. *Langenbecks Arch Surg* 2008; 393 (4):449-57
151. Ruchholtz S. [External quality management in the clinical treatment of severely injured patients]. *Unfallchirurg* 2004; 107 (10):835-43
152. Ruchholtz S, Lefering R, Debus F, Mand C, Kuhne C, Siebert H. [TraumaregisterTraumaNetwork DGU und TraumaRegister DGU. Success by cooperation and documentation]. *Chirurg* 2013; 84 (9):730-8
153. Ruchholtz S, Lefering R, Paffrath T, Oestern HJ, Neugebauer E, Nast-Kolb D, Pape HC, Bouillon B. Reduction in mortality of severely injured patients in Germany. *Dtsch Arztebl Int* 2008; 105 (13):225-31
154. Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen (Hrsg). Bedarfsgerechtigkeit und Wirtschaftlichkeit. Gutachten 2000/2001. Band II Qualitätsentwicklung in Medizin und Pflege. Kurzfassung. 2001
155. Sack C, Lutkes P, Gunther W, Erbel R, Jockel KH, Holtmann GJ. Challenging the holy grail of hospital accreditation: a cross sectional study of inpatient satisfaction in the field of cardiology. *BMC Health Serv Res* 2010; 10 :120
156. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996; 312 (7023):71-2
157. Schafer I, Kuver C, Gedrose B, von Leitner EC, Treszl A, Wegscheider K, van den Bussche H, Kaduszkiewicz H. Selection effects may account for better outcomes of the German Disease Management Program for type 2 diabetes. *BMC Health Serv Res* 2010; 10 :351
158. Schäfer RD. Ärztliche Qualitätssicherung heute. *Z Evid Fortbild Qual Gesundhwes (ZEFQ)* 2008; (102):594-7
159. Schmitt J, Petzold T, Eberlein-Gonska M, Neugebauer EA. [Requirements for quality indicators. The relevance of current developments in outcomes research for quality management]. *Z Evid Fortbild Qual Gesundhwes* 2013; 107 (8):516-22
160. Schneider A, Broge B, Szecsenyi J. Müssen wir messen, um (noch) besser werden zu können. *Z Allg Med* 2003; (79):547-52
161. Schröder P, Reiter A, Boy O, Fischer B, Döbler K. Qualitätsindikatoren der BQS als Monitoring - Instrument zur Leitlinienimplementierung am Beispiel ausgewählter

- Qualitätsindikatoren bei Mammakarzinom und Schenkelhalsfraktur. Z Evid Fortbild Qual Gesundheitswes 2009; 103 (1):17-25
162. Schulze RU, Grams A, Smektala R. Leitlinienkonforme Versorgung. Unfallchirurg 2008; 111 (2):65-70
  163. Schwarzbach M, Bonninghoff R, Harrer K, Weiss J, Denz C, Schnulle P, Birck R, Post S, Ronellenfisch U. Effects of a clinical pathway on quality of care in kidney transplantation: a non-randomized clinical trial. Langenbecks Arch Surg 2010; 395 (1):11-7
  164. Scott I. What are the most effective strategies for improving quality and safety of health care? Intern Med J 2009; 39 (6):389-400
  165. Scottish Intercollegiate Guidelines Network. Critical appraisal: Notes and checklists. 2012 [Letzter Zugriff am: 13.02.2014] <http://www.sign.ac.uk/methodology/checklists.html>
  166. Sens B. BÄK/KBV/AWMF-Curriculum „Ärztliches Qualitätsmanagement“: Resumée aus drei Jahren praktischer Umsetzung einer zielorientierten Qualifikation. Z Arztl Fortbild Qualitätssich 2000; 94 (8):689-94
  167. Shekelle P, Wachter R, Pronovost P, Schoelles K, McDonald K, Dy S, Shojania K, Reston J, Berger Z, Johnsen B, Larkin J, Lucas S, Martinez K, Motala A, Newberry S, Noble M, Pfoh E, Ranji S, Rennke S, Schmidt E, Shanman R, Sullivan N, Sun F, Tipton K, Treadwell J, Tsou A, Vaiana M, Weaver S, Wilson R, Winters B (Hrsg). Making Health Care Safer II: An Updated Critical Analysis of the Evidence for Patient Safety Practices. Comparative Effectiveness Review No. 211. AHRQ Publication No.13. Agency for Healthcare Research and Quality. Rockville, MD: 2013
  168. Shojania KG, Grimshaw JM. Evidence-based quality improvement: the state of the science. Health Aff (Millwood) 2005; 24 (1):138-50
  169. Siebig J. [The role of benchmarking seen from the perspective of the Federal Joint Committee (G-BA)]. Z Evid Fortbild Qual Gesundheitswes 2011; 105 (5):408-11
  170. Simoes E, Boukamp K, Mayer ED, Schmahl FW. Gibt es Belege für den Impact qualitätssichernder/-fördernder Verfahren in anderen Ländern? Gesundheitswesen 2004; 66 (6):370-9
  171. Snyder C, Anderson G. Do quality improvement organizations improve the quality of hospital care for Medicare beneficiaries? JAMA 2005; 293 (23):2900-7
  172. Sozialgesetzbuch Fünftes Buch - Gesetzliche Krankenversicherung. § 137f Strukturierte Behandlungsprogramme bei chronischen Krankheiten in der Fassung des Gesetzes vom 21. Juli 2004 (BGBl. I S. 1791). 2013 [Letzter Zugriff am: 02.12.2013] [http://www.gesetze-im-internet.de/sgb\\_5/\\_137f.html](http://www.gesetze-im-internet.de/sgb_5/_137f.html)
  173. SQUIRE Guidelines (Standards for QUality Improvement Reporting Excellence). SQUIRE overview of the items in the SQUIRE checklist. Final revision - 29.04.2008. 2008 [Letzter Zugriff am: 22.01.2014] [http://squire-statement.org/guidelines/using\\_the\\_guidelines/](http://squire-statement.org/guidelines/using_the_guidelines/)
  174. Squires DA. The U.S. health system in perspective: a comparison of twelve industrialized nations. Issue Brief (Commonw Fund) 2011; 16 :1-14
  175. Stoffel MP, Barth C, Lauterbach KW, Baldamus CA. Evidence-based medical quality management in dialysis - Part I: Routine implementation of QiN, a German quality management system. Clin Nephrol 2004; 62 (3):208-18

176. Stoffel MP, Barth C, Lauterbach KW, Baldamus CA. Evidence-based medical quality management in dialysis - Part II: Improvement of hemodialysis adequacy. Clin Nephrol 2004; 62 (3):219-25
177. STOP. Schmerztherapie in der Onkologischen Pädiatrie. Ein Qualitätsmanagement- und Fortbildungsprogramm. 2014 [Letzter Zugriff am: 14.02.2014] <http://www.stiftung-lichtblicke.de/gstperkl.htm>
178. Stuart EA. Matching methods for causal inference: A review and a look forward. Stat Sci 2010; 25 (1):1-21
179. Sunol R, Vallejo P, Thompson A, Lombarts MJ, Shaw CD, Klazinga N. Impact of quality strategies on hospital outputs. Qual Saf Health Care 2009; 18 Suppl 1 :i62-i68
180. Sutton M, Nikolova S, Boaden R, Lester H, McDonald R, Roland M. Reduced mortality with hospital pay for performance in England. N Engl J Med 2012; 367 (19):1821-8
181. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. BMJ Qual Saf 2013; (0):1-9
182. Thalau J. Vergleichende Qualitätsmessung in der medizinischen Rehabilitation. In: Stock J, Szecsenyi J (Hrsg). Stichwort: Qualitätsindikatoren. Bonn: KomPart-Verl.-Ges.; 2007.
183. Thomeczek C, Birkner B, Everz D, Gass S, Gramsch E, Hellmann A, Herholz H, Oesingmann U, Gibis B, Stobrawa F, Weidringer JW, Ollenschläger G. Die Checkliste Q-M-A. Ein Instrument zur Bewertung von Qualitätsmanagementsystemen in der ambulanten Versorgung. Gesundh Wesen 2003; 65 (10):585-92
184. TraumaNetzwerk DGU der Deutschen Gesellschaft für Unfallchirurgie. Homepage. 2014 [Letzter Zugriff am: 14.02.2014] <http://www.dgu-traumanetzwerk.de/>
185. TraumaRegister DGU der Deutschen Gesellschaft für Unfallchirurgie. Homepage. 2014 [Letzter Zugriff am: 14.02.2014] <http://www.traumaregister.de/>
186. TÜV Rheinland. Zertifizierung Akutschmerztherapie. 2014 [Letzter Zugriff am: 14.02.2014] [http://www.tuv.com/de/deutschland/gk/managementsysteme/medizin\\_gesundheitswesen/zerertifizierung\\_akutschmerztherapie/zertifizierung\\_akutschmerztherapie.jsp](http://www.tuv.com/de/deutschland/gk/managementsysteme/medizin_gesundheitswesen/zerertifizierung_akutschmerztherapie/zertifizierung_akutschmerztherapie.jsp)
187. Vagts DA, Martin J, Dahmen K. Qualitätsmanagement im Krankenhaus - Prozesse und Ergebnisse kontinuierlich verbessern. Anesthesiol Intensivmed Notfallmed Schmerzther 2008; 43 (2):156-60
188. van Emmerich C, Metzinger B. Qualitätssicherung mit Routinedaten aus Sicht der Deutschen Krankenhausgesellschaft. dasKrankenhaus 2010; (12): 77-1182
189. Veit C, Hertle D, Bungard S, Trümner A, Ganske V, Meyer-Hofmann B. Pay-for-Performance im Gesundheitswesen: Sachstandsbericht zu Evidenz und Realisierung sowie Darlegung der Grundlagen für eine künftige Weiterentwicklung. Gutachten im Auftrag des Bundesministeriums für Gesundheit. BQS Institut für Qualität und Patientensicherheit: Hamburg. 2012 [Letzter Zugriff am: 23.12.2013] <http://bqs-institut.de/aktuelles/298-2012-08-20.html>

190. Vereinbarung. zur Qualitätssicherung und Qualitätsmanagement in der stationären  
Vorsorge und Rehabilitation nach § 137d Abs. 1 und 1a SGB V zwischen den  
Spitzenverbänden der Krankenkassen und den für die Wahrnehmung der Interessen der  
Vorsorge- und Rehabilitationseinrichtungen maßgeblichen Spitzenorganisationen vom  
1.4.2004. 2004 [Letzter Zugriff am: 20.03.2012]  
[http://www.bdpk.de/media/file/97.Vereinbarung\\_nach\\_%C2%A7\\_137d\\_Abs.\\_1und\\_1a\\_SGB\\_V.pdf](http://www.bdpk.de/media/file/97.Vereinbarung_nach_%C2%A7_137d_Abs._1und_1a_SGB_V.pdf)
191. Vogel P, Vassilev G, Kruse B, Cankaya Y. Morbidity and Mortality conference as part of  
PDCA cycle to decrease anastomotic failure in colorectal surgery. *Langenbecks Arch Surg*  
2011; 396 (7):1009-15
192. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP. The  
Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)  
statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008; 61  
(4):344-9
193. von Hodenberg E, Eder S, Grunebaum P, Melichercik J. Kardiologie – Qualitätssicherung  
in Deutschland Pro und Contra. *Dtsch Med Wochenschr* 2009; 134 Suppl 6 :S198-S199
194. Wallwiener M, Brucker SY, Wallwiener D. Multidisciplinary breast centres in Germany: a  
review and update of quality assurance through benchmarking and certification. *Arch  
Gynecol Obstet* 2012; 285 (6):1671-83
195. Willms G, Bramesfeld A, Pottkamper K, Broge B, Szecsenyi J. [Current challenges of  
external quality assurance in the German healthcare system]. *Z Evid Fortbild Qual  
Gesundhwes* 2013; 107 (8):523-7
196. Wissenschaftliches Institut der AOK (WIdO). QMK - Qualitätsmodell Krankenhaus. 2014  
[Letzter Zugriff am: 14.02.2014] <http://www.wido.de/qualitaetsmodellkh.html>
197. World Health Organization (Hrsg). WHO - World Alliance for Patient safety. WHO Draft  
Guidelines for adverse event reporting and learning systems. From information to action.  
Genf: WHO Press; 2005 [Letzter Zugriff am: 21.01.2014]  
[http://www.who.int/patientsafety/events/05/Reporting\\_Guidelines.pdf](http://www.who.int/patientsafety/events/05/Reporting_Guidelines.pdf)
198. Zentralstelle der Deutschen Ärzteschaft zur Qualitätssicherung in der Medizin (Hrsg).  
Beurteilung klinischer Messgrößen des Qualitätsmanagements - Qualitätskriterien und -  
Indikatoren in der Gesundheitsversorgung. Konsensuspapier der Bundesärztekammer, der  
Kassenärztlichen Bundesvereinigung und der AWMF. Jena: Urban & Fischer Verlag; 2001
199. Zernikow B, Hasan C, Hechler T, Huebner B, Gordon D, Michel E. Stop the pain! A  
nation-wide quality improvement programme in paediatric oncology pain control. *Eur J  
Pain* 2008; 12 (7):819-33

## 7 Anhang

### 7.1 Suchstrategie

#### 1 Hauptrecherche für Deutschland und international

Recherchequellen:

Literaturdatenbank Medline über <http://www.pubmed.org>

Cochrane Library - Cochrane Database of Systematic Reviews

Google, Google Scholar

Internetseiten (ausgewählte themenrelevante Institutionen)

1.1 Medline über <http://www.pubmed.org>

Recherchezeitpunkt: 15.12.2013

Anzahl der Treffer: 1282 (Einzeltrefferliste einsehbar)

Davon relevant: 20

Suchstrategie:

Nr.	Suchfrage	Anzahl
#5	#1 AND #2 AND #3 Filters: Publication date from 2011/12/01 to 2013/12/15; English; German	1082
#4	#1 AND #2 AND #3 Search (((Quality Assurance, Health Care[MeSH Terms] OR quality management OR national quality campaigns OR Quality Assurance Program* OR Quality management Program*)) AND (Effectiveness OR Efficacy OR efficiency OR Evaluation OR Sustainability OR patient-Outcome OR Health Impact Assessments OR quality improvement)) AND (((inpatient OR hospitals [MeSH Terms])))	10801
#3	inpatient OR hospitals [MeSH Terms]	253764
#2	Effectiveness OR Efficacy OR efficiency OR Evaluation OR Sustainability OR patient-Outcome OR Health Impact Assessments OR quality improvement	2405424
#1	Quality Assurance, Health Care[MeSH Terms] OR quality management OR national quality campaigns OR Quality Assurance Program* OR Quality management Program*	387051

Von Pubmed hinterlegte Detailauflösung der Schlagworte:

((("quality assurance, health care"[MeSH Terms] OR (quality[All Fields] AND ("organization and administration"[MeSH Terms] OR ("organization"[All Fields] AND "administration"[All Fields]) OR "organization and administration"[All Fields] OR "management"[All Fields] OR "disease management"[MeSH Terms] OR ("disease"[All Fields] AND "management"[All Fields]) OR "disease management"[All Fields])) OR (("federal government"[MeSH Terms] OR ("federal"[All Fields] AND "government"[All Fields]) OR "federal government"[All Fields] OR "national"[All Fields]) AND quality[All Fields] AND campaigns[All Fields]) OR (quality assurance program[All Fields] OR quality assurance programme[All Fields] OR quality assurance programming[All Fields] OR quality assurance programs[All Fields]) OR (quality management program[All Fields] OR quality management programme[All Fields] OR quality management programmes[All Fields] OR quality management programs[All Fields])) AND (Effectiveness[All Fields] OR Efficacy[All Fields] OR ("evaluation studies"[Publication Type] OR "evaluation studies as topic"[MeSH Terms] OR "evaluation"[All Fields]) OR Sustainability[All Fields] OR patient-Outcome[All Fields] OR ("health impact assessment"[MeSH Terms] OR ("health"[All Fields] AND "impact"[All Fields] AND "assessment"[All Fields]) OR "health impact

assessment"[All Fields] OR ("health"[All Fields] AND "impact"[All Fields] AND "assessments"[All Fields]) OR "health impact assessments"[All Fields]) OR ("quality improvement"[MeSH Terms] OR ("quality"[All Fields] AND "improvement"[All Fields]) OR "quality improvement"[All Fields])) AND (("inpatients"[MeSH Terms] OR "inpatients"[All Fields] OR "inpatient"[All Fields]) OR "hospitals"[MeSH Terms]) AND (("2011/12/01"[PDAT] : "2013/12/15"[PDAT]) AND (English[lang] OR German[lang]))

*Zusatzrecherche in Medline für Überblick zu QS-/QM-Programmen (ohne Evaluation):*

Recherchezeitpunkt: 15. 12.2013

Anzahl der Treffer: 156 Treffer (Einzeltrefferliste einsehbar)

Davon relevant: 3 für international

Gewählte Schlagworteingabe:

(Quality Assurance Program\*[Title] OR Quality management Program\*[Title]) Filters: Publication date from 2003/12/01 to 2013/12/01

Von Pubmed hinterlegte Detailauflösung der Schlagworte:

quality assurance program [Title] OR quality assurance programme[Title] OR quality assurance programmes[Title] OR quality assurance programming[Title] OR quality assurance programs[Title] OR (quality management program[Title] OR quality management programs[Title])) AND ("2003/12/01"[PDAT] : "2013/12/01"[PDAT])

## 1.2 Cochrane Library

Recherchezeitpunkt: 21.12.2013

Anzahl der Treffer: 88 Treffer (Einzeltrefferliste einsehbar)

Davon relevant: 0 (3 relevante: Dubletten)

Suchstrategie:

Nr.	Suchfrage	Anzahl
#3	#1 AND #2 Limit 2003 – 2013 Cochrane Database of Systematic Reviews	88
#2	(Effectiveness or Efficacy or Evaluation or Sustainability or patient-Outcome or Health Impact Assessments or quality improvement)	5683
#1	("Quality Assurance" or "quality management" or "national quality campaigns" or "Quality Assurance Program*" or "Quality management Program*")	90

## 1.3 Google Scholar: (keine Einzeltrefferdokumentation)

Recherchezeitpunkt: 21.05.2013

Anzahl der Treffer: 551 Treffer

Davon relevant: 47

Gewählte Schlagworteingabe:

Qualitätssicherungsprogramm\* OR Qualitätsmanagementprogramm\* OR Zertifizierungsprogramm\* - keine Einschränkungen

## **2. Zusatzrecherche nur für Deutschland**

Recherchequellen: Medline über <http://www.pubmed.org>

Recherchezeitpunkt: 11.05.2013

Anzahl der Treffer: 467 Treffer (Einzeltrefferliste einsehbar)

Gewählte Schlagworteingabe (keine zeitliche Einschränkung):

Nr.	Suchfrage	Anzahl
	Search (Quality Assurance Program OR Quality management Program) AND Germany AND (clinic OR clinical OR hospital OR hospitals OR in-patient OR inpatient)	467

*Nachrecherche* mit geänderter Strategie:

Recherchezeitpunkt: 15.12.13

Anzahl der Treffer: 369 Treffer (Einzeltrefferliste einsehbar)

Davon relevant: 0

Suchstrategie:

Nr.	Suchfrage	Anzahl
	(Quality Assurance, Health Care[MeSH Terms] OR quality management OR national quality campaigns OR Quality Assurance Program* OR Quality management Program*) AND (Effectiveness OR Efficacy OR Evaluation OR Sustainability OR patient-Outcome OR Health Impact Assessments OR quality improvement) AND (inpatient OR hospitals[MeSH Terms]) AND German* Filters: published in the last 10 years	369
	(Quality Assurance, Health Care[MeSH Terms] OR quality management OR national quality campaigns OR Quality Assurance Program* OR Quality management Program*) AND (Effectiveness OR Efficacy OR Evaluation OR Sustainability OR patient-Outcome OR Health Impact Assessments OR quality improvement) AND (inpatient OR hospitals[MeSH Terms]) AND German*	587
	(Quality Assurance, Health Care[MeSH Terms] OR quality management OR national quality campaigns OR Quality Assurance Program* OR Quality management Program*) AND (Effectiveness OR Efficacy OR Evaluation OR Sustainability OR patient-Outcome OR Health Impact Assessments OR quality improvement) AND (inpatient OR hospitals[MeSH Terms])	10894

## 7.2 Evidenztabelle – Studiendarstellung Deutschland

PICO-Schema<sup>18</sup>

Population	Intervention	Kontrolle	Outcome
stationärer Sektor oder klinikähnliche Einrichtungen	QS/QM-Programm (siehe E1)	keine Intervention oder Vergleichsgruppe zu Beginn der Intervention, wenn noch keine Wirkung zu erwarten (siehe E1)	Wirksamkeit hinsichtlich Parametern der Ergebnis-, Struktur- oder Prozessqualität

Legende :

+	Angabe vorhanden, trifft zu
–	nicht erfolgt
k.A.	keine Angabe
n.a	nicht anwendbar
KI	Konfidenzintervall
COI	Conflict of Interest
S3	Stufe 3 in der Klassifikation von Leitlinien der AWMF (höchste Kategorie)

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko										
Krahwinkel et al., 2011 [111]  <b>1</b>	Vorher-Nachher-Studie ohne Kontrollgruppe, retrospektiv	Peer-Review unter Nutzung von Qualitätsindikatoren aus Routinedaten  Vergleich der Daten vor und nach Einführung des Verfahrens im Jahr 2001  Helios Kliniken GmbH  fächerübergreifend	2000 bis 2008 bzw. 2010  rund 100.000 aus 16 Konzern-Kliniken für die 3 ausgewählten Krankheiten  Routinedaten	Ergebnisqualität  Endpunkt: Krankenhaussterblichkeit für Herzinfarkt, Herzinsuffizienz und Pneumonie Jahr 2000 12,9%, 17,4% und 13,1% -bis zum Jahr 2010 Absinken auf Werte von 9,1% für Herzinfarkt, 6,4% für Herzinsuffizienz und 8,4% für Pneumonie mit Unterschreiten der bundesweiten Krankenhaussterblichkeiten des Statistischen	Verzerrungsrisiko: hoch  <table border="1"> <tbody> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung</td> <td>+</td> </tr> <tr> <td>(„keine“, obwohl Angestellte von Helios)</td> <td></td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>k.A.</td> </tr> <tr> <td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveau etc.)</td> <td>-</td> </tr> <tr> <td>Confounder berücksichtigt/diskutiert, ggf.</td> <td>k.A.</td> </tr> </tbody> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung	+	(„keine“, obwohl Angestellte von Helios)		Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveau etc.)	-	Confounder berücksichtigt/diskutiert, ggf.	k.A.
Angaben zur Finanzierung oder Interessenkonfliktserklärung	+														
(„keine“, obwohl Angestellte von Helios)															
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.														
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveau etc.)	-														
Confounder berücksichtigt/diskutiert, ggf.	k.A.														

<sup>18</sup> EBM-Terminologie zur Konkretisierung einer spezifischen Frage für die systematische Literaturanalyse [114]. Die Buchstaben haben folgende Bedeutungen: P=Population, I=Intervention, C=Control, O=Outcome

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension	Studienlimitation/ Verzerrungsrisiko																			
				Ergebnis der Evaluierung																				
				Bundesamtes  Effekt: ja positiv	<table border="1"> <tr><td>Adjustierung erfolgt</td><td></td></tr> <tr><td>Beobachtungszeitraum adäquat</td><td>+</td></tr> <tr><td>Population repräsentativ</td><td>+</td></tr> <tr><td>vergleichbare Gruppen</td><td>+</td></tr> <tr><td>Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten</td><td>k..A.</td></tr> <tr><td>adäquate Kontrollgruppe</td><td>n.a.</td></tr> <tr><td>Interventionszeitpunkt klar beschrieben</td><td>+</td></tr> </table>	Adjustierung erfolgt		Beobachtungszeitraum adäquat	+	Population repräsentativ	+	vergleichbare Gruppen	+	Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k..A.	adäquate Kontrollgruppe	n.a.	Interventionszeitpunkt klar beschrieben	+					
Adjustierung erfolgt																								
Beobachtungszeitraum adäquat	+																							
Population repräsentativ	+																							
vergleichbare Gruppen	+																							
Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k..A.																							
adäquate Kontrollgruppe	n.a.																							
Interventionszeitpunkt klar beschrieben	+																							
2  Nimptsch et al., 2013 [129]	Vorher-Nachher-Studie ohne Kontrollgruppe, retrospektiv	Vergleich der Daten vor und nach Einführung des Verfahrens  QM-System mit Messung von Qualitätsindikatoren auf Basis von administrativen Daten, monatliches Feedback an die Häuser, öffentliche Berichterstattung der Daten, halbjährliche Qualitätskonferenzen und Analyse und Verbesserung durch ein Peer-Review-Verfahren  Helios Kliniken GmbH  fächertübergreifend	Januar 2003 bis Dezember 2012 mit Follow-up über 5 Jahre. Beginn der Beobachtung in Abhängigkeit Eintritt in den Klinikkonzern.  18 Kliniken 12.000 bis 32.000 Fälle je nach Krankheit  Routinedaten	Ergebnisqualität  Endpunkt: Standardisierte Mortalitätsrate (SMR)  2 Subgruppen gesondert betrachtet : Subgruppe der Kliniken mit initial suboptimalen Werten: signifikante Verbesserung der SMR bis unter den Bundesdurchschnitt.  Bei der Subgruppe der Häuser mit initial bereits guten Werten: kein signifikanter Unterschied  -Herzinfarkt: 754 Fälle in Subgruppe suboptimal → Reduktion vorher-nachher SMR von 1,37 auf 0,55.  -Herzinsuffizienz: 2.093 Fälle in Subgruppe suboptimal → Reduktion vorher-nachher SMR von 1,37 auf 0,55  -Schlaganfall: 1.709 Fälle in Subgruppe initial suboptimal → Reduktion vorher-nachher SMR von 1,48 auf 0,92  -Pneumonie: 2152 Fälle in Subgruppe initial suboptimal →	Verzerrungsrisiko: hoch  Limitation: k.A., ob Subgruppen a priori definiert wurden. Positiv: Berücksichtigung von Confoundern wie Komplexität des Falls, Case-mix und Faktoren wie Verlegungen, Dauer des stationären Aufenthalts (Komorbidität nicht berücksichtigt).  „Clinical complexity level“ war in der Gruppe nach Einführung des QM bei Pneumonie und Herzinfarkt signifikant höher, die Dauer des stationären Aufenthalts bei Herzinsuffizienz und Pneumonie signifikant verkürzt und zum Teil vorübergehend erhöhte Verlegungsraten, sodass eine Beeinflussung der Ergebnisse in Betracht gezogen werden muss.  Die Autoren erhalten dauerhafte Fördermittel von Helios. Helios war nicht an der Auswertung beteiligt.	<table border="1"> <tr><td>Angaben zur Finanzierung oder Interessenkonfliktserklärung</td><td>+</td></tr> <tr><td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td><td>n.a.</td></tr> <tr><td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveau etc.)</td><td>+</td></tr> <tr><td>vergleichbare Gruppen</td><td>(-)</td></tr> <tr><td>Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td><td>+</td></tr> <tr><td>Interventionszeitpunkt klar beschrieben</td><td>+</td></tr> <tr><td>Beobachtungszeitraum adäquat</td><td>(+)</td></tr> <tr><td>Population repräsentativ</td><td>+</td></tr> <tr><td>Erfassung/Berücksichtigung von inkompletten</td><td>k.A.</td></tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung	+	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	n.a.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveau etc.)	+	vergleichbare Gruppen	(-)	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	+	Interventionszeitpunkt klar beschrieben	+	Beobachtungszeitraum adäquat	(+)	Population repräsentativ	+	Erfassung/Berücksichtigung von inkompletten	k.A.
Angaben zur Finanzierung oder Interessenkonfliktserklärung	+																							
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	n.a.																							
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveau etc.)	+																							
vergleichbare Gruppen	(-)																							
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	+																							
Interventionszeitpunkt klar beschrieben	+																							
Beobachtungszeitraum adäquat	(+)																							
Population repräsentativ	+																							
Erfassung/Berücksichtigung von inkompletten	k.A.																							

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko																
				Reduktion vorher-nachher SMR von 1.28 auf 0.65  Effekt: ja positiv	oder fehlenden Daten																
Rink O et al, 2012 [147; 148]  3 + 4	Vorher-Nachher-Studie ohne Kontrollgruppe, retrospektiv	Peer-Review  Vergleich der Daten vor und nach Einführung des Verfahrens  Gemeinnütziger Verein Initiative Qualitätsmedizin (IQM) - aus Verfahren Helios entsprungen (siehe oben)  fächerübergreifend	2009 bis 2011  Herzinsuffizienz an 13 Kliniken mit 244 Fällen;  Herzinfarkt in 14 Kliniken mit 275 Fällen;  Hirminfarkt an 26 Kliniken mit 377 Patientenverläufen  Routinedaten	Ergebnisqualität  -Endpunkt: Mortalität  Reduktion bei Herzinfarkt von 12,2 % in 2008 auf 6,2% in 2010 und weiter auf 5,3% in 2011  bei Herzinsuffizienz von 19,8% vor dem Review auf 10,5% in 2010 und 11,7% in 2011  bei Schlaganfall von 15,8% in 2008 (oder 2009?) auf 14,2% in 2010 und 9,1% in 2011  Effekt: ja positiv	Verzerrungsrisiko: hoch <table border="1" data-bbox="1503 520 2058 951"> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung?</td> <td>+ Angestellte von Helios</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>k.A.</td> </tr> <tr> <td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td>-</td> </tr> <tr> <td>Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td> <td>(+)</td> </tr> <tr> <td>Interventionszeitpunkt klar beschrieben</td> <td>(+)</td> </tr> <tr> <td>Beobachtungszeitraum adäquat</td> <td>+</td> </tr> <tr> <td>Population repräsentativ</td> <td>+</td> </tr> <tr> <td>vergleichbare Gruppen</td> <td>+</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung?	+ Angestellte von Helios	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	-	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	(+)	Interventionszeitpunkt klar beschrieben	(+)	Beobachtungszeitraum adäquat	+	Population repräsentativ	+	vergleichbare Gruppen	+
Angaben zur Finanzierung oder Interessenkonfliktserklärung?	+ Angestellte von Helios																				
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.																				
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	-																				
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	(+)																				
Interventionszeitpunkt klar beschrieben	(+)																				
Beobachtungszeitraum adäquat	+																				
Population repräsentativ	+																				
vergleichbare Gruppen	+																				
[153] Ruchholtz S et al, 2008  5	„Vorher-Nachher-Studie“ mit Vergleich zu Beginn und im Verlauf der Interventionszeit, retrospektiv, keine Kontrollgruppe	TraumaRegister der Deutschen Gesellschaft für Unfallchirurgie - anonym erhobene Daten vom gesamten Behandlungsprozess von Unfallpatienten mit Einführung von strukturiertem Feedback 1999  QM-System	Zeitraum von 1999 bis 2005  105 Kliniken mit 11.013 Fällen  Registerdaten	Ergebnisqualität  Endpunkt: Mortalitätsraten  1999 Anteil der tatsächlichen Mortalität versus der prognostisch erwarteten 22,8% versus 22,0%; 2005 Anteil 18,7% versus 22,1%  Effekt: ja positiv	Verzerrungsrisiko: sehr hoch <table border="1" data-bbox="1503 1050 2058 1374"> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung</td> <td>+ keine</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>k.A.</td> </tr> <tr> <td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td>(+)</td> </tr> <tr> <td>Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td> <td>(+)</td> </tr> <tr> <td>Beobachtungszeitraum adäquat</td> <td>+</td> </tr> <tr> <td>Population repräsentativ</td> <td>+</td> </tr> <tr> <td>vergleichbare Gruppen?</td> <td>+</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung	+ keine	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	(+)	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	(+)	Beobachtungszeitraum adäquat	+	Population repräsentativ	+	vergleichbare Gruppen?	+		
Angaben zur Finanzierung oder Interessenkonfliktserklärung	+ keine																				
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.																				
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	(+)																				
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	(+)																				
Beobachtungszeitraum adäquat	+																				
Population repräsentativ	+																				
vergleichbare Gruppen?	+																				

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko																		
6 Rucholtz S et al., 2004 [151]	„Vorher-Nachher Studie“ mit Vergleich zu Beginn und im Verlauf der Interventionszeit, retrospektiv, keine Kontrollgruppe	Siehe oben TraumaRegister DGU	1.1.1999 bis 31.12.2001  Daten von 6936 Patienten des Registers und per Fragebogen Erhebung der Veränderungen aus 21 Kliniken	Ergebnisqualität und Prozessqualität  In 85 % der Krankenhäuser wurden seit Einführung des TraumaRegisters qualitätsverbessernde Maßnahmen als Folge der Rückmeldung aus dem Datenregister eingeführt, insbesondere im Rahmen der Schockraumversorgung  Endpunkt: Zeitdauer diagnostischer Maßnahmen: signifikante Verkürzung (p<0.05)  Zeitdauer bis zur Not-Operation bzw. Aufnahme auf Intensivstation signifikante verkürzt (p<0.05) Beatmungsdauer und Aufenthaltsdauer Intensivstation signifikant verkürzt.  Tatsächliche versus erwartete Letalität: Im gesamten Beobachtungszeitraum tatsächliche Letalität niedriger, jedoch im Verlauf kein signifikanter Unterschied.  Effekt positiv: Prozessqualität ja Ergebnisqualität nein	Verzerrungsrisiko: sehr hoch <table border="1" data-bbox="1503 448 2056 863"> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung</td> <td>+ keine</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>n.a.</td> </tr> <tr> <td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td>+</td> </tr> <tr> <td>vergleichbare Gruppen</td> <td>k.A.</td> </tr> <tr> <td>Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td> <td>k.A.</td> </tr> <tr> <td>Interventionszeitpunkt klar beschrieben</td> <td>+</td> </tr> <tr> <td>Beobachtungszeitraum adäquat</td> <td>+</td> </tr> <tr> <td>Population repräsentativ</td> <td>+</td> </tr> <tr> <td>Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten</td> <td>k.A.</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung	+ keine	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	n.a.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+	vergleichbare Gruppen	k.A.	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	k.A.	Interventionszeitpunkt klar beschrieben	+	Beobachtungszeitraum adäquat	+	Population repräsentativ	+	Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.
Angaben zur Finanzierung oder Interessenkonfliktserklärung	+ keine																						
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	n.a.																						
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+																						
vergleichbare Gruppen	k.A.																						
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	k.A.																						
Interventionszeitpunkt klar beschrieben	+																						
Beobachtungszeitraum adäquat	+																						
Population repräsentativ	+																						
Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.																						
7 Zernikow B et al, 2008 [199]	Prospektive, multizentrische Kohortenstudie mit Kontrollgruppe (Vergleich zu Beginn und im Verlauf der Interventions-	QM-Programm: Schmerztherapie in der pädiatrischen Onkologie (STOP)-Programm mit Fortbildungsmaßnahmen, Audit, strukturiertes Feedback alle 6 Monate, basierend auf Qualitätsindikatoren zur	1997 bis 2000  8 Abteilungen mit QM (224 Patienten) versus 64 Abteilungen ohne QM  Zusätzlich erhobene Daten	Ergebnisqualität und Prozessqualität  Endpunkt: Unterschied in Schmerzkontrolle innerhalb der QM-Gruppe im Verlauf vom ersten zum letzten	Verzerrungsrisiko: hoch  Vergleichbarkeit der Gruppen QM versus non-QM ist auf Grund der fehlenden detaillierten Angaben der Population der non-QM-Abteilungen nur sehr eingeschränkt beurteilbar  Vergleich der Outcomes Schmerzkontrolle und Schmerzmittel-																		

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm) <b>Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet</b>	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) <b>Datengrundlage</b>	Endpunktdefinition/ Qualitätsdimension  <b>Ergebnis der Evaluierung</b>	Studienlimitation/ Verzerrungsrisiko														
	zeit)	Schmerzkontrolle (selbst definiert auf Literaturbasis und Expertenmeinung)  Vergleich von Abteilungen mit und ohne QM Programm  Initiator: QM-Gruppe der Deutschen Gesellschaft für pädiatrische Hämatologie und Onkologie	und Krankenakten	Drittel der Studie:  signifikante Reduktion der Häufigkeit von starken Schmerzen $p < 0.001$ )  Verbrauch Opiode/Non-Opiode kein signifikanter Unterschied.  Selbsteinschätzung von Pat. und Eltern bzgl. Schmerzkontrolle: kein signifikanter Unterschied in der Schmerzkontrolle, signifikant mehr Unzufriedenheit bzgl. Schmerzkontrolle bei schmerzhaften Prozeduren ( $p < 0.027$ ) in der QM-Gruppe. Wissenszuwachs der Ärzte: signifikanter Unterschied zugunsten QM-Gruppe  Effekt: inkonsistent für beide Qualitätsdimensionen	verbrauch erfolgte lediglich innerhalb der QM-Gruppe im zeitlichen Verlauf ohne Kontrollgruppe <table border="1" data-bbox="1503 421 2058 826"> <tr> <td>Angaben zur Finanzierung oder Interessenkonflikterklärung</td> <td>+ viele Angaben zu COI</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>(+)</td> </tr> <tr> <td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveau etc.)</td> <td>+</td> </tr> <tr> <td>Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td> <td>(-)</td> </tr> <tr> <td>Beobachtungszeitraum adäquat</td> <td>+</td> </tr> <tr> <td>Population repräsentativ</td> <td>(+)</td> </tr> <tr> <td>vergleichbare Gruppen</td> <td>(+)</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonflikterklärung	+ viele Angaben zu COI	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	(+)	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveau etc.)	+	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	(-)	Beobachtungszeitraum adäquat	+	Population repräsentativ	(+)	vergleichbare Gruppen	(+)
Angaben zur Finanzierung oder Interessenkonflikterklärung	+ viele Angaben zu COI																		
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	(+)																		
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveau etc.)	+																		
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	(-)																		
Beobachtungszeitraum adäquat	+																		
Population repräsentativ	(+)																		
vergleichbare Gruppen	(+)																		
Stoffel MP et al, 2004 Part I [175]  8	„Vorher-Nachher Studie“ mit Vergleich zu Beginn und im Verlauf der Interventionszeit, prospektiv?, keine Kontrollgruppe	QiN – Qualität in der Nephrologie: externes QM-System mit Datenerhebung in einer zentralen Datenbank für terminale Nierenerkrankungen und 3monatigen strukturierten einrichtungsspezifischen Bericht in pseudonymisierter Form bzgl. der Ergebnisse der Qualitätsindikatoren und Daten zum Benchmark  Vergleich zwischen Daten 1999 bei Einführung des Programms und Daten 2003  QiN Träger: 2 unabhängige Institutionen der Universität Köln (Klinik für Nephrologie und Institut für Gesundheitsökonomie und Klini-	2.Quartal 1999 bis 2.Quartal 2003  11 Dialyseeinheiten mit 600 Patienten zu Beginn, am Ende 7000 in 111 Dialyseeinrichtungen  Datengrundlage Routine-daten	Prozessqualität  <u>Geplante</u> Endpunkte des Programms:  -Lebensqualität (1x/Jahr) über validierte Messbögen -Morbidität (Häufigkeit und Dauer von Krankenhausaufenthalten) -Mortalität -Daten zur Prozessqualität (Beurteilungsparameter für die der Dialyseeffizienz), Blutdruck, Laborparameter zur renalen Anämie, zum Calcium/Phosphat-Haushalt, Ernährung/Inflammation)  <u>Dargestellte</u> Endpunkte:	Verzerrungsrisiko: sehr hoch  keine Kontrollgruppe, freiwillige Teilnahme, Selektionbias  keine Angaben zur Ergebnisqualität, nur zur Prozessqualität, nur Surrogatparameter  Daten zur Mortalität, Lebensqualität und Morbidität wurden nicht dargestellt <table border="1" data-bbox="1503 1161 2058 1383"> <tr> <td>Angaben zur Finanzierung oder Interessenkonflikterklärung</td> <td>(+) Finanzierung ja, zu COI nein</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>+</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonflikterklärung	(+) Finanzierung ja, zu COI nein	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	+										
Angaben zur Finanzierung oder Interessenkonflikterklärung	(+) Finanzierung ja, zu COI nein																		
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	+																		

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko											
		sche Epidemiologie)		<p>Qualitätsindikatoren für Dialyседurchführung (Dauer der Dialyse von 4,22 auf 4,37 h, Kt/V 1,31 auf 1,50) und für Anämie von 10,6 auf 11,5 g/dl statistisch signifikant (p&lt;0.05) verbessert</p> <p>Für Blutdruck diskret signifikante Verbesserung (142/77 auf 138/74 mmHg)</p> <p>Für die anderen Parameter keine Unterschiede</p> <p>Effekt: inkonsistent</p>	<table border="1"> <tr> <td data-bbox="1503 344 1962 421">Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td data-bbox="1962 344 2063 421">+</td> </tr> <tr> <td data-bbox="1503 421 1962 472">Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td> <td data-bbox="1962 421 2063 472">(+) </td> </tr> <tr> <td data-bbox="1503 472 1962 497">Beobachtungszeitraum adäquat</td> <td data-bbox="1962 472 2063 497">+</td> </tr> <tr> <td data-bbox="1503 497 1962 523">Population repräsentativ</td> <td data-bbox="1962 497 2063 523">+</td> </tr> <tr> <td data-bbox="1503 523 1962 549">vergleichbare Gruppen</td> <td data-bbox="1962 523 2063 549">-</td> </tr> </table>		Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	(+)	Beobachtungszeitraum adäquat	+	Population repräsentativ	+	vergleichbare Gruppen	-
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+															
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	(+)															
Beobachtungszeitraum adäquat	+															
Population repräsentativ	+															
vergleichbare Gruppen	-															
<p>Stoffel MP et al, 2004 Part II [176]</p> <p>9</p>	<p>retrospektive Kohortenstudie im Vergleich zwischen 2 Subgruppen</p> <p>"Vorher-Nachher-Studie" mit Vergleich zu Beginn und im Verlauf der Interventionszeit</p>	<p>QiN s.o Vergleich der Qualität zwischen Einrichtungen mit früher (zu Beginn (Q2/99) versus später (Q1/02) Programmteilnahme</p>	<p>1999 bis 2002</p> <p>40 „frühe“ versus 11 „späte“ Dialyseedeinrichtungen</p>	<p>Prozessqualität</p> <p>marginal schlechtere Werte der „späten“ Gruppe (keine Signifikanzwerte dargestellt) der Mittelwerte der Beurteilungsparameter für die Dialyseedeffizienz (Zeitpunkt Q1/2002) zwischen den Gruppen:</p> <p>-URR (%) (Urea Reduction ratio) 70.2 versus 69.2 -Kt/V: 1,50 versus 1,46 -Dauer (h) Dialyse 4,47 versus 4,31</p> <p>Signifikante Verbesserung aller o.g. Werte in <u>beiden</u> Gruppen von 1/2002 bis 4/2002</p> <p>Effekt: inkonsistent</p>	<p>Verzerrungsrisiko: sehr hoch</p> <p>keine adäquate Kontrollgruppe, freiwillige Teilnahme, Selektionsbias</p> <p>keine Angaben zur Ergebnisqualität, nur zur Prozessqualität, nur Surrogatparameter. Statistisch signifikante Werte, die klinisch fraglich relevant</p> <p>Daten zur Mortalität, Lebensqualität und Morbidität wurden nicht dargestellt</p> <table border="1"> <tr> <td data-bbox="1503 1038 1962 1187">Angaben zur Finanzierung oder Interessenkonfliktserklärung</td> <td data-bbox="1962 1038 2063 1187">(+) Finanzierung ja, zu COI nein</td> </tr> <tr> <td data-bbox="1503 1187 1962 1264">Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td data-bbox="1962 1187 2063 1264">k.A.</td> </tr> <tr> <td data-bbox="1503 1264 1962 1340">Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td data-bbox="1962 1264 2063 1340">+</td> </tr> <tr> <td data-bbox="1503 1340 1962 1386">Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td> <td data-bbox="1962 1340 2063 1386">(+) </td> </tr> </table>		Angaben zur Finanzierung oder Interessenkonfliktserklärung	(+) Finanzierung ja, zu COI nein	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	(+)		
Angaben zur Finanzierung oder Interessenkonfliktserklärung	(+) Finanzierung ja, zu COI nein															
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.															
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+															
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	(+)															

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko														
					<table border="1"> <tr><td>Kontrollgruppe adäquat</td><td>-</td></tr> <tr><td>Beobachtungszeitraum adäquat</td><td>+</td></tr> <tr><td>Population repräsentativ</td><td>-</td></tr> <tr><td>vergleichbare Gruppen</td><td>k.A.</td></tr> </table>	Kontrollgruppe adäquat	-	Beobachtungszeitraum adäquat	+	Population repräsentativ	-	vergleichbare Gruppen	k.A.						
Kontrollgruppe adäquat	-																		
Beobachtungszeitraum adäquat	+																		
Population repräsentativ	-																		
vergleichbare Gruppen	k.A.																		
<p>Wallwienner M et al, 2012 UND Brucker SY et al, 2008 [22; 194]</p> <p>10 + 11</p>	<p>„Prospektive, multizentrische Interventionsstudie (3 Phasen)“</p> <p>„Vorher-Nachher-Studie“ mit Vergleich zu Beginn und im Verlauf der Interventionszeit, prospektiv, keine Kontrollgruppe</p>	<p>Qualitätssicherungsprogramm: Bildung von zertifizierten, multidisziplinären Brustzentren und Benchmarking-Programm – Ranking auf Basis von Qualitätsindikatoren</p> <p>Datenkollektion und- analyse durch unabhängige externe Organisation Westdeutsches Brust-Centrum (WBC) beauftragt durch Deutsche Krebsgesellschaft und Deutsche Gesellschaft für Senologie</p> <p>Zentren erhalten halbjährlich Feedback-Berichte ihrer eigenen Leistung und der Leistung im Vergleich mit den anderen Teilnehmern in anonymer Form</p> <p>Trägerschaft: Deutsche Krebsgesellschaft und Deutsche Gesellschaft für Senologie</p>	<p>2003 bis 2010 Auswertungszeitraum, Studie noch fortlaufend.</p> <p>Teilnehmende Zentren aus Deutschland, Schweiz, Österreich und deutschsprachigem Italien</p> <p>Ansteigende Teilnahme von 59 auf 210 Zentren oder 5.994 auf 34.678 Brustkrebsfälle</p> <p>Zusätzlich erhobene Daten für eigene Datenbank</p>	<p>Prozessqualität</p> <p>Erhebung von 173 Parametern, zusammengefasst in 11 Qualitätsindikatoren-Sets für die Messung der Prozessqualität in Konkordanz mit Leitlinie. Prozess Qualitätsindikatoren als Surrogatparameter für Ergebnisqualität</p> <p>Verbesserung aller Qualitätsindikatoren im Verlauf:</p> <p>Auswahl: -präoperative histologische Bestätigung der Diagnose: 58 auf 96% -Leitliniengerechte endokrine Therapie: 27 auf 97% - Leitliniengerechte adjuvante oder neoadjuvante Chemotherapie: 32 auf 78%</p> <p>Effekt: ja positiv</p>	<p>Verzerrungsrisiko: sehr hoch</p> <p>Selektionsbias möglich, da freiwillige Teilnahme</p> <p>Wenig Angaben zur Auswertung keine Angaben zu den Begleitvariablen als mögliche Confounder (Alter, Komorbidität) bzw. Adjustierungsverfahren (Stratifizierung etc.)</p> <p>Kausalität auf Grund des Studiendesigns nicht belegt. Multifaktorielle Genese der Verbesserung anzunehmen (verbesserte Dokumentation etc.)</p> <p>Nur Surrogatparameter der Prozessqualität (Qualitätsindikatoren auf Leitlinien-Basis)</p> <p>Limitation der Autoren: geringe Anzahl an Teilnehmern zu Beginn des Programms. S3 Leitlinie war zu Beginn des Programms noch nicht vorhanden, wurde erst 2004 publiziert</p> <table border="1"> <tr><td>Angaben zur Finanzierung oder Interessenkonfliktserklärung</td><td>+</td></tr> <tr><td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td><td>k.A.</td></tr> <tr><td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td><td>(+)</td></tr> <tr><td>Confounder berücksichtigt/diskutiert, Adjustierung erfolgt</td><td>k.A.</td></tr> <tr><td>Beobachtungszeitraum adäquat</td><td>+</td></tr> <tr><td>Population repräsentativ</td><td>+</td></tr> <tr><td>vergleichbare Gruppen</td><td>k.A.</td></tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung	+	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	(+)	Confounder berücksichtigt/diskutiert, Adjustierung erfolgt	k.A.	Beobachtungszeitraum adäquat	+	Population repräsentativ	+	vergleichbare Gruppen	k.A.
Angaben zur Finanzierung oder Interessenkonfliktserklärung	+																		
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.																		
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	(+)																		
Confounder berücksichtigt/diskutiert, Adjustierung erfolgt	k.A.																		
Beobachtungszeitraum adäquat	+																		
Population repräsentativ	+																		
vergleichbare Gruppen	k.A.																		
Brucker	Prospektive, multi-	Qualitätssicherungsprogramm durch	2003 bis 2009	Prozessqualität	Verzerrungsrisiko: sehr hoch														

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko														
SY et al, 2011 [23] Subgruppenanalyse für Radiotherapie  12	zentrische Interventionsstudie (3 Phasen) - Subgruppenanalyse für Radiotherapie  „Vorher-Nachher-Studie“ mit Vergleich zu Beginn und im Verlauf der Interventionszeit, keine Kontrollgruppe (k.A. zu pro- oder retrospektiv)	Bildung von zertifizierten, multidisziplinären Brustzentren und Benchmarking-Programm – Ranking auf Basis von Qualitätsindikatoren  siehe oben	Auswertungszeitraum, Studie noch fortlaufend  Teilnehmende Zentren aus Deutschland, Schweiz, Österreich und deutschsprachigem Italien  Ansteigende Teilnahme von 59 auf 210 Zentren oder 5.994 auf 34.678 Brustkrebsfälle	Verbesserung aller Qualitätsindikatoren zur Radiotherapie im Verlauf.  Auswahl: -Radiotherapie nach Brusterhalten-der Operation von 21 auf 89% -Radiotherapie nach Mastektomie von 10 auf 88% -Radiotherapie nach DCIS (Duktales Carcinoma in situ) von 65 auf 75%  Effekt: ja positiv	siehe oben <table border="1"> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung</td> <td>-</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>k.A.</td> </tr> <tr> <td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td>s.o.</td> </tr> <tr> <td>Confounder berücksichtigt/diskutiert?, Adjustierung erfolgt</td> <td>k.A.</td> </tr> <tr> <td>Beobachtungszeitraum adäquat</td> <td>+</td> </tr> <tr> <td>Population repräsentativ</td> <td>+</td> </tr> <tr> <td>vergleichbare Gruppen</td> <td>n.a.</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung	-	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	s.o.	Confounder berücksichtigt/diskutiert?, Adjustierung erfolgt	k.A.	Beobachtungszeitraum adäquat	+	Population repräsentativ	+	vergleichbare Gruppen	n.a.
Angaben zur Finanzierung oder Interessenkonfliktserklärung	-																		
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.																		
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	s.o.																		
Confounder berücksichtigt/diskutiert?, Adjustierung erfolgt	k.A.																		
Beobachtungszeitraum adäquat	+																		
Population repräsentativ	+																		
vergleichbare Gruppen	n.a.																		
Blum K et al, 2006 [14; 51]  13 + 14	Querschnittsstudie (keine Kontrollgruppe für DMP-Auswirkungen; nur Kontrollgruppe für strukturellen Vergleich)	Krankenhäuser mit Teilnahme am Disease Management Programm Brustkrebs im Vergleich zu Nicht-DMP-Häusern oder im Veränderung im Verlauf  Initiative des Sachverständigenrates für die Konzertierte Aktion im Gesundheitswesen, nachfolgend gesetzliche Grundlage durch SGB V die § 137f und 137g  Seit 2004 Gemeinsamer Bundesausschuss zuständig	2002 bis Mitte 2005  Beteiligung an Erhebung 176 DMP-Krankenhäuser (Rücklaufquote von 45%)  Insgesamt antworteten 93 Nicht-DMP-Häuser (Ausschöpfung von 33,7%)  zusätzlich erhobene Daten	Prozessqualität und Strukturqualität  Befragung  Struktureller Vergleich DMP-versus Nicht-DMP-Häuser (Betten, Fallzahlen etc.)  Keine Evaluationsergebnisse zur Wirksamkeit, nur Erhebung zu Veränderungen/Auswirkungen im Verlauf der DMP-Häuser (keine Kontrollgruppe)  Angabe bei einer Mehrheit von jeweils knapp zwei Drittel der Häuser gar keine Änderungen der Versorgung durch das DMP Brustkrebs Bei weiteren 15% nur wenige Auswirkungen bei den genannten Parametern Bei jeweils ca. 10-15% Angabe von mittelmäßigen bis sehr großen	Verzerrungsrisiko: hoch <table border="1"> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung? (DKI beauftragt DKG)</td> <td></td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>k.A.</td> </tr> <tr> <td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td>-</td> </tr> <tr> <td>Confounder berücksichtigt/diskutiert?, Adjustierung erfolgt</td> <td>wenig</td> </tr> <tr> <td>Beobachtungszeitraum adäquat</td> <td>n.a.</td> </tr> <tr> <td>Population repräsentativ</td> <td>+</td> </tr> <tr> <td>vergleichbare Gruppen</td> <td>n.a.</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung? (DKI beauftragt DKG)		Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	-	Confounder berücksichtigt/diskutiert?, Adjustierung erfolgt	wenig	Beobachtungszeitraum adäquat	n.a.	Population repräsentativ	+	vergleichbare Gruppen	n.a.
Angaben zur Finanzierung oder Interessenkonfliktserklärung? (DKI beauftragt DKG)																			
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.																		
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	-																		
Confounder berücksichtigt/diskutiert?, Adjustierung erfolgt	wenig																		
Beobachtungszeitraum adäquat	n.a.																		
Population repräsentativ	+																		
vergleichbare Gruppen	n.a.																		

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko
				Veränderungen. Erfasst wurden bei der Befragung u.a Veränderungen bei diagnostischen oder therapeutischen Standards, bei operativen Verfahren und bei der Verwendung von evidenzbasierten Leitlinien  Effekt: inkonsistent	
Lehmkuhl et al., 2011 [116]  15	Vorher-Nachher-Studie mit geplanter Kontrollgruppe, (prospektiv für Gruppe 1, retrospektiv für Gruppe 2)	Gruppe (G) 1: Qualitätsverbesserung in der postoperativen Schmerztherapie (QUIPS) – Benchmarking- Projekt (Qualitätsindikatoren: Datenerhebung, -analyse von Patientenangaben zur Schmerztherapie mit Feedback an die Kliniken)  Gruppe (G) 2: Initiative Schmerzfreie Klinik mit Zertifikat QM Akutschmerz (durch TÜV Rheinland vergeben)  Gruppe (G) 3: Kontrollgruppe – keine Beteiligung an beiden QM/QS-Maßnahmen  Vergleich vor und nach Einführung der Intervention Gruppe 2	Anonymisierte Datenerhebung und Auswertung  G1: Q 1/2007 bis 4/2008 Traumatologie, Orthopädie, Allgemeinchirurgie mit 26125 Datensätzen  G2: 7 Kliniken Kardio-, Viszeral-, Allgemein-, plastische und Unfallchirurgie, Orthopädie, Gynäkologie, Urologie mit 2505 Datensätzen (1064 vor und 1441 nach Einführung im Zeitraum 18 Monate vor und 18 Monate nach Einführung - 2007 bis 2009)  G3: 434 Patienten 2005 bis 2007 für 4 Wochen - Messung der schmerztherapeutischen Ergebnisqualität aus Allgemein-, Viszeral- und Unfallchirurgie  Datensatz: standardisierte Erhebung mittels validiertem Fragebogen mit Qualitätsindikatoren basierend auf S3 Leitlinie	Ergebnisqualität  patientenseitige Bewertung bzgl. 1)postoperativer Schmerzen 2)Nebenwirkungen 3)schmerzbedingte Einschränkungen 4) Patientenzufriedenheit (Skala 1-15) mit Schmerztherapie  im Verlauf bei G1 und G2 (G3 kein Nachher-Vergleich)  G1: signifikante Verbesserung 1)nein 2)nein 3)ja 4)ja mit 11.9 auf 12.4 Mittelwert  G2: 1)ja 2)nein 3)ja 4)ja mit 13.2 auf 13.7 Mittelwert  G3: Patientenzufriedenheit mit 8.2 Mittelwert (Skala 1-10)  Effekt: inkonsistent	Verzerrungsrisiko hoch  Limitationen: - wegen der geringen Fallzahl in G3 kein Vergleich mit G1+ G2 - In G3 keine Follow-up Erhebung erfolgt - G3 Kontrollgruppe vor der Intervention 2005-2007 -Patientenzufriedenheit wurde in G1/2 mit anderer Skala gemessen, bei dieser in etwa gleich hohe Ausprägung - G3 daher keine adäquate Kontrollgruppe  - Keine Darstellung der Populationscharakteristik der Kliniken und Patienten (Alter, Komorbidität, ASA (American Society of Anaesthesiology score), Operationsart, andere QS-Maßnahmen etc.)

Angaben zur Finanzierung oder Interessenkonfliktserklärung	+ keine
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	+
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+
vergleichbare Gruppen	k.A.
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	k.A.
Interventionszeitpunkt klar beschrieben	+
Beobachtungszeitraum adäquat	+
Population repräsentativ	+
Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	+
adäquate Kontrollgruppe	-

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko																				
			zusätzlich erhobene Daten																						
16 Gastmeier 2002 [70]	„Vorher-Nachher-Studie“ mit Vergleich zu Beginn und im Verlauf der Interventionszeit, prospektiv, mit Kontrollgruppe	QM zur Reduktion von nosokomialen Infektionen in 4 Krankenhäusern mit und 4 Kontrollen ohne QM  Fokus operative Medizin  QM: Qualitätszirkel mit Entwicklung von QS-Instrumenten und Erhebung von nosokomialen Ereignissen  BMG geförderte Studie	26 Wochen Zeitraum mit 3 Beobachtungszeiträumen (OP1, OP2, OP3)  und  2 Interventionsvergleichen:  Zeitverlauf 1 zwischen Beobachtungszeiträumen OP1/OP2  Zeitverlauf 2 zwischen Beobachtungszeiträumen OP1/OP3  Vergleich der Verbesserung zwischen Interventions- und Kontrollgruppe 11.295 Patientenfälle insgesamt  zusätzlich erhobene Daten	Ergebnisqualität  Primärer Endpunkt: Ereignisrate an nosokomialen Infektionen (NI)  Vergleich der Verbesserung zwischen Interventions- und Kontrollgruppe - bei Zeitverlauf 1 zwischen Beobachtungszeiträumen OP1/OP2 signifikanter Unterschied: Odds Ratio 0.75 (95% KI 0,58-0,97)  - bei Zeitverlauf 2 zwischen Beobachtungszeiträumen OP1/OP3 nicht-signifikanter Unterschied: Odds Ratio 0.78 (Konfidenzintervall 95% 0,60-1,01)  Effekt: inkonsistent	Verzerrungsrisiko moderat  Da Pat. im Kontrollklinik signifikant älter und höherer ASA Scores (American Society of Anaesthesiology Score), mit mehr Infektionen bei Aufnahme und längerem Aufenthalt auf Intensivstation ist das Ergebnis noch aussagekräftiger  <table border="1"> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung</td> <td>-</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>k.A.</td> </tr> <tr> <td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td>+</td> </tr> <tr> <td>vergleichbare Gruppen</td> <td></td> </tr> <tr> <td>Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td> <td>+</td> </tr> <tr> <td>Interventionszeitpunkt klar beschrieben</td> <td>k.A.</td> </tr> <tr> <td>Beobachtungszeitraum adäquat</td> <td>+</td> </tr> <tr> <td>Population repräsentativ</td> <td>+</td> </tr> <tr> <td>Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten</td> <td>k.A.</td> </tr> <tr> <td>adäquate Kontrollgruppe</td> <td>—</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung	-	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+	vergleichbare Gruppen		Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	+	Interventionszeitpunkt klar beschrieben	k.A.	Beobachtungszeitraum adäquat	+	Population repräsentativ	+	Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.	adäquate Kontrollgruppe	—
Angaben zur Finanzierung oder Interessenkonfliktserklärung	-																								
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.																								
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+																								
vergleichbare Gruppen																									
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	+																								
Interventionszeitpunkt klar beschrieben	k.A.																								
Beobachtungszeitraum adäquat	+																								
Population repräsentativ	+																								
Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.																								
adäquate Kontrollgruppe	—																								
17 Meissner W et al, 2008 [123]	Vorher-Nachher-Studie mit Vergleich zu Beginn und im Verlauf der Interventionszeit, prospektiv, keine Kontrollgruppe	Qualitätsverbesserung in der postoperativen Schmerztherapie (QUIPS) – Benchmarking- Projekt (Qualitätsindikatoren-Datenerhebung, Datenanalyse von Patientenangaben zur Schmerztherapie mit Feedback an die Kliniken)  Zu Beginn Förderung durch BMG, im Verlauf durch Fachgesellschaften der Chirurgie und Anästhesie	Zwischen Oktober 2003 bis September 2006  Teilnahme von 30 Abteilungen von 6 Krankenhäusern  Baseline Mitte 2004  Follow-up: 12.389 Datensätze  zusätzlich erhobene Daten	Ergebnisqualität  Maximale Schmerzintensität postoperativ gemessen mit der Numerischen Rating-Skala (NRS) mit 3 Messpunkten (Baseline, in der Mitte und um Ende):  Werte Vergleich Baseline versus letzte Messung (in Publikation E-Tabelle 1 u 2): - 2 Kliniken signifikante Besserung mit Schmerzreduktion auf NRS	Verzerrungsrisiko: hoch  p-Werte explorativ Confounder (Operation, Komorbidität etc.) nicht berücksichtigt, Selektionsbias wahrscheinlich  <table border="1"> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung (Erstautor viele Angaben bei COI und Letztautor Präsident der tragenden Fachgesellschaft)</td> <td>+</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>n.a.</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung (Erstautor viele Angaben bei COI und Letztautor Präsident der tragenden Fachgesellschaft)	+	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	n.a.																
Angaben zur Finanzierung oder Interessenkonfliktserklärung (Erstautor viele Angaben bei COI und Letztautor Präsident der tragenden Fachgesellschaft)	+																								
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	n.a.																								

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko															
				<p>(Klinik 1+6: Mittelwert NRS 4,4 auf 3,4 und 5,1 auf 3,9) - 2 Kliniken (Klinik 3+4) keine Änderung – nicht signifikante. Reduktion auf NRS - 2 Kliniken (Klinik 2+5) signifikante Verschlechterung mit Anstieg NRS 5,1 auf 5,5 und 4,4 auf 5,4</p> <p>nur zu Beginn in 5 von 6 Kliniken signifikante Verbesserung, nicht mehr im Verlauf</p> <p>Effekt: inkonsistent</p>	<table border="1"> <tr> <td data-bbox="1503 344 1966 421">Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td data-bbox="1966 344 2063 421">+</td> </tr> <tr> <td data-bbox="1503 421 1966 448">adäquate Gruppen</td> <td data-bbox="1966 421 2063 448">k.A..</td> </tr> <tr> <td data-bbox="1503 448 1966 496">Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td> <td data-bbox="1966 448 2063 496">k.A..</td> </tr> <tr> <td data-bbox="1503 496 1966 523">Interventionszeitpunkt klar beschrieben</td> <td data-bbox="1966 496 2063 523">+</td> </tr> <tr> <td data-bbox="1503 523 1966 550">Beobachtungszeitraum adäquat</td> <td data-bbox="1966 523 2063 550">+</td> </tr> <tr> <td data-bbox="1503 550 1966 577">Population repräsentativ</td> <td data-bbox="1966 550 2063 577">+</td> </tr> <tr> <td data-bbox="1503 577 1966 635">Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten</td> <td data-bbox="1966 577 2063 635">k.A.</td> </tr> </table>		Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+	adäquate Gruppen	k.A..	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	k.A..	Interventionszeitpunkt klar beschrieben	+	Beobachtungszeitraum adäquat	+	Population repräsentativ	+	Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+																			
adäquate Gruppen	k.A..																			
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	k.A..																			
Interventionszeitpunkt klar beschrieben	+																			
Beobachtungszeitraum adäquat	+																			
Population repräsentativ	+																			
Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.																			
<b>Systematischer Review</b>																				
Mayer et al., 2012 [121]  18	Systematischer Review	Gutachten des Kompetenz-Centrum Qualitätssicherung/ Qualitätsmanagement im Auftrag des GKV Spitzenverbands zur Aufarbeitung des Internen QM in der medizinischen Versorgung zur Umsetzung und Nutzenfrage im stationären und ambulanten Sektor national und international.	27 Volltexte als themenrelevante Treffer  Bis zum 1.11.2011 Recherchezeitraum  Datenbanken Pubmed, Internetsuchmaschinen und Handsuche  Keine methodische Studieneinschränkungen	Alle Qualitätsdimensionen  Fazit der Autoren: Keine aussagekräftigen Studien zu den Effekten der Einführung von internem QM in bundesdeutschen Kliniken  Effekt: inkonsistent	Keine methodische Studieneinschränkung, kaum Effektstärken (Signifikanz angegeben).  Verzerrungsrisiko: gering für den extrahierten Themenbereich des Gutachtens der methodischen Durchführung der Literaturanalyse, nicht der eingeschlossenen Einzelstudien (sämtliche gefundenen Studien von niedriger Qualität)  nur Daten vom stationären Bereich extrahiert:  Literaturbelege: unkontrollierte Studien - Vogel P et al, 2011 - Zernikow B et al, 2008 - Brand S et al., 2005 - Ruchholtz S et al, 2004 - Ruchholtz S et al, 1997 - Behrens S et al, 2002  kontrollierte Studien: - Lehmkuhl D et al, 2011 - Gastmeier P et al, 2002															

### 7.3 Evidenztabelle – Studiendarstellung international

PICO-Schema<sup>19</sup>

Population	Intervention	Kontrolle	Outcome
stationärer Sektor oder klinikähnliche Einrichtungen	QS/QM Programm (siehe E2)	keine Intervention oder Vergleichsgruppe zu Beginn der Intervention, wenn noch keine Wirkung zu erwarten (siehe E2)	Wirksamkeit hinsichtlich Parametern der Ergebnis-, Struktur- oder Prozessqualität

Legende :

+	Angabe vorhanden, trifft zu
-	nicht erfolgt
k.A.	keine Angabe
n.a	nicht anwendbar
KI	Konfidenzintervall
COI	Conflict of Interest

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko				
Canovas et al., 2009 [33]  19	RCT Vorher-Nachher-Vergleich  Spanien, Murcia	3 Gruppen randomisierte Zuteilung von 10 allgemeinmedizinischen Gesundheitszentren:  -Gruppe G1: selbstorganisiertes internes QM nach Schulung am Bsp. Verbesserung Behandlung unkomplizierter respiratorischer Infekte mit Leitlinien-Adhärenz  -Gruppe G2: externe QS-Erhebung mit Feedbackbericht ohne weitere Maßnahmen  -Gruppe G3: keinerlei QS/QM-Maßnahmen	Erste Messung vor und zweite Messung 6 Monate nach der Intervention  Erste Messung 401 bzw. zweite Messung 396 Krankenakten  Erste Messung: G1: 51 G2: 250 G3: 100  Analysejahr: k.A.	Prozessqualität  Endpunkte: Verbesserung der Kriterien 1) Körperliche Untersuchung 2) Dokumentation der Verschreibung 3) keine Verschreibung von Antibiotika, Antihistaminika und Glukokortikoiden  Erster Endpunkt: G1: Verbesserung; p=0.02 G2: Verbesserung; nicht signifikant G3: Verbesserung; p=0.01	Verzerrungsrisiko: moderat  wenig Angaben zu QM-Maßnahmen und Studienpopulation, geringe Fallzahl Angabe zur staatlichen Finanzierung, nicht zu COI des Autors  <table border="1" style="width: 100%;"> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung</td> <td>+ zur Finanzierung</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>+</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung	+ zur Finanzierung	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	+
Angaben zur Finanzierung oder Interessenkonfliktserklärung	+ zur Finanzierung								
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	+								

<sup>19</sup>EbM-Terminologie zur Konkretisierung einer spezifischen Frage für die systematische Literaturanalyse [114]. Die Buchstaben haben folgende Bedeutungen: P=Population, I=Intervention, C=Control, O=Outcome

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko		
		Vergleich Einfluss von interner QS versus externer QS mit Feedback versus keine QS-Maßnahme in allen drei Gruppen  Studie durch öffentliche Forschungsgelder gefördert	zusätzlich erhobene Daten	Zweiter Endpunkt: G1: Verbesserung; p=0.05 G2: Verbesserung; nicht signifikant G3: Verbesserung; nicht signifikant  Dritter Endpunkt: G1: Verbesserung; p=0.004 G2: Verbesserung; nicht signifikant G3: Verbesserung; nicht signifikant  Effekt: ja	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.) vergleichbare Gruppen Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt Interventionszeitpunkt klar beschrieben Beobachtungszeitraum adäquat Population repräsentativ Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten Randomisierung Concealment of Allocation	+ k.A. n.a. + (+) (+) k.A. +	
Snyder C, JAMA. 2005 [171]  20	Vorher-Nachher-Studie mit Kontrollgruppe, retrospektiv  USA	Einfluss von 4 übergeordneten staatlichen QM-Programmen in 5 Bundesstaaten in den USA auf Krankenhäuser mit freiwilliger Teilnahme (n= ca. 152) versus Krankenhäuser ohne QM (n= ca. 195) Bundesstaat: Utah, Nevada, Washington, Maryland, NewYork  Medicare (staatliche Krankenversicherung in USA) beauftragt Institutionen zur Qualitätssicherung. Verschiedene Methoden wie Schulungsmaterialien, Datenerhebung, Feedbackberichte zu Qualitätsindikator-Ergebnissen, Hilfe bei der Erstellung von Behandlungspfaden, SOP	1998 Baseline vor Einführung und 2000/2001 Follow-up  700 bis 800 Krankenakten pro Bundesstaat und pro Erhebung  Beginn der Einführung 1999  zusätzlich erhobene Daten auf Basis Krankenakten	Prozessqualität 15 Qualitätsindikatoren (basierend auf Leitlinien) aus 5 Krankheiten: Herzinfarkt, Pneumonie, Vorhofflimmern, Herzinsuffizienz, Schlaganfall  Kein signifikanter Unterschied in der Verbesserung von 14 von 15 Qualitätsindikatoren (nur bei einem Qualitätsindikator „Screening für Pneumokokkenimpfung“ Unterschied zugunsten der teilnehmenden Kliniken)  Gleiches Ergebnis auch für Subgruppenanalysen für einzelne Bundesstaaten nur explorativ  Effekt: nein	Verzerrungsrisiko: moderat  Subgruppenanalysen für einzelne Bundesstaaten nur explorativ ausgewertet, da zu geringe Fallzahlen  Auswertung im Auftrag der AHRQ	Angaben zur Finanzierung oder Interessenkonfliktserklärung Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.) vergleichbare Gruppen Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt Interventionszeitpunkt klar beschrieben Beobachtungszeitraum adäquat Population repräsentativ Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	+ keine n.a. + + + + + k.A.

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko																		
Gagneux 1998 [69]  21	Vorher-Nachher-Studie ohne Kontrollgruppe, prospektiv  Frankreich	Internes QS-Programm mit Erhebung von Qualitätskriterien, Feedback, Leitlinien-Anwendung, Standardisierung (SOPs), Schulung vor und nach Einführung im Jahr 1993  Setting: Traumazentrum/Notaufnahmestelle der Orthopädie, Unfallchirurgie und plastischen Chirurgie in Frankreich  Verantwortlicher Träger: k.A.	18 Monate 1993 und 1995  1995 n=1187 Fälle 1993 n= 2056 Fälle  zusätzlich erhobene Daten auf Basis Krankenakten	Prozessqualität  Auftreten von Diagnose-, Behandlungs- oder Organisationsfehler, die zu einem kritischen, unerwünschten Ereignis führen könnten  Signifikante (p<0.001) Verbesserung von 204 (9,9%) 1993 auf 71 (5,9%) 1995  Effekt: ja	Verzerrungsrisiko: moderat  keine Kontrollgruppe  <table border="1"> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung?</td> <td>-</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>-</td> </tr> <tr> <td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td>+</td> </tr> <tr> <td>vergleichbare Gruppen</td> <td>+</td> </tr> <tr> <td>Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td> <td>+</td> </tr> <tr> <td>Interventionszeitpunkt klar beschrieben</td> <td>+</td> </tr> <tr> <td>Beobachtungszeitraum adäquat</td> <td>+</td> </tr> <tr> <td>Population repräsentativ</td> <td>+</td> </tr> <tr> <td>Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten</td> <td>k.A.</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung?	-	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	-	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+	vergleichbare Gruppen	+	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	+	Interventionszeitpunkt klar beschrieben	+	Beobachtungszeitraum adäquat	+	Population repräsentativ	+	Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.
Angaben zur Finanzierung oder Interessenkonfliktserklärung?	-																						
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	-																						
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+																						
vergleichbare Gruppen	+																						
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	+																						
Interventionszeitpunkt klar beschrieben	+																						
Beobachtungszeitraum adäquat	+																						
Population repräsentativ	+																						
Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.																						
Guillamondegui et al, 2012 [85]  22	„Vorher-Nachher-Studie“ mit Vergleich zu Beginn und im Verlauf der Interventionszeit, prospektiv, keine Kontrollgruppe  USA	QS-Programm: National Surgical Quality Improvement Program (NSQIP) in der Tennessee Surgical Quality Collaborative  Teilnahme: 10 Krankenhäuser  Register mit risikoadjustierten Prozess- und Outcomedaten mit Feedback für weitere QS-Maßnahmen  Initiator: American College of Surgeons Gründung der Initiative in Tennessee Mai 2008  Vergleich zwischen Analysezeitraum 1 und 2	Analysezeitraum 1: Januar bis Dezember 2009  Analysezeitraum 2: Januar bis Dezember 2010  14.205 chirurgische Fälle in Analysezeitraum 1 und 14.901 chirurgische Fälle Analysezeitraum 2  Registerdaten	Ergebnisqualität  Daten von 20 Kategorien mit  primärer Endpunkt: postoperative Komplikationen  sekundärer Endpunkt: 30-Tages Mortalität  signifikante Verbesserung in 5 Kategorien:  -oberflächliche Wundinfektion (19%, p=0.0005), -Beatmung länger als 48 Stunden (15%, p=0.012), -Prothesen-/Transplantatversagen (60%, p<0.0001) -Akutes Nierenversagen (25%,	Verzerrungsrisiko: moderat  Umfassende Berücksichtigung von möglichen Confoundern (Risikofaktoren, soziodemografische Daten etc.) mit transparenter Darstellung: Vergleichbarkeit zwischen den 2 Gruppen)  Limitation: kaum Angaben zu den Maßnahmen, die die Kliniken als Konsequenz aus den Daten ziehen (Benchmarking, QS-Verfahren, Feedback, Etablierung SOPs etc.) und der Art und Weise des Feedbacks an die Kliniken  <table border="1"> <tr> <td>Angaben zur Finanzierung oder Interessenkonfliktserklärung</td> <td>+ keine</td> </tr> <tr> <td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td> <td>k.A.</td> </tr> <tr> <td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td> <td>+</td> </tr> <tr> <td>vergleichbare Gruppen</td> <td>+</td> </tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung	+ keine	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+	vergleichbare Gruppen	+										
Angaben zur Finanzierung oder Interessenkonfliktserklärung	+ keine																						
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	k.A.																						
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+																						
vergleichbare Gruppen	+																						

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko																				
				<p>p=0.023), -Wundheilungsstörung (34%, p=0.011)</p> <p>signifikante Verschlechterung in 3 Kategorien -Tiefe Beinvenenthrombose/Thrombophlebitis (34,9%, p=0.0126), -Pneumonie (23,1%, p=0.0023), - Harnwegsinfekte (41,8%, p&lt;0.00001)</p> <p>alle übrigen Kategorien: kein signifikanter Unterschied</p> <p>Effekt: inkonsistent</p>	<table border="1"> <tr><td>Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td><td>+</td></tr> <tr><td>Interventionszeitpunkt klar beschrieben</td><td>+</td></tr> <tr><td>Beobachtungszeitraum adäquat</td><td>+</td></tr> <tr><td>Population repräsentativ</td><td>+</td></tr> <tr><td>Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten</td><td>k.A.</td></tr> <tr><td>adäquate Kontrollgruppe</td><td>n.a.</td></tr> </table>	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	+	Interventionszeitpunkt klar beschrieben	+	Beobachtungszeitraum adäquat	+	Population repräsentativ	+	Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.	adäquate Kontrollgruppe	n.a.								
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	+																								
Interventionszeitpunkt klar beschrieben	+																								
Beobachtungszeitraum adäquat	+																								
Population repräsentativ	+																								
Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.																								
adäquate Kontrollgruppe	n.a.																								
Cong et al., 2012 [41]  <b>23</b>	Vorher-Nachher-Studie ohne Kontrollgruppe, prospektiv  China	Vergleich Wirksamkeit Einführung eines QM zu Dekubitalgeschwüren vor und nach Einführung in einem Lehr-Krankenhaus (3000 Betten)  QM: Erhebung, Rückmeldung, Audit, Fortbildung, Diskussion in Qualitätskonferenzen, Erstellung evidenzbasierter Handlungsanweisungen für präventive Maßnahmen, multidisziplinäres Wundteam	1. Erhebung April 2009 2. Erhebung Okt 2009 (6 Monate nach Implementierung)  2913 Fälle vorher 3019 Fälle nachher  zusätzlich erhobene Daten	Ergebnisqualität  Auftreten von Dekubitalgeschwüren  insgesamt: nicht signifikant Ereignisrate 1,8 auf 1,4% (p=0.186) Subgruppenanalyse ohne Stadium I: 0,8 auf 0,6 (p=0.226)  Effekt: nein	Verzerrungsrisiko: moderat  <table border="1"> <tr><td>Angaben zur Finanzierung oder Interessenkonfliktserklärung</td><td>-</td></tr> <tr><td>Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen</td><td>n.a.</td></tr> <tr><td>Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)</td><td>+</td></tr> <tr><td>vergleichbare Gruppen</td><td>+</td></tr> <tr><td>Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt</td><td>+</td></tr> <tr><td>Interventionszeitpunkt klar beschrieben</td><td>+</td></tr> <tr><td>Beobachtungszeitraum adäquat</td><td>(+)</td></tr> <tr><td>Population repräsentativ</td><td>+</td></tr> <tr><td>Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten</td><td>k.A.</td></tr> <tr><td>adäquate Kontrollgruppe</td><td>-</td></tr> </table>	Angaben zur Finanzierung oder Interessenkonfliktserklärung	-	Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	n.a.	Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+	vergleichbare Gruppen	+	Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	+	Interventionszeitpunkt klar beschrieben	+	Beobachtungszeitraum adäquat	(+)	Population repräsentativ	+	Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.	adäquate Kontrollgruppe	-
Angaben zur Finanzierung oder Interessenkonfliktserklärung	-																								
Verblindung (z.B. bei der Auswertung oder bei der Datenerhebung) bei nicht objektiven Variablen	n.a.																								
Angaben zur statistischen Auswertung zu statistischen Maßzahlen (p-Werte, KI, Signifikanzniveaus etc.)	+																								
vergleichbare Gruppen	+																								
Confounder berücksichtigt/diskutiert, ggf. Adjustierung erfolgt	+																								
Interventionszeitpunkt klar beschrieben	+																								
Beobachtungszeitraum adäquat	(+)																								
Population repräsentativ	+																								
Erfassung/Berücksichtigung von inkompletten oder fehlenden Daten	k.A.																								
adäquate Kontrollgruppe	-																								

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko
Systematische Reviews					
Simoes et al., 2004 [170]  24	Systematischer Review	Effektivität und Anwendbarkeit von QM/QS international (Zertifizierung, Leitlinien, Register, Audit und Feedback, Einsatz von mit Qualitätsindikatoren)	18 untersuchte Länder. Systematische Recherche in Datenbanken von Medline, Cochrane, des International Network of Agencies for Health Technology Assessments, DIMDI, Internet und Handsuche mit Einschluss von (bevorzugt) RCT	<p>Alle Qualitätsdimensionen</p> <p>Insgesamt limitierte Evidenz für Effektivität - positive Auswirkungen auf Outcome nicht belegt.</p> <p>Leitlinien: inkonsistente Datenlage zur Auswirkung auf ärztliche Verhaltensänderung oder verbesserte gesundheitliche Ergebnisse in der Versorgung</p> <p>Peer-Review, Audit, externe QS-Messung, Feedbackstrategien: inkonsistente Datenlage zur Auswirkung auf ärztliche Verhaltensänderung oder verbesserte gesundheitliche Ergebnisse in der Versorgung</p> <p>Total Quality Management (TQM) und Continuous Quality Improvement (CQI) : kein Beleg zur verbesserten Qualität der Patientenversorgung</p> <p>Akkreditierung (Synonym Zertifizierung): inkonsistente Datenlage zur verbesserten Versorgungsqualität</p> <p>Register-Datenbanken: Insgesamt scheint der Nutzen von Registern und Datenbanken als Grundlage übergreifender Analysen belegbar, sofern entsprechende Voraussetzungen (z.B. Datenlink, staatliche Kontrolle, Vollständigkeitsanspruch) erfüllt werden.</p>	<p>Verzerrungsrisiko: moderat des Reviews, nicht der Einzelstudien (Qualität der Einzelstudien kaum dargestellt - Einschluss von „bevorzugt“ RCT – Verzerrungsrisiko moderat vermutet)</p> <p>Limitation: keine Trefferzahl, kein Flowchart, keine festen Ein- und Ausschlusskriterien, limitierte Darstellung des Details, des Studiendesigns und der Qualität der Einzelstudien. Keine Differenzierung zwischen ambulanten und stationären Setting.</p> <p>Keine Angaben zu Interessenkonflikten</p> <p>Literaturbelege:</p> <p><u>Leitlinien:</u></p> <ul style="list-style-type: none"> <li>- Thomas L, Cullum N, McColl E et al. Guidelines in professions allied to medicine. Cochrane Review. The Cochrane Library 2003; 3</li> <li>- Grimshaw J, Russel I. Effects of guidelines on medical practice. A systematic review of rigorous evaluations. Lancet 1993; 342: 1317–1322</li> <li>- Van der Weijden T, Grol R, Knottnerus J. Feasibility of a national cholesterol guideline in daily practice. A randomized controlled trial in 20 general practices. Int Jour Qual Health Care 1999; 11: 131–137</li> <li>- Lichtman JH, Roumanis SA, Radford MJ et al. Can practice guidelines be transported effectively to different settings? Results from a multi-center interventional study. Joint Comm Jour Qual Improvem 2001: 42–53</li> <li>- Klazinga N, Lombarts K, van Everdingen J. Quality management in medical specialties: the use of channels and dikes in improving health care in The Netherlands. Jt Comm Jour Qual Improv 1998; 24: 240–250</li> </ul> <p><u>Peer-Review, Audit, externe QS-Messung, Feedbackstrategien:</u></p> <ul style="list-style-type: none"> <li>- Weiss KB, Wagner R. Performance measurement through audit, feed-back, and profiling as tools for improving clinical care. Chest 2000;118: 53S–58</li> <li>- Thomson O'Brien MA, Oxman AD, Davis DA et al. Audit and feedback versus alternative strategies: effects on professional practice and health care outcomes. Cochrane Review. The Cochrane Library 2003; 3</li> <li>- O'Connell D, Henry D, Tombins R. Randomized controlled trial of effect of feedback on general practitioners' prescribing in Australia. BMJ 1999; 318: 507–511</li> <li>- Winkens RA, Pop P, Grol RP et al. Effects of routine individual feedback over nine years on general practitioners' requests for tests. BMJ 1996; 312: 490</li> <li>- Irwin P, Rutledge Z, Rudd AG. The feasibility of a national audit of stroke. Br Jour Clin Govern 2001; 6: 27–33</li> <li>- Hearnshaw H, Harker R, Cheater F et al. A study of the methods used to select review criteria for clinical audit. Health Technol Assess 2002;</li> <li>- Lombarts MJ, Klazinga NS. A policy analysis of the introduction and dissemination of external peer review (visitatie) as a means of professional self-regulation amongst medical specialists in the Netherlands in the period 1985–2000. Health Policy 2001; 58: 191–21</li> </ul>

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko
					<p><b>Total Quality Management (TQM) und Continuous Quality Improvement (CQI) :</b></p> <ul style="list-style-type: none"> <li>- Shortell SM, Levin DZ, O'Brien JL et al. Assessing the evidence on CQI: is the glass half empty or half full? Hosp Health Serv Adm 1995; 40: 4–24</li> <li>- Casparie AF, Sluijs EM, Wagner C et al. Quality systems in Dutch health care institutions. Health Policy 1997; 42: 255–267</li> <li>- Kerr D, Bevan H, Gowland B et al. Redesigning cancer care. BMJ 2002; 324: 164–166</li> <li>- Shortell SM, Bennett CL, Byck GR. Assessing the impact of continuous quality improvement on clinical practice: what it will take to accelerate progress. Milbank Quart 1998; 76: 593–624</li> <li>- Blumenthal D, Kilo C. A report card on continuous quality improvement. Milbank Quart 1998; 76: 625–648</li> </ul> <p><b>Akkreditierung (Synonym Zertifizierung):</b></p> <ul style="list-style-type: none"> <li>- Dean Beaulieu N, Epstein AM. National committee on quality assurance health-plan accreditation: predictors, correlates of performance, and market impact. Med Care 2002; 40: 325–337</li> <li>- Viswanathan HN, Salmon JW. Accrediting organizations and quality improvement. Am Jour Manag Care 2000; 6: 1117–1130</li> <li>- Collopy BT, Campbell JC, Williams JW et al. Acute Health Clinical Indicator Project. Final Report. Vol. 1. ACHS Care Evaluation Program in Association with Monash University Department of Epidemiology and Preventive Medicine. 1999</li> <li>- Frolich A, Christensen M. Accreditation of hospitals. A review of international experiences. Ugeskr Laeger 2002; 164: 4412 (engl. Abstract)</li> <li>- Montagu D. Accreditation and other external quality assessment systems for healthcare. Review of experience and lessons learned. Department for International Development (DFID), Health Systems Resource Centre (HSRC) 2003.</li> </ul> <p><b>Register-Datenbanken:</b></p> <ul style="list-style-type: none"> <li>- Asplund K, Hulter Asberg K, Norrving B et al. Riks-stroke – a Swedish national quality register for stroke care. Cerebrovasc Dis 2003; 15: 5–7</li> <li>- Or Z. Improving the performance of health care systems: from measures to action (a review of experiences in four OECD countries). Directorate for Education, Employment, Labour and Social Affairs/Employment, Labour and Social Affairs Committee, Labour Market and Social Policy – Occasional Papers No. 57, DEELSA/ELSA/WD 2002;</li> <li>- Havelin LI, Engesaeter LB, Espehaug B et al. The Norwegian arthroplasty register: 11 years and 73,000 arthroplasties. Acta Orthop Scand 2000; 71: 337–353</li> </ul>
Conry MC et al., 2012 [42]  25	Systematischer Review	Effektivität aller Interventionen zur Verbesserung der Versorgungsqualität in Krankenhäusern (keine Einschränkung)  <i>(Publikation daher nur partiell zu Einschlusskriterien der vorliegenden Arbeit passend – nur Teile verwendet)</i>	13.195 Treffer  Datenbanken: PsychInfo, Medline, PubMed, EmBase and CinNahl. 2000 2010 Recherchezeitraum  Von 120 potentiellen Volltexten wurden 20 einge-	Alle Qualitätsdimensionen  Überwiegend positive Auswirkungen mit Verbesserungen auf unterschiedlichste Endpunkte wie Patientenzufriedenheit und Prozessparameter bei moderater bis niedriger Qualität der Einzelstudien mit überwiegend Prä-Post-Vergleichen. Einziger RCT mit hoher Studien-	Verzerrungsrisiko: gering des Reviews, nicht der Einzelstudien - siehe Spalte links (Evidenzgrad der Einzelstudien) COI-Angabe vorhanden (keine)  Trefferzahl, Flowchart, Qualitätsbewertung nach GRADE dargestellt, Effektstärken (Signifikanz, KI) kaum dargestellt. Details der Studien durch GRADE-Schema nicht ersichtlich (da zusammengefasst in Hauptkategorien) wie Vorhandensein Kontrollgruppe der Einzelstudien, retro-oder prospektiv etc.

Studie (Autor, Jahr, Referenz)	Studiendesign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdimension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko
			geschlossen in den Review	<p>qualität zeigt keine eindeutige Wirksamkeit der Intervention (Schulung und Förderung der Aufnahme einer Begleitperson bei Geburt in 10 Krankenhäusern) hinsichtlich vermehrter Aufnahme.</p> <p>Literaturbelege und Qualitätsbewertung (hier nur die Studien zu Programmen extrahiert):</p> <ul style="list-style-type: none"> <li>- Brown H et al., 2007– high (RCT)</li> <li>- Kipp KM et al, 2001 – very low</li> <li>- Scott IA et al., 2000 -moderate</li> <li>- Mehta RH et al, 2002 -moderate</li> <li>- Halm EA et al, 2004-moderate</li> <li>- Choma NN et al, 2009-moderate</li> <li>- Smith H et al, 2004- moderate</li> <li>- Meehan TP et al, 2001-moderate</li> <li>- van Zyl DG et al, 2004 - low</li> </ul>	<p>Jegliche QS-Intervention von kleineren Maßnahmen wie Zusammenlegung von Stationen oder Befragungen bis hin zu übergeordneten Programmen inkludiert worden</p> <p>Angabe des Landes der Studien fehlend</p> <p>Funding from the European Community's Seventh Framework Programme</p>
<p>Mayer et al., 2012 [121]</p> <p>(siehe oben - hier internationaler Teil des Gutachtens dargestellt)</p>	<p>Gutachten des Kompetenzzentrum Qualitätssicherung/Qualitätsmanagement im Auftrag des GKV Spitzenverbands zur Aufarbeitung des Internen QM in der medizinischen Versorgung zur Umsetzung und Nutzenfrage im stationären und ambulanten Sektor national und</p>	<p>Umfassendes Gutachten der KCQ im Auftrag des GKV Spitzenverbands zur Aufarbeitung des Internen QM in der medizinischen Versorgung zur Umsetzung und Nutzenfrage im stationären und ambulanten Sektor national und international</p>	<p>49 Volltexte als themenrelevante Treffer</p> <p>Bis zum 1.11.2011 Recherchezeitraum Pubmed und Internetsuchmaschinen und Handsuche</p> <p>Keine methodische Studieneinschränkung</p>	<p>Alle Qualitätsdimensionen</p> <p>Heterogene Datenlage, Untersuchungen aus den unterschiedlichsten Gesundheitssystemen, basierend auf unterschiedlichsten QM-Ansätzen mit verschiedenen Zielparametern. Limitierte methodische Güte der Einzelstudien überwiegend, mehrheitlich nicht zur Ableitung von kausalen Zusammenhängen geeignet.</p> <p>Die Studien mit höherwertigem Design können keine konsistenten Ergebnisse zu Patientennutzen ableiten, fast durchgehend trotz verbesserter Prozessqualität keine Optimierung des Patientenoutcomes</p> <p>Effekt: inkonsistent</p>	<p>Verzerrungsrisiko: gering für den extrahierten Themenbereich des Gutachtens, nicht der Einzelstudien</p> <p>Überwiegender Teil der gefundenen Studien von niedriger Qualität</p> <p>Keine methodische Studieneinschränkung, kaum Effektstärken angeben</p>

Studie (Autor, Jahr, Referenz)	Studiende- sign	Intervention und Vergleich (QS/QM-Programm)  Zusatzangaben: Initiator/verantwortlicher Träger, Teilnahme/Fachgebiet	Analysezeitraum der rekrutierten Daten/ Interventionsdauer  N= xx (Datensätze der Auswertung) Datengrundlage	Endpunktdefinition/ Qualitätsdi- mension  Ergebnis der Evaluierung	Studienlimitation/ Verzerrungsrisiko
	international				

#### 7.4 Kriterien für die Klassifikation der Studienbewertung

Verzerrungsrisiko	angewandte Kriterien
sehr hoch	„Vorher-Nachher-Studie" mit Vergleich zu Beginn und im Verlauf der Interventionszeit, retrospektiv, keine Kontrollgruppe
hoch	retrospektive Vorher-Nachher-Studie ohne Kontrollgruppe oder nicht adäquat kontrollierte Vorher-Nacher-Studie
moderat	-Vorher-Nachher-Studie mit Kontrollgruppe oder schlechter RCT oder sehr gute, unkontrollierte Kohorte bei dargestellter Vergleichbarkeit der Vorher-Nacher-Studiengruppen unter Berücksichtigung von Confoundern/Adjustierung, prospektiv etc.  -mittelmäßiger systematischer Review
gering	-guter RCT oder sehr gute Vorher-Nachher-Studie /Kohortenstudie mit Kontrollgruppe  -guter systematischer Review

##### Studientypen der identifizierten, relevanten Studien:

- Vorher-Nachher-Studie mit Vergleich zu Beginn und im Verlauf der Interventionszeit, retrospektiv, keine Kontrollgruppe
- Vorher-Nachher-Studie ohne Kontrollgruppe, retrospektiv oder prospektiv
- Vorher-Nachher-Studie mit Kontrollgruppe, retrospektiv oder prospektiv
- Kohortenstudie mit Kontrollgruppe, retrospektiv oder prospektiv
- Querschnittsstudie mit Kontrollgruppe
- RCT

## 7.5 Checkliste zur Studienbewertung

Kriterien zur Qualitätsbewertung von nicht-randomisierten Studien, Kohortenstudien, (kontrollierten) Vorher-Nachher-Studien (modifiziert nach [144; 145])

Selection of subjects	
Representativeness of the intervention cohort	
- truly representative	<input type="checkbox"/>
- no description	<input type="checkbox"/>
Ascertainment of intervention	
- secure record	<input type="checkbox"/>
- other / no description	<input type="checkbox"/>
Demonstration that outcome of interest was not present at start of study	
a)yes	<input type="checkbox"/>
b)no	<input type="checkbox"/>
Comparability	
Comparability of cohorts on the basis of the design or analysis	
a)study controls for e.g. age, sex, comorbidity	<input type="checkbox"/>
b)study controls for any additional factors	<input type="checkbox"/>
Outcome	
Assessment of outcome	
a)independent blind assessment	<input type="checkbox"/>
b)record linkage	<input type="checkbox"/>
c)self-report	<input type="checkbox"/>
d)other / no description	<input type="checkbox"/>
Was follow up long enough for outcomes to occur	
a)yes	<input type="checkbox"/>
b)no	<input type="checkbox"/>
Adequacy of lost-to-follow up	
a)complete follow up: all subjects accounted for	<input type="checkbox"/>
b)description of those lost	<input type="checkbox"/>
c)no statement	<input type="checkbox"/>

Assessment of how researchers dealt with confounding	
Method for <i>identifying</i> relevant confounders described: yes no If yes, describe the method used:	<input type="checkbox"/> <input type="checkbox"/>
Relevant confounders described: yes no	<input type="checkbox"/> <input type="checkbox"/>
Method used for controlling for confounding - At design stage: matching Variables on which subjects matched:  - At analysis stage e.g. stratification, multivariable regression, propensity scores (matching), propensity scores (multivariable regression) List confounders controlled for:	<input type="checkbox"/> <input type="checkbox"/>
Description of statistical analysis and results Report of precision of statistical results as confidence intervall, p-value etc.	<input type="checkbox"/> <input type="checkbox"/>
Others	
Disclosure of Conflict of Interest or financial support yes no	<input type="checkbox"/> <input type="checkbox"/>
Blinding of data analysis yes no	<input type="checkbox"/> <input type="checkbox"/>
Clearly defined point in time when the intervention occurred (prä-post-design) yes no	<input type="checkbox"/> <input type="checkbox"/>

*Die Bewertung von RCT oder sytematischen Übersichtsarbeiten erfolgte nach den Checklisten des Scottish Intercollegiate Guidelines Network [165].*

## 7.6 Dokumentation Volltextsichtung für die primäre Forschungsfrage

### 1.1 Deutschland – systematische Recherche

1. Neuderth S, Lukaszczik M, Musekamp G, Gerlich C, Saupe-Heide M, Lobmann R, Vogel H. Entwicklung von Verfahrensweisen und Instrumenten zur externen Qualitätssicherung in Einrichtungen der stationären Vorsorge und Rehabilitation für Mütter und Väter einschließlich Mutter-/Vater-Kind-Einrichtungen. Rehabilitation (Stuttg) 2013;52(1):10-9	A1
2. Squires DA. Explaining high health care spending in the United States: an international comparison of supply, utilization, prices, and quality. Issue Brief (Commonw Fund) 2012;10:1-14	A1
3. Squires DA. The U.S. health system in perspective: a comparison of twelve industrialized nations. Issue Brief (Commonw Fund) 2011;16:1-14	A1
4. Vogel P, Vassilev G, Kruse B, Cankaya Y. Morbidity and Mortality conference as part of PDCA cycle to decrease anastomotic failure in colorectal surgery. Langenbecks Arch Surg 2011;396(7):1009-15	A1
5. Gogol M, Schmidt D, Dettmer-Flugge A, Vaske B. Pneumonie im Alter. Z Gerontol Geriatr 2011;44(4):235-9	A2
6. Biesheuvel C, Weigel S, Heindel W. Mammography Screening: Evidence, History and Current Practice in Germany and Other European Countries. Breast Care (Basel) 2011;6(2):104-9	A3
7. Brucker SY, Wallwiener M, Kreienberg R, Jonat W, Beckmann MW, Bamberg M, Wallwiener D, Souchon R. Optimizing the quality of breast cancer care at certified german breast centers: a benchmarking analysis for 2003-2009 with a particular focus on the interdisciplinary specialty of radiation oncology. Strahlenther Onkol 2011;187(2):89-99	E1
8. De Allegri M, Schwarzbach M, Loerbroks A, Ronellenfitsch U. Which factors are important for the successful development and implementation of clinical pathways? A qualitative study. BMJ Qual Saf 2011;20(3):203-8	A1
9. Kostuj T, Smektala R. Qualitätssicherung mithilfe von Routinedaten. Unfallchirurg 2010;113(12):1047-2	A1
10. Tsilimparis N, Haase O, Wendling P, Kipfmüller K, Schmid M, Engemann R, Schwenk W. Laparoskopische „Fast-track“-Sigmaresektion bei Divertikulitis in Deutschland. Ergebnisse einer prospektiven Qualitätssicherungsmaßnahme. Dtsch Med Wochenschr 2010;135(36):1743-8.	A3

11. Kostuj T, Schulze-Raestrup U, Noack M, Buckup K, Smektala R. Mindestmengen in der Kniegelenkendoprothetik. <i>Chirurg</i> 2011;82(5):425-32	A2
12. Helbig M, Helbig S, Kahla-Witzsch HA, Kroll T, May A. Certifying a university ENT clinic using the ISO 9001:2000 international standard. <i>Int J Health Care Qual Assur</i> 2010;23(3):268-76	A1
13. von Hodenberg, Eder S, Grunebaum P, Melichercik J. Kardiologie – Qualitätssicherung in Deutschland Pro und Contra. <i>Dtsch Med Wochenschr</i> 2009;134 Suppl 6:S198-S199	A1
14. Schwarzbach M, Bonninghoff R, Harrer K, Weiss J, Denz C, Schnulle P, Birck R, Post S, Ronellenfitsch U. Effects of a clinical pathway on quality of care in kidney transplantation: a non-randomized clinical trial. <i>Langenbecks Arch Surg</i> 2010;395(1):11-7	A1
15. du Bois, Strutas D, Buhrmann C, Traut A, Ewald-Riegler N, Kommos S, Gomez R, Harter P. Impact of treatment guidelines and implementation of a quality assurance program on quality of care in endometrial cancer. <i>Onkologie</i> 2009;32(8-9):493-8	A2
16. Hawighorst H, Becker G, Hodapp N, Wenz F. Erste Ergebnisse der Qualitätssicherung durch Ärztliche Stellen in der Radiotherapie am Beispiel Baden-Württemberg: Teil 1. <i>Strahlenther Onkol</i> 2009;185(8):493-9	A2
17. Kommos S, Harter P, Traut A, Strutas D, Riegler N, Buhrmann C, Gomez R, du Bois. Compliance to consensus recommendations, surgeon's experience, and introduction of a quality assurance and management program: influence on therapy of early-stage ovarian carcinoma. <i>Int J Gynecol Cancer</i> 2009;19(4):787-93	A2
18. Braumann C, Guenther N, Wendling P, Engemann R, Germer CT, Probst W, Mayer HP, Rehnisch B, Schmid M, Nagel K, Schwenk W. Multimodal perioperative rehabilitation in elective conventional resection of colonic cancer: results from the German Multicenter Quality Assurance Program 'Fast-Track Colon II'. <i>Dig Surg</i> 2009;26(2):123-9	A3
19. Schulze RU, Grams A, Smektala R. Leitlinienkonforme Versorgung. <i>Unfallchirurg</i> 2008;111(2):65-70	A5
20. Zernikow B, Hasan C, Hechler T, Huebner B, Gordon D, Michel E. Stop the pain! A nation-wide quality improvement programme in paediatric oncology pain control. <i>Eur J Pain</i> 2008;12(7):819-33	E1
21. Ollenschläger G, Lelgemann M, Kopp I. Nationale VersorgungsLeitlinien - Nutzung im Qualitätsmanagement unter besonderer Berücksichtigung von Klinischen Behandlungspfaden und Regionalen Leitlinien. <i>Med Klin (Munich)</i> 2007;102(7):565-9	A1
22. Stausberg J, Bartels C, Bobrowski C. Gewinnung von Managementinformationen	A1

aus der externen vergleichenden Qualitätssicherung. Med Klin (Munich) 2007;102(7):507-14	
23. Glattacker M, Jackel WH. Evaluation der Qualitätssicherung - aktuelle Datenlage und Konsequenzen für die Forschung. Gesundheitswesen 2007;69(5):277-83	A5
24. du Bois, Rochon J, Lamparter C, Pfisterer J. Das Qualitätssicherungsprogramm der AGO Organkommission OVAR (QS-OVAR): Versorgungsstruktur und Realität in Deutschland 2001. Zentralbl Gynakol 2005;127(1):9-17	A2
25. Stoffel MP, Barth C, Lauterbach KW, Baldamus CA. Evidence-based medical quality management in dialysis--Part I: Routine implementation of QiN, a German quality management system. Clin Nephrol 2004;62(3):208-18	E1
26. Harter M, Sitta P, Keller F, Metzger R, Wiegand W, Schell G, Stieglitz RD, Wolfersdorf M, Felsenstein M, Berger M. Stationäre psychiatrisch-psychotherapeutische Depressionsbehandlung. Nervenarzt 2004;75(11):1083-91	A2
27. Eich HT, Staar S, Gossmann A, Engert A, Franklin J, Sieber M, Krug B, Diehl V, Lackner KJ, Muller RP. The HD12 panel of the German Hodgkin Lymphoma Study Group (GHSg): a quality assurance program based on a multidisciplinary panel reviewing all patients' imaging. Am J Clin Oncol 2004;27(3):279-84	A3
28. Eich HT, Staar S, Gossmann A, Hansemann K, Skripnitchenko R, Kocher M, Semrau R, Engert A, Josting A, Franklin J, Krug B, Diehl V, Muller RP. Centralized radiation oncologic review of cross-sectional imaging of Hodgkin's disease leads to significant changes in required involved field-results of a quality assurance program of the German Hodgkin Study Group. Int J Radiat Oncol Biol Phys 2004;58(4):1121-7	A3
29. Paeger A. Qualitätsmanagement in der medizinischen Rehabilitation - Vergleich internationaler und nationaler Konzepte. Z Arztl Fortbild Qualitätssich 2002;96(1):17-24	A1
30. Kawski S, Koch U. Qualitätssicherung in der Psychosomatischen Rehabilitation. Psychother Psychosom Med Psychol 1999;49 (9-10):316-25	A1
31. Keitel W. Qualitätssicherung in der Rheumatologie. Z Arztl Fortbild Qualitätssich 1999;93(2):141-4	A1
32. Bergmann KO, Michalcik M. Risk-Management als Mittel der Qualitätssicherung im Krankenhaus. Z Arztl Fortbild Qualitätssich 1999;93(1):57-62	A1
33. Rau R. Qualitätsmanagement im interdisziplinären Fachkrankenhaus. Z Rheumatol 1998;57(6):413-9	A1
34. Jackel WH, Gerdes N. Qualitätsmanagement in der medizinischen Rehabilitation. Z Rheumatol 1998;57(5):345-50	A2 (A5)
35. Farin E, Jackel WH. Qualitätssicherung und Qualitätsmanagement in der medizini-	A1

schen Rehabilitation. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitschutz 2011;54(2):176-84	
36. Obi N, Waldmann A, Schafer F, Schreer I, Katalinic A. Impact of the Quality assured Mamma Diagnostic (QuaMaDi) programme on survival of breast cancer patients. Cancer Epidemiol 2011;35(3):286-92	A4
37. Marx R. Qualitätsanalysen und Qualitätssicherung in der kardiologischen Rehabilitation. Dtsch Med Wochenschr 2009;134 Suppl 6:S228-S229	A1
38. Lukasczik M, Gerlich C, Musekamp G, Saupe-Heide M, Lobmann R, Vogel H, Neuderth S. Externe Qualitätssicherung in Einrichtungen der stationären Vorsorge und Rehabilitation für Mütter und Väter einschließlich Mutter-/Vater-Kind-Einrichtungen: Einrichtungsvergleichende Analysen im Bereich Ergebnisqualität. Gesundheitswesen 2013	A2
39. Bassler M, Nosper M, Follert P, Bowering L, Polak U. Datenquellen für eine kontinuierliche Qualitätsverbesserung in der medizinischen Rehabilitation. Rehabilitation (Stuttg) 2007;46(3):155-63	A1
40. Neuderth S, Saupe-Heide M, Bruckner U, Gross B, Wenderoth N, Vogel H. Wissenschaftliche Begleitung der Weiterentwicklung des Visitationsverfahrens für Reha-Kliniken der Deutschen Rentenversicherung Bund - Das Projekt „Visit II“. Rehabilitation (Stuttg) 2012;51(3):171-80	A5
41. Ronellenfitsch U, Rossner E, Jakob J, Post S, Hohenberger P, Schwarzbach M. Clinical Pathways in surgery: should we introduce them into clinical routine? A review article. Langenbecks Arch Surg 2008;393(4):449-57	A1
42. Helbig M, Kahla-Witzsch HA, Helbig S, Knecht R, Gstottner W. Qualitätsmanagement nach DIN EN ISO 9001:2000. HNO 2006;54(12):922-8	A1
43. Vagts DA, Martin J, Dahmen K. Qualitätsmanagement im Krankenhaus - Prozesse und Ergebnisse kontinuierlich verbessern. Anasthesiol Intensivmed Notfallmed Schmerzther 2008;43(2):156-60	A1
44. Gulker H. Qualitätssicherung in der interventionellen Kardiologie. Dtsch Med Wochenschr 2009;134 Suppl 6:S192-S194	A2
45. Simoes E, Boukamp K, Mayer ED, Schmahl FW. Gibt es Belege für den Impact qualitätssichernder/-fördernder Verfahren in anderen Ländern? Gesundheitswesen 2004;66(6):370-9	E2
46. Riskin L, Campagna JA. Quality assessment by external bodies: intended and unintended impact on healthcare delivery. Curr Opin Anaesthesiol 2009;22(2):237-41	A1
47. Kliche T. Versorgungsstrukturen und Qualitätssicherung für Prävention und Gesundheitsförderung in Deutschland. Bundesgesundheitsblatt Gesundheitsforschung	A1 (A4)

Gesundheitsschutz 2011;54(2):194-206	
48. Stoffel MP, Barth C, Lauterbach KW, Baldamus CA. Evidence-based medical quality management in dialysis--Part II: Improvement of hemodialysis adequacy. Clin Nephrol 2004;62(3):219-25	E1
49. Deck R, Raspe H. Regionale Qualitätssicherung in der medizinischen Rehabilitation. Qualitätsgemeinschaft medizinische Rehabilitation in Schleswig-Holstein - Hauptstudie und Ergebnisse aus orthopädischen Kliniken. Rehabilitation (Stuttg) 2006;45(5):272-81	A1
50. Lubke N, Meinck M. Aktuelle Bestandsaufnahme externer Qualitätssicherungsverfahren im Bereich geriatrischer Rehabilitation. Rehabilitation (Stuttg) 2008;47(1):39-48	A1
51. Glattacker M, Muller E, Jackel WH. Qualitätssicherung und Leitlinien in der Rehabilitation bei muskuloskelettalen Krankheiten. Z Rheumatol 2008;67(7):575-82	A1
52. Sens B. BÄK/KBV/AWMF-Curriculum "Ärztliches Qualitätsmanagement": Resümee aus drei Jahren praktischer Umsetzung einer zielorientierten Qualifikation. Z Arztl Fortbild Qualitätssich 2000;94(8):689-94	A1
53. Kawski S, Koch U. Qualitätssicherung in der medizinischen Rehabilitation in Deutschland. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2004;47(2):111-7	A1
54. Zeymer U, Vogt A, Neuhaus KL. Qualitätssicherung in der Klinik - Beispiel kardiologische Techniken. Z Arztl Fortbild Qualitätssich 1999;93(4):281-6 54	A1
55. Höffken K, Junginger T, Enghofer E, Bussar-Maatz R, Schmitt-Thomas B, Kreienberg R. Das Qualitätssicherungsprogramm der Deutschen Krebsgesellschaft. Onkologe 2001;7(3):243-9.	A1
56. Klosterhuis H. Aktuelle Strategien der Deutschen Rentenversicherung in der Qualitätssicherung der medizinischen Rehabilitation. Prävention und Rehabilitation, 2013;20 (Nr.4/2008):184-92.	A1
57. van Emmerich C, Metzinger B. Qualitätssicherung mit Routinedaten aus Sicht der Deutschen Krankenhausgesellschaft. dasKrankenhaus 2010; (12): 77-1182	A1
58. Schäfer RD. Ärztliche Qualitätssicherung heute. Z Evid Fortbild Qual Gesundheitswesen (ZEFQ) 2008;(102):594-7.	A1
59. Busse R, Nimptsch U, Mansky T. Measuring, monitoring, and managing quality in Germany's hospitals. Health Aff (Millwood ) 2009;28(2):w294-w304	A2
60. Nimptsch U, Mansky T. Quality measurement combined with peer review improved German in-hospital mortality rates for four diseases. Health Aff (Millwood) 2013;32(9):1616-23	E1

## 1.2 Deutschland - Handsuche

1. Eberlein-Gonska M. [What is evidence-based concerning quality management? Thoughts about an apparently simple question]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2011;54(2):148-53	A1
2. Hulscher ME, Laurant MG, Grol RP. Process evaluation on quality improvement interventions. Qual Saf Health Care 2003;12(1):40-6	A1
3. Blumenstock G. [Assessing the quality of quality indicators]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2011;54(2):154-9 PM:21290268,	A2
4. Krahwinkel W, Rink O, Liebetrau M, Gunther M, Schuler E, Kuhlen R. [10 years of Peer Reviewing: improving treatment by quality indicators from administrative data]. Dtsch Med Wochenschr 2011;136(41):2083-8	E1
5. Ruchholtz S, Lefering R, Paffrath T, Oestern HJ, Neugebauer E, Nast-Kolb D, Pape HC, Bouillon B. Reduction in mortality of severely injured patients in Germany. Dtsch Arztebl Int 2008;105(13):225-31	E1
6. Beckmann U, Klosterhuis H, Mitschele A. Qualitätsentwicklung durch Qualitätssicherung - Erfahrungen aus zehn Jahren Qualitätssicherung der Rehabilitation. DAng-Vers 2005;(52):431-8.	A5
7. Brüggemann S. Das Reha-Leitlinienprogramm der BfA. Z ärztl Fortbild Qual Gesundheitswes 2005;(99):47-50.	A1
8. Farin E, Jackel WH, Schalaster V. [The quality assurance programme of the statutory health insurance funds in medical rehabilitation: results and further developments]. Gesundheitswesen 2009;71(3):163-74	A2
9. Klosterhuis H. [Rehabilitation guidelines as tool of quality assurance in rehabilitation]. Z Arztl Fortbild Qualitätssich 2005;99(1):41-5	A1
10. Katalinic A, Bartel C, Raspe H, Schreer I. Beyond mammography screening: quality assurance in breast cancer diagnosis (The QuaMaDi Project). Br J Cancer 2007;96(1):157-61	A4
11. Waldmann A, Adrich S, Eisemann N, Fauteck H, Grande-Nagel I, Schaefer FK, Vorkefeld M, Warnecke C, Weidner K, Wilke C, Katalinic A. [Structure and process quality of quality-assured mamma diagnostics in Schleswig-Holstein]. Rofo 2012;184(2):113-21	A4
12. Wallwiener M, Brucker SY, Wallwiener D. Multidisciplinary breast centres in Germany: a review and update of quality assurance through benchmarking and certification. Arch Gynecol Obstet 2012;285(6):1671-83	E1
13. Eberlein-Gonska M, Rink O. Peer Review -Fortschreiten in einem lernenden System. Deutsches Ärzteblatt 2013; 110 (16):A760-A766.	A2

14. Rink O. Das IQM-Peer-Review-Verfahren. Verbesserung der Ergebnisqualität, Strategie und Ergebnisse. Präz Gesundheitsf 2012; (8):22-8.	E1
15. Rink O. [The IQM peer review procedure - results of the "initiative qualitätsmedizin"]. Z Evid Fortbild Qual Gesundheitswes 2012;106(8):560-5	E1
16. Siebig J. [The role of benchmarking seen from the perspective of the Federal Joint Committee (G-BA)]. Z Evid Fortbild Qual Gesundheitswes 2011;105(5):408-11	A1
17. Blum, Offermann, DKI. DMP BRUSTKREBS - Krankenhäuser sehen wenig Positives. Deutsches Ärzteblatt 2006; 103(41):A 2678-A 2680	E1
18. Blum K, Offermann M, Deutsches Krankenhausinstitut (DKI). Auswirkungen der DMP Brustkrebs auf die Krankenhäuser. 2006.	E1
19. Mayer ED, Boukamp KB. Internes Qualitätsmanagement in der medizinischen Versorgung. Gutachten des Kompetenz-Centrum Qualitätssicherung/Qualitätsmanagement (KCQ) des GKV-Spitzenverbandes und der Gemeinschaft der Medizinischen Dienste der Krankenversicherung. 2012 [letzter Zugriff: 23.12.2013] <a href="http://www.kcqq.de/106-640.php">http://www.kcqq.de/106-640.php</a>	E1
20. Geraedts M, Selbmann HK. [Benchmarking in health care: conclusions and recommendations]. Z Evid Fortbild Qual Gesundheitswes 2011;105(5):412-6	A1
21. Bundesärztekammer und Kassenärztliche Bundesvereinigung. Qualitätssicherung und Qualitätsverbesserung in der Medizin. Grundlagen einer bedarfsgerechten Gesundheitsversorgung: gemeinsame Bestandsaufnahme von BÄK und KBV über die Aktivitäten der Spitzenorganisationen der ärztlichen Selbstverwaltung auf dem Gebiet der Qualitätssicherung in der Medizin von 1955 bis 1995 / [Red.: Günter Ollenschläger]. München: Zuckschwerdt; 1996.	A5
22. Ollenschläger G. Qualitätssicherung-wem nützt das? Arztebl Sachsen 1999;(9):418-21	A5
23. Brucker SY, Schumacher C, Sohn C, Rezai M, Bamberg M, Wallwiener D. Benchmarking the quality of breast cancer care in a nationwide voluntary system: the first five-year results (2003-2007) from Germany as a proof of concept. BMC Cancer 2008;8:358	E1
24. Drabik A, Sawicki PT, Muller D, Passon A, Stock S. [The methods within the evaluation of disease management programmes in control-group designs using the example of diabetes mellitus - a systematic literature review]. Gesundheitswesen 2012;74(8-9):496-501	A1
25. Petzold T, Steinwitz A, Schmitt J, Eberlein-Gonska M. [Evaluation of external quality assurance in accordance with sect. 137 SGB V at the Carl Gustav Carus University hospital in Dresden]. Z Evid Fortbild Qual Gesundheitswes.	A5

2013;107(8):541-7	
26. Veit C, Bungard S, Hertle D, Grothaus FJ, Kötting J, Arnold N. [Potentials of cooperative quality management initiatives: BQS Institute projects, January 2010 - July 2013]. Z Evid Fortbild Qual Gesundheitswes. 2013;107(8):534-40	A1
27. Albrecht M, Loos S, Otten M. [Cross-sectoral quality assurance in ambulatory care]. Z Evid Fortbild Qual Gesundheitswes. 2013;107(8):528-33	A1
28. Willms G, Bramesfeld A, Pottkämper K, Broge B, Szecsenyi J. [Current Challenges of external quality assurance in the German healthcare system]. Z Evid Fortbild Qual Gesundheitswes. 2013;107 (8):523-7. doi: 10.1016/j.zefq.2013.10.015. Epub 2013 Nov 4. German..	A1
29. Schmitt J, Petzold T, Eberlein-Gonska M, Neugebauer EA. [Requirements for quality indicators. The relevance of current developments in outcomes research for quality management]. Z Evid Fortbild Qual Gesundheitswes. 2013;107(8):516-22. doi: 10.1016/j.zefq.2013.09.014. Epub 2013 Nov 4.	A1
30: Schäfer I, Küver C, Gedrose B, von Leitner EC, Treszl A, Wegscheider K, van den Bussche H, Kaduszkiewicz H. Selection effects may account for better outcomes of the German Disease Management Program for type 2 diabetes. BMC Health Serv Res. 2010 Dec 31;10:351.	A1
31. Abteilung Allgemeinmedizin und Versorgungsforschung Universitätsklinikum Heidelberg. Evaluation des DMP Diabetes mellitus Typ 2 im Rahmen der ELSID-Studie. Abschlussbericht für den AOK-Bundesverband. 2012 [letzter Zugriff: 02.12.2013] <a href="http://www.aok-gesundheitspartner.de/bund/dmp/evaluation/elsid/index.html">http://www.aok-gesundheitspartner.de/bund/dmp/evaluation/elsid/index.html</a>	A4
32. Miksch A, Laux G, Ose D, Joos S, Campbell S, Riens B, Szecsenyi J. Is there a survival benefit within a German primary care-based disease management program? Am J Manag Care. 2010 Jan;16(1):49-54.	A4
33. Pilz M, Shamiyeh A, Hörmandinger K et al. Die Implementierung der S3-Leitlinie Akutschmerztherapie über die „Initiative Schmerzfreie Klinik“ – Verbesserung der Ergebnisqualität in einer Viszeralchirurgischen Klinik. Perioperative Medizin. 2010; 2(1):50-56	A2
34. AQUA - Institut für angewandte Qualitätsförderung im Gesundheitswesen GmbH. Qualitätsreport 2012. 2013 [letzter Zugriff: 30.11.2013] <a href="https://www.sqg.de/themen/qualitaetsreport/index.html">https://www.sqg.de/themen/qualitaetsreport/index.html</a>	A5
35. Maass C, Schleiz W, Weyermann M, Drösler SE. [Are hospital administrative data suitable for external quality assurance? Comparison of quality indicators based on separate statutory data collections (BQS) and hospital administrative data].	A2

Dtsch Med Wochenschr. 2011 Mar;136(9):409-14.	
36. Gastmeier P, Bräuer H, Forster D, Dietz E, Daschner F, Rüden H. A quality management project in 8 selected hospitals to reduce nosocomial infections: a prospective, controlled study. Infect Control Hosp Epidemiol. 2002 Feb;23(2):91-7.	E1
37. Lehmkuhl D, Meissner W, Neugebauer EA. [Evaluation of the "initiative pain-free clinic" for quality improvement in postoperative pain management. A prospective controlled study]. Schmerz. 2011 Sep;25(5):508-15	E1
38. Meissner W. [QUIPS: quality improvement in postoperative pain management]. Z Evid Fortbild Qual Gesundheitswes. 2011;105(5):350-3	A2
39. Klosterhuis H (2010): Reha-Qualitätssicherung der Rentenversicherung – eine kritische Bestandsaufnahme. RVaktuell, Jg. 57, H. 8, S. 260-268.	A5
40. Böhmer AB, Poels M, Simanski C, Trojan S, Messer K, Wirtz MD, Neugebauer EA, Wappler F, Joppich R. [Quality management in acute pain therapy: results from a survey of certified hospitals]. Schmerz. 2012 Aug;26(4):425-30	A5
41. Qualitätsverbesserung in der postoperativen Schmerztherapie (QUIPS). Homepage QUIPS. 2014 [letzter Zugriff: 26.01.2014] <a href="http://www.quips-projekt.de">http://www.quips-projekt.de</a>	A1
42. Meissner W, Mescha S, Rothaug J, Zwacka S, Goettermann A, Ulrich K, Schleppe A. Quality improvement in postoperative pain management: results from the QUIPS project. Dtsch Arztebl Int. 2008 Dec;105(50):865-70.	E1

### 1.3 International – systematische Recherche

1. Aiken LH, Sermeus W, Van den Heede K, Sloane DM, Busse R, McKee M, Bruyneel L, Rafferty AM, Griffiths P, Moreno-Casbas MT, Tishelman C, Scott A, Brzostek T, Kinnunen J, Schwendimann R, Heinen M, Zikos D, Sjetne IS, Smith HL, Kutney-Lee A. Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. BMJ 2012;344:e1717	A1
2. Benzer JK, Young GJ, Burgess JF, Jr., Baker E, Mohr DC, Charns MP, Kaboli PJ. Sustainability of Quality Improvement Following Removal of Pay-for-Performance Incentives. J Gen Intern Med 2013	A1
3. Canovas JJ, Hernandez PJ, Botella JJ. Effectiveness of internal quality assurance programmes in improving clinical practice and reducing costs. J Eval Clin Pract 2009;15(5):813-9	E2
4. Cima R, Dankbar E, Lovely J, Pendlimari R, Aronhalt K, Nehring S, Hyke R, Tynedale D, Rogers J, Quast L. Colorectal surgery surgical site infection reduction pro-	A2

gram: a national surgical quality improvement program--driven multidisciplinary single-institution experience. <i>J Am Coll Surg</i> 2013;216(1):23-33	
5. Cong L, Yu J, Liu Y. Implementing a continuous quality improvement program for reducing pressure prevalence in a teaching hospital in China. <i>J Wound Ostomy Continence Nurs</i> 2012;39(5):509-13	E2
6. Conry MC, Humphries N, Morgan K, McGowan Y, Montgomery A, Vedhara K, Panagopoulou E, Mc GH. A 10 year (2000-2010) systematic review of interventions to improve quality of care in hospitals. <i>BMC Health Serv Res</i> 2012;12:275	E2
7. Daudelin DH, Kulick ER, D'Amore K, Lutz JS, Barrientos MT, Foell K. The Massachusetts Emergency Medical Service Stroke Quality Improvement Collaborative, 2009-2012. <i>Prev Chronic Dis</i> 2013;10:E161	A4
8. Esse T, Serna O, Chitnis A, Johnson M, Fernandez N. Quality compensation programs: are they worth all the hype? A comparison of outcomes within a Medicare advantage heart failure population. <i>J Manag Care Pharm</i> 2013;19(4):317-24	A4
9. Farias M, Friedman KG, Powell AJ, de Ferranti SD, Marshall AC, Brown DW, Kulik TJ. Dynamic evolution of practice guidelines: analysis of deviations from assessment and management plans. <i>Pediatrics</i> 2012;130(1):93-8	A1
10. Flodgren G, Pomey MP, Taber SA, Eccles MP. Effectiveness of external inspection of compliance with standards in improving healthcare organisation behaviour, healthcare professional behaviour or patient outcomes. <i>Cochrane Database Syst Rev</i> 2011;(11):CD008992	A2
11. Fuchshuber PR, Greif W, Tidwell CR, Klemm MS, Frydel C, Wali A, Rosas E, Clopp MP. The power of the National Surgical Quality Improvement Program--achieving a zero pneumonia rate in general surgery patients. <i>Perm J</i> 2012;16(1):39-45	A1
12. Gagneux E, Lombrail P, Vichard P. Trauma emergency unit: long-term evaluation of a quality assurance programme. <i>Qual Health Care</i> 1998;7(1):12-8	E2
13. Guillamondegui OD, Gunter OL, Hines L, Martin BJ, Gibson W, Clarke PC, Cecil WT, Cofer JB. Using the National Surgical Quality Improvement Program and the Tennessee Surgical Quality Collaborative to improve surgical outcomes. <i>J Am Coll Surg</i> 2012;214(4):709-14	E2
14. Harter P, Muallem ZM, Buhrmann C, Lorenz D, Kaub C, Hils R, Kommos S, Heitz F, Traut A, du Bois. Impact of a structured quality management program on surgical outcome in primary advanced ovarian cancer. <i>Gynecol Oncol</i> 2011;121(3):615-9	A2
15. Higashi T, Nakamura F, Saruki N, Sobue T. Establishing a quality measurement system for cancer care in Japan. <i>Jpn J Clin Oncol</i> 2013;43(3):225-32	A1
16. Jha AK, Joynt KE, Orav EJ, Epstein AM. The long-term effect of premier pay for	A2

performance on patient outcomes. <i>N Engl J Med</i> 2012;366(17):1606-15	
17. Kim SM, Jang WM, Ahn HA, Park HJ, Ahn HS. Korean National Health Insurance value incentive program: achievements and future directions. <i>J Prev Med Public Health</i> 2012;45(3):148-55	A1
18. Muething SE, Goudie A, Schoettker PJ, Donnelly LF, Goodfriend MA, Bracke TM, Brady PW, Wheeler DS, Anderson JM, Kotagal UR. Quality improvement initiative to reduce serious safety events and improve patient safety culture. <i>Pediatrics</i> 2012;130(2):e423-e431	A2
19. O'Neill HJ, Coe LJ, Magdon-Ismail Z, Schwamm LH. Implementing a state-based stroke quality improvement collaborative: the Massachusetts experience. <i>Crit Pathw Cardiol</i> 2012;11(3):114-22	A2
20. Renzi C, Sorge C, Fusco D, Agabiti N, Davoli M, Perucci CA. Reporting of quality indicators and improvement in hospital performance: the P.Re.Val.E. Regional Outcome Evaluation Program. <i>Health Serv Res</i> 2012;47(5):1880-901	A2
21. Ruchholtz S. [External quality management in the clinical treatment of severely injured patients]. <i>Unfallchirurg</i> 2004;107(10):835-43	E1
22. Ruchholtz S, Lefering R, Debus F, Mand C, Kuhne C, Siebert H. [Traumaregister-TraumaNetwork DGU(R) und TraumaRegister DGU(R). Success by cooperation and documentation]. <i>Chirurg</i> 2013;84(9):730-8	A5
23. Scott I. What are the most effective strategies for improving quality and safety of health care? <i>Intern Med J</i> 2009;39(6):389-400	A2
24. Shahian DM, Meyer GS, Mort E, Atamian S, Liu X, Karson AS, Ramunno LD, Zheng H. Association of National Hospital Quality Measure adherence with long-term mortality and readmissions. <i>BMJ Qual Saf</i> 2012;21(4):325-36	A1
25. Sutton M, Nikolova S, Boaden R, Lester H, McDonald R, Roland M. Reduced mortality with hospital pay for performance in England. <i>N Engl J Med</i> 2012;367(19):1821-8	A1

#### 1.4 International - Handsuche

1. Albert JM, Das P. Quality indicators in radiation oncology. <i>Int J Radiat Oncol Biol Phys</i> 2013;85(4):904-11	A1
2. Grol R, Baker R, Moss F. Quality improvement research: understanding the science of change in health care. <i>Qual Saf Health Care</i> 2002;11(2):110-1	A1
3. Grimshaw J, McAuley LM, Bero LA, Grilli R, Oxman AD, Ramsay C, Vale L, Zwarenstein M. Systematic reviews of the effectiveness of quality improvement strategies and programmes. <i>Qual Saf Health Care</i> 2003;12(4):298-303	A1

4. Ovretveit J, Gustafson D. Evaluation of quality improvement programmes. Qual Saf Health Care 2002;11(3):270-5	A5
5. Lugtenberg M, Burgers JS, Westert GP. Effects of evidence-based clinical practice guidelines on quality of care: a systematic review. Qual Saf Health Care 2009;18(5):385-92	A1
6. Groene O, Skau JK, Frolich A. An international review of projects on hospital performance assessment. Int J Qual Health Care 2008;20(3):162-71	A2
7. Pitches DW, Mohammed MA, Lilford RJ. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. BMC Health Serv Res 2007;7:91	A1
8. Kruis AL, Smidt N, Assendelft WJ, Gussekloo J, Boland MR, Rutten-van Mölken M, Chavannes NH. Integrated disease management interventions for patients with chronic obstructive pulmonary disease. Cochrane Database Syst Rev. 2013 Oct 10;10	A4
9. Snyder C, Anderson G. Do quality improvement organizations improve the quality of hospital care for Medicare beneficiaries? JAMA. 2005 Jun 15;293(23):2900-7.	E2
10. Wagner C, Gulacsi L, Takacs E, Outinen M. The implementation of quality management systems in hospitals: a comparison between three countries. BMC Health Serv Res 2006;6:50	A2
11. Donabedian A. The effectiveness of quality assurance. Int J Qual Health Care 1996;8(4):401-7	A1

#### Legende

Ausschlussgründe (siehe auch Kapitel 2.1.1 Studienauswahl)	
A1	Keine Ergebnisse einer Evaluation des QS/QM-Programms
A2	Ergebnisse vorhanden, aber keine Evaluationsergebnisse eines <i>übergeordneten</i> Programms, singuläre QS-/QM-Intervention, Einzelfaktoren, keine Evaluation zur <i>Wirksamkeit</i>
A3	Ergebnisse vorhanden, aber kein <i>QS/QM</i> -Programm (z.B. Behandlungsprogramm mit Evaluation von Behandlungsinterventionen)
A4	nicht stationär/kein Krankenhaus
A5	Methodik (anderes Studiendesign als a priori festgelegt)
Einschlussgründe (siehe auch Kapitel 2.1.1 Studienauswahl)	
E1	Intervention QS/QM-Programm mit Evaluationsergebnissen zur Wirksamkeit - <i>Deutschland</i>
E2	Intervention QS/QM-Programm mit Evaluationsergebnissen zur Wirksamkeit - <i>international</i>
E3	Allgemeine Hintergrundinformationen
E4	Instrumente für QS/QM mit Evaluationsergebnissen zur Wirksamkeit

## 7.7 Abbildungsverzeichnis

Abbildung 1 Grundelemente komplexer Qualitätsmanagementprogramme .....	1
Abbildung 2 Zusammenhang der Begriffe des Qualitätsmanagements .....	5
Abbildung 3 Integration von Leitlinien in Qualitätsmanagement-Modelle .....	9
Abbildung 4 Flussdiagramm zum Prozess der Studienaushwahl.....	22
Abbildung 5 Anteil der Primär-Publikationen aus den einzelnen Fachgebieten.....	28
Abbildung 6 Verteilung der in den Studien gemessenen Qualitätsdimensionen.....	28
Abbildung 7 Verzerrungsrisiko der Primärstudien .....	29
Abbildung 8 Wirksamkeitsnachweis aller Studien .....	31

## 7.8 Tabellenverzeichnis

Tabelle 1: Beispiele für QM/QS-Instrumente und Verfahren .....	2
Tabelle 2 Historischer Überblick ausgewählter Qualitätssicherungsaktivitäten .....	11
Tabelle 3 Ein- und Ausschlusskriterien der Studienselktion.....	24
Tabelle 4 Übersicht der evaluierten Verfahren in Deutschland .....	54
Tabelle 5 Übersicht der evaluierten Verfahren international .....	62

## **8 Lebenslauf**

Mein Lebenslauf wird aus Gründen des Datenschutzes in der elektronischen Fassung meiner Arbeit nicht veröffentlicht.