

P4P: Aktuelle Einschätzung, konzeptioneller Rahmen und Handlungsempfehlungen



M. Schrappe

Dieser Artikel basiert auf einem Vortrag vom 23.10.2013 auf dem 12. Deutschen Kongress für Versorgungsforschung in Berlin.

Version 1.0.0. vom 31.3.2014

Der Artikel wird fortlaufend aktualisiert (s. Versionsgeschichte).

Es wird in regelmäßigen Abständen eine Druckversion erstellt, die Web-Version ist jedoch die aktuelle und primäre Quelle.

Zitate unter Angabe der Web-Adresse: <http://www.matthias.schrappe.com/texte/p4p/>

M. Schrappe

P4P: Aktuelle Einschätzung, konzeptioneller Rahmen und Handlungsempfehlungen

Inhaltsverzeichnis

Inhaltsverzeichnis

Zusammenfassung

Textfassung

1. Einleitung
 - 1.1. *Pay for Performance* - Aktualität des Themas
 - 1.2. Begriffsbestimmung und Definition
 - 1.3. Technische Umsetzung und Methodik der Qualitätsmessung
 - 1.4. Zum Begriff des Indikators
 - 1.5. Konzeption und Gesundheitssystem
2. Langfristige Effekte und Weiterentwicklung
 - 2.1. Anfänglicher Optimismus
 - 2.2. HQIP in den USA
 - 2.3. HQIP: Gesundheitspolitischer Kontext in den USA
 - 2.4. *Value-Based Purchasing*
 - 2.5. Ambulanter Sektor: QOF-Programm in Großbritannien
 - 2.6. Synopse: Unerwünschte Nebeneffekte von P4P
 - 2.6.1. Systematik, Zugang zur Versorgung
 - 2.6.2. Professionelle und institutionelle Effekte, Datenqualität
 - 2.6.3. Qualität auf Sektor- und Systemebene, Kosteneffektivität
3. Motivation, Organisation, System
 - 3.1. Kontext von P4P: Vier Domänen
 - 3.2. Professional Bureaucracy: die Expertenorganisation
 - 3.2.1. Das Konzept der Expertenorganisation
 - 3.2.2. Einschränkungen und Schwächen
 - 3.3. "It's Not A Rocket Science!"
 - 3.3.1. System- und Komplexitätstheorie
 - 3.3.2. Doppelte Komplexität

- 3.4. Konzeptioneller Rahmen: Komplexe professionelle Systembürokratie
- 3.5. Verhaltensänderung in der "komplexen professionellen Systembürokratie"
 - 3.5.1. Einführung, lerntheoretische Modelle
 - 3.5.2. Soziale Wahrnehmung
 - 3.5.3. Organisatorischer Wandel
 - 3.5.4. Lernen im Kontext, Zusammenfassung
- 4. Ökonomie
 - 4.1. Einführung
 - 4.2. Höhe der monetären Bewertung, Informationsasymmetrie
 - 4.3. Ökonomische Grundannahmen: *behavioural economics*
- 5. Integration in bestehende Vergütungssysteme
 - 5.1. Dominierende Vergütungsanreize im deutschen Gesundheitssystem
 - 5.2. Einzelleistungsvergütung und sektorale Pauschalen
 - 5.3. Von der transsektoralen zur Populationspauschale
- 6. Politische Verantwortung
 - 6.1. Governance im Gesundheitssystem
 - 6.2. P4P: Handlungsfelder der Politik (1)
 - 6.3. P4P: Handlungsfelder der Politik (2)
- 7. Empfehlung für die zukünftige Nutzung von P4P
 - 7.1. Allgemeines
 - 7.2. "P4P Is Here To Stay"
 - 7.3. Conceptual Framework - ein Rahmenkonzept ist Bedingung!
 - 7.4. P4P als *Feedback*-Instrument: Voraussetzungen
 - 7.5. Organisation und System beachten!
 - 7.6. Ökonomie und Vergütungssystem
 - 7.7. Politische Verantwortung

Literatur

Abkürzungen

Stichworte

Versionen

Impressum

Zusammenfassung

1. Aktualität und Rahmenkonzept

Die vorliegende Arbeit geht von der Darstellung der bisherigen wissenschaftlichen Evaluationsstudien zu *Pay for Performance* (P4P) bzw. zur Qualitäts-orientierten Vergütung aus. Trotz der Tatsache, dass die vorliegenden Evaluationsergebnisse zwar eine gewisse Wirksamkeit von P4P, aber keine durchschlagenden Ergebnisse zeigen, sind in den USA (*Value-Based Purchasing*-Programm (VBP)) und in Großbritannien (*Quality and Outcome Framework* (QOF)) Elemente der Qualitäts-orientierten Vergütung im Gesundheitswesen eingeführt worden. Diese Programme verfolgen einen langfristigen und umfassenden Ansatz. Ebenso wie im Koalitionsvertrag von CDU/CSU und SPD vom November 2013 als politischer Wille erkennbar, wurde in diesen Ländern ein als paradigmatisch zu bezeichnender Wandel eingeleitet, der das Ziel hat, statt der Leistungsmenge verstärkt die Qualität der Leistung zu vergüten.

Um einen Hintergrund für die weitergehende Analyse der wissenschaftlichen Daten und für die Weiterentwicklung von P4P zu schaffen, wird in dieser Arbeit ein **Rahmenkonzept** entwickelt, das die Ableitung von Empfehlungen für die Einführung und Umsetzung von P4P auch im Kontext des deutschen Gesundheitswesens erlaubt. Dieses Rahmenkonzept wird über mehrere Stufen entwickelt und enthält Bezüge zum organisationstheoretischen Konzept der Expertenorganisation und zur Komplexitätstheorie, untersucht die Implikationen aus den grundlegenden Modellen der Verhaltensänderung mit Schwerpunkt soziale Rollen, organisatorischer Wandel und Lernen im Kontext, bezieht ökonomische Aspekte (*principal agent*-Theorie, Verhaltensökonomie) sowie die Wechselwirkungen mit anderen Vergütungssystemen mit ein und schließt mit einer Einschätzung der gesundheitspolitischen Optionen ab, die sich in einem komplexen System wie dem Gesundheitswesen mit seiner korporatistischen Struktur eröffnen.

Zu Beginn wird der **Begriff** *Pay for Performance* (P4P) bzw. Qualitäts-orientierte Vergütung (Endpunkt Qualität) unter Einbeziehung des verwandten *Value-Based Purchasing* (Endpunkt Effizienz) untersucht und verfeinert. P4P ist ein Feedback-Verfahren, das Wettbewerbselemente einsetzt, um Qualitätsdefizite in Gesundheitsversorgung und Prävention günstig zu beeinflussen. Es ist aus zwei Systematiken zusammengesetzt und basiert auf (1) einer Qualitätsmessung durch definierte Indikatoren und (2) einer nachvollziehbaren Kopplung der Qualität oder Effizienz (*value*) der Versorgung an Vergütungsbestandteile. P4P wird in der Regel nicht als einziges, alleinstehendes Vergütungssystem, sondern integriert in andere

Vergütungssysteme (z.B. DRG-System) eingesetzt. Es gehört zu den Instrumenten des sog. Qualitätswettbewerbes und setzt direkte finanzielle Anreize ein, während *public reporting* (s. Qualitätssicherung nach §137 SGB V) seine Wirkung über indirekte finanzielle Anreize (Marktvorteil wegen besserer Qualität) erzielt.

Die Entwicklung in den USA ist vom Krankenhausbereich ausgegangen (*Premier Hospital Quality Incentive Demonstration Project* (HQIP)), in Großbritannien stand von Anfang an die ambulante Versorgung im Mittelpunkt. Seit 1.10.2013 ist die Teilnahme aller Krankenhäuser in den USA verpflichtend, die Medicare-Patienten behandeln. Bis 2017 wird der Anteil des Erlöses, der über Qualitätsindikatoren verteilt wird, auf 2% des Gesamterlöses ansteigen. Als Indikatoren, die fünf ausgewählte Krankheitsbilder betreffen, werden Prozessindikatoren, Patienten-Erfahrungen, Patientensicherheitsindikatoren, *Hospital-Acquired Conditions* und in ausgewählten Fällen risikoadjustierte Ergebnisindikatoren eingesetzt. Im ersten Jahr wurden nur Prozessindikatoren und solche auf der Basis von Patienten-Erfahrungen (im Verhältnis von 70 zu 30%) verwendet, ab dem 2. Jahr (2014) werden die 5 Gruppen je 20% gewichtet. Auch im *Quality and Outcome Framework* (QOF) in Großbritannien stehen Prozessindikatoren im Vordergrund, die entsprechend des Einsatzes im ambulanten Sektor konsequent auf die Behandlung und Prävention chronischer Erkrankungen ausgerichtet sind. In beiden Ländern sind die jetzigen Programme das Ergebnis einer rund 15jährigen Entwicklung weitgehender Konzeptionen, die jeweils das gesamte Gesundheitssystem teilweise auch unter Einbeziehung des *Public Health*-Bereichs umfassen. In den USA hat diese Konzeption eine stringente **system- bzw. komplexitätstheoretische Ausrichtung** und wurde vom *Institute of Medicine* in "*Crossing the Quality Chasm*" im Jahr 2001 veröffentlicht.

In der hier vorliegenden Arbeit wird diese komplexitätstheoretische Ausrichtung aufgenommen und um das Konzept der **Expertenorganisation** (*professional bureaucracy*) ergänzt, das im Jahr 1979 erstmalig von Mintzberg vorgestellt worden war. Obwohl die Konzepte im ersten Fall einen systemischen, im zweiten Fall einen organisatorischen Fokus aufweisen, ergänzen sie sich in weiten Bereichen hervorragend, insbesondere durch die Tendenz zur Selbstorganisation, das hohe Maß der Autonomie ihrer Mitglieder und die Verdecktheit der gleichwohl vorhandenen internen Regeln. Besonders ergibt die Analyse jedoch Übereinstimmungen im Bereich der Innovation und im Umgang mit Unsicherheit. Beide Konzepte weisen einerseits eine hohe Innovationsbereitschaft auf, sind andererseits in ihrer Reaktion auf die Umwelt jedoch verhalten und meiden daraus resultierende Innovationsansätze ("Innovations-Paradoxon"). Noch entscheidender für den Bereich Qualität und Patientensicherheit ist

aber die Beobachtung, dass beide Konzepte eine weitgehende Toleranz von Unsicherheit aufweisen ("intrinsische Unsicherheit"). Die Expertenorganisation erreicht Koordination über Standardisierung in der Ausbildung und durch den Einsatz von *peers*, die Ärzte im Nachtdienst sind jedoch allein und müssen "durchkommen". Das komplexe System weist Unsicherheit, Spannung und Paradoxie sogar als konstituierendes Element seiner Struktur aus. Die Toleranz gegenüber Unsicherheit ist einerseits gewiss ein Vorteil, denn sie schützt vor voreiliger Festlegung und irreführender Linearität, andererseits weist sie durch die Innovationsresistenz gegenüber Umweltreizen und somit auch externen Qualitätserwartungen schwerwiegende Defizite auf. Um in der vorliegenden Arbeit diese Synergien beider Konzepte kenntlich zu machen, wird der Arbeitsbegriff "**komplexe professionelle Systembürokratie**" verwendet.

Von P4P wird erwartet, dass es eine **Verhaltenveränderung** bewirkt, die aufseiten der Leistungserbringer und Gesundheitsberufe in einer Verbesserung der Versorgungsqualität resultiert. P4P ist ein Feedback-Verfahren, das einen externen finanziellen Anreiz setzt und den lerntheoretischen Konzepten zuzurechnen ist. Folgt man dem Maslow'schen Konzept, setzt P4P an sehr basalen Bedürfnissen an, eine Veränderung sozialer Rollen oder gar Wertschätzung wird nicht angesprochen. Allerdings hat es in den letzten Jahrzehnten mehrere Versuche gegeben, mit Modellen der sozialen Wahrnehmung, die über den individuellen Ansatz der lerntheoretischen Konzept hinausgehen und z.B. Haltungen und Einstellungen thematisieren, die Veränderungsresistenz des Gesundheitswesens zu überwinden. In diesem Zusammenhang wurde auch auf das Konzept des Professionalismus gesetzt, das interne Motivation, Altruismus und Autonomie in den Mittelpunkt stellt; es existieren hier Parallelen zur Diskussion z.B. um die Einführung von Leitlinien oder der *Evidence-based Medicine*. Zusammenfassend muss man jedoch festhalten, dass aus diesen Ansätzen keine Konzeption entwickelt werden konnte, die tragfähige Lösungen für die virulenten Probleme bietet. Der Grund ist darin zu suchen, dass diese Ansätze zu kurz greifen, die Einbeziehung von Konzepten des organisatorischen Wandels und des Kontext-bezogenen Lernens (soziales Marketing) ist notwendig. Das Rahmenkonzept beinhaltet folglich ein eindeutiges Plädoyer für ein Organisations- und Kontext-bezogenes Vorgehen bei der Einführung von Innovationen wie P4P im Gesundheitswesen, die auch einer politischen Steuerung (*direction pointing*, s.u.) und eines langfristigen *commitment* der beteiligten Akteure bedürfen.

Dieser Ansatz führt zu **ökonomischen Festlegungen**, ohne die P4P nicht erfolgsversprechend einzuführen ist. Bei den Überlegungen zur Höhe der P4P-Vergütung müssen Opportunitätskosten und Diskontierung mit einbezogen werden,

insbesondere wenn man eine Risikoaversion der Leistungserbringer und gegebenenfalls die Unsicherheit einer Zahlung mit berücksichtigt. Letztere ist vor allem relevant, wenn man die Zahlung an eine relative Position in einer Rankingliste ("die besten fünf") koppelt, weil die einzelnen Einrichtungen erst spät einschätzen können, ob sich ihre Investition in Qualität über die P4P-Vergütung "auszahlt". Gestaffelte absolute Grenzwerte und die Vergütung relativer Positionsverbesserungen sind daher vorzuziehen. Die *principal agent* - Theorie verweist auf die notwendige Abgrenzung von P4P und der Einzelleistungsvergütung, letztere ist vorzuziehen, wenn die Informationsasymmetrie Qualitäts-relevanter Leistungen z.B durch EBM aufgehoben ist (z.B. Blutkultur vor Antibiotika-Gabe bei der Pneumonie). P4P sollte auf Bereiche beschränkt werden, bei denen die Leistungserbringer einen deutlichen Wissensvorsprung bzgl. des Zustandekommens der Versorgung aufweisen können, z.B. in der Behandlung von chronischen Erkrankungen. Die Einbeziehung der Erkenntnisse der *behavioural economics* (Verhaltensökonomie) mit ihren Elementen *framing*, *isolation effect* und Überschätzung relativer Risiken erbringt schon erste Empfehlungen für das weitere Vorgehen. Neben der oben bereits genannten Risikoaversion (Zeitnähe, Verlässlichkeit und Nachvollziehbarkeit der Zahlungen) sind kleinere, häufigere Zahlungen mit on/off-Charakteristik wirkungsvoller als größere aber nur selten gezahlte Summen (z.B. im Rahmen des Gesamterlöses eines Krankenhauses), die Einbehaltung von Vergütungsbestandteilen der stärkere Anreiz als die zusätzliche Vergütung (Verlust-Aversion), die Indikatoren sollten "unverbraucht" sein und nicht schon vorher Gegenstand anderer Systeme (z.B. *Public Reporting*) gewesen sein, weil dann kein weiterer Effekt mehr auftritt (*ceiling*), und ein P4P-Programm sollte nicht gleichzeitig mit mehreren anderen Regelungen eingesetzt werden (*isolation effect*). Insbesondere bei Ergebnisindikatoren muss die Risikoadjustierung optimal gestaltet werden, weil sich ansonsten bei Erkrankungen mit niedriger Inzidenz vor allem kleinere Einrichtungen wegen des Morbiditäts- und Komorbiditätsrisikos nicht an dem P4P-Programm beteiligen bzw. alternativ Risikoselektion betreiben. Daher sind Prozessindikatoren, die keiner Risikoadjustierung bedürfen, zu präferieren, evtl. ergänzt durch einige Ergebnis- (z.B. adjustierte Mortalität) und Strukturindikatoren.

Da - wie bereits angemerkt - P4P grundsätzlich in **bestehende Vergütungssysteme** (z.B. DRG) "eingebettet" wird, sind die Wechselwirkungen mit diesen Systemen von großer Bedeutung für dessen Wirkung. Das deutsche Gesundheitssystem ist sehr Mengen-orientiert, stark sektoral gegliedert, wenig auf Prävention ausgerichtet und vor allem fokussiert auf Akuterkrankungen statt auf chronische und Mehrfacherkrankungen. Beabsichtigt man nun, durch P4P dem Mengenanreiz einen Qualitäts-Anreiz

entgegenzusetzen ("Qualität statt Menge"), dann ist dies gerade bei den Vergütungssystemen mit dem größten Mengenanreiz (Einzelleistungsvergütung und DRG) besonders schwierig, weil dort die Opportunitätskosten am höchsten sind ("ein Fall geht noch"). Es besteht sogar die Gefahr, dass bei Einzelleistungsvergütung und sektoralen Pauschalen der **Mengenanreiz** durch P4P verstärkt wird, und zwar wenn folgende Bedingungen zusammentreffen: (1) es handelt sich um Leistungen, bei denen die Möglichkeit zur Mengenausweitung besteht (z.B. Endoprothetik), (2) es werden Indikatoren mit geringer Sensitivität bzgl der Qualitätsprobleme verwendet, wie es bei Indikatoren auf der Basis administrativer Daten der Fall ist, (3) mit den resultierenden Qualitätsdaten wird eine Mengenausweitung begründet und in den Verhandlungen mit den Kostenträgern durchgesetzt, und eventuell wird zusätzlich (4) eine aktive Risikoselektion betrieben, weil die Einrichtung aufgrund ihrer Größe mit eigenen Daten eine Risikobewertung ihrer Patienten betreiben kann.

Sieht man allerdings vom Mengenanreiz ab, können bei Einzelleistungsvergütung und sektoraler Pauschalierung durchaus interessante Einsatzmöglichkeiten für P4P darin bestehen, die Koordination der Behandlung und überhaupt die Behandlung von chronischen Erkrankungen zu verbessern. Man kann den Nachteil chronischer Erkrankungen aus dem nicht-operativen Bereich, der dadurch entsteht, dass der primäre ökonomische Anreiz bei diesen Vergütungsformen eher auf operativ zu behandelnde Akuterkrankungen gerichtet ist, versuchen auszugleichen und hätte damit eines der dringenden Qualitätsprobleme in Deutschland aufgegriffen. Gleiches gilt für auch für andere Themen wie Patientensicherheit (z.B. Indikatoren zur Einführung und sinnvollen Handhabung von Instrumenten wie CIRS). In den genannten Fällen sind Prozessindikatoren sinnvoll einzusetzen, die keiner Risikoselektion bedürfen. In Kombination mit höhergradig pauschalierenden Vergütungssystemen (integrierte transsektorale Versorgung, Erkrankungspauschalen, *Managed Care*), die mehr Gewicht auf die Koordination der Behandlung und die Versorgung von Patienten mit chronischen, multiplen Erkrankungen legen, kann man mit einem gezielten Einsatz von P4P sinnvoll eingreifen und Schwerpunkte in der Qualitätsentwicklung setzen (s. QOF-Projekt in Großbritannien). Gegen den Anreiz zur Risikoselektion kann man durch die Wahl von Prozessindikatoren und/oder eine adäquate Risikoadjustierung von Ergebnisindikatoren gegensteuern. Die Risikoadjustierung findet jedoch ihre Grenzen bei Erkrankungen mit geringerer Fallzahl, bei denen es zu einer Benachteiligung von kleineren Einrichtungen kommt. Ähnlich es es durch einen adäquaten Einsatz von P4P grundsätzlich möglich, den vor allem durch Diskontierungsaspekte gehemmten Einsatz der Prävention zu fördern, ohne dabei den Grundanreiz der jeweiligen Vergütung zu

verändern.

Das Rahmenkonzept wird abgerundet durch die **politische Ebene**.

Expertenorganisation und Komplexitätstheorie weisen den Institutionen und den professionellen Strukturen der Selbstorganisation eine wichtige Rolle zu, gerade die korporatistischen Strukturen der Selbstorganisation, in Deutschland unter dem Begriff der Selbstverwaltung zusammengefasst, sind in den letzten beiden Jahrzehnten immer wichtiger geworden, parallel zum Wandel des politischen Grundverständnisses weg von einem hierarchischen Modell zu Konzepten wie dem *Governance*-Konzept. Dieser Wandel ist durchaus als funktional zu bezeichnen, entspricht er doch auf der organisatorischen Ebene der Expertenautonomie, auf der Systemebene der Komplexität des Gesundheitssystems, in Bezug auf die notwendige Verhaltensänderung den Kontext-bezogenen Theorien und hinsichtlich der ökonomischen Grundannahmen den verhaltensökonomischen Erkenntnissen. Es werden jedoch vier Punkte herausgearbeitet, in denen es keine Alternative zur Übernahme der Verantwortung durch die politische Ebene gibt: die Richtung muss vorgegeben werden (*direction pointing* ist als Begriff in "*Crossing the Quality Chasm*" eingeführt worden), strategische Ziele müssen gesetzt werden, weil die Auswahl der Indikatoren weder zufällig noch aus Opportunität erfolgen darf, potentielle negative Auswirkungen müssen kontrolliert werden, und es müssen Rahmenbedingungen geschaffen werden, die den Hintergrund für die Initiierung des politischen Prozesses, für die Implementierung, die Umsetzung und die Evaluation bilden.

Zusammenfassung

2. Empfehlungen und Résumé

Aus den vorliegenden empirischen Daten und den aktuellen Konzepten (VBP, QOF) werden vor dem Hintergrund des in dieser Arbeit entwickelten Rahmenkonzeptes insgesamt 31 Empfehlungen abgeleitet, die teils grundsätzliche, teils technische und praktische Fragen betreffen. In einigen Fällen lassen sich Empfehlungen entwickeln, die auf den ersten Blick nicht mit den gängigen Vorstellungen kongruent sind; diese sind hier in der Zusammenfassung gekennzeichnet.

Die drei ersten Empfehlungen betreffen allgemeine Fragen (s. Kap 7.1., 7.2. und 7.3.):

Empfehlung 1: Umfassendes Rahmenkonzept

Pay for Performance (P4P) ist kein isoliert einzusetzendes Instrument, sondern nur im Zusammenhang mit zahlreichen Kontextfaktoren (z.B. anderen Vergütungselementen) wirksam. Die Implementierung, Weiterentwicklung und Evaluation von P4P kann folglich nur vor dem Hintergrund eines umfassenden Rahmenkonzeptes erfolgen, das sowohl die professionellen als auch die organisatorischen, ökonomischen und Systemfaktoren umfasst. Eine erkennbare Wirksamkeit von P4P ist nur dann zu erwarten, wenn durch eine vorangegangene strategische Analyse die Ansatzpunkte im Gesundheitssystem identifiziert werden konnten, die von P4P beeinflussbar sind.

Empfehlung 2: P4P ist als langfristige Entwicklung zu etablieren

Es gibt keine Alternative zum Einsatz von P4P, allerdings muss es kritisch, differenziert und integriert in andere Entwicklungen im Gesundheitswesen implementiert werden, denn in der derzeitigen Situation kann auf kein Instrument verzichtet werden, das die Qualitäts- bzw. Sicherheitsprobleme günstig zu beeinflussen verspricht. Aber P4P ist keine *magic bullet*, mit dem alle Defizite des Gesundheitssystems geheilt werden können, es ersetzt keine genuine Weiterentwicklungsperspektive auf Systemebene (Beispiel DRG-System und Sektorproblematik).

Empfehlung 3: Rahmenkonzept für die Implementierung und die Umsetzung von P4P

P4P sollte nicht als rein lerntheoretisch begründetes, einfaches Feedback-Instrument mit Belohnungskomponente verstanden und eingesetzt werden, dies wäre "zu kurz gesprungen". Stattdessen müssen sowohl Konzepte des organisatorischen Wandels und Kontext-bezogene Veränderungsstrategien als auch ökonomische, vergütungstechnische und politische Weichenstellungen mit einbezogen werden, erst dann kann von einer realistischen Erfolgsschance ausgegangen werden. Aus ökonomischer Sicht sind vor allem Opportunitätskosten, Diskontierung bzw. Zeitachse und Aspekte der Risikoaversion zu bewerten, außerdem ein optimales *framing*. Auf der *policy*-Ebene geht es in erster Linie um das *direction pointing* (Interpretationsebene für die anstehenden Entwicklungen) und um strategische Überlegungen bzgl. der optimalen Einsatzgebiete der P4P-Vergütung. Weiterhin müssen günstige Rahmenbedingungen geschaffen und eventuelle negative Nebeneffekte kontrolliert werden.

In der zweiten Gruppe von Empfehlungen werden technische Aspekte des Feedback einschließlich des finanziellen Anreizes adressiert. Es ist wenig überraschend, dass die Auswahl der Indikatoren im Vordergrund steht, umfassend Vorhersagefunktion, Betonung von Patientensicherheit und Patienten-Erfahrungen, die Wahl zwischen administrativen und klinischen sowie von Prozess- von Ergebnisindikatoren, die Risikoselektion und Stichproben als Mittel gegen ein *gaming*.

Empfehlung 4: Die Wahl von Indikatoren mit Vorhersagefunktion erfordert eine strategische Problemanalyse

Am Anfang der Implementierung eines P4P-Programmes steht die Entscheidung, ob einzelne (Qualitäts-relevante) Leistungen eingekauft werden sollen oder ob Indikatoren mit Vorhersagefunktion für die Qualitätsprobleme eines Versorgungsbereiches genutzt werden sollen. Im ersten Fall etabliert man eine Einzelleistungsvergütung, im zweiten Fall kann man von P4P im eigentlichen Sinne sprechen, weil die Problemlösungskompetenz der Leistungsanbieter vor Ort anerkannt wird. Eine vorherige Problemanalyse ist jedoch unverzichtbar.

Empfehlung 5: Patientensicherheits-Indikatoren und Patientenerfahrungen integrieren

Im Vordergrund der internationalen P4P-Programme stehen Patientensicherheits- und Indikatoren auf der Basis von Patientenerfahrungen. Neben strategisch positionierten Indikatoren zu verstärkter Integration, Behandlung chronischer Erkrankungen und zur Stärkung der Prävention sollten diese auch in Deutschland besonders beachtet werden.

Empfehlung 6: Administrative Indikatoren haben schlechte Sensitivität, klinische Indikatoren sind keine Diagnosen

Administrative Indikatoren sollten wegen ihrer mangelnden Sensitivität nur in P4P-Konzepten mit Einzelleistungsvergütungs-Charakter verwendet werden (s. Empfehlung 4). Wenn man auf administrative Indikatoren setzt, läuft man die Gefahr einer verstärkten Mengenausweitung, außerdem wird der Qualitätsgedanke geschwächt, weil diese Indikatoren nicht sensitiv genug sind, also die Ereignisse, die abgebildet werden sollen, nur unvollständig wiedergeben. In der Expertenorganisation wird Qualitätserfassung als *management-business* angesehen. Bei der Verwendung von klinischen Indikatoren ist der Erhebungsaufwand einzuplanen (und in der Höhe der Anreize zu berücksichtigen), weiterhin ist auf den Unterschied von epidemiologischen Falldefinitionen (s. nosokomiale Infektionen) zu klinischen Diagnosen hinzuweisen.

Empfehlung 7: Prozessindikatoren sind entscheidend, Ergebnisindikatoren können ergänzen

Im Mittelpunkt sollten wie in den großen internationalen Referenzprojekten (USA, UK) Prozessindikatoren stehen, da bei Ergebnisindikatoren die Problematik der Risikoselektion zu gewichtig ist, und selbst wenn die Risikoadjustierung perfekt entwickelt wäre, kleine Einrichtungen systematisch schlechter gestellt werden. Für die Zukunft sind jedoch Ergebnisindikatoren nicht völlig ausgeschlossen (in erster Linie Outcome-relevante Patientensicherheits-Indikatoren (z.B. Katheterinfektionen).

Empfehlung 8: Risikoselektion bedeutet Unterversorgung

Besonders bei zugrundeliegenden Vergütungsansätzen (z.B. Pauschalen), die bereits ihrerseits zur Risikoselektion neigen, ist eine zusätzliche P4P-bedingte Risikoselektion geeignet, eine relevante Unterversorgung zu verursachen. Zumindest soweit Ergebnisindikatoren eingesetzt werden, ist daher eine möglichst optimale Risikoadjustierung notwendig. Prozessindikatoren bedürfen im allgemeinen keiner Risikoadjustierung.

Empfehlung 9: Sobald Indikatoren finanziell relevant werden, sind effektive Stichproben notwendig

Eine P4P-Einführung ohne die gleichzeitige Implementierung von Stichproben zur Kontrolle des *gaming* ist nicht sinnvoll.

Empfehlung 10: Die organisatorischen Voraussetzungen einer verbesserten Integration der Versorgung fördern

Da in einem stark sektorierten Gesundheitssystem Verantwortungsbereich und Anreiz nur schwer in Deckung zu bringen sind, müssen die Indikatoren (Prozessindikatoren) auch organisatorische Aspekte einer verstärkten Integration (IT, gemeinsame Verantwortung für die Patienten) mit einbeziehen.

Empfehlung 11: Feedback nur zeitnah gestalten!

Der Feedback muss zeitnahe gestaltet sein, am besten außerhalb der normalen Erlösvergütung, so dass die P4P-bezogene Vergütung klar als Anreiz erkennbar ist und nicht von Zeitferne und Unsicherheit geschmälert wird.

Empfehlung 12: Ein *ceiling*-Effekt muss bedacht und verhindert werden

Ein *ceiling*-Effekt tritt vor allem bei Übergang in Einzelleistungsvergütung und bei nicht aktuellen Indikatoren auf.

In der vorliegenden Arbeit werden Prozessindikatoren zumindest für den Einstieg in eine P4P-Vergütung klar präferiert, die nach strategischen Überlegungen ausgewählt bzw. entwickelt werden und sich unter Berücksichtigung organisatorischer Aspekte einer verstärkten Integration besonders auf die Behandlung chronischer Erkrankungen und die Prävention beziehen. Die nächste Gruppe von Empfehlungen richtet sich auf organisatorische und systemische Faktoren, die bei der Implementierung und Praxis von P4P zu beachten sind. Ein *motivation crowding out* ist nicht zwingend zu befürchten, der organisatorische Wandel führt neben den Kontext-bezogenen Modellen als Konzept der Verhaltensänderung, und last but not least: man sollte diskutieren, ob man nicht die Effizienz statt der Qualität allein zur Zielgröße von P4P wählt.

Empfehlung 13: Multiprofessionelle und multidimensionale Implementierung ist zu bevorzugen

Die monoprofessionelle Beschränkung des Professionalismus-Konzeptes auf interne Motivation, Altruismus und Autonomie hat zu keiner tragfähigen Strategie geführt. Die

Implementierung von P4P sollte einem multiprofessionellen Ansatz folgen und primär die Rollenverständnisse der beteiligten Berufsgruppen weiterzuentwickeln versuchen, ohne die Miteinbeziehung von Konzepten des organisatorischen Wandels und des Kontextlernens wird jedoch kein Erfolg zu erzielen sein.

Empfehlung 14: Die Organisation als Ganzes ansprechen, den organisatorischen Wandel fördern

P4P setzt die Organisation unter "internen Stress", es müssen jetzt nämlich zusätzlich zu Kosten und Mengen-bezogenen Erlösen auch Erlös-wirksame Qualitätsinformationen intern bearbeitet werden, und zwar sowohl hinsichtlich der Abgrenzung von Verantwortung als auch von Aufgabenstellungen und in der Führungsarbeit.

Empfehlung 15: *Motivation Crowding Out* - im Blick behalten, aber nicht überschätzen

Es ist nach den empirischen Ergebnissen und aus theoretischer Sicht unklar, ob sich bei P4P externe und interne Motivation gegenseitig schwächen. Das Problem sollte begleitend untersucht werden (Befragungen etc.).

Empfehlung 16: *Value* (Effizienz) statt allein Qualität als Zielgröße für P4P-Programm in Betracht ziehen

Das VBP-Programm in den USA setzt auf *value* (Effizienz) als Zielgröße für P4P und damit evtl. eine bessere Akzeptanz erreichen, denn die Kosten der Versorgung sind ein etabliertes Kriterium für die Außenbeziehungen der Gesundheitseinrichtungen.

Empfehlung 17: Attraktoren des komplexen Systems diskutieren, Attraktoren nutzen

Es ist lohnenswert, sich getreu der Konzepte der System- bzw. Komplexitätstheorie Gedanken über mögliche Attraktoren zu machen und sie bei der Implementierung von P4P-Programmen zu nutzen.

Selbst wenn P4P als perfektes, "erweitertes" Feedback imponiert, kann dieses Instrument im isolierten Einsatz keine grundlegenden Umkehrungen bestehender Anreize im Gesundheitswesen bewirken, sondern bedarf zusätzlicher ökonomischer (und außerdem politischer) Weichenstellungen. Diese Weichenstellungen betreffen neben einfachen ökonomischen Überlegungen (z.B. Opportunitätskosten) vor allem Konsequenzen aus der *principal agent* Theorie (Informationsasymmetrie) und der Verhaltensökonomie (z.B. Risikoaversion) sowie Wechselwirkungen von P4P mit dem dominierenden Vergütungssystem, in das P4P-Elemente integriert werden (z.B. DRG-System). Ganz im Vordergrund stehen Fragen der Informationsasymmetrie, die für das Konzept von P4P als durchaus kritisch anzusehen sind und die Berechtigung des P4P-Ansatzes in Abgrenzung von der Einzelleistungsvergütung potentiell in Frage stellen ("*Is P4P really Fee for Service?*").

Empfehlung 18: Höhe der P4P-Zahlungen muss besonders im DRG-System Opportunitätskosten berücksichtigen

Die Höhe des Qualitäts-bezogenen Erlösanteiles muss den Messaufwand, die Opportunitätskosten und die Diskontierung berücksichtigen und hängt daher stark von der Art des dominierenden Vergütungssystems ab. Bei Einzelleistungsvergütung und DRGs sind hohe Opportunitätskosten anzusetzen, die Höhe der P4P-Vergütung muss entsprechend hoch angesetzt werden.

Empfehlung 19: Kleine und häufigere Zahlungen sind zu präferieren

Kleine, häufigere und Ereignis-bezogene P4P-Zahlungen mit on/off-Charakteristik sind größeren integrierten Zahlungen vorzuziehen

Empfehlung 20: Relative Position, relative Verbesserungen und absolute Grenzwerte kombinieren

Die monetäre Kopplung der P4P-Vergütung kann nicht alleine aufgrund der relativen Position auf einer Rankingliste vorgenommen werden, sondern muss ebenso gestaffelte Grenzwerte und relative Positionsverbesserungen mit einbeziehen, damit auch die *poor performer* einen realistischen Anreiz zur Qualitätsverbesserung haben.

Empfehlung 21: Besonders bei Prozessindikatoren Übergang zur Einzelleistungsvergütung beachten!

Konsequenzen aus der *principal-agent* Theorie und damit Überschneidungen zur Einzelleistungsvergütung gehören zu den schwierigsten Fragestellungen, mit denen sich das P4P-Konzept auseinandersetzen hat. Die wichtigste Einsicht besteht darin, dass die Einzelleistungsvergütung überlegen ist, wenn es sich um gut bekannte, wissenschaftlich abgesicherte Prozessindikatoren handelt. P4P ist nur sinnvoll, wenn eine Informationsasymmetrie besteht, d.h. das eigentliche Wissen um die Gestaltung der qualitativ angestrebten Leistung bei den Leistungserbringern liegt. Dies ist z.B. bei Indikatoren aus dem Grenzbereich zwischen Ergebnis- und Prozessindikatoren (z.B. zur Prävention vermeidbarer Komplikationen), bei der Behandlung chronischer oder Mehrfacherkrankungen oder bei prozessualen Patientensicherheits-Indikatoren der Fall. Bei Ergebnisindikatoren ist die Informationsasymmetrie meist kein Problem, hier stellt sich allerdings mit Macht die Problematik der Risikoselektion. Strukturindikatoren sind als Investitionszuschuss anzusehen.

Empfehlung 22: Indikatoren müssen aktuell sein und regelmäßig gewechselt werden

Um Aktualität und Wechsel zu gewährleisten, ist die Delegation an ein Institut wie dem geplanten Qualitätsinstitut sinnvoll

Empfehlung 23: Keine unhinterfragte Kombination mit *public reporting*, beide Instrumente sind nicht synergistisch

Entgegen der bisherigen Ansicht ist die Kombination mit *public reporting (dual use)* äußerst kritisch zu sehen (*ceiling*, Aufhebung der Informationsasymmetrie, komplementäre Wirkung hinsichtlich Wettbewerbsdichte), eher ist ein differenzierter Einsatz beider Instrumente sinnvoll. P4P ist im Vergleich zu *public reporting* besonders wirksam in Bereichen (Regionen, Krankheitsgruppen), in denen die Wettbewerbsdichte

gering ist. Ein nach Wettbewerbsdichte differenzierter Einsatz nach Fachgebieten ist ein interessanter Ansatz, der in der Zukunft zu diskutieren wäre.

Empfehlung 24: Berücksichtigung von Risiko- und Verlustaversion sowie optimales *framing* erforderlich

Die Einführung von P4P muss auf die Risiko- und Verlustaversion der Einrichtungen Bezug nehmen, der Diskontierung entgegenzutreten und die positiven Aspekte (Qualitätsverbesserung, professionelle Autonomie) in den Vordergrund stellen. Das *framing* kann nicht ohne gesellschaftliche bzw. politische Rahmenbedingungen geschehen, die die Notwendigkeit und die Perspektiven eines Wechsels der Vergütungslogik als allgemeinen Konsens wiedergeben.

Empfehlung 25: Verlässliche und nachvollziehbare Zahlungen

Das Verständnis für das Zustandekommen der P4P-Zahlungen verbessert die Wirksamkeit des Feedback.

Empfehlung 26: Integration in andere Vergütungssysteme kritisch und differenziert vornehmen

P4P kann grundsätzlich nicht die dominanten Anreize der Vergütungssysteme ausgleichen, in die es "eingebettet" wird. Dies gilt in Deutschland vor allem für das DRG-System im Krankenhausbereich. Allerdings ist es möglich, eine vorgeschaltete strategische Analyse vorausgesetzt, durch einen differenzierten Einsatz von P4P Bereiche zu fördern, die vom DRG-System vernachlässigt werden. Hier ist insbesondere an die Versorgung chronisch, mehrfach erkrankter Patienten zu denken, weiterhin an die Koordination über die Sektorgrenzen hinweg und die Prävention

Die Koordination einer Gesellschaft kommt nicht ohne eine hierarchische Ebene aus, die Markt, Organisationen und Gemeinschaft in dieser Aufgabe ergänzt. Die korporatistischen Strukturen der Selbstverwaltung sind in den letzten beiden Jahrzehnten immer wichtiger geworden, parallel zu einem Wandel des politischen Grundverständnisses weg vom hierarchischen Verständnis hin zu Konzepten wie *Governance*, das das Zusammenwirken von staatlichen, privaten und zivilgesellschaftlichen Strukturen in den Mittelpunkt stellt. Diese Verbreiterung der peripheren Regelungskompetenz mindert jedoch nicht die Notwendigkeit politisch gesetzter Rahmenbedingungen, neben dem *direction pointing* spielt hier die strategische Analyse und das daraus resultierende Setzen strategischer Ziele der Gesundheitspolitik eine große Rolle, weiterhin die Kontrolle eventueller negativer Auswirkungen, und viertens die Etablierung von Rahmenbedingungen.

Empfehlung 27: *direction pointing* als elementares Mittel zur Implementierung nutzen

Direction pointing weist wichtige koordinative und unterstützende Eigenschaften auf, ohne die eine Innovation wie ein P4P-Programm nicht erfolgreich eingeführt und umgesetzt werden kann. *Direction pointing* setzt einen aktiven und sichtbaren Einsatz der politischen Ebene voraus, der Qualität als Richtschnur der gesundheitspolitischen Agenda betont, der alle Ebenen des Gesundheitssystems einbezieht, öffentliche Foren schafft, die mit der Thematik beschäftigten Fachgesellschaften fördert, die Ausbildung entsprechend differenziert und Kontakte zu Einrichtungen der Zivilgesellschaft schafft.

Empfehlung 28: Strategische Ziele setzen I - kein "weiter so"

Aufgrund der Analysen dieser Arbeit kann eine Weiternutzung von langjährig verwendeten Indikatoren nicht empfohlen werden, genauso wenig wie ein *dual use* mit Indikatoren aus dem *public reporting* (s. Empfehlung 23), weil P4P in dieser Konstellation nach den vorliegenden Daten keine Wirksamkeit aufweist. Da der Einsatz von Ergebnisindikatoren wegen der Risikoselektion und der strukturellen Benachteiligung kleinerer Einrichtungen nicht unproblematisch ist, bleiben Prozessindikatoren, die nach einer vorangegangenen Analyse des Handlungsbedarfs ausgewählt werden.

Empfehlung 29: Strategische Ziele setzen II - Prozessindikatoren Problemorientiert auswählen

Nicht alle Prozessindikatoren sind sinnvoll bei P4P einzusetzen (s. Empfehlung 21). Es müssen diejenigen Prozessindikatoren identifiziert werden, bei denen die Informationsasymmetrie nicht aufgehoben ist (s.o.) und die übergeordneten Zielen in der strukturellen Weiterentwicklung des deutschen Gesundheitssystems dienen: Verbesserung der Versorgung von chronisch und mehrfach erkrankten Patienten hohen Alters, Abbau der Mengenorientierung, Verbesserung der Koordination, Förderung der Prävention und Integration der Sektoren. Diese Zielsetzung kann nur auf der politischen Ebene unter "hierarchischer Verantwortung" erfolgen.

Empfehlung 30: Verschlechterung des Zugangs und Datenmanipulation kontrollieren

Die Politik verantwortet die *surveillance* und Kontrolle etwaiger negativer Auswirkungen einer P4P-Einführung wie Zugangsprobleme (sozial benachbarte Gruppen, Kontinuität, Risikoselektion) und Verzerrung durch Datenmanipulation. Frühzeitig im Implementierungsprozess müssen Evaluationen geplant und angestoßen werden.

Empfehlung 31: Rahmenbedingungen als politische Aufgabe

Explizite Rahmenbedingungen bilden den Hintergrund für die Initiierung des politischen Prozesses, für die Implementierung, die Umsetzung und die Evaluation. Die Rahmenbedingungen betreffen die Koordination der Zuständigkeiten im Implementierungsprozess genauso wie das Setzen von Meilensteinen für die Akteure, die an der Einführung von P4P beteiligt sind, die Antizipation von Fehlentwicklung und hemmenden Faktoren, und die Integration in die anderen Entwicklungen des Gesundheitssystems, mit denen P4P in Wechselwirkung tritt.

Résumé: Die "komplexe professionelle Systembürokratie", so der verwendete Arbeitsbegriff, ist in hohem Maße dazu in der Lage, Unsicherheit zu tolerieren ("intrinsische Unsicherheit"), folglich sind externe Versuche, das System auf explizite Qualitätskriterien festzulegen, nur schwer zum Erfolg zu führen. Wenngleich eine große Innovationsnähe vorhanden ist, ist doch die Innovationsbereitschaft als Reaktion auf Umweltreize beschränkt. In dieser Situation hat P4P als Feedback-Instrument nur dann eine Chance, wenn die Einführung gut mit den ökonomischen Rahmenbedingungen abgestimmt und eine langfristige politische Begleitung gegeben ist. Bereits die Auswahl der Qualitätsindikatoren erfordert größte Umsicht, denn P4P muss mit "unverbrauchten" Indikatoren an der Start gehen, die zudem auch noch häufig gewechselt werden müssen - sonst ist keine Wirkung zu erzielen. Auch ein *dual use* mit *public reporting* ist nicht sinnvoll, anders als oft angenommen wird, sind beide Instrumente nicht synergistisch. Ergebnisindikatoren sind wegen der Notwendigkeit zur Risikoadjustierung und der Problematik kleiner Gruppen in kleinen Einrichtungen schwer zu etablieren. Aber auch bei Prozessindikatoren ist Vorsicht geboten, denn es besteht bei schon lange in Gebrauch befindlichen und "allseits bekannten" Prozessindikatoren die Gefahr, dass sich daraus eine reine Einzelleistungsvergütung entwickelt, die dann besser primär einzuführen wäre. Prozessindikatoren sind allerdings sinnvoll einzusetzen bei der Behandlung chronischer Erkrankungen, in der Verbesserung der Koordination des Systems und im Patientensicherheitsbereich; hierzu sind jedoch vorangehende strategische Überlegungen notwendig, die die Schwachstellen des jeweiligen Gesundheitssystems identifizieren. Indikatoren auf der Basis von Patientenerfahrungen sind ebenso zu diskutieren wie die Frage, ob man nicht die Effizienz (Qualität zu Kosten, *value*) als Zielgröße für P4P einsetzt, so wie es in den USA der Fall ist. Vorsicht ist bei der Verwendung von administrativen Daten geboten, da sie ein Sensitivitätsproblem haben, dem wichtigsten Validitäts-Kriterium eines Indikators; es besteht hier die Gefahr der Mengenausweitung. Epidemiologische Falldefinitionen auf der Basis klinischer Daten, so wie schon seit geraumer Zeit bei der Erhebung nosokomialer Infektionen üblich, sind eindeutig vorzuziehen. Die Höhe der P4P-Vergütung muss die Opportunitätskosten berücksichtigen und ist bei Einzelleistungs- und DRG-Vergütung besonders hoch anzusetzen. Die monetäre Kopplung an die Qualitätskriterien sollte so gestaltet sein, dass auch die schlechteren Einrichtungen einen Anreiz verspüren, sich zu beteiligen (z.B. relative Positionsverbesserungen als Zielpunkt). Die Zahlungen müssen wegen der Tendenz zur Diskontierung und Risikoaversion zeitnah, verlässlich, nachvollziehbar und getrennt von anderen Zahlungsströmen erfolgen. Eine Einführung von P4P kann nicht erfolgreich sein, wenn man nicht die Risiko- und Verlustaversion der Einrichtungen in Rechnung stellt,

insbesondere da sonst die Diskontierung ein zu großes Gewicht erhält und die positiven Aspekte (Qualitätsverbesserung, professionelle Autonomie) in den Hintergrund rücken. P4P ist nie allein am Start, es ist immer in andere Vergütungssysteme integriert, und - auch wenn man es wünschen möchte - P4P kann grundsätzlich nicht die dominanten Anreize dieser Vergütungssysteme ausgleichen, in die es "eingebettet" ist. Dies gilt in Deutschland vor allem für das DRG-System im Krankenhausbereich. Allerdings ist es möglich, eine vorgeschaltete strategische Analyse vorausgesetzt, durch einen differenzierten Einsatz von P4P Bereiche zu fördern, die vom vorherrschenden Vergütungssystem, z.B. dem DRG-System, vernachlässigt werden. Hier ist insbesondere an die Versorgung chronisch, mehrfach erkrankter Patienten zu denken, weiterhin an die Koordination über die Sektorgrenzen hinweg und die Prävention. Die Einführung von P4P verlangt ein optimales *framing*, das auf die Risiko- und Verlustaversion der Akteure abgestimmt ist, der Diskontierung entgegentritt und die positiven Aspekte (Qualitätsverbesserung) in den Vordergrund stellt. Die politische Ebene hat im Sinne des *Governance*-Konzeptes der Peripherie viel Autonomie und Handlungsfreiheit zugestanden, spielt bei der Einführung einer neuen Vergütungslogik wie P4P jedoch eine entscheidende Rolle. In erster Linie handelt es sich um ein effektives *direction pointing*, einen aktiven und sichtbaren Einsatz der politischen Ebene, der Qualität als Richtschnur der gesundheitspolitischen Agenda betont, alle Ebenen des Gesundheitssystems einbezieht, öffentliche Foren schafft, die mit der Thematik beschäftigten Fachgesellschaften fördert, die Ausbildung entsprechend differenziert und Kontakte zu Einrichtungen der Zivilgesellschaft schafft. Ein "weiter so" wird nicht funktionieren, vor allem weil eine einfache Fortschreibung der Qualitätssicherung der vergangenen Jahre bei der Einführung von P4P nicht sinnvoll ist, weil P4P unter diesen Bedingungen keine Wirksamkeit entfaltet. Auch das *framing* kann nicht ohne gesellschaftliche bzw. politische Rahmenbedingungen geschehen, die die Notwendigkeit und die Perspektiven eines Wechsels der Vergütungslogik als allgemeinen Konsens erscheinen lassen. Die Bevölkerung sieht in der politischen Ebene die Gewährleistung dafür, dass unerwünschte Begleiterscheinungen eines solchen neuen Vergütungsansatzes rechtzeitig erkannt und korrigiert werden. Sollte P4P in Deutschland kommen, dann darf nicht wieder das geschehen, was in der Vergangenheit schon mehrfach der Fall war: dass mit der Evaluationsforschung gar nicht oder zu spät begonnen wurde. Evaluation gehört in das *conceptual framework* genauso wie alle anderen Fragen, die hier behandelt wurden. Doppelte Komplexität: ein komplexes Instrument in einem komplexen System.

Textfassung

1. Einleitung

1.1. *Pay for Performance* - Aktualität des Themas

Qualitäts-bezogene Vergütung bzw. *Pay for Performance* (P4P) ist in die Jahre gekommen, einerseits, die Anfangseuphorie ist einer differenzierteren und realistischeren Einschätzung gewichen. Andererseits setzen gerade jetzt große und unterschiedlich strukturierte Gesundheitssysteme wie das der USA, Großbritanniens und Deutschlands auf die Bindung der Vergütung an die erbrachte Qualität, um die als einseitig empfundene Mengenkopplung der Vergütung zu relativieren. So beabsichtigt der Koalitionsvertrag der Großen Koalition vom 27.11.2013, Qualität über den Mehrerlösausgleich bei der Vergütung zu berücksichtigen (s. Tableau 1). In den USA hat Medicare ein umfassendes *Value-based Purchasing* Programm aufgelegt, das auf den in den letzten 10 Jahren gesammelten Erfahrungen aufbaut (CMS 2011A, Belmont et al. 2011, Ryan et al. 2012, Chien und Rosenthal 2013). In Großbritannien setzt der *National Health Service* das 2004 in Kraft getretene *Quality and Outcome Framework* mit jährlich neu justierten Indikatoren fort (NHS 2013B).

Tableau 1: Koalitionsvertrag 27.11.2013, S. 78: "In einer Qualitätsoffensive werden wir die Qualität der stationären Versorgung verbessern. Qualität wird als weiteres Kriterium für Entscheidungen der Krankenhausplanung gesetzlich eingeführt (§ 1 KHG). (...) Gute Qualität muss sich für die Krankenhäuser auch finanziell lohnen. Die Menge soll künftig nur da berücksichtigt werden, wo sie entsteht. Das heute bestehende System der Mehrleistungsabschläge wollen wir dabei differenzieren: Leistungen mit nachgewiesener hoher Qualität können von Mehrleistungsabschlägen ausgenommen werden, für besonders gute Qualität sind Zuschläge möglich. Umgekehrt sollen bei unterdurchschnittlicher Qualität für einzelne Leistungen auch höhere Abschläge möglich sein. Die Qualität soll dabei risikoadjustiert und anhand wesentlicher Indikatoren gemessen werden."

Diese aktuelle Hinwendung zu P4P resultiert nicht aus den kurzfristigen (z.B. Campbell et al. 2007, Lindenauer et al. 2007) und langfristigen Ergebnissen der Evaluationsstudien (z.B. Campbell et al. 2009, Ryan et al. 2012B), denn diese haben die anfänglich hochfliegenden Erwartungen, zumindest auf den ersten Blick, nicht erfüllt.

Der Grund für die gestiegene Bedeutung des Themas P4P im Gesundheitswesen besteht vielmehr darin, dass sich der Blick auf P4P geändert hat: P4P ist kein *Magic Bullet*, das alle Probleme um die Qualität der Gesundheitsversorgung schlagartig behebt (Sorbero et al. 2006), sondern ein Vergütungsinstrument unter vielen, das gut abgestimmt mit anderen Vergütungselementen und eingebettet in eine langjährig entwickelte Qualitäts-orientierte Strategie zur Weiterentwicklung des Gesundheitswesens seine Wirksamkeit erzielen kann. Diese Sichtweise erfordert für die Implementierung von P4P jedoch mehr als ein paar technische Details und eine Verständigung über Indikator-Spezifikationen, so wie es in Deutschland gerne gesehen wird. Für eine wirksame Implementierung von P4P ist vielmehr ein konzeptioneller Rahmen (*conceptual framework*) notwendig, der zum Beispiel über die Wechselwirkung mit anderen Vergütungselementen oder über die Integration in eine generelle Qualitäts-Strategie Auskunft gibt. Ein solches Rahmenkonzept wird in diesem Artikel entwickelt, um daraus in einem zweiten Schritt die Bedingungen der Umsetzung und realistische Annahmen über die Auswirkungen abzuleiten (s. Kapitel 3, insbes. 3.4.).



Die Ausgangssituation für die Einführung von P4P im Gesundheitswesen bestand und besteht darin, dass sich die Qualitätsprobleme in der Gesundheitsversorgung der entwickelten Länder trotz mannigfaltiger Bemühungen zur Qualitätsverbesserung und hoher finanzieller Aufwendungen nicht beheben lassen. Auch die deutsche

Gesundheitsversorgung weist - entgegen aller Beteuerungen - erhebliche Qualitätsmängel auf, sowohl hinsichtlich der wichtigsten Kerndaten (z.B. OECD-Datensatz) als auch in punkto *Infection Control* und Patientensicherheit (Behnke et al. 2013, Gastmeier und Geffers 2008, Ott et al. 2013, Schrappe et al. 2008), hinsichtlich der Varianz der Ergebnisse im regionalen Vergleich (Aqua-Institut 2013) oder der prozessualen Parameter wie z.B. Koordination der Behandlung (Schoen et al. 2005, zuletzt Schoen et al. 2011). Es lag also nahe, auch wettbewerbliche Elemente auf ihre Fähigkeit hin zu untersuchen (s. Abb. 1, s. "Qualitätswettbewerb" Kap. 4.3.), ob sie diesem Defizit abhelfen können; der Sachverständigenrat (s. Tableau 2) hatte dies bereits im Jahr 1997 gefordert (SVR 1997, Kurzfassung, Nr. 137).

Tableau 2: Die Gutachten des Sachverständigenrates zur Begutachtung der Entwicklung im Gesundheitswesen (früher: ... für die Konzertierte Aktion im Gesundheitswesen; kurz **SVR Gesundheit**) nach §142 SGB V sind unter www.svr-gesundheit.de kostenlos als pdf downzuladen. Der SVR wird vom Bundesministerium für Gesundheit (BMG) berufen, das die Gutachten "unverzüglich" an den Bundestag und den Bundesrat weiterzuleiten hat.

Einerseits versuchte man, die Transparenz der Qualitätsinformationen zu verbessern (Qualitätsberichte, *Public Reporting/Public Disclosure*), um Patienten und Zuweisern eine informierte Wahl zu ermöglichen, u.U. unterstützt durch finanzielle Anreize für das reine Berichten der Qualitätsdaten (*Pay for Reporting*). Andererseits wurden direkte finanzielle (positive und negative) Anreize im Sinne einer Qualitäts-bezogenen Vergütung bzw. *Pay for Performance* (P4P) diskutiert, die dann in den USA bis zu einem Effizienz-basierten Vergütungssystem fortentwickelt wurden, bei dem die Qualitätsparameter in Beziehung zu den Kosten adressiert werden (*Value-Based Purchasing*, s. IOM 2001, S. 181ff). Die Nicht-Finanzierung von Patientensicherheits-relevanten "*Never Events*" (z.B. bestimmte nosokomiale Infektionen) sind als Sonderfall von P4P anzusehen (sog. *Non-Payment for Non-Performance*", s. z.B. Pronovost et al. 2008).

(weiter: 1. Einleitung, 1.2. Begriffsbestimmung und Definition)

1.2. Begriffsbestimmung und Definition

In diesem Zusammenhang soll kurz auf den Begriff P4P eingegangen werden. Die Definition des Sachverständigenrates von 2008 (SVR 2008, Nr. 732) als “Vergütungssystem, das die Qualität der Leistungserbringer in den Mittelpunkt stellt” (vgl. Kasten) kann als Ausgangspunkt dienen, man muss aber in mindestens zweierlei Richtungen über Weiterungen nachdenken und in einem dritte Sinne eine Abgrenzung vornehmen:

► Vor allem in den USA hat man von Beginn an Definitionen verwendet, die sich nicht nur auf die gesonderte Vergütung von Qualität beziehen, sondern die Effizienz (das Verhältnis von Qualität und den angefallenen Kosten, *value*) in den Mittelpunkt stellen (Dudley und Rosenthal 2006, Tompkins et al. 2009, aber auch schon Wenzel 1992), der auch in die Bezeichnung für das umfassende P4P-Programm von Medicare “*Hospital Inpatient Value-Based Purchasing*” Eingang gefunden hat (CMS 2011A). In Analogie wäre hier eigentlich der Begriff “*Pay for Value (P4V)*” angezeigt, ist aber nicht gebräuchlich.

Tableau 3: Was ist *Pay for Performance (P4P)*?

P4P oder Qualitäts-bezogene Vergütung “ist ein Vergütungssystem, das die Qualität der Leistungserbringer in den Mittelpunkt stellt” (SVR 2008, Nr. 732). In den USA wird zunehmend “Value” vergütet, also nicht nur die Qualität, sondern das Verhältnis von Qualität und den eingesetzten Kosten (strenggenommen *Pay for Value (P4V)*, so im umfassenden “*Hospital Inpatient Value-Based Purchasing*” (VBP)-Programm von Medicare (CMS 2011A)). Ökonomisch stammt der Begriff aus der *Principal-Agent-Theory* (vgl. Prendergast 1999). Die entscheidende Abgrenzung ist gegenüber Mengenorientierten Vergütungssystemen vorzunehmen (z.B. DRG). Die hier verwendete Definition basiert auf der expliziten Qualitätsmessung durch Indikatoren und der nachvollziehbaren Kopplung an monetäre Vergütungsbestandteile. (Definition s. Text)

► Die zweite Weiterung des Begriffes stammt aus der ökonomischen Theorie (*Principal-Agent Theory*) und betrachtet das Thema unter dem Management-Blickwinkel. Der Auftraggeber (*Principal*) gibt dem *Agent*, der einen Informationsvorsprung hat (z.B. einem Arzt, einem Krankenhaus), den Auftrag, in seinem Sinne eine Tätigkeit (eine Behandlung) auszuführen, die er selbst (der *Principal*) nicht ausführen kann, und die er auch nicht vollständig kontrollieren kann. P4P ist in diesem Verständnis ein Mittel des

Principal, sich der Dienste des *Agent* in seinem Sinne zu sichern, oder zumindest eine größere Chance dazu zu haben, dass der *Agent* in seinem Sinne handelt (vgl. Prendergast 1999, Wodchis et al. 2007, weitere Implikationen s. Kap. 4.2.).

► In Hinblick auf die Diskussion in Deutschland erscheint die Abgrenzung von Vergütungsbestandteilen des DRG-Systemes von Bedeutung. In einem Gutachten des BQS-Institutes (Veit et al. 2012) werden z.B. die Regelungen zur Fallzusammenführungen bei Wiederaufnahme nach Krankenhausbehandlung als Qualitäts-orientierte Vergütung bezeichnet. Sicherlich besteht hier ein Zusammenhang zu Qualitätsproblemen (z.B. Wiederaufnahme nach Komplikationen), andererseits besteht der Sinn dieser Regelung in der Verhinderung einer Mengenausweitung und nicht primär in einer Qualitätsverbesserung.

In der vorliegenden Arbeit wird daher eine **spezifische Definition** verwendet, die von der Qualität und der Sicherheit der Versorgung ausgeht und

- die explizite Messung von Qualität mittels definierter Indikatoren verbunden mit
- einer nachvollziehbaren Kopplung dieser Indikatoren an monetäre

Vergütungsbestandteile

in den Mittelpunkt stellt. P4P besteht also aus zwei Systemen, zum einen aus einem System der Qualitätsmessung, zum anderen aus einem System der finanziellen Bewertung. Vergütungssysteme, die auf der Effizienz der Versorgung basieren, also auf dem Verhältnis von Qualität und den aufgewendeten Kosten, sind ebenfalls als P4P zu verstehen, soweit sie Qualität (und die Kosten) explizit messen und eine nachvollziehbare Kopplung an die Vergütung aufweisen. P4P betrifft fast immer nur einen Teil der Gesamtvergütung und ist somit in andere Vergütungssysteme eingebettet. In Erweiterung der SVR-Definition wird in der vorliegenden Arbeit daher folgende Definition verwendet:

***Pay for Performance (P4P)* im Gesundheitswesen basiert auf einer Qualitätsmessung durch definierte Indikatoren und koppelt die Qualität der Versorgung an Vergütungsbestandteile mit dem Ziel, die Qualität der Gesundheitsversorgung zu verbessern. P4P kann sich auch auf Qualität im Verhältnis zu den Kosten (Effizienz, *value*) beziehen.**

Tableau 4: Kunst kommt von Können, wird immer gesagt. Etymologisch stimmt das (Kluge 2002), im Zusammenhang mit der Diskussion um Qualität und Sicherheit in der Gesundheitsversorgung steht der Satz jedoch für das Gegenüber von implizitem ("Kunst") und explizitem ("Können") Qualitätsverständnis: Können kann man messen, Kunst entzieht sich der Quantifizierung. Natürlich gibt es implizite, intuitiv erfahrbare Qualitätsaspekte. Es ist aber trotzdem gesellschaftlich konsentiert, dass die medizinische Behandlung regelgebunden und in der Wahl ihrer Mittel nicht von allen Regeln frei ist. Die Arbeit der Gesundheitsberufe ist also daher keine Kunst, sie entspricht nicht der Eingebung des Malers vor seiner leeren Leinwand. Da man allerdings weiß, wie schwierig diese Arbeit ist, wieviel Können sie erfordert, kann man in einem zweiten Sinne natürlich schon sagen, jemand ist fachlich so gut, dass er als ein Künstler in seinem Fach zu betrachten ist, dass sie ist eine richtige Künstlerin sei. Beide haben aber sicherlich nichts dagegen, ihren Status durch Zahlen zu hinterlegen.

Das Interesse in der Gesundheitspolitik an P4P fällt in eine Zeit, das darf nicht übersehen werden, in der grundsätzliche Vorbehalte gegen jegliche nachvollziehbare und valide Messung der Qualität im Gesundheitswesen (wieder) zunehmen, in der grundlegende Zweifel an den wissenschaftlichen Zahlen zur Häufigkeit von Problemen der Patientensicherheit wieder hoffähig werden (Anonymous 2014), alles einhergehend mit einer allgemeinen Kritik an der "Bürokratisierung" der Gesundheitsversorgung und der vorgeblichen Standardisierung bzw. "Ökonomisierung" durch die Evidenz-basierte Medizin (arte 17.9.2013). Qualität ließe sich gar nicht oder nur sehr schwer messen, die methodischen Probleme seien zu zahlreich und im Grunde nicht lösbar. Es wird auch wieder von der "Kunst der Qualitätsmessung" (und nicht deren exakten Methodik) gesprochen (Anonymous 2011). Wenn "Kunst" im Spiel ist, das kennen wir aus der Patientensicherheitsdiskussion um den Begriff des "Kunstfehlers", dann bezieht man sich auf ein implizites Qualitätsverständnis, das dem Patienten und der Öffentlichkeit nicht zugänglich ist (s. Tableau 4).

(weiter: 1. Einleitung, 1.3. Umsetzung und Methodik)

1.3. Technische Umsetzung und Methodik der Qualitätsmessung

Die notwendige Auseinandersetzung mit P4P, die Wertung der bisherigen Ergebnisse und die Integration des Instrumentes P4P in eine Strategie der Weiterentwicklung des Gesundheitssystems bezieht naturgemäß die technischen, methodischen, konzeptionellen und Gesundheitssystem-seitigen Faktoren ein, die die Anwendung und den Nutzen von P4P beeinflussen. Wie wichtig dieserlei Faktoren sind, führt ein einfaches Gedankenexperiment vor Augen: wenn es im deutschen Gesundheitswesen zusätzliche finanzielle Mittel in der Größenordnung von 100 Mrd. € pro Jahr gäbe (also weit mehr als die Hälfte der gesamten GKV-Ausgaben), die ausschließlich Qualitätsorientiert verwendet werden würden - wären wir wirklich sicher, dass sich eine deutliche Qualitätsverbesserung einstellt? Natürlich wären die Erwartungen hoch, aber wahrscheinlich wäre eine deutliche Verbesserung gar nicht zu beobachten. Warum ist dies so? Unwillkürlich wird man an den Satz von Janet Corrigan, der Chefin des *National Quality Forums* der USA aus dem Jahr 2009 erinnert: *“The current care system cannot do the job. Trying harder will not work. Changing systems of care will.”* (Corrigan und McNeill 2009).

Man muss sich also mit der Frage auseinandersetzen, inwieweit man von einem einzelnen Instrument wie P4P eine Änderungswirkung erwarten kann, die den Beharrungskräften des Systems gewachsen ist, die diese sogar übertrifft. Und wenn man hier (mit gutem Recht) skeptisch ist: welche Rahmenbedingungen beeinflussen am stärksten die Wirkung von P4P, bzw. unter welchen Rahmenbedingungen wäre eine Wirkung besonders gut zu beobachten? Es geht also nicht nur um die Faktoren, die für einen sinnvollen und erfolgreichen Einsatz des Instrumentes P4P Voraussetzung und Umfeld bilden, sondern auch um die Erwartungen, die man realistischer Weise an ein solches Instrument stellen kann, wenn es in einem derart komplexen System wie einem nationalen Gesundheitswesen eingesetzt wird. Die betreffenden Faktoren sind ausserordentlich zahlreich und in ihren Wechselwirkungen vielfältig, so sehr dass ihre Wirkung - selbst bei scheinbar optimaler Kenntnis der Einzelfaktoren - nicht direkt als Summe bzw. logisches Konstrukt dieser Vielzahl dargestellt werden kann (s.u. Komplexitätstheorie). Die Komplexität des Systems dupliziert sich noch zusätzlich auf der Ebene der Organisationen des Gesundheitswesens (s.u. Expertenorganisation), ganz zu schweigen von den Annahmen, die man zur Frage der Verhaltensänderung im Gesundheitswesen trifft (s. Kapitel 3.5.), den ökonomischen Grundannahmen (Kapitel 4), zur Einbettung in andere bzw. übergeordnete Vergütungssysteme (Kapitel 5) und zur Aufgabe der Gesundheitspolitik als Instanz der Gewährleistung und Sicherstellung (Kapitel 6).

In Kapitel 7 werden auf dieser Basis die bisherigen Ergebnisse der Evaluationsstudien (Kap. 2) und die zahlreichen technischen sowie methodischen Fragen diskutiert und daraus Empfehlungen abgeleitet. Um sich dem Thema jedoch zunächst anzunähern, werden hier die wichtigsten Einflussfaktoren in vier Gruppen eingeteilt und dargestellt, sowohl auf der Ebene der gegenwärtigen Diskussion (technische Umsetzung und Methodik der Qualitätsmessung), als auch mit einem weitergefassten Fokus (Konzeption der Umsetzung und Systemfaktoren):

- Technische Umsetzung
- Methodik der Qualitätsmessung
- Konzeption der Umsetzung
- Faktoren des Gesundheitssystems.

Die ersten beiden Gruppen werden auch in der aktuellen Diskussion häufig genannt, kurz soll hier auf folgende Sachverhalte eingegangen werden:

► **Technische Umsetzung:** Freiwilligkeit, Art der monetären Bewertung, Grenzwerte bei der Auslösung der Zahlung, wer ist der Adressat der qualitäts-orientierten Vergütung, der einzelne Arzt oder die Organisation (das Krankenhaus, das Praxisnetz) (zusammenfassende Darstellung s. Tableau 5)? Diese technischen Fragen sind jedoch nur auf den ersten Blick rein technischer Natur, sondern sie werden stark von unserem Grundverständnis des Zustandekommens von Gesundheits- bzw. Versorgungsleistungen beeinflusst. Um ein Beispiel zu nennen: Kommt der alte Spruch "Ich, Ihr Arzt, bin Ihre Prognose" der Wahrheit näher als die Annahme, dass die medizinische Behandlung das Ergebnis einer Team- bzw. Organisationsleistung ist? Muss man also den einzelnen Arzt oder das Team für Qualität vergüten? Hinter den vordergründig "nur" technischen Fragen verbergen sich also weitergehende organisations-theoretische Überlegungen, so dass sie nicht ohne die Klärung solcher grundlegenden Annahmen beantwortet werden können (s. auch "konzeptionelle Faktoren", im weiteren dann insbes. Kap. 3 und Kap 4).

Tableau 5: Technische und methodische Fragen bei P4P
(SVR 2008, Nr. 733):

- Teilnahme freiwillig oder verpflichtend?
- wie werden Qualitätsindikatoren monetär bewertet,
- Zuschläge, Abschläge oder beides? Wie sieht die Refinanzierung der Zuschläge aus?
- wie werden die Zu- bzw. Abschläge ausgelöst (ab einem gewissen Grenzwert, ab einer relativen Position im Vergleich zu den Wettbewerbern, oder durch die relative Verbesserung zur vorangegangenen Erhebung),
- wer ist Adressat der Zu-/Abschläge (einzelner Arzt, Team, Institution: Einzel- oder Gruppenmotivation),
- gibt es Ausnahmeregelungen für bestimmte Patienten (z.B. wegen fehlender Compliance),
- Struktur-, Prozess- oder Ergebnisindikatoren?
- klinische *Surveillance*- oder administrative Daten oder beides?
- Risikoadjustierung der Indikatoren,
- welche Ziele werden verfolgt: Akuterkrankungen oder Populations-bezogene Qualitätsziele oder beides,
- Kombination mit anderen Instrumenten wie z.B. *public reporting*?
- wer übernimmt die Kosten für die Datenerhebung,
- wie sieht die technische Umgebung aus, insbesondere wie sind die IT-Voraussetzungen gestaltet.

► **Methodik der Qualitätsmessung:** In erster Linie geht es hier um die Auswahl der **Indikatoren für die Qualitätsmessung**. Indikatoren (s. Tableau 6) sagen Qualität bzw. Qualitätsprobleme voraus, die Rückverlegungsrate auf Intensivstation steht z.B. für die Qualität der intensivmedizinischen Versorgung (s. Abb. 2). Im Kapitel "[7.3. Feedback - Voraussetzungen](#)" wird darauf eingegangen, an dieser Stelle sei lediglich darauf hingewiesen, dass immer noch zwei Fragen die Diskussion dominieren:

- Prozess- versus Ergebnis-Indikatoren und
- administrative (Routine- oder Sozial-) Daten versus klinische Daten.

Kurz zusammengefasst: **Prozessindikatoren** sind weniger relevant, während **Ergebnisindikatoren** (wie der Name schon sagt) mehr über das Outcome aussagen. Allerdings müssen Letztere risikoadjustiert werden, was zu kleinen Gruppen führt, kleine Krankenhäuser benachteiligt, die ein einziges negatives Ereignis schlecht ausgleichen können (Davidson et al. 2007, Scott und Ward 2006) und bei mangelhafter Risikoadjustierung die Anbieter zur Risikoselektion veranlasst (s. Nicholson et al. 2008). Die statistische Aussagekraft ist daher bei Prozessindikatoren, die nicht oder in geringerem Maße risikoadjustiert werden müssen (s. allerdings Mehta et al. 2008), deutlich besser; hinzu kommt, dass diese dem grundsätzlichen Präventionsgedanken

des Qualitätsmanagements sehr viel näher kommen, weil "das Kind noch nicht in den Brunnen gefallen ist" (sog. *bad apple*-Problematik, s. z.B. Cannon 2006). Es ist daher wenig verwunderlich, dass Prozessindikatoren nicht nur in der Vergangenheit z.B. im HQIP-Programm von Premier und CMS eingesetzt wurden (Lindenauer et al. 2007, s.u.), sondern auch im aktuellen *Value-Based Purchasing*-Programm (VBP) von Medicare (Tompkins et al. 2009, CMS 2012) und im ambulanten *Quality and Outcome Framework* (QOL) des NHS in Großbritannien (NHS 2013A) eine große Rolle spielen.

Tableau 6: Was ist ein Indikator? Diese zentrale Frage gehört zu den größten Problemen in der gegenwärtigen Qualitätsdiskussion, weil der ungenaue Gebrauch dieses Konzeptes die Diskussion erschwert. Indikatoren stellen nicht zwingend selbst Qualitätsprobleme dar, sondern sagen solche voraus (Ampelfunktion). Wie die JCAHO es schon vor mehr als 20 Jahren sagte: "Ein Indikator ist ein Maß, welches zum Monitoring und zur Bewertung der Qualität wichtiger Leitungs-, Management-, klinischer und unterstützender Funktionen genutzt werden kann, die sich auf das Behandlungsergebnis des Patienten auswirken. Ein Indikator ist kein direktes Maß der Qualität. Es ist mehr ein Werkzeug, das zur Leistungsbewertung benutzt werden kann, das Aufmerksamkeit auf potentielle Problembereiche lenken kann, die einer intensiven Überprüfung innerhalb einer Organisation bedürfen können." (JCAHO 1991, in der Übers. v. Sens et al. 2003). Wenn wir also die *readmission-rate* ansehen, dann geht es nicht darum, ob diese "zu hoch" oder "o.k." ist (kein direktes Qualitätsmaß), sondern darum, ob diese valide die Qualität der intensivmedizinischen Behandlung in Zusammenarbeit mit den Normalpflegestationen vorhersagen (Reliabilität der Erfassung vorausgesetzt). Natürlich gibt es (gute) Indikatoren, die selbst Qualitäts-relevant sind (aber das ist die Ausnahme!), z.B. nosokomiale intravenöse Katheterinfektionen, die selbst ein Qualitätsproblem darstellen, gleichzeitig und in erster Linie aber die Qualität der Versorgung einer ganzen Station abbilden.

Über die Nutzung **administrativer Daten** versus **klinischer Daten** wird ähnlich erbittert gestritten. Erstere haben den Vorteil einer leichteren Erhebbarkeit, sie haben aber gegenüber klinischen und Beobachtungsdaten (Kontrolle der Reliabilität der Erhebung durch Stichproben vorausgesetzt) klare Defizite in der Sensitivität, der alles entscheidenden Eigenschaft von Indikatoren (s. z.B. Azaouagh und Stausberg 2008, Calderwood et al. 2014, Miller et al. 2001, Pawlson et al. 2007, Powell et al. 2001, s. auch SVR 2008 Nr.654ff; zur Bedeutung der Sensitivität s.u., zusammenfassende Darstellung SVR 2008, Nr. 645ff). Der Grund liegt darin, dass sie primär für Vergütungszwecke und nicht für Qualitätsmessungen entwickelt und erhoben werden. Allerdings muss bei **klinischen Daten** sehr differenziert mit der Messmethodik umgegangen werden, da z.B. durch direkte Beobachtung erhobene klinische Daten

deutlich höhere Ergebnisse erbringen als Daten aus Krankenakten (sog. *chart review*) (s.u.). Es gibt jedoch langjährige Erfahrungen mit epidemiologischen Falldefinitionen auf dem Gebiet der Erhebung nosokomialer Infektionen, die sich klinische Daten stützen (*National Nosocomial Infection Surveillance*, NNIS, der *Centers of Disease Control*, CDC). Anders als oft angenommen, handelt es sich dabei nicht um klinische Diagnosen, sondern zu rein epidemiologischen Zwecken entwickelte Instrumente (vgl. Talbot et al. 2013). Fasst man nun alle Erfahrung und alle Studien zusammen: erstens sollte man Indikatoren im Set verwenden, also mehrere Indikatoren miteinander kombinieren, zweitens ist es sicher richtig, Prozessindikatoren mit einigen Ergebnisindikatoren zu kombinieren bzw. Prozessindikatoren zu verwenden, die besonders Ergebnis-relevant sind (z.B. Rückverlegungsrate auf Intensivstation) und drittens: natürlich administrative Daten verwenden, die bereits vorliegen, wenn es möglich ist, denn der Messaufwand ist eine kritische Größe: "*Indicators of performance do not measure quality, people do!*" (Kazandijan et al. 1995).

weiter : 1. Einleitung, 1.4. Zum Begriff des Indikators

1.4. Zum Begriff des Indikators

Um Parameter, die als Indikatoren in Frage kommen, weiter zu charakterisieren, stehen Zuverlässigkeit (**Reliabilität**: wird der Parameter richtig bestimmt?) und **Validität** (werden die Qualitätsprobleme identifiziert, die vorliegen, wird also das gemessen, was gemessen werden soll?) im Vordergrund (s. Abb. 2). Auch in unserer aktuellen deutschen Diskussion werden Reliabilität und Validität oft nicht korrekt getrennt. Wenn etwa bemängelt wird, Indikatoren seien deswegen nicht geeignet (nicht valide, erfüllen nicht ihre Vorhersagefunktion), weil sie wegen Dokumentationsmängeln (nicht reliabel) falsche Ergebnisse erbringen, dann ist hier zunächst eine Aussage zur Reliabilität und nicht zur Validität gemacht (z.B. Petzold et al. 2013). Natürlich kann ein nicht reliabler Indikator kaum valide sein; können jedoch die Probleme in der Reliabilität behoben werden, dann kann durchaus eine hervorragende Validität vorliegen. Und natürlich haben unsere Erhebungsmethoden z.B. in der Qualitätssicherung nach §137 SGB V massive Reliabilitätsprobleme, basieren sie doch weitgehend auf freiwilligen Selbstauskünften, die nur schlecht durch unabhängige und unangemeldete Stichproben "geeicht" sind. Besonders gut ist der Einfluss der **Messmethodik** auf das Ergebnis der Erhebungen bei den unerwünschten Ereignissen auf dem Gebiet der Patientensicherheit dokumentiert. Durch freiwillige Meldung werden nur 1% der relevanten, mit negativen Konsequenzen für den Patienten verbundenen *adverse events* entdeckt (Classen et al. 2011).

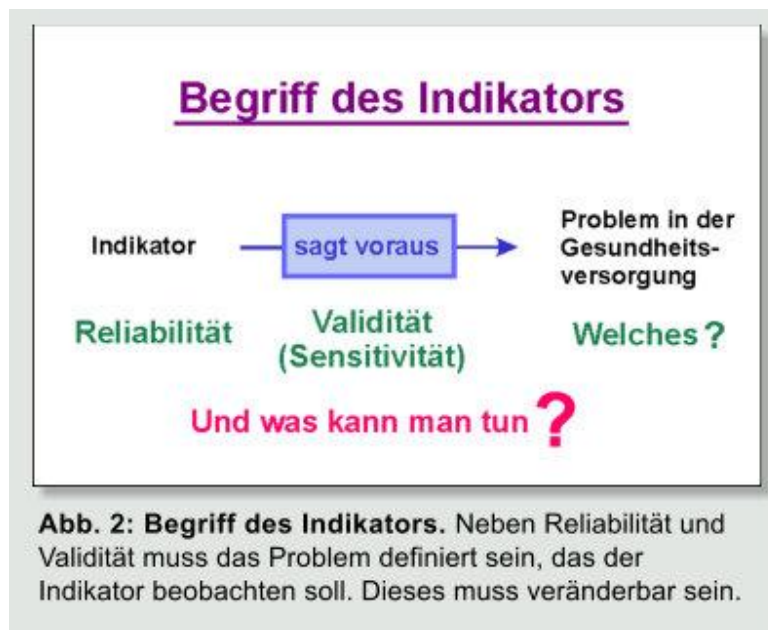


Abb. 2: Begriff des Indikators. Neben Reliabilität und Validität muss das Problem definiert sein, das der Indikator beobachten soll. Dieses muss veränderbar sein.

Bei der **Validität** (zusammenfassende Darstellung der Anforderungen s. Schrappe

2010) wird der Sensitivität der Vorzug gegenüber der Spezifität gegeben (man will alle Qualitätsprobleme erkennen und hat nichts dagegen einzuwenden, "umsonst gerufen worden zu sein", vgl. Tableau 7 und Abb. 3). Aber auch valide Indikatoren können im konkreten Messzusammenhang nutzlos sein (Abb. 2), so befinden wir uns bei §137-Indikatoren schon nahe am Optimum, eine weitere Verbesserung der Indikatoren (nicht der Versorgung) ist kaum noch möglich (sog. **ceiling-Effekt**) (AQUA 2013). Dies kann auch die Konsequenz haben, dass wir einen positiven Effekt von P4P deswegen nicht erkennen, weil wir auf dem Hintergrund der intensiven Beschäftigung mit dem Thema gleichzeitig zu einer Verbesserung unserer Messmethodik (scheinbare Zunahme von Qualitätsdefiziten) kommen und hiermit den Erfolg (tatsächliche Abnahme von Qualitätsdefiziten) verdecken. Auch muss, wie von Szescenyi und Mitarbeitern immer wieder hervorgehoben wird, eine hohe **Relevanz und Umsetzbarkeit** von Indikatoren gegeben sein (Willms et al. 2013). Es hat wenig Sinn, einen Indikator zu implementieren, der auf Qualitätsdefizite verweist, die nicht abänderbar sind (z.B. außerhalb des institutionellen Kompetenzrahmens liegen), sondern es sollten "Qualitätspotentiale" existieren.

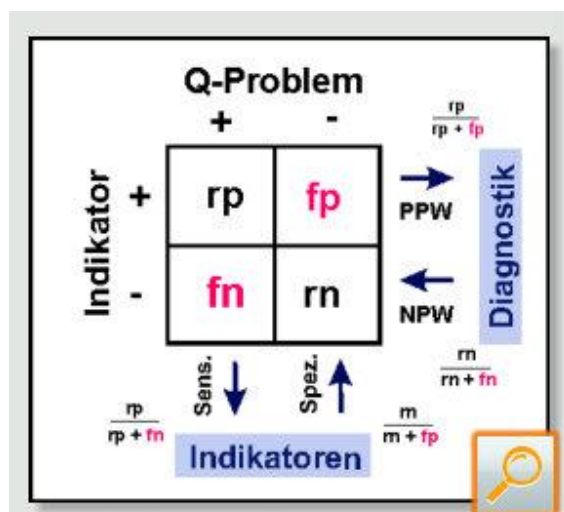


Abb. 3: Gebrauch von Indikatoren in Abgrenzung zu diagnostischen Tests. Bei Indikatoren sind vor allem die falsch-negativen Befunde wichtig (Sensitivität), während bei der Diagnostik die Rate von falsch-positiven Ergebnissen niedrig sein muss (PPW, pos. prädiktiver Wert).

weiter: 1. Einleitung, 1.5. Konzeption und Gesundheitssystem

Tableau 7: Der Unterschied zwischen Indikatoren und diagnostischen Parametern in der **ärztlichen (Differential)Diagnostik** ist besonders wichtig (s. Abb. 3). Eine gute diagnostische Methode muss in erster Linie eine therapeutische Entscheidung tragen können, daher darf sie keine Erkrankung vorspiegeln, obwohl sie gar nicht vorliegt. Sie sollte also wenig falsch-positive Ergebnisse erbringen (hohe "Spezifität" wird dann gesagt, es handelt sich aber im Grunde um einen hohen Positiven Prädiktiven Wert (PPW), da man allein von der Kenntnis des Testergebnisses ausgeht). Anderenfalls würde eine falsche Therapie durchgeführt bzw. durch Unterlassen weiterer Untersuchungen die wirklich vorliegende Erkrankung nicht erkannt. Aus diesem Grund wird von ärztlicher Seite immer wieder die mangelnde "Spezifität" von Indikatoren ins Feld geführt ("war doch gar nicht so") (z.B. Albrecht et al. 2013). Dabei wird aber der grundlegende Unterschied zu diagnostischen Test verkannt; anders in der Diagnostik dürfen Indikatoren durchaus in einem gewissen Umfang falsch-positive Ergebnisse erbringen, wenn sie "dafür" aber alle Qualitätsdefizite erkennen, die tatsächlich vorliegen (keine falsch-negativen Ergebnisse).

Beispiel: In einer Auswertung der lokalen §137-Qualitätssicherung kommt eine QM-Arbeitsgruppe zu dem Ergebnis, dass dem Ansprechen der Indikatoren in den meisten Fällen gar kein Qualitätsdefizit zugrundeliegt (viele falsch-positive Ergebnisse, niedrige "Spezifität") (Petzold et al. 2013). Interessant wäre es aber gewesen nachzuschauen, wie hoch der Anteil von (a priori) bekannten Qualitätsdefiziten ist, der von den Indikatoren entdeckt wird (Sensitivität). Bei niedriger Sensitivität wären die Indikatoren wirklich nutzlos, denn sie machen nicht valide auf Probleme aufmerksam, die vorliegen. Wenn sie dagegen (mit vertretbarem Aufwand) alle Qualitätsprobleme identifizieren, ist die Validität hoch, auch wenn es zusätzlich einige falsch-positive Ergebnisse geben sollte. Die Übertragung der EBM-Anforderungen für diagnostische Methoden auf Indikatoren ist aus dieser Sicht diskussionswürdig (Schmitt et al. 2013).

1.5. Konzeption und Gesundheitssystem

Die anhaltende Diskussion um Qualitätsindikatoren und Instrumente wie *public reporting* sowie P4P in Deutschland ist noch immer sehr von den oben geschilderten technischen und methodischen Fragen bestimmt, was einerseits mit den spezifischen Interessenlagen der Beteiligten zusammenhängt, andererseits aber auch die Wichtigkeit dieser Fragen hervorhebt. Wenn die Qualität der Versorgung wettbewerbsrelevant werden soll, muss zu diesen Problemen Einvernehmen hergestellt werden, zumindestens insoweit, als dass normative Regelungen Fuß fassen können. Jedoch kann man nicht davon absehen, dass zusätzlich die konzeptionellen und Systemfaktoren von noch weitreichenderer Bedeutung sind, denn sie spiegeln unser grundsätzliches Verständnis des Gesundheitssystems und des Verhaltens seiner Akteure wieder. Sie zeigen nicht nur die Bedeutung der Rahmenbedingungen für Instrumente wie P4P auf, deren genügende Beachtung dann zu den gewünschten Ergebnissen kommt. Vielmehr bilden sie, wie oben bereits angeführt, den Erwartungshorizont für die Wirkung von Interventionen wie der Einführung von P4P. Umsetzungs- und Gesundheitssystem-seitige Faktoren, hier nur in Beispielen genannt, bilden daher den Schwerpunkt der vorliegenden Darstellung und werden im Folgenden in einer ersten Näherung dargestellt:

Gliederung: erfolgreiche Umsetzung von P4P

- Technische Umsetzung (Kap. 1.3.)
- Methodik der Qualitätsmessung (1.3.-1.4.)
- Konzeption der Umsetzung
- Faktoren des Gesundheitssystems

► Die **Konzeption der Umsetzung** von P4P umfasst in erster Linie die Annahmen zur internen und externen Motivation, die beim Einsatz finanzieller Anreize im Gesundheitswesen zugrundegelegt werden, die organisationstheoretischen Konzepte zur Führung von Angehörigen der Gesundheitsberufe (Stichwort Expertenorganisation), überhaupt Managementkonzepte für Organisationen im Gesundheitswesen und -viertens - das Umgehen mit der hohen Komplexität, die für das Gesundheitswesen typisch ist. Zu berücksichtigen ist dabei, dass sich das Gesundheitswesen gegenüber äußeren Einflüssen bislang sehr erfolgreich abschotten konnte: Veränderungen werden selten als Chance, meist als Bedrohung gesehen, Stabilität und Abschottung sind die Regel. Andererseits sind die (gesundheitspolitischen) Interventionen oft singulärer Natur, sie drehen nur an einer Schraube, und sind insofern in ihrer Wirksamkeit beschränkt (s. politische Verantwortung, Kap. 6.3).

► *Last not least* sind natürlich **Faktoren des Gesundheitssystems** zu berücksichtigen: kann man, um ein Beispiel zu nennen, den ja nicht unerheblichen Mengenanreiz eines DRG-Systemes wirklich durch qualitätsbezogene Vergütungsanreize egalisieren oder wenigstens positiv beeinflussen, so wie es im Koalitionsvertrag (S. 78) von CDU/CSU und SPD vom 27.11.2013 durch die Kopplung an den Mehrerlösausgleich insinuiert wird (s. Tableau 1). Sind durch P4P die Koordinationsdefizite der konkurrierenden Sektoren, die ja für zahlreiche Qualitätsdefizite verantwortlich zeichnen, oder der Mengenanreiz pauschalierender Vergütungssysteme zu überwinden? Hier ist Skepsis angezeigt. Andererseits wird auf diese Weise auch deutlich, wo die Zukunft von *Pay for Performance* liegt: im gezielten Einsatz in sinnvoller Kombination mit anderen Vergütungsinstrumenten (s. Kapitel 5).

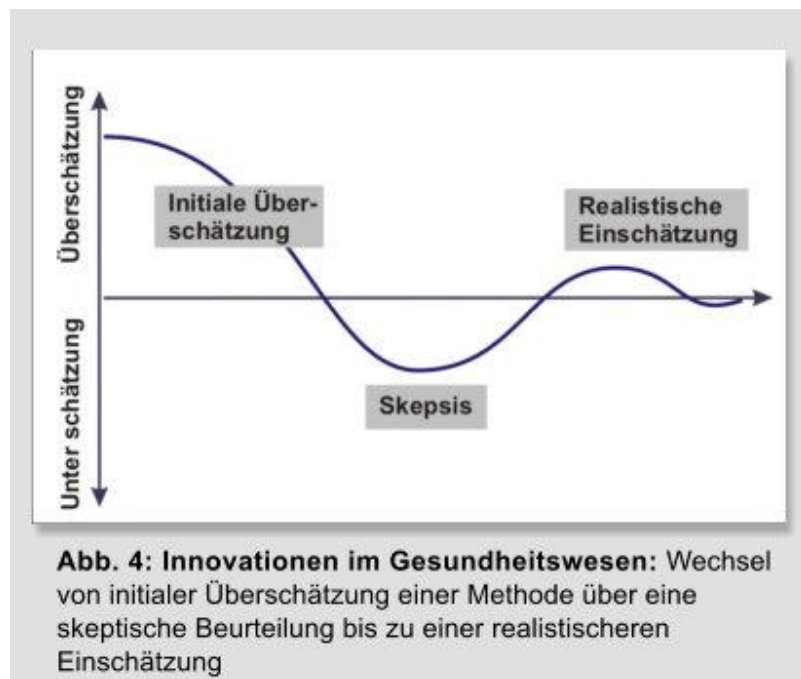
In der Folge wird für das Verständnis dieser Faktoren im Hinblick auf gesundheitspolitische bzw. Vergütungs-bezogene Interventionen ein Rahmenkonzept entwickelt, aus dem realistische Anforderungen, Erwartungen und Kontextfaktoren abgeleitet werden können. Zunächst wird im nächsten Kapitel jedoch auf die bislang vorliegenden Evaluationsergebnisse der verschiedenen internationalen P4P-Projekte eingegangen.

Weiter: 2. Langfristige Effekte und Weiterentwicklung, 2.1. Anfänglicher Optimismus

2. Langfristige Effekte und Weiterentwicklung

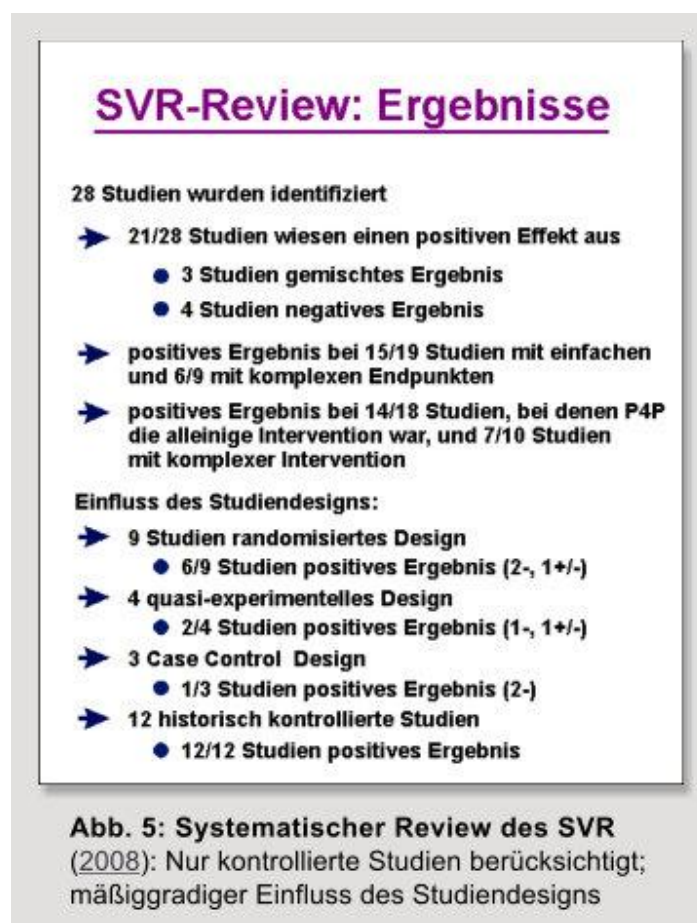
2.1. Anfänglicher Optimismus

Wir kennen das Phänomen von anderen Methoden im Gesundheitswesen: anfangs ist alles Gold. Es folgt ein Auf-und-Ab von Ernüchterung, dann wieder positiver Einschätzung, nochmals Skepsis usw., bis man dann nach einer gewissen Zeit zu einer adäquaten Beurteilung der Wirksamkeit der Methode kommt, die der jeweiligen Anwendung und ihrem Kontext angemessen ist (s. Abb. 4). So verhielt es sich auch bei P4P (Rosenthal et al. 2007); die früheren Studien und Reviews kamen zu besseren Ergebnissen als die späteren Evaluationen, und aus letzteren ergaben sich wiederum Hinweise, die in den USA und in Großbritannien zur Entwicklung weiterführender Programme wie dem *Value-Based Purchasing*-Programm von Medicare (CMS 2011B) und dem *Quality and Outcome Framework* des NHS führten (NHS 2013A)). Auch Einzelstudien neueren Datums kamen wieder zu ermutigenderen Ergebnissen (z.B. Bardach et al. 2013, Calikoglu et al. 2013, Casale et al. 2007), Petersen et al. 2013, allerdings sind in alle diese neueren Projekte die Erfahrungen aus vorangegangenen "Gehversuchen" eingeflossen.



In seinem frühen Systematischen Review (Studien bis 2006) zur kurzfristigen Wirksamkeit von P4P auf definierte Versorgungsziele konnte der SVR Gesundheit 28 kontrollierte Studien identifizieren, wovon 21 Studien einen positiven Effekt zeigten (SVR 2008, Nr. 738). Jedoch war bereits in diesem frühen Review auffällig, dass zwar 100%

aller methodisch schwächeren Studien einen durchschlagenden Erfolg zeigten, dies jedoch nur bei 6/9 Studien mit methodisch stringenterem, randomisiertem Design der Fall war. Diese Verschiebung der Effektmaße hin zu den methodisch schwächeren Studien könnte als ein Zeichen für die mangelnde interne Validität der dem Review zugrundeliegenden Studien gewertet werden: methodisch schwächere Studien überschätzen ja gerne das Ergebnis (s. Abb. 5). Der SVR empfahl daher nur vorsichtig “die schrittweise Einführung von Elementen dieser Vergütungsform mit Pilotierung und intensiver Evaluation” (SVR 2008, Kurzfassung, Nr. 132).



Die weitere Entwicklung gab dieser vorsichtigen Einschätzung zunächst recht. Im Systematischen Review von Van Herck et al. (2010) wurde auf der Basis von 128 Studien positive Ergebnisse (Wirksamkeit von P4P) überwiegend bei klinischen, eindimensionalen Endpunkten (z.B. Impfrate) gefunden, bei komplexen Endpunkten (Koordination, Kontinuität der Behandlung, Patientenzentriertheit, Kosteneffektivität) waren die Ergebnisse hingegen widersprüchlich. Auch diese Konstellation muss vorsichtig stimmen, denn eindimensionale Endpunkte sind gerade bei Interventionen wie P4P in der Praxis weit weniger adäquat als komplexe Endpunkte. Zwei kleinere,

methodisch aber sehr strenge Systematische Reviews zeigten geringe positive Auswirkungen (Mehrotra et al. 2009, 8 Studien, ausschließlich HQIP-Krankenhäuser) an einem Teil der untersuchten Endpunkte (Scott et al. 2011, Cochrane-Review, 7 Studien). (Die älteren Systematischen Reviews sind an dieser Stelle nicht nochmals aufgeführt, vgl. SVR 2008; die Systematischen Reviews von Gillem et al. (2012) und Langdown et al. (2013) waren nur auf das britische QOF-Programm ausgerichtet). Ein Meta-Review, der 22 Systematische Reviews umfasste, kam zu keinen definitiven Ergebnissen (Eijkenaar et al. 2013).

Tableau 8: Das BGS Institut für Qualität und Patientensicherheit GmbH ist im Jahr 2008 aus der Bundesgeschäftsstelle Qualitätssicherung (BQS) hervorgegangen. Gesellschafter sind die Bundesärztekammer, die Deutsche Krankenhausgesellschaft e.V., Bundesverbände der Krankenkassen und der Verband der privaten Krankenversicherung e.V. (s. <http://www.bqs-institut.de/bqs-institut/gesellschafter.html>, 20.11.2013)

Im Jahr 2012 wurde ein Review des BQS Institut für Qualität und Patientensicherheit GmbH (s. Tableau 8) veröffentlicht, das vom Bundesministerium für Gesundheit beauftragt war (Veit et al. 2012). Dieser Review basierte auf einer weiter gefassten Operationalisierung des Begriffes P4P, indem er sich auf das "Leistungsniveau" und nicht auf die Qualität als Zielparameter bezog, und nutzte außerdem Einschlusskriterien für die untersuchten Studien, die auch Fallserien und Fallberichte einschlossen. Die Wertung der Studienlage und die Empfehlungen bleiben zurückhaltend, die langfristigen Effekte werden jedoch zutreffend und kritisch dargestellt.

Weiter: 2. Langfristige Effekte, 2.2. HQIP in den USA

2.2. HQIP in den USA

Vor allem aber waren es die langfristigen Ergebnisse der beiden großen Studien, die in den USA (im Krankenhausbereich) und in Großbritannien (im ambulanten Bereich) durchgeführt wurden, die die Grenzen eines isolierten und voraussetzungslosen P4P-Einsatzes vor Augen führten, gleichzeitig aber wichtige Hinweise für die weitere Entwicklung und einen integrierten Einsatz gaben. Auf diesen Erfahrungen basiert vor allem das *Value-based Purchasing* Programm (VBP) von Medicare, an dem alle Akutkrankenhäuser der USA teilnehmen (s. gesundheitspolitischer Hintergrund in den USA).

Das größte P4P-Projekt im Krankenhausbereich war das *Premier Hospital Quality Incentive Demonstration Project* (HQIP) von Premier und der *Centers of Medicare and Medicaid Services* (CMS) (Grossbart 2006, Lindenauer et al. 2007; s. Tableau 9). Im Zweijahreshorizont konnte hier bei 266 Krankenhäusern, die P4P kombiniert mit *Public Reporting* einsetzten, an Prozess-Indikatoren zu fünf Krankheitsgruppen und an einem *Composite*-Indikator eine recht deutliche Verbesserung gegenüber einer Kontrollgruppe, die lediglich *Public Reporting* allein einsetzte, beobachtet werden (Lindenauer et al. 2007). Allerdings gab es bereits zu diesem Zeitpunkt bezüglich der Indikatoren zum Herzinfarkt eine große unabhängige Studie, die eine positive Wirkung von P4P in Abrede stellte (Glickman et al. 2007). In der Folge wurde die initial positive Wertung von HQIP durch drei große Untersuchungen mit langfristiger Perspektive relativiert:

► In einer ersten Untersuchung (Werner et al. 2011), die einen um zwei Jahre längeren Zeitraum als die Studie von Lindenauer et al. umfasste, war beim *Composite*-Indikator kein Vorteil der P4P-Häuser mehr festzustellen, allerdings konnten die P4P-Häuser die Verbesserung zwei Jahre schneller als die Kontrollen umsetzen (nach drei Jahren hatten 56% der P4P-Häuser einen Score von 90% erreicht gegenüber 32% der Kontrollen). Gerade bei P4P-Krankenhäusern, die hohe Zahlungen erhalten hatten (da sie viele Medicare-Patienten behandelt haben), war der Unterschied zu Kontrollhäusern, die bei Teilnahme an HQIP gleich hohe Zahlungen erhalten hätten, besonders groß, ein indirekter Hinweis auf die Bedeutung der Höhe des Anreizes. Außerdem waren die Qualitätsverbesserungen in Märkten mit geringer ausgeprägter Konkurrenz (regional wenige Häuser vorhanden, oder: Monopolstellung durch Konzentration) deutlich höher als in Märkten mit ausgeprägter Konkurrenz, da hier Qualitätsverbesserungen allein durch *Public Reporting* vorher nicht stattgefunden hatten. Auch waren die Verbesserungen bei Krankenhäusern mit einer gesunden finanziellen Grundausstattung stärker als bei finanziell schwach aufgestellten Häusern (Werner et al. 2011).

Tableau 9: Das *Premier Hospital Quality Incentive Demonstration Project* (HQIP) wurde seit 2003 von den *Centers of Medicare and Medicaid* (CMS) und der *Premier Healthcare Alliance*, einer Vereinigung von 2900 Krankenhäusern zum Zweck der Qualitätsverbesserung (2013), betrieben (Lindenauer et al. 2007). Die Teilnahme war freiwillig. Alle teilnehmenden 266 Krankenhäuser (63% der in Frage kommenden) hatten bereits an einem *Benchmarking* Projekt von Premier teilgenommen (*Public Reporting*). Bis zu 2% des Budgets wurde nach den Ergebnissen von 34 Prozessindikatoren als Zu- (bestes Dezil) oder Abschlag (schlechtestes Dezil) umverteilt (relativer Anreiz). Ab dem vierten Jahr (2007) wurde auch das Erreichen eines Performance-Medians finanziell vergütet (absoluter Anreiz) (Werner et al. 2011). Die Indikatoren gehörten zu klassischen Sets der JCAHO, der AHRQ und des NQF, sie betrafen ambulant erworbene Pneumonie, Herzinfarkt, chronische Herzinsuffizienz, Bypass-OP und Knie/Hüft-TEP. Insgesamt wurden in den ersten 5 Jahren 48 Mill. \$ umverteilt. Das HQIP diente zur Vorbereitung des *Medicare Value-Based Purchasing Program*, das seit 2012 läuft.

► Die Erkrankungs-bezogene Langfristauswertung des HQIP durch Ryan et al. (2012B) zeigte dann über einen 6 Jahres-Zeitraum (2003-2009) keine Verbesserung der P4P-Häuser (P4P plus *Public Reporting*) gegenüber gematchten Kontrollhäusern (*Public Reporting* allein); alle beobachteten Krankenhäuser wurden besser in ihren Qualitätsdaten, ohne dass sich die beiden Gruppen jedoch unterschieden. Allerdings war bei den Interventionshäusern eine deutliche Abnahme der Varianz der Ergebnisse zu beobachten. Anders als vermutet, verbesserten sich die anfangs schlechteren Häuser nicht stärker als die Top-Häuser. Dies war auch zu beobachten, nachdem man 2006 begonnen hatte, neben der relativen Position der Häuser (an der Spitze Zuschläge, am "schlechten Ende" Abschläge) zusätzlich auch das Überschreiten eines definierten Schwellenwertes (Median der Performance in den Vorjahren aller Häuser) zu belohnen, damit auch die im Mittelfeld stehenden Krankenhäuser einen Anreiz haben, sich zu verbessern (zur monetären Kopplung s. Kap. 4.1., Abb. 11).

► In der dritten Studie konnte ebenfalls keine Verbesserung festgestellt werden, und zwar im Vergleich zwischen den 252 HQIP-Krankenhäusern und 3363 Kontrollhäusern (*public reporting* allein) anhand der 30-Tage Mortalität, also eines wichtigen Outcome-Parameters. Auch hier zeigte sich keine Veränderung bei den “*poor performers*” (Jha et al. 2012).

(Fortsetzung: 2.3. HQIP: Gesundheitspolitischer Kontext in den USA)

2.3. HQIP: Gesundheitspolitischer Kontext in den USA

Jenseits der Datenlage lässt sich die eigentliche Bedeutung des *Premier Hospital Quality Incentive Demonstration Project* (HQIP) jedoch erst in seinem langfristigen gesundheitspolitischen Zusammenhang abschätzen. Die Wurzeln gehen bis in die 90er Jahre des letzten Jahrhunderts zurück und weisen gleichzeitig weit in die Zukunft (zur Gesetzgebung s. Tableau 10). HQIP war letztlich das Pilotprojekt und die Vorbereitung für das *Value-based Purchasing* (VBP)-Programm, eine der zentralen Säulen der umfassenden Neuorientierung, in dem sich das Gesundheitswesen der USA derzeit befindet, und wozu folgende Elemente gehören (abgesehen von der Ausweitung der Versicherungspflicht) (Belmont et al. 2011):

- *“Never Events”*: Fortsetzung der *Non-Payment for Non-Performance* Politik,
- *Value-based Purchasing* (VBP)-Programm für sämtliche Bereiche des US-Gesundheitswesens,
- *Bundled Payment*: Form der “integrierten Versorgung” für 10 Erkrankungen bzw. Krankheitsgruppen,
- *Accountable Care Organizations*: Zusammenschlüsse (“Netze”) von Anbietern, die mindestens 5000 Medicare-Versicherte versorgen,
- *Health Homes* für Medicaid-Versicherte,
- Verstärkung der *Public Reporting* Programme.

Bemerkenswert ist dabei aber nicht allein die umfassende Breite des Ansatzes, sondern die Tatsache, dass man hier nicht nur ein eindeutiges Bekenntnis zu Qualität und Sicherheit als Merkmale der Gesundheitsversorgung abgibt, und dass man dabei von einem sehr ausdifferenzierten Grundverständnis des Gesundheitssystems der USA ausgeht. Von diesem Grundverständnis, das im IOM-Report *“Crossing the Quality Chasm - A New Health System for the 21st Century”* ausführlich dargelegt ist (IOM 2001), werden einerseits die notwendigen Veränderungen zur Verbesserung der Gesundheitsversorgung abgeleitet, andererseits aber auch ein Weg zur Implementierung dieser Veränderungen einschließlich einer Abschätzung der Größe der zu erwartenden Effekte entwickelt. Man bezieht sich dabei explizit auf die Erkenntnisse der **Systemtheorie** und beschreibt das Gesundheitssystem als sog. komplexes System. Komplexe Systeme weisen, anders als komplizierte Systeme, keine erkennbaren Regeln auf und sind aus unabhängigen Akteuren, Institutionen und Systemfaktoren zusammengesetzt, die gleichzeitig derartig in Wechselwirkung stehen, dass nicht vorhersehbare Entwicklungen an der Tagesordnung sind und das System äußerst anpassungsfähig, innovativ und kreativ agiert (Plsek 2003). Zusammengehalten wird dieses System von verborgenen Regeln und “Schatten-Systemen” sowie Zuständen mit

bedingter Stabilität, die Orientierung geben (sog. Attraktoren).

Tableau 10: Wichtigste Gesetzgebung mit gesundheitspolitischen Schwerpunkt in den USA aus den letzten Jahren:

Balanced Budget Act (1997)

Medicare Prescription Drug, Improvement and Modernization Act (MMA) (2003)

Deficit Reduction Act (2005)

Tax Relief and Health Care Act (2006)

Patient Protection and Affordable Care Act (2010)

Weiter unten wird genauer auf die Frage eingegangen, ob dieses Sichtweise für die Interpretation unserer Kenntnisse über ein Instrument wie P4P sowie für dessen Weiterentwicklung nützlich ist. Hier sei zunächst nur hervorgehoben, dass (a) hier der seltene Fall vorzuliegen scheint, dass Gesundheitspolitik auf der Basis einer dezidiert dargelegten Grundkonzeption betrieben wird, und dass (b) es sich bei dem genannten Grundverständnis um eine Konzeption handelt, die äußerst langfristig und in sehr stabiler Form entwickelt wurde (IOM 2001, IOM 2007, CMS 2011B, CMS 2012, Chien und Rosenthal 2013).

Für die inhaltliche Federführung sind hier insbesondere das *National Quality Forum* (NQF), die *Joint Commission on Accreditation of Healthcare Organizations* (JCAHO) und die *Centers of Medicare and Medicaid Services* (CMS), die Nachfolgeorganisation der HCFA, und die *Agency for Healthcare Research and Quality* (AHRQ) zu nennen:

► Ausgangspunkt war die Schaffung des sog. **HEDIS-Datensatzes** (*Health Effectiveness Data and Information Set*) anlässlich der Öffnung der HMO's (Managed Care, *Health Maintenance Organizations*) für Medicare-Versicherte durch das *Balanced Budget Act* 1997, der Informationen über die Qualität der Managed Care Organisationen zur Verfügung stellte.

► In der Folge wurde vom *Institute of Medicine* im Jahr 1999 der bekannte **Report "To Err Is Human"** zur Patientensicherheit veröffentlicht, der diese spezielle und grundlegende Thematik bereits sehr intensiv und sachgerecht im Licht organisatorischer und systemischer Zusammenhänge diskutierte (Kohn et al. 1999). Zwei Jahre später folgte der IOM-Report "**Crossing the Quality Chasm**" (IOM 2001), der auf der Basis der System- und Komplexitätstheorie eine umfassende Analyse der Situation in den USA sowie ein umfassendes Zukunftskonzept vorstellte, das neben Führungsverantwortung, neuen Management-strategien, gezieltem Einsatz von Informationstechnologie und

Evidence-Based Medicine vor allem eine Reform der Vergütungssysteme ganz im Vordergrund stellte. Dies geschah in der Erkenntnis, dass manche Vergütungsansätze einer Qualitätsverbesserung geradezu entgegenstehen, wenn sie z.B. nur das Leistungsvolumen, nicht aber Qualität oder Effizienz vergüten (z.B. Einzelleistungsvergütung) (IOM 2001, S. 181ff).

► Als Konsequenz der angestoßenen Diskussionen wurde 2003 im *Medicare Prescription Drug, Improvement and Modernization Act* (MMA) und 2005 im *Deficit Reduction Act* ein **Public Reporting**-Programm unter der Bezeichnung *Hospital Inpatient Quality Reporting Program* (Hospital-IQR) eingeführt (IOM 2007, CMS 2012, s. unten), alsbald im Jahr 2006 im *Tax Relief and Health Care Act* ergänzt durch **Pay for Reporting**-Elemente.

► Bereits im *Deficit Reduction Act* (in Kraft seit dem 8.2.2006) wurden Einschränkungen in der DRG-Vergütung bei Auftreten von Komplikationen beschlossen, sozusagen ein **Non-Payment for Non-Performance-Programm** (Graves und McGowan 2008, Pronovost et al. 2008). Ab 2015 ist zusätzlich ein 1%-Abzug für die Krankenhäuser im Quartil mit den meisten dieser “*Never Events*” geplant (Belmont et al. 2011).

► Auf den Erfahrungen des Hospital-IQR und der *No-Payment Policy* baute dann wiederum das HQIP auf, seinerseits Pilotprojekt für das VBP-Programm, das im Krankenhausbereich bereits im Jahr 2012 gestartet ist und in der Perspektive auf alle Bereiche des Gesundheitswesens ausgedehnt wird (CMS 2012, s. unten).

Weiter: 2. Langfristige Effekte, 2.4. *Value-based Purchasing*

2.4. Value-Based Purchasing

► Weiterhin zeigte sich das gesteigerte öffentliche Interesse an der Qualitäts- und Sicherheitsfrage daran, dass der Congress die CMS damit beauftragte, die Grundlagen für ein **“Value-Based Purchasing”-Programm (VBP)** (s. Tableau 11) als umfassenderes Vergütungssystem zu entwickeln, das der Qualität der Versorgung Rechnung trägt und nicht mehr allein die Leistungsmenge (Rosenthal 2008). Das VBP-Programm wurde im Jahr 2010 im Rahmen des *Patient Protection and Affordable Care Act* beschlossen und ist seit Ende 2012 implementiert, seit 1.10.2013 sind die finanziellen Regelungen im Krankenhausbereich aktiv (Brown et al. 2014, Chien und Rosenthal 2013).

Das VBP-Programm entspricht insofern einer konzeptionellen Weiterentwicklung des P4P-Konzeptes, als dass nicht mehr allein die Qualität einer erbrachten Leistung (*“quality-based purchasing”*, Cannon 2006), sondern primär das Verhältnis von Qualität und finanziellem Aufwand, also die Effizienz der Versorgung, Gegenstand des finanziellen Anreizes ist (CMS 2011A, CMS 2011B, Ryan et al. 2012). Dieser Effizienz-Gedanke wird als *Value* (*“quality in relation to the cost of care”*, Tompkins et al. 2009) bezeichnet und wird bereits seit Anfang der 90er Jahre in den USA intensiv diskutiert (Wenzel 1992), gegenwärtig ist *Value* auch der Kernbegriff der gegen die Überversorgung mit nutzlosen bzw. schädlichen Verfahren gerichteten *Choosing Wisely* Initiative der *National Physicians Alliance* in den USA (Morden et al. 2014). Zudem geht es nicht mehr primär um *Performance*, sondern um *Purchasing*, im deutschen Sprachgebrauch am ehesten mit “Kontrahieren” zu übersetzen.

Tableau 11: Value-based Purchasing

(VBP) gehört zu den Konzepten des Qualitätswettbewerbes:

- *Public Reporting*
- *Pay for Reporting*
- *Pay for Performance (P4P)*
- *Non-Payment for Non-Performance*
- *Value-based Purchasing (VBP)*

VBP geht insofern über P4P hinaus, als dass primär die Effizienz der Versorgung, also die erreichte Qualität in Relation zu den aufgebrauchten Mitteln, Vergütungsgegenstand ist.

Inhaltlich bezieht sich das VBP-Programm auf die sechs “Qualitätsdomänen” der *National Quality Strategy* des *Department of Health and Human Services (HHS)*

(VanLare et al. 2012, s. Abb. 3), die einerseits Patientensicherheit und Klinische Versorgung betonen, andererseits aber mindestens gleichwertig die Patienten- und "Caregiver"-Erfahrungen (einschließlich der *Patient-Reported Outcomes*), die Koordination der Behandlung und die "Population or Community Health" unter Einschluss des Zugangs zur Versorgung (*Access to Care*) sowie letztendlich die Effizienz und Kostenreduktion in den Vordergrund stellen.

Das aktuelle VBP-Programm ist sehr breit aufgestellt, neben dem Krankenhaussektor (Hospital-VBP) ist auch die Versorgung chronisch Nierenkranker mit einem eigenen Programm vertreten (QIP: *End-Stage Renal Disease Bundled-Payment and Quality Incentive Program*), und auch die ambulante Versorgung wird über den *Physician Value-Based Payment Modifier* adressiert, der erste finanzielle Auswirkungen im Jahr 2017 haben wird (VanLare et al. 2012, Tompkins et al. 2009, CMS 2011A, CMS 2011B, Chien und Rosenthal 2013).

Alle Akutkrankenhäuser, die Medicare-Patienten behandeln, sind zur Teilnahme am Hospital-VBP verpflichtet. Derzeit werden bis zu 1% der Erlöse in die Zu- und Abschläge einbezogen, dieser Wert soll bis 2017 auf 2% steigen, zu diesem Zeitpunkt werden dann jährlich knapp 2 Mrd. \$ nach Qualitätskriterien verteilt (Ryan et al. 2012C, zu weiteren Daten s. auch Brown et al. 2014). Es werden weitgehend die bereits aus dem Hospital-IQR Programm bekannten Indikatoren eingesetzt (s. Tableau 12), in erster Linie

- Prozessindikatoren und
- Patienten-Erfahrungen (im ersten Jahr nur diese beiden), weiterhin
- Patientensicherheitsindikatoren und
- *Hospital-Acquired Conditions*,
- nur in ausgewählten Fällen risikoadjustierte Ergebnisindikatoren.

Im ersten Jahr wurden allein Prozess- und Patienten-seitige Indikatoren im Verhältnis von 70 zu 30% verwendet, ab dem 2. Jahr (2014) werden alle 5 Gruppen je 20% gewichtet (Ryan et al. 2012C). Bemerkenswert ist wiederum die große Bedeutung der Patientenerfahrungen sowie der große Anteil von insgesamt 60%, den die Patientensicherheitsindikatoren ausmachen, zu denen ja auch die *Hospital-Acquired Conditions* und *Mortality*-Indikatoren hinzurechnen sind. Ergebnisse können naturgemäß zum jetzigen Zeitpunkt noch nicht vorliegen, allerdings gibt es bereits jetzt Modellrechnungen über die finanziellen Auswirkungen. Es wird diskutiert, ob diese ausreichen, um wirklich Verbesserungen erreichen zu können (Werner et al. 2012).

Tableau 12: Indikatoren des Hospital Value-Based Purchasing

Programms der CMS (n. Ryan et al. 2012C)

(1) Clinical process

- **Acute myocardial infarction:** fibrinolytic therapy received within 30 minutes of hospital arrival; primary percutaneous coronary intervention received within 90 minutes of hospital arrival
- **Heart failure:** discharge instructions
- **Pneumonia:** blood cultures performed in the emergency department before initial antibiotic received in hospital; initial antibiotic selection for community-acquired pneumonia in immunocompetent patients
- **Infections associated with health care:** prophylactic antibiotic received within 1 hour before surgical incision; prophylactic antibiotic selection for surgical patients; prophylactic antibiotics discontinued within 24 hours after surgery end time; cardiac surgery patients with controlled 6 a.m. postoperative serum glucose levels
- **Surgery:** surgery patients taking a beta-blocker before arrival who received a beta-blocker during the perioperative period; recommended venous thromboembolism prophylaxis ordered; appropriate venous thromboembolism prophylaxis received no more than 24 hours before surgery and up to 24 hours after surgery

(2) Patient experience

- Communication with nurses
- Communication with physicians
- Responsiveness of hospital staff
- Pain management; communication about medicines
- Hospital cleanliness and quietness
- Discharge instructions
- Overall rating of hospital

(3) Patient mortality

- Mortality at 30 days, acute myocardial infarction
- Mortality at 30 days, heart failure
- Mortality at 30 days, pneumonia

(4) Hospital-acquired conditions

- Foreign object retained after surgery
- Air embolism
- Blood incompatibility
- Pressure ulcer stages 3 and 4
- Falls and trauma (includes fracture, dislocation, intracranial injury, crushing injury, burn, electric shock)
- Infections associated with vascular catheters
- Urinary tract infections associated with catheters
- Manifestations of poor glycemic control

(5) Patient safety

- Composite measure of complication and patient-safety indicators (pressure ulcer, iatrogenic pneumothorax, bloodstream infections related to central venous catheters, postoperative hip fracture, postoperative pulmonary embolism or deep vein thrombosis, postoperative wound dehiscence, and accidental puncture or laceration)
- Mortality for selected medical conditions (composite)

Weiter: 2. Langfristige Effekte, 2.5. Ambulanter Sektor: QOF

2.5. Ambulanter Sektor: QOF-Programm in Großbritannien

Das *Quality and Outcome Framework* (QOF) des National Health Service (NHS) in Großbritannien stellt das zweite, wissenschaftlich breit evaluierte P4P-Projekt dar. Es betrifft den ambulanten Primärarzt-Sektor (*General Practitioner*) und bezieht sich ebenso wie das HQIP in erster Linie auf Prozessindikatoren. Ähnlich wie beim VBP-Programm des CMS hat auch der NHS dem QOL eine umfassende Strategie hinterlegt (Anonymous 1999, Campbell et al. 2007, NHS 2013B), die (ohne eine derart explizit systemtheoretische Fundierung wie in den USA) fünf Domänen umfasst und parallel zu entsprechenden Frameworks für *Public Health* und *Adult Social Care* implementiert wurde:

- Domain 1: *Preventing people from dying prematurely*
- Domain 2: *Enhancing quality of life for people with long-term conditions*
- Domain 3: *Helping people to recover from episodes of ill health or following injury*
- Domain 4: *Ensuring that people have a positive experience of care*
- Domain 5: *Treating and caring for people in a safe environment; and protecting them from avoidable harm*

Das QOL startete im Jahr 2004 nach umfänglichen Diskussionen zwischen NHS und der britischen Ärzteschaft, die nur nach der Bereitschaft der englischen Regierung, das QOL-Programm mit erheblichen zusätzlichen Mitteln auszustatten, zu einem erfolgreichen Ende geführt werden konnten (Roland 2004). Im Jahr 2013/2014 waren insgesamt 123 Indikatoren (NHS 2013A) in Kraft, die sich wie folgt zusammensetzen:

- 95 klinische Indikatoren (Tableau 13),
- *Public Health Domain* (18 Indikatoren): Primärprävention KHK (2), Bluthochdruck (1), Adipositas (1), Rauchen (5), Screening Cervix-Ca. (4), Kindergesundheit (1), Müttergesundheit (1), Kontrazeption (3),
- *Quality and Productivity Domain* (9 Indikatoren, z.B. Teilnahme an Peer Review-Verfahren, Leitlinien, Patientensicherheit),
- *Patient Experience Domain* (1 Indikator).

Über diese Indikatoren werden pro Praxis bis zu 1050 Punkte verteilt, die entsprechend vergütet werden. Es fällt sofort auf, dass das Spektrum der klinischen Indikatoren sehr viel breiter und sehr viel mehr auf chronische Erkrankungen ausgerichtet ist als beim HQIP-Programm, wo ja nur 5 klinische (v.a. akute) Krankheitsbilder einbezogen wurden; dieser Umstand ist für die allgemeine GP-Versorgung sicherlich adäquat. Es handelt sich fast ausschließlich um Prozessindikatoren. Für ein Teil der klinischen Indikatoren wird die Bereitstellung von Daten für Erkrankungsregister (z.B. chronische

Herzinsuffizienz) gesondert vergütet. Weiterhin fällt die große Bedeutung der Public Health- sowie der organisatorischen Indikatoren auf (zu Beginn einschließlich IT-Ausstattung), außerdem ein hoch bewerteter *Composite*-Indikator zu *Patient-Reported Outcomes*. Ziel des Programms ist eine strukturierte, Team-basierte und auf Evidenz-gestützte Qualitätsindikatoren ausgerichtete Versorgung; seit 2009 liegt die Indikatorenentwicklung in den Händen des *National Institute for Health and Clinical Excellence* (NICE) (Gillem et al. 2012).

Tableau 13: Klinische QOF-Indikatoren 2013/2014 (NHS 2013A):

Vorhofflimmern (4 Indikatoren)
Sek. Prävention der KHK (6)
Chronische Herzinsuffizienz (6)
Bluthochdruck (5)
Periphere arterielle Verschlusskrankheit (4)
Schlaganfall und transitorisch-ischämische Attacke (7)
Diabetes mellitus (16)
Schilddrüsenunterfunktion (2)
Asthma (4)
Chronisch-obstruktive Lungenerkrankung (6)
Demenz (3)
Depression (2)
Psychische Erkrankungen (10)
Krebs (2)
Chronische Niereninsuffizienz (4)
Epilepsie (3)
Lernstörungen (2)
Osteoporose: sekundäre Fraktur-Prävention (3)
Rheumatoide Arthritis (4)
Palliativversorgung (2)

Ein weiterer deutlicher Unterschied zu HQPI ist die Einführung von *Exception Rules* (s. Tableau 14), die es unter bestimmten Bedingungen (s. Kasten) erlauben, Patienten aus dem P4P-Programm herauszunehmen, insbesondere wenn Patienten-seitige Gründe für ein Nicht-Erreichen des angestrebten Ziels (z.B. Compliance bzgl. Impfungen) vorliegen (NHS 2013A). Diese Regelung soll einer Risikoselektion und Schlechterstellung (bzw. niedrigerer Attraktivität) von Praxen in sozial unterprivilegierten Wohngebieten entgegenwirken.

Die zusätzlich ausgeschütteten Mittel sind sehr hoch, bereits im ersten Jahr konnte eine Praxis bis zu 42.000 Engl. Pfd. (etwa € 50.000.-) zusätzlich einnehmen, wenn alle 1050 Punkte erreicht wurden. In der ersten Periode der Umsetzung wurde 1 Mrd. Engl. Pfd. Qualitäts-bezogen verteilt, das entspricht einer 20%-igen Erhöhung des Budget für die *Family Practitioners* im United Kingdom (Roland 2004). Im Median wurden 83,4% der

klinischen Qualitätsziele und im Median insgesamt 1003 Punkte erreicht, 230 der 8105 auswertbaren Praxen erreichten die maximale Punktzahl von 1050. Für im Median 6% der Patienten kam ein *exception reporting* zur Anwendung. In den Jahren 2009-2010 kamen die Praxen auf im Durchschnitt 937 Punkte bei einem "Punktwert" von 130 Engl. Pfd. (Gillem et al. 2012).

Tableau 14: Exception rules (2013/14) (NHS 2013A):

Patients may be excepted if they meet the following criteria:

- A. Patients who have been recorded as refusing to attend review (...)
- B. Patients for whom it is not appropriate to review the chronic disease parameters due to particular circumstances (...)
- C. Patients newly diagnosed or who have recently registered with the contractor (...)
- D. Patients who are on maximum tolerated doses of medication whose levels remain sub-optimal.
- E. Patients for whom prescribing a medication is not clinically appropriate (...)
- F. Where a patient has not tolerated medication.
- G. Where a patient does not agree to investigation or treatment (informed dissent) (...)
- H. Where the patient has a supervening condition which makes treatment of their condition inappropriate (...)
- I. Where an investigative service or secondary care service is unavailable.

Ähnlich wie beim HQIP-Programm kann auch hier zwischen den frühen Evaluationen, die eine gewisse Wirksamkeit zeigten, und den langfristigen Untersuchungen unterschieden werden, die den Optimismus etwas dämpften. In der Evaluation von Campbell et al. (2007) war in einer *interrupted time-series* Analyse die Qualitätsverbesserung insbesondere bei den Krankheitsbildern Asthma und Typ 2-Diabetes von dem langjährigen säkularen Trend einer langsamen auch ohne Intervention sichtbaren Verbesserung deutlich abzugrenzen, dieser Effekt war aber in den Jahren danach nicht mehr nachweisbar (Campbell et al. 2009). Dieser Effekt wurde als *ceiling*-Effekt interpretiert, eine Anhebung der Grenzwerte empfohlen (Doran und Roland 2010). Vor allem aber waren die nicht direkt vom P4P-Ansatz betroffenen Diagnosen von dem Verbesserungseffekt ausgeschlossen. Während die Patientenorientierte Outcomes z.B. hinsichtlich Zugang und Erreichbarkeit besser wurden, war für die Kontinuität der Behandlung (vor allem hinsichtlich "persönlicher Arzt") eine Verschlechterung zu verzeichnen. Besondere Aufmerksamkeit erlangte eine Studie, in der Kaiser Permanente in Californien im Primarztbereich zwei Indikatoren aus dem P4P-Programm herausnahm, die auch im NHS gleichzeitig aktiv waren (Screening des Augenhintergrunds bei Diabetes mellitus, Screening für Zervix-Carcinom), und zeigte,

dass die erreichten Verbesserungen wieder auf den Ausgangsbereich zurückfielen (Lester et al. 2010).

Zwei Systematische Reviews, die sich ausschließlich auf Studien zum QOL-Programm bezogen, erhärteten diese Ergebnisse (Gillem et al. 2012, Langdown et al. 2013). Die Verbesserungen werden ähnlich wie im HQIP zeitlich vorgezogen, Team-Arbeit und Veränderung der Rollen gefördert - andererseits beschreiben Patienten eine Verschlechterung in der persönlichen Kontinuität der Betreuung als "Kosten" für die bessere Erreichbarkeit, nicht angereizte Krankheitsbilder profitieren nicht, und es gibt Befürchtungen, dass soziale Unterschiede, trotz *exception rules*, verstärkt werden.

Weiter: 2. Langfristige Effekte, 2.6. Unerwünschte Nebeneffekte

2.6. Synopse: Unerwünschte Nebeneffekte von P4P

2.6.1. Systematik, Zugang zur Versorgung

Bei jedem Vergütungs- und Anreizsystem muss mit unerwünschten Nebeneffekten gerechnet werden (Brown et al. 2014). Auch bei P4P sind zahlreiche solcher Wirkungen bereits früh beschrieben worden (SVR 2008, Nr. 741ff (s. Abb. 5), Roland 2004) und werden hier zunächst systematisch und synoptisch aufgeführt, um im späteren Verlauf in die Analyse und die Empfehlungen eingearbeitet werden zu können. Ganz entscheidend werden dabei die Frage der professionellen Motivation der Angehörigen der Gesundheitsberufe, die mögliche Verstärkung sozialer Ungleichheit, der Anreiz zur Risikoselektion und die Frage der Mengenausweitung sein, vor allem in Kombination mit anderen Vergütungssystemen. Zu beachten ist, dass sich für eine große Zahl von unerwünschten Effekten Studienergebnisse finden lassen, es aber einen gewissen Bias zugunsten ungünstiger Ergebnisse gibt, man muss die Ergebnisse daher wirklich in ihrem Kontext sehen. In Erweiterung der Darstellung des SVR (s. Abb. 5) kann man die Nebeneffekte folgendermaßen systematisieren:

- Zugang zur Versorgung einschl. Kontinuität und Risikoselektion
- Professionelle Faktoren
- Institutionelle Ebene
- Datenqualität
- Qualität der Versorgung im gleichen Sektor
- Qualität der Versorgung auf Systemebene
- Kosteneffektivität der Versorgung

► Der **Zugang** zur Versorgung mit P4P-bezogenen Leistungen (s. Tableau 15) kann im Einzelfall besser werden (Wodchis et al. 2007), vor allem soweit es sich, wie beim QOF-Programm, um eine versteckte Einzelleistungsvergütung handelt (s. auch dort). Im Allgemeinen wird aber eine Verschlechterung der Versorgung für sozial, ethnisch oder aus anderen Gründen benachteiligten Gruppen befürchtet (Rosenthal und Dudley 2007). Eine Benachteiligung **ethnischer Gruppen** konnte in den USA bislang nicht belegt werden (Blustein et al. 2011, Casalino und Elster 2007, Ryan 2010), wenngleich immer wieder betont wird, dass Aufmerksamkeit angebracht ist. Bezüglich allgemeiner **sozialer Unterschiede** ist die Tendenz noch kritischer (Alshamsan et al. 2010, Casalino und Elster 2007, Jha et al. 2010, Langdown und Peckham 2013, McLean et al. 2006, Millett et al. 2007). Ein **exception reporting** ist häufiger in Regionen mit niedrigem Einkommen (QOL-Programm in Großbritannien (Doran et al. 2006), in einem Systematischen Review konnte jedoch eine leichte Verbesserung dieser Differenz nachgewiesen werden (Gillam et al. 2012). Besondere Bedeutung kommt der Analyse

von Ryan (2013) über das erste Jahr des VBP-Programms von Medicare zu, in der ein inverser Zusammenhang zwischen sozialer Ungleichheit und Zahlungen durch das VBP-Programm nahegelegt wird. Eine Studie berichtet über Nachteile von **weiblichen** Patientinnen (Millett et al. 2007).

Gliederung: Unerwünschte Nebeneffekte von P4P

- Zugang zur Versorgung einschließlich Kontinuität und Risikoselektion (Kap. 2.6.1.)
- Professionelle Faktoren (Kap. 2.6.2.)
- Institutionelle Ebene (Kap. 2.6.2.)
- Datenqualität (Kap. 2.6.2.)
- Qualität der Versorgung im gleichen Sektor (Kap. 2.6.3.)
- Qualität der Versorgung auf Systemebene (Kap. 2.6.3.)
- Kosteneffektivität der Versorgung (Kap. 2.6.3.)

Zur Zugangsproblematik gehört auch die **Kontinuität der Versorgung** (Roland 2004). Während sich allgemein die *patient reported outcomes* unter P4P verbessern, wird die Kontinuität in einigen Studien und Befragungen schlechter (Campbell et al. 2009, Gillem et al. 2012, Maisey et al. 2008). Patienten kommen "nicht zu ihrem Recht" (Chew-Graham et al. 2013), ob dies an arbeitsteiligeren Konzepten in der ärztlichen Versorgung, die Übernahme ärztlicher Tätigkeiten durch Pflegende oder anderen Ursachen liegt, ist noch nicht abschließend geklärt. Ein weiterer und im Kontext von P4P äußerst wichtiger Aspekt des Zugangs zur Versorgung ist eine mögliche **Risikoselektion**, weil auf diese Weise für bestimmte Patienten bzw. Patientengruppen die Versorgung erschwert sein kann (Brown et al. 2014, Casalino und Elster 2007, Shen 2003). In den USA wurde keine Risikoselektion beobachtet (Ryan und Blustein 2012C). Eine Risikoselektion gehört prinzipiell zu den unmittelbaren Wirkungen eines Finanzierungssystems, das entweder die umfassende Behandlung von Erkrankungsgruppen oder Populationen (z.B. *Disease Management* oder *Managed Care*) in den Mittelpunkt stellt (s. Kap. 5) oder auf den Behandlungserfolg (Ergebnisindikatoren) setzt; bei Verwendung von Prozessindikatoren ist sie weniger stark ausgeprägt (Mehta et al. 2008). Die Tendenz zur Risikoselektion stellt daher eines der entscheidenden Argumente bei der Auswahl der verwendeten Indikatoren und der Integration von P4P in bestehende Vergütungssysteme dar (s. auch Kap. 7).

Zusammenfassend kann festgehalten werden, dass das Thema einer Verstärkung sozialer Benachteiligung durch P4P besonderer Beachtung bedarf, ebenso muss die Kontinuität der Versorgung und die Gefahr einer Risikoselektion mit daraus folgender Zugangsbeschränkung beachtet werden. Die Kontrolle aller drei Formen ist eine wichtige Aufgabe der politischen Instanzen bei Implementierung von P4P.

Weiter: 2.6. Unerwünschte Nebeneffekte, 2.6.2. Professionelle und institutionelle Ebene,
Datenqualität

2.6. Synopse: Unerwünschte Nebeneffekte von P4P

2.6.2. Professionelle und institutionelle Effekte, Datenqualität

► **Professionelle Faktoren** beziehen sich in erster Linie auf die professionelle Motivation der unterschiedlichen Gesundheitsberufe, die Berufsbilder und die Koordination der Arbeit auf Teamebene (s. Tableau 16). Grundsätzlich muss bei jedem Vergütungssystem, das im beruflichen Kontext angewandt wird, mit Auswirkungen auf Motivation, Rolle und Zusammenarbeit gerechnet werden (Prendergast 1999, s. auch Kap. 3.5.1.). Zur Motivation der Ärzte besteht die wichtigste Befürchtung darin, dass die interne Motivation (professionelle Einstellung), durch die externe Motivation i.S. einer Qualitäts-bezogene Vergütung gemindert wird ("*crowding out*"; Berenson et al. 2013, Cassel und Jain 2012). Empirisch ist dies jedoch nicht zu belegen, wie Untersuchungen z.B. im QOL-Projekt in Großbritannien zeigen (McDonald et al. 2007), und auch von theoretischer Seite gibt es Gegenargumente (Prendergast 1999, Staehle 1999, S. 242). Dieses häufig geäußerte Argument ist also nicht so durchschlagend, wie man es auf den ersten Blick meint. Allerdings kommt es zu Verschiebungen in den beruflichen Rollen, insbesondere ändert sich im Rahmen einer P4P-Einführung die Arbeitsteilung zwischen Ärzten und Pflege. Durch die Übernahme vorher ärztlicher Tätigkeiten durch die Pflege kann es dort zu zusätzlicher Belastung und Unzufriedenheit kommen (Kurtzman et al. 2011, Maisey et al. 2008, McDonald et al. 2007). Eventuell wird die Teamarbeit erschwert (Maisey et al. 2008), wenngleich im Allgemeinen eine Verbesserung berichtet wird.

Gliederung: Unerwünschte Nebeneffekte von P4P

- Zugang zur Versorgung einschließlich Kontinuität und Risikoselektion (Kap. 2.6.1.)
- Professionelle Faktoren (Kap. 2.6.2.)
- Institutionelle Ebene (Kap. 2.6.2.)
- Datenqualität (Kap. 2.6.2.)
- Qualität der Versorgung im gleichen Sektor (Kap. 2.6.3.)
- Qualität der Versorgung auf Systemebene (Kap. 2.6.3.)
- Kosteneffektivität der Versorgung (Kap. 2.6.3.)

► Auf der **institutionellen Ebene** (s. Tableau 16) kommt es durch eine P4P-Einführung zu "organisatorischem Stress". In der internen Steuerung wird jetzt nicht nur über Mengen und Kosten (bzw. interne Leistungsverrechnung) diskutiert, sondern es muss über die Qualität der Leistungserbringung gesprochen werden, da sie jetzt vergütungsrelevant ist (Corrigan und McNeill 2009). Dies beginnt mit der Erhebung der Qualitätsdaten, eine Ergänzung oder gar Restrukturierung der IT-Ausstattung wird oft notwendig. Zentral ist dann die Frage, ob die Organisation oder die einzelnen für die

Behandlung verantwortlichen Ärzte oder Abteilungen Adressaten der P4P-Vergütung sind (Rosenthal und Dudley 2007). Die Vergütung muss dabei mit dem fachlichen Verantwortungsbereich kongruent sein. Während im Allgemeinen die Organisation die Zahlungen empfängt (ohne dass sie weitergeleitet werden), gibt es Hinweise darauf, dass (relativ kleine) Zahlungen direkt an die Behandler sehr effektiv sind (Chung et al. 2010, Torchiana et al. 2013). Das Konzept der Ärztlichen Leitung (*medical leadership*) wird ein bedeutender Faktor (Frolich et al. 2007), da hier diese Diskussionen entscheidend mit gestaltet werden. Durch die organisatorische Kleinteiligkeit wird die Situation im ambulanten Bereich noch komplizierter, weil die Verantwortung für die Qualität der Leistung auf verschiedene Akteure verteilt ist (Pham et al. 2007).

Tableau 16: Unerwünschte Effekte von P4P. Professionelle und institutionelle Faktoren einschl. Datenqualität:

- Professionelle Faktoren
 - Motivation
 - Rollenverständnis der Gesundheitsberufe
 - Teamarbeit
- Institutionelle Faktoren
 - IT-Ausstattung
 - Verantwortung für Qualität
 - Adressaten der Zahlungen
 - Konzept des *medical leadership*
- Datenqualität
 - Technische Erhebungsprobleme
 - Gaming

► Die **Datenqualität** betrifft in diesem Zusammenhang in erster Linie die Reliabilität (Zuverlässigkeit) der Daten (Begriffe und Abgrenzung zur Validität s.o.). Unter mangelnder Reliabilität versteht man die zwischen Personen und zu unterschiedlichen Zeiträumen inkonstante Erfassung der Daten, vor allem aber auch Verzerrungen durch Mängel der Dokumentation und Fälschung der Daten (sog. *gaming*, s. Tableau 16). Während die erste Form durch Verbesserung der Definition und der Spezifikationen der verwendeten Indikatoren anzugehen ist, ist dies im Fall des *gaming* schwieriger. Der Grund besteht darin, dass dem *gaming* eine skeptische oder ablehnende Grundhaltung gegenüber der Qualitätsmessung zugrundeliegt, und zwar sowohl bei den Gesundheitsberufen als auch auf der Ebene der Organisation. Es gibt Studien, in denen Ärzte in Befragungen ein *gaming* offen zugeben (Maisey et al. 2008). Da in diesen Fällen Qualitätsmessung und -vergleich weder zum professionellen Selbstverständnis noch zur Qualitätskultur der Einrichtung gehört, werden Ausflüchte und Erklärungen gesucht (Mears und Webley 2010). Man unterscheidet analog zur Steuerhinterziehung

eine partiell und eine komplett ablehnende Haltung, bei ersterer wird bis zu einem gewissen Punkt Compliance geübt und erst darüber hinaus Datenmanipulation betrieben, bei der zweiten Haltung werden von vorneherein die Daten verfälscht. In den USA gibt es im Bereich der Datenerhebung zur Infektionsepidemiologie die Einrichtung eines "klinischen Vetos" (*clinician veto*) gegen die Erhebung von nosokomialen Infektionen, in denen das nach infektionsepidemiologischen Kriterien erhobene Ereignis durch klinische Einschätzungen in Frage gestellt wird (und meist trotzdem antibiotisch behandelt wird); die Infektionsepidemiologen warnen dort vor einer Gefährdung der für infektiologische Zwecke optimierten Erhebungsmethodik (Talbot et al. 2013). In jedem Fall ist die Frage der Qualitäts- bzw. Sicherheitskultur, einer der Eckpfeiler der Organisationskultur und damit zentraler Baustein des organisatorischen Zusammenhalts, bei der Beurteilung des *gaming* von zentraler Bedeutung (s. organisatorisches Lernen).

Zusammenfassend sind auf professioneller und institutioneller Ebene Veränderungen der beruflichen Rollen beschrieben, die Unsicherheit erzeugen. Verlässliche Daten zur Infragestellung der internen Motivation der Ärzte durch den externen finanziellen Anreiz gibt es kaum. Vergütungs-relevante Qualitätsaspekte stellen große institutionelle Herausforderungen für die interne Steuerung dar. Die Reliabilität der Daten betrifft vor allem das *gaming*, begleitet von zentralen Defiziten im Hinblick auf die Organisations- und Sicherheitskultur.

Weiter: 2.6. Unerwünschte Nebeneffekte, 2.6.3. Qualität auf Sektor- und Systemebene, Kosteneffektivität

2.6. Synopse: Unerwünschte Nebeneffekte von P4P

2.6.3. Qualität auf Sektor- und Systemebene, Kosteneffektivität

► Die **Qualität auf der Ebene des gleichen Sektors** ist gefährdet, wenn die Anreizwirkung eines Instrumentes wie P4P trotz erhaltenem Zugang zur Versorgung und professioneller sowie institutioneller Akzeptanz nicht zur gewünschten Verbesserung der Versorgung führt, ohne dass dies auf methodische Ursachen wie z.B. inadäquate Kopplung von Qualität und Vergütung zurückzuführen wäre (s. Tableau 16). Einer der hier relevanten Mechanismen wird als **teaching to the test** bezeichnet. Man versteht darunter die einseitige Ausrichtung der Aktivitäten auf den Indikator unter Vernachlässigung der anderen Aufgaben, um "gute Ergebnisse" zu erzielen (Brown et al. 2014, Casalino und Elster 2007). Durch dieses Verhalten wird die "Stellvertreterfunktion" des Indikators, der als gut messbarer Parameter Qualitätsprobleme in einem Versorgungsbereich vorhersagen soll, untergraben (s. Kap. 1.4.). Es handelt sich hierbei um eine dem Indikator-konzept immanente Gefahr, denn Indikatoren machen Qualitätsprobleme explizit und lösen Vermeidungsreaktionen aus. Bei P4P und anderen Formen des "Qualitätswettbewerbs" (zur Systematik s. Abb. 1) ist diese Gefahr noch größer, denn P4P tendiert dazu, finanzielle Ressourcen von nicht angereizten Leistungen zu angereizten Leistungen zu allozieren (Ryan 2010). Allerdings ist dieses Phänomen zumindest in den amerikanischen Programmen nicht in großem Umfang zu beobachten gewesen (Ryan und Blustein 2012). Ganz entscheidend ist die Frage, ob ein **spill over** zu anderen, nicht angereizten Leistungsbereichen nachzuweisen ist, und ob sich durch P4P im Sinne der lernenden Organisation eine **bleibende Verbesserung** der Versorgung erreichen lässt, auch wenn das P4P-Programm beendet wird. Im QOL-Programm in Großbritannien konnte gezeigt werden, dass nicht angereizte Leistungsbereiche sich nicht verbessern, obwohl dies eigentlich angenommen worden war (Gillem et al. 2012, Langdown und Peckham 2013). Zumindest in einer Studie konnte nachgewiesen werden, dass nach Beendigung von P4P-Anreizen die Qualität der von P4P betroffenen Leistung wieder zurückgeht (Lester und Peckham 2010).

► Die **Qualität auf Systemebene** wird hier von der potentiellen Qualitätsverschlechterung auf der Ebene des gleichen Sektors abgegrenzt, um die Gefährdung allgemeiner Gesundheitsziele (adäquate Behandlung chronisch und mehrfach erkrankter Patienten, Mengenanreiz, Versorgungsengpässe) zu kennzeichnen (s. Tableau 16.1). Man kann das **ceiling**, den ausbleibenden Verbesserungseffekt durch mangelndes Verbesserungspotential, auch als methodisches Problem ansehen, aber es hat in jedem Fall systemische Auswirkungen, weil das Instrument in Verruf gerät und Ressourcen für die Messung und gesonderte Vergütung von Indikatoren verwendet

werden, bei denen gar keine Möglichkeit einer Verbesserung mehr besteht (Bhattacharyya et al. 2009, Doran und Roland 2010). *Ceiling* tritt besonders dann auf, wenn die für P4P verwendeten Indikatoren bereits lange im Gebrauch sind und auch schon für andere Anreizsysteme (in erster Linie *Public Reporting*) verwendet wurden (Ryan und Blustein 2011). Insbesondere bei ungenügend risikoadjustierten P4P-Konzepten auf Basis von Ergebnisindikatoren kann es zu einer **Benachteiligung kleiner Einrichtungen** kommen (Davidson et al. 2007), so dass in ländlichen Regionen sogar die Gefahr einer Unterversorgung entsteht. Ist die Abgrenzung zur Einzelleistungsvergütung ungenügend, kann der Mengenanreiz, z.B. bereits bestehend durch das DRG-System, noch verstärkt werden. Wenn P4P nicht zielgerichtet und strategisch eingesetzt wird, kann die Versorgung von Patientengruppen, die von besonderem Interesse sind (z.B. die Versorgung von **multimorbiden chronisch Kranken**) unter P4P sogar leiden (Cannon 2006), insbesondere falls sich darunter die Tendenz zur Risikoselektion verbirgt (s.o.). In einer bemerkenswerten Stellungnahme von H. Luft wird darauf hingewiesen, dass die Datenmenge und -verfügbarkeit, die für das Funktionieren von P4P notwendig sind (insbesondere für die Risikoadjustierung), eine ernsthafte Bedrohung des **Datenschutz** darstellen kann (Luft 2012) .

Tableau 16.1: Unerwünschte Effekte von P4P. Sektor- und Systemebene

- Gleicher Sektor
 - Teaching to the Test*
 - Spill over*
 - Bleibende Verbesserung
- System
 - ceiling*
 - Benachteiligung kleiner Einrichtungen
 - Mengenanreiz
 - Versorgung chronisch Kranker
 - Datenschutz

► Die **Kosteneffektivität** von P4P ist letztendlich auch als Systemparameter anzusprechen, denn wenn die Kosten für die Durchführung von P4P derart hoch sind, dass sie an anderer Stelle sinnvoller zur Qualitätsverbesserung hätten eingesetzt werden können, ist das Ziel einer besseren und kostenbewussteren Gesundheitsversorgung nicht erreicht (Roland 2004). Die Kosteneffektivität ist aber sehr wenig untersucht (Cannon 2006), in Großbritannien wird eine gemäßigte Kosteneffektivität berichtet (Gillem et al. 2012).

Weiter 3. Motivation ..., 3.1. Kontext von P4P: Vier Domänen

3. Motivation, Organisation, System

3.1. Kontext von P4P: Vier Domänen

Auf den ersten Blick kann man es nicht anders sagen - die bisherigen Erfahrungen mit *Pay for Performance*, insbesondere die beiden großen Programme in den USA (*Value-based Performance*) und in Großbritannien (*Quality and Outcome Framework*), lassen die Bäume nicht in den Himmel wachsen. Die an Qualitäts- und Patientensicherheitsindikatoren geknüpfte Vergütung führt nur zu quantitativ und zeitlich limitierten Verbesserungen, die Effekte sind reversibel, es gibt keinen *spin-off* für andere Erkrankungen außerhalb der Programme. Die organisatorischen Strukturen und die Arbeitsteilung der Berufsgruppen können teilweise günstig beeinflusst werden, aber ein *breakthrough* (Pearson et al. 2007) auf der Ebene des Gesundheitssystems sieht anders aus. Allerdings ist zumindestens in Deutschland bislang nicht die Frage beantwortet worden, ob die Erwartungen, die an P4P im Sinne eines *one fits it all*-Effektes gerichtet waren, überhaupt realistisch waren, und wie denn ein weiter gespannter konzeptioneller Rahmen, der den Hintergrund für die Generierung von Hypothesen über Wirksamkeit und Umfeldbedingungen des Einsatzes bilden könnten, aussehen könnte (Frolich et al. 2007). Um einen solchen konzeptionellen Rahmen zu konstruieren, wird in der vorliegenden Arbeit von folgenden vier Punkten ausgegangen:

- ▶ Ein **conceptual framework** muss die Vielfalt der Faktoren, die auf das Verhalten der Leistungsanbieter einwirken und den Effekt von P4P mit beeinflussen, berücksichtigen und gleichzeitig das Gesundheitssystem als Ganzes in seiner Reaktion auf P4P mit einbeziehen. Bevor man über den Nutzen von P4P abschließend urteilt, muss man über ein realistisches Bild zu den Kontextfaktoren seiner Einführung und Praxis verfügen, insbesondere weil sonst die Gefahr besteht, dass man mit P4P ein wirksames Prinzip von der weiteren Entwicklung ausschließt, obwohl die vermeintliche Wirkungslosigkeit durch inadäquate Einsatzbedingungen hervorgerufen wird.
- ▶ Auf der Grundlagen einer derart abgeleiteten Aussage über die erwartenden Effekte (unter spezifischen Umsetzungsbedingungen) können **Zielerwartungen** formuliert und die Evaluation geplant werden. Realistische Vorstellungen zur wissenschaftlichen Evaluation und deren möglichen Ergebnisse sind notwendig, damit nicht voreilig von "fehlendem Nutznachweis" gesprochen wird, wenn die Studien zu P4P widersprüchliche oder "einerseits - andererseits"-Ergebnisse erbringen; schon an dieser Stelle soll angemerkt werden, es handelt sich dabei nicht um Medikamentenstudien sondern komplexe Versorgungsforschungsvorhaben.

► Dieses Vorgehen muss die Tatsache reflektieren, dass sich das Konzept einer Qualitäts- oder Effizienz-orientierten Vergütung auf der Seite der Gesundheitspolitik einer steigenden Beliebtheit erfreut. Wie eingangs bereits ausgeführt, spricht davon in Deutschland z.B. der Koalitionsvertrag vom 27.11.2013, aber auch der enorme, langfristige Ausbau von P4P-Programmen z.B. in den USA und in Großbritannien mit ihren sehr unterschiedlich strukturierten Gesundheitssystemen ist beeindruckend. Die **politischen Entscheider** müssen daher von realistischen Umsetzungsbedingungen und Erwartungen ausgehen können; der konzeptionelle Rahmen sollte dabei vor allem auch die Langfristigkeit der Entwicklung herausstellen (s. Kapitel 7 “Empfehlungen für die zukünftige Nutzung von P4P”).



► Bei der Entwicklung und Ausgestaltung eines *conceptual framework* darf P4P nicht als isoliertes Problem betrachtet werden, sondern ist im **Zusammenhang mit anderen Themen** im Gesundheitswesen zu sehen: Umsetzung von Leitlinien (Grimshaw et al. 2004), Einführung von *Evidence-Based Medicine*-gestützten klinischen Behandlungsmethoden (Bero et al. 1998), erfolgreicher Einsatz von *Infection Control*-Maßnahmen wie einer adäquaten Händedesinfektion (Erasmus 2010), Auswirkungen des institutionellen Qualitätsmanagements (Groene 2011), Verbesserungen auf dem Gebiet der Patientensicherheit durch die Instrumente, die in den letzten Jahren auf diesem Gebiet entwickelt worden sind (Wachter 2010). Diesen Themen ist eine

manchmal erschütternde Langsamkeit in der Umsetzung der notwendigen Veränderungen gemeinsam, obwohl sie auf einer recht plausiblen Grundidee beruhen, die jedoch nicht Motivation genug zu sein scheint. Diese Tatsache ist um so erstaunlicher, als dass die genannten Instrumente und Konzepte für alle Beteiligten große Vorteile bei eigentlich nur geringem Aufwand versprechen (Beispiele mit exemplarischer Nennung):

- **Leitlinien:** Vereinfachung der Routine, Freisetzen von Zeit für komplizierte Patienten, Erleichterung der Aus- und Fortbildung für Berufsanfänger, Ausbildung von professionellen Standards;
- **Evidence-Based Medicine:** Verstärkung des akademischen Anspruchs, sowohl der Ärzte als auch der anderen, in der Akademisierung befindlichen Berufsgruppen, Legitimierung der Tätigkeit, Wissensbasis in der Ausbildung;
- **Infection Control:** Vermeidung von häufigen, vermeidbaren Komplikationen und Verbesserung der Behandlungsergebnisse, finanzielle Ersparnisse und Imagegewinn;
- **Qualitätsmanagement:** Lernen, schrittweise Verbesserung der Strukturen, strategische Beweglichkeit, Anpassung an Innovationsprozesse;
- **Patientensicherheit:** Entlastung im Bereich der Fehlerproblematik, geringeres Risiko medikolegaler Konsequenzen, finanzielle Entlastung und Sicherheit durch Erhalt der Versicherbarkeit;
- **P4P:** Qualität statt Menge heißt letztlich, nicht mehr dem Mengenanreiz ("Hamsterrad") ausgeliefert zu sein, sondern eine qualitativ hochstehende Versorgung umsetzen zu können.

Diese Diskrepanzen können nur so erklärt werden, dass sie das Fehlen eines konzeptionellen Rahmens widerspiegeln. Gleichzeitig stellen sie einen wichtigen Informationshintergrund für dessen Entwicklung dar, der gesamte Kontext dieser Themen und Instrumente ist mit in die Betrachtung einzubeziehen. Im vorliegenden Artikel wird dabei, in einem sicherlich vereinfachenden Bild, von einem Zusammenwirken von vier **Kontextdomänen** ausgegangen (s. Abb. 6):

- die Domäne der Professionen,
- die Ebene der Organisationen bzw. Institutionen,
- die wissenschaftliche Akzeptanz und Absicherung und
- die gesellschaftlichen (politisch-ökonomischen) Grundannahmen.

Für die Analyse entscheidend ist die Einsicht, dass Anpassungen auf der Ebene einer einzelnen Kontextdomäne das Problem nicht lösen. Zwar können Interventionsstudien

gelegentlich nachweisen, dass eine Berücksichtigung einzelner Faktoren zu einer Verbesserung der Umsetzung führt, aber eine nachhaltige Strategie, die auch situationsunabhängig und strukturübergreifend wirksam wäre, konnte sich daraus bislang nicht entwickeln. Die Organisationen im Gesundheitswesen sind durch eine starke professionelle Bindung der Leistungsträger geprägt (s. unten), die Bindung an die Organisation selbst ist schwach. Aber auch die gesellschaftliche Domäne, sei sie jetzt mehr als politische verstanden oder als Ebene des ökonomischen Diskurses, kann allein nicht "durchregieren". Gleiches gilt für die wissenschaftliche Ebene. Einerseits weisen die oben genannten Beispiele einen sehr unterschiedlichen Grad der wissenschaftlichen Absicherung auf, so können z.B. *Infection Control*-Maßnahmen wie die adäquate Händedesinfektion als hochgradig abgesichert gelten, für die positiven Effekte von Qualitätsmanagement oder eben P4P ist dies jedoch nicht der Fall. Vor allem aber besteht auch der wissenschaftliche Kontext nicht voraussetzungslos, sondern ist mit dem Objekt inhaltlich verbunden. Dies zeigt sich besonders deutlich am Beispiel der *Evidence-Based Medicine*, deren positivistische Grundannahmen (lineares Model, Identifikation isolierter Einzelfaktoren etc.) historisch zwar dringend geboten erscheinen und bei biomedizinischen Interventionen auch weiterhin geboten sind, bei Themen wie P4P oder anderen komplexen Interventionen jedoch an ihre Grenzen stoßen (Broom et al. 2009).

Um Veränderungsstrategien im Gesundheitswesen zu entwickeln und einen Horizont für realistisch zu erwartende Veränderungen zu beschreiben, wird in den nächsten drei Kapiteln ein konzeptioneller Rahmen entwickelt, der von der Beschreibung der Einrichtungen des Gesundheitswesens als Expertenorganisation (*professional bureaucracy*) ausgeht und zentrale Botschaften der System- bzw. Komplexitätstheorie mit einbezieht.

Weiter: 3. Motivation ..., 3.2. Professional Bureaucracy, 3.2.1. Konzept

3.2. Professional Bureaucracy: die Expertenorganisation

3.2.1. Das Konzept der Expertenorganisation

Die hier genutzte Unterscheidung von vier Kontextdomänen, die die Umsetzung und Wirksamkeit eines Instrumentes wie P4P beeinflussen, betont besonders die Rolle der professionellen Faktoren, die für Organisationen wie die des Gesundheitswesens typisch sind (s. Abb. 6a). Derartige Organisationen bezeichnet man als Expertenorganisationen oder *professional bureaucracies* (Lega und DePietro 2005, Mintzberg 1979). Neben Krankenhäusern und anderen Gesundheitseinrichtungen gehören Universitäten (s. Tableau 17), Beratungsunternehmen, Schulen und soziale Dienste sowie Assoziationen wie z.B. Rechtsanwaltskanzleien zu dieser Gruppe von Organisationen.

Tableau 17: Eine klassische Form der Expertenorganisation ist die Universität: "***Business of a university is not university business.***" (Chandler 1999). Das Management einer Universität (*university business*) steht im Konflikt mit dem Wesen oder der Aufgabe einer Universität (*business of a university*), Forschung und Lehre zu betreiben, verbunden mit der sog. akademischen Freiheit. Allerdings gibt es Umstände, in denen ein aktives Management unabdingbar ist und es Konflikte zwischen diesen beiden Logiken gibt, z.B. wenn die finanzielle Basis der Universität bedroht ist, weil öffentliche Gelder ausbleiben, oder wenn sich die hochschulpolitischen Gegebenheiten ändern.

In erster Linie zeichnen sie sich durch eine hochgradige **Autonomie** von Experten (z.B. Universitätsprofessoren, Chefärzte) aus, die den wertschaffenden "operativen Kern" der Organisation bilden. Die Tätigkeiten der Experten sind sehr komplex, so dass sie nicht von Externen oder internen Managern analysiert und standardisiert werden können. Meist handelt es sich um den Dienstleistungsbereich, die Leistungen, die die Experten erbringen, werden meist stark nachgefragt, sodass die Experten über eine große berufliche Mobilität verfügen. Sie verfügen weiterhin über die eigentlichen externen Kundenbeziehungen und gestalten die Produkt-Markt-Beziehungen in maßgeblicher Form (z.B. im Krankenhaus gehen die Patienten zum Chefarzt oder zu Ärzten, die von ihm supervidiert werden, und nicht zum Geschäftsführer, um sich behandeln zu lassen). In der Konsequenz können sie sich daher dem direkten Durchgriff der (mittleren und oberen) Managementebene entziehen. Auch das Wachstum der Organisation vollzieht sich horizontal um die Fachgebiete der Experten herum und/oder durch weitere Spezialisierung der Funktionen (z.B. Innere Medizin (ein Chefarzt) wird zu Gastroenterologie und Kardiologie mit zwei Chefarzten).

Natürlich ist in einer solchen Struktur ein hohes Maß von **Koordination** notwendig. Diese Koordination innerhalb der Expertenorganisation wird durch Standardisierung, die Internalisierung von Standards während einer langen Ausbildungs- und Trainingsphase und eine inhaltliche (Mit)Kontrolle durch externe Peers bzw. Fachgesellschaften gewährleistet (McCulloch 2006). Dieser Aspekt betont die inhaltliche Nähe zum Begriff des **Professionalismus** (zur Definition s. Freidson 2001, auch Relman 2007, Tableau 20.1.). Das interne Management spielt dagegen nur eine sehr geringe Rolle: *“one salutes the man, not the the stripes”* (Mintzberg 1979). So stammen die relevanten Arbeitszeugnisse eines ärztlichen Mitarbeiters immer vom Chefarzt, also dem Abteilungsleiter, ein (seltenes) Zeugnis des Geschäftsführers der Klinik hat für die Karriere eines Arztes kaum Relevanz. Die professionellen Experten sehen sich in erster Linie ihren fachlichen Kollegen gegenüber verantwortlich, die Karrieren werden primär über die Professionen geplant und ermöglicht.

Tableau 18: Eigenschaften der Expertenorganisation (EO):

1. Die EO gehört zu den funktionalen Organisationen mit fachlich begründeter Arbeitsteilung.
2. Große Autonomie der Experten
3. Experten verfügen über eigene Kundenbeziehung und gestalten die Produkt-Markt-Beziehung.
4. Horizontales Wachstum um die Fachgebiete oder durch weitere Spezialisierung
5. Koordination durch Standardisierung und Training
6. Diagnose (*pigeonholing*) ist die zentrale Leistung und löst standardisierte Aufgabenbewältigung aus.
7. Unsicherheit wird in einem gewissen Maße toleriert.
8. Experten verfügen über horizontale Vernetzung und gleichzeitig vertikale Leitungsaufgaben.
9. Das zentrale Management ist schwach ausgeprägt.
9. Die Aufgaben des Managements sind Außenbezug, Konfliktlösung und Support.
10. Die wichtigste Schwäche der EO besteht (wie bei jeder funktionalen Organisation) in der Innovationsresistenz besonders bzgl. Prozess- und Strukturinnovation.
11. Weitere Schwächen sind die mangelnde Ausbildung eines Qualitäts- und Risikomanagements sowie die potentielle Überdehnung der Angebotsmacht (Verlust der Akzeptanz).

Eine Voraussetzung für die Wirksamkeit dieser Standardisierung ist die **funktionale “Diagnose”** im Vorfeld der Aufgabenerfüllung (*“pigeonholing”*, das Einsortieren in Schubladen (Mintzberg 1979)); es muss zuerst klar sein, worum es geht, bevor dann die Organisation entsprechend eines standardisierten Vorgehens aktiv wird.

Die Experten haben eine **umfangreiche horizontale Vernetzung** auf der Arbeitsebene (Kollegialität), nehmen aber gleichzeitig auch zahlreiche “vertikale” Leitungsaufgaben

wahr (wie z.B. die eigenständige Gestaltung der Kundenbeziehungen). Leitungsgremien werden kollegial besetzt, Managementfunktionen "mit den eigenen Leuten". Das nichtprofessionelle Management ist im Vergleich zu anderen Organisationen unbedeutend und beschränkt sich zunächst auf die Aufrechterhaltung der Supportfunktionen: die *power of expertise* (die fachliche Qualifikation der Experten) ist stärker als die *power of office* (die Managementkompetenz in einer "normalen" bürokratischen Hierarchie) (Mintzberg 1979). Allerdings gibt es in manchen Organisationen neben dem professionellen System klassische bürokratische Linien-Hierarchien anderer Berufsgruppen, die stark top-down mit klaren Führungsstrukturen organisiert sind (im Krankenhaus z.B. die Pflege, an der Universität nicht-wissenschaftliche Mitarbeiter usw.). Die Koexistenz der partizipativen Expertenstruktur und der klassischen Hierarchie kann zu Konflikten Anlass geben und macht unter Umständen eine Abbildung der unterschiedlichen Systeme auf der Leitungsebene notwendig (Chandler 1999).

Allerdings kann eine Expertenorganisation, entgegen dem ersten Eindruck, nicht vollständig auf Managementfunktionen verzichten. Neben den Supportfunktionen (z.B. IT-Ausstattung, Bibliotheken, Beschaffungs- und Abrechnungsfragen) müssen "**nicht-professionelle**" **Manager** z.B. im Konfliktfall schlichtend eingreifen, wenn die intraprofessionellen Mechanismen nicht ausreichen. Außerdem werden sie für die Gestaltung der Außenbeziehungen der Organisation benötigt, soweit sie die Gesamtinstitution betreffen (z.B. gegenüber politischen Gremien oder Geldgebern). Es wird dabei allerdings unbedingt erwartet, dass sie die Autonomie der Experten schützen und in der Gesamtbilanz genügend Ressourcen zur Verfügung stellen.

Weiter: 3. Motivation ..., 3.2. Professional Bureaucracy (2)

3.2. Professional Bureaucracy: die Expertenorganisation

3.2.2. Einschränkungen und Schwächen

Bei aller Standardisierung und *pigeonholing* der Klienten/Patienten wird **Unsicherheit** akzeptiert: die Lehrer stehen allein vor der Klasse, die Ärzte machen allein Bereitschaftsdienst. Diese Unsicherheit kann ein Maß annehmen, das in normalen Bürokratien nicht akzeptiert werden würde. Der Grad der Unsicherheit kann sich auch soweit erhöhen, dass die Gesamtorganisation bedroht wird, und zwar z.B. dann, wenn einzelne oder mehrere Experten ihre kollegialen Kontrolle durch die externen Professionen vernachlässigen und durch Qualitäts- bzw. Sicherheitsmängel auffallen (*failure of professionalism*, s. Marshall et al. 2010, s. Tableau 19). Eine externe Kontrolle von Qualität und Sicherheit ist sehr schwer möglich, eine dazu notwendige Ausweitung des nicht-professionellen Managements mit Ausbildung eines Qualitäts- und Risikomanagements wird von den Experten primär abgelehnt und kommt nur auf ausserordentlich großen externen Druck zustande.

Tableau 19: Schwächen der Expertenorganisation

- (1) *Failure of Professionalism*: professionelle Kontrolle funktioniert nicht
- (2) *Means-Ends-Inversion*: Mißbrauch der Angebotsmacht auf Expertenebene
- (3) Innovationsresistenz als Bürokratie

Eine zweite Gefahr besteht darin, dass Experten die Organisation nur noch als Plattform für den Verkauf ihrer Leistungen ansehen und ihre Angebotsmacht in einem Maße mißbrauchen ("alle Patienten brauchen diese Therapie"), dass die Akzeptanz in der Umgebung in Gefahr gerät (sog "**means-ends-inversion**", also die Verkehrung von Mittel und Zweck, s. Mintzberg 1979). Gerade bei Klienten, oder - um im Gesundheitsbereich zu bleiben - Patienten, die sich nicht wegen einer einzelnen Akuterkrankung, sondern einer chronischen Mehrfacherkrankung an die Organisation wenden, ist eine solche Haltung dysfunktional, wenn nämlich weniger der "Absatz" der Einzelleistung, sondern die Kooperation mehrerer Leistungen angezeigt ist.

Eine weitere Schwäche einer Expertenorganisation resultiert daraus, dass sie eine "**Bürokratie**" ist und bleibt. Sie hat zwar gegenüber die "Maschinenbürokratie" den Vorteil, dass sie sehr hochspezialisierte Leistungen "auf Expertenniveau" anbieten kann und auch sehr innovationsnahe ist, sie basiert aber wie diese auf einer funktionalen Arbeitsteilung. Diese Art der Arbeitsteilung führt "in ruhigen Zeiten" zu großen

Spezialisierungsvorteilen, ist aber immer dann überfordert, wenn sie sich in instabilen Umwelten bewähren muss, wenn also unvorgesehene Ereignisse eintreten. Innovationen auf der Fachebene können von Expertenorganisationen immer erwartet werden, Prozess- und Strukturinnovationen in Zeiten des Wandels sind jedoch nicht die Stärke dieser Organisationsform.

Tableau 20: Vergleich

Manager	Ärzte
Populationen	Individuelle Patienten
Prozesse	Ergebnisse
Gestaltete Strukturen	Autonomie
Lineare Problemlösung als Standard	<i>muddling through</i> im klinischen Alltag
Hochgradige <i>upwards</i> -Verantwortlichkeit	Lockere Kontrolle durch Profession
Primäre Loyalität: Organisation	Primäre Loyalität: Fach, Profession
Außenreferenz als Aufgabe (Marketing)	Informelle Selbstbestimmung bedrohlich
Politik als Ansprechpartner normal	Politik "uneinsichtig" bzgl. Fachlichkeit

Der **Konflikt im Gesundheitswesen zwischen den Experten, meist der Ärzteschaft, und dem Management** ist ein schon fast als klassisch zu beschreibendes Beispiel für die der Expertenbürokratie innewohnenden Widersprüche (Edwards 2003, Edwards 2005), s. Tableau 20). Während Manager ihre Aufmerksamkeit eher auf Patienten-Populationen statt auf individuelle Patienten richten, gestaltete Strukturen und lineare Problemlösungen vorziehen und bei primärer Organisationsloyalität zumindest partiell den Austausch mit instabilen Umwelten als ihre Aufgabe ansehen, sind Ärzte dem individuellen Patienten verpflichtet, achten auf ihre professionelle Autonomie und sind Außenbedingungen, die über das professionelle Umfeld hinausgehen, skeptisch gegenüber eingestellt (s. Kasten). Ärzte müssen täglich *trade offs* eingehen, sie sind verpflichtet, sich in unsicheren Umgebungen zurechtzufinden (z.B. überfüllte Notfallambulanz) (Marshall et al. 2010). Der Konflikt wird teilweise sehr scharf ausgetragen (Zitat: "*In the case of bacon and eggs, it has been said the chicken is involved but the pig is committed*" - Ärzte sind verpflichtet, Hilfe zu leisten (sind *committed*), das Management schafft die adäquate Umgebung und ist *involved* (Mintzberg 1997)). Aus Managementsicht ist die primäre Verbundenheit der Ärzte mit der Profession und erst sekundär mit der Organisation das Hauptproblem und führt zu Vorwürfen, wenn der Chefarzt am Ende der Sitzung sagt "zurück zur Arbeit", was ja auch tatsächlich eine Herabsetzung des Managements darstellt. Der Konflikt hat sich um die Position des Ärztlichen Direktors (*medical director*) zugespitzt, der regelmäßig

von den Kollegen als zu Management-lastig angesehen wurde, was aber letztendlich dazu geführt hat, dass die Ärzteschaft in den meisten Einrichtungen des Gesundheitswesens keine Person mehr mit Letztverantwortung positionieren konnte und alle verantwortlichen Positionen von Managern mit anderer Provenienz besetzt wurden (Mintzberg 1997).

Tableau 20.1: Definition Professionalismus
(s. Freidson 2001, Relmann 2007)

- Hochgradige Spezialisierung
- Spezifisches Wissen und Können
- Zertifizierung durch die Profession
- Exklusive Eigengerichtsbarkeit
- Geschützte Stellung auf dem Arbeitsmarkt
- Hohe Priorität von professionellen Werten

Verstärkend kommt die **Veränderungsresistenz** der Expertenorganisation hinzu. Von außen an die Organisation herangetragene Impulse wie solche durch veränderte Finanzierungsbedingungen oder politische Rahmenveränderungen werden von den Experten primär als Autonomie-einschränkend erlebt und nicht als zu lösendes und evtl. lösbares Problem. Obwohl zahlreiche Untersuchungen gezeigt haben, dass gerade professionelle Faktoren wie Vorbildfunktion, Einbeziehung der Entscheider und Feedback-Mechanismen positive Prädiktoren für die Veränderung der Organisation und das Bewältigen von Krisen darstellen (z.B. Greco und Eisenberg 1993, Bero et al. 1998, Grimshaw et al. 2004), ist die Chance auch von den Professionen und den Expertenorganisationen nicht umfassend genutzt worden, kohärente Organisationskonzepte haben sich daraus nicht entwickelt.

Letztendlich ist der **Professionalismus** (Definition s. Tableau 20.1.) in einer Sackgasse verfangen und taugt insofern nicht als abschließendes Konzept für ein modernes System. Gut lässt sich dies am Beispiel der *Evidence-Based Medicine* zeigen (vergleiche hierzu Timmermanns und Mauck 2005), die ja sehr auf die Wissenschaftlichkeit der Entscheidungsprozesse abhebt und sicher ein wichtiges Element einer modernen Medizin darstellt. Die Gegenposition, die auf der Erfahrungsgebundenheit des ärztlichen Berufes aufbaut (Handwerk/"Kunst"), sieht allerdings durch EBM die professionelle Autonomie bedroht, ein Konflikt, der bislang nicht gelöst wurde. Wenig wurde bislang beachtet, dass sich beide Positionen in einem Punkt einig sind, nämlich der (mono)professionellen Sichtweise. Aus diesem Grund können sie sich einerseits nicht verständigen und verhindern damit aber auf der anderen Seite gemeinsam multiprofessionelle und multidimensionale Implementierungs- und

Veränderungsstrategien (Timmermanns und Mauck 2005).

Weiter: 3. Motivation ..., 3.3. *"It's not a rocket science!"*

3.3. “It’s not a rocket science!”

3.3.1. System- und Komplexitätstheorie

Auf der organisatorischen Ebene allein lässt sich folglich ein umfassendes Rahmenkonzept nicht entwickeln. Diese Situation hat dazu geführt, dass in den 90er Jahren des letzten Jahrhunderts systemtheoretische Ansätze bemüht wurden, um das Verhalten von Individuen und Institutionen gerade im Hinblick auf Innovationen im Gesundheitswesen zu beschreiben (s. Abb. 6b). So nutzte das *Institute of Medicine* in seinen Reports “*To Err Is Human*” (Kohn et al. 1999) und “*Crossing the Quality Chasm*” (IOM 2001) einen explizit systemtheoretischen Ansatz und unterlegte damit das gesamte P4P- bzw. *Value-based Purchasing*-Programm für die Versorgung von Medicare-Patienten in den USA (s. Tableau 21). Der Schwerpunkt lag dabei auf dem **Begriff der Komplexität** (Plsek 2001C). Eine parallel erscheinende, kurze Artikelserie zur Komplexitätstheorie von P. Plsek und T. Greenhalgh im *Brit. Med. J.* im Jahr 2001 spielte ebenfalls eine große Rolle (Plsek und Greenhalgh 2001, Wilson und Holt 2001, Plsek und Wilson 2001, Fraser und Greenhalgh 2001). Es soll an dieser Stelle dahingestellt bleiben, ob die Akzeptanz dieses Konzeptes der Komplexität (*complexity*) als Systemeigenschaft nicht zumindestens teilweise durch seine sprachliche Nähe zum Begriff “Kompliziertheit” (*complicacy*) zu erklären ist (s.u. Tableau 23), verbunden mit der Entlastung der individuellen Verantwortung gegenüber einem undurchschaubaren, “komplexen” System, das im Alltag des Gesundheitswesens als undurchschaubar und unbeeinflussbar erlebt wird. Denn es gibt keinen Zweifel: die Komplexitätstheorie steht in heftigem Widerspruch zu den linearen Modellen, die gerade in der Medizin und Gesundheitsversorgung vorherrschen, gerade auch in der klinischen Versorgung und im Verständnis organisatorischer Prozesse.

Tableau 21: Quality as a system property (Institute of Medicine 2001, S. 4f)

“The committee is confident that Americans can have a health care system of the quality they need, want, and deserve. But we are also confident that this higher level of quality cannot be achieved by further stressing current Systems of care. The current care Systems cannot do the job. Trying harder will not work. Changing Systems of care will.

(...) Members of the health care workforce are already trying hard to do their jobs well. In fact, the courage, hard work, and commitment of doctors, nurses, and others in health care are today the only real means we have of stemming the flood of errors that are latent in our health care systems.

Health care has safety and quality problems because it relies on outmoded systems of work. Poor designs set the workforce up to fall, regardless of how hard they try. If we want safer, higher-quality care, we will need to have redesigned systems of care, including the use of information technology to support clinical and administrative processes.”

Daher sei hier die **Definition von Komplexität** und die Abgrenzung zum Begriff der Kompliziertheit vorangestellt (s.u. Tableau 23). Ein komplexes System besteht aus zahlreichen, in ihrer Zahl schwankenden, nicht-linear per multiplen Feedback miteinander verbundenen Teilen, die zu in Zeit und Stärke unvorhersehbare Ereignissen führen, einzelne, nicht-explicite und veränderbare interne Regeln kennen und zur Selbstorganisation, Adaptation an die Umwelt sowie zu Lernprozessen in der Lage sind (s. auch Richardson 2008). Entsprechend der systemtheoretischen Provenienz ist das System größer als die Summe der Einzelteile, wobei kleine Veränderungen sehr große Effekte aufweisen können ("Sensibilität gegenüber Anfangsfehlern", das Schlagen des berühmten Schmetterlingsflügels). Anders als einem linearem "Maschinenmodell", das durch Eindeutigkeit, Trend zum Reduktionismus, Vorhersehbarkeit und dem Versuch der Spannungsreduktion charakterisiert ist, sind einem komplexen System gerade Spannung, Angst, Unsicherheit und Paradoxien konstitutiv zu eigen. Es sind weiterhin sog. **Attraktoren**, Konstruktionen von Zwischen- und Endzuständen höherer Stabilität, vorhanden, zu denen das System sich hinorientiert, die aber von außen nicht sichtbar sind.

Tableau 22: Die alte Metapher vom Ferrari und dem Urwald ist hervorragend geeignet, den **Begriff der Komplexität** zu erklären. Ein Ferrari ist ein zweifelsfrei ein kompliziertes System, dessen konstituierende Regeln nicht sofort erkennbar sind. Mit entsprechendem Aufwand ist es jedoch trotzdem zu durchdringen. Anders beim Urwald: die Metapher zielt darauf ab, dass komplexe Systeme keine erlernbaren Regeln aufweisen, nach denen die Folgen (und die Wirksamkeit) bestimmter Interventionen sicher vorhersagbar sind. Man kann sie nur "auf Sicht", zeitnah und durch indirekte Erkenntnisse über ihr Funktionieren beeinflussen. Der größte Fehler ist es hier, mit starken, einfachen Interventionen einen durchschlagenden Effekt erreichen zu wollen. Dieser frustrane Versuch ist dennoch häufig im Gesundheitswesen zu beobachten

Komplizierte Systeme (Beispiel Ferrari, s. Tableau 22) sind dagegen zwar schwer zu verstehen, man kann jedoch trotzdem die Regeln erlernen, und vor allem hat es Sinn, nach den Regeln zu suchen; bei komplexen Systemen wird man sie dagegen nicht finden. Ein Computer ist zweifelsohne eine komplizierte Struktur, aber wer hat das Internet erfunden? - das Internet kann ebenso wie das Wetter als paradigmatisches Beispiel für komplexe Systeme gelten.

Tableau 22.1: Eigenschaften komplexer Systeme

- bestehen aus zahlreichen Teilen
- Zahl der Teile veränderlich
- Teile sind interdependent
- nicht-linear mit multiplem Feedback verbunden
- interne Regeln nicht-explizit
- Ereignisse in Zeit und Stärke nicht vorhersehbar
- Neigung zur Selbstorganisation
- Adaptation an Umwelt und Lernen möglich
- Sensibilität gegenüber Anfangsfehlern
- Akzeptanz von Paradoxon und Unsicherheit
- Orientierung an Zwischen- und Endzuständen
relativer Stabilität (Attraktoren)

Der geschilderte Ansatz hat natürlich enorme Auswirkungen auf Bereiche wie die klinische Versorgung von Patienten, Management-Konzepte und die (Klinische) Wissenschaft. Insbesondere die Behandlung chronischer Erkrankungen, z.B. die Versorgung eines Diabetes-Patienten, kann als komplexes System verstanden werden, bei dem derart viele Einflussfaktoren ineinandergreifen, dass lineare Modelle an ihre Grenzen kommen (Wilson und Holt 2001). Auch organisatorische Umgebungen wie z.B. ein Operationssaal sind wegen ihrer hochgradigen *Interrelatedness* (man könnte hier am ehesten von "Vernetzung" sprechen, s. Kannampalli et al. 2011) als komplexes System aufzufassen, ebenso aber auch als ein kompliziertes System, bei dem es bestimmte Faktoren zu spezifizieren gilt, die das System sicherer machen (Matern et al. 2006). Im Management wird eine komplizierte Situation mittels Experten (sie kennen die Regeln) gelöst, in einer komplexen Situation können Experten jedoch fehl am Platz sein, wenn die ins Spiel gebrachten Regeln beim Blick auf das komplexe System nur hinderlich sind (die sog. "todsicheren" Lösungen, s. Baker 2001).

Weiter: 3. Motivation ..., 3.3. "It's not a rocket science", 3.3.2. Doppelte Komplexität

3.3. "It's not a rocket science!"

3.3.2. Doppelte Komplexität

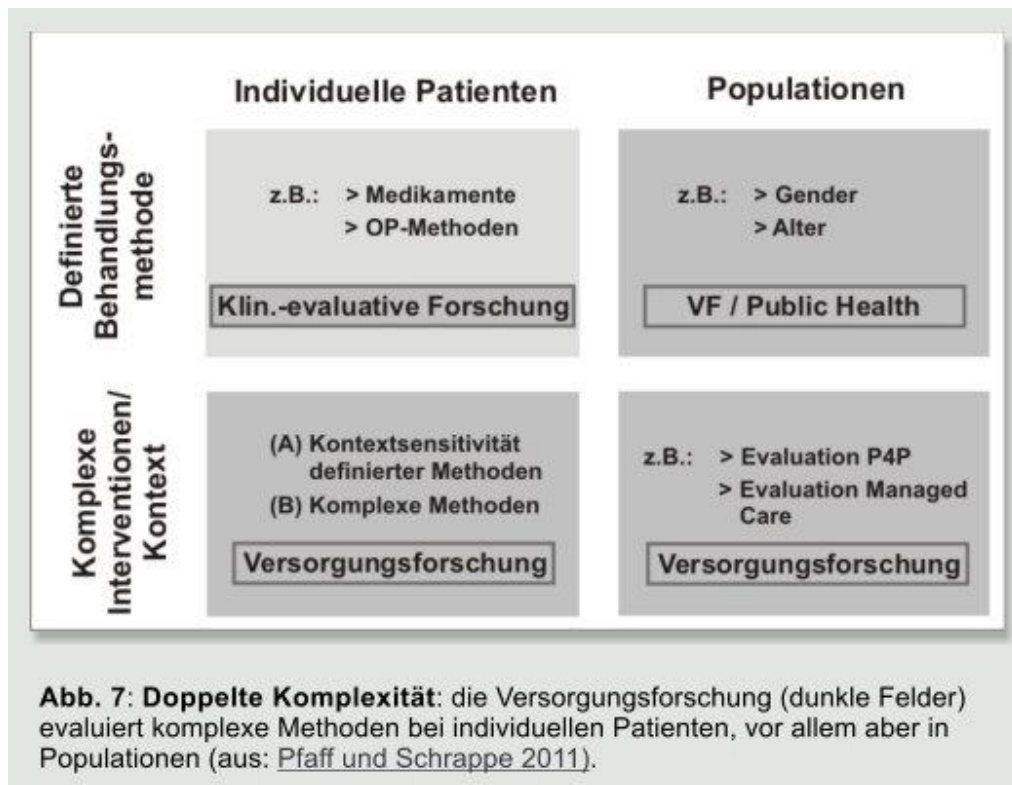
Heute hat die Betrachtung von Qualität und Sicherheit als Systemeigenschaften weite Verbreitung gefunden. So hat z.B. Marc Chassin (2013) die wunderbare Metapher "*It's not a rocket science*" geschaffen: eine Rakete ist sicherlich kompliziert, das Gesundheitswesen aber komplex. Ein Problem wie die Händedesinfektion sei nicht mit "einfachen Methoden" anzugehen (s.a. Erasmus 2010), für ausbleibende Verbesserungen gäbe es (a) sehr viele ursächliche und beteiligte Faktoren, (b) jede Ursache bedürfe einer eigenen Intervention, und (c) für jedes Krankenhaus bzw. Gesundheitseinrichtung gelten andere Faktoren (Chassin 2013). Um die Bedeutung der Komplexität zu unterlegen, zieht er den Vergleich mit der Nuklearindustrie heran: *"Imagine a protocol that is as essential to the safety of a nuclear power plant as hand hygiene is to preventing infections in hospitals— it is inconceivable that workers in the power plant would exhibit a compliance rate of only 40 percent."* Die Komplexität im Gesundheitswesen kann also noch nicht verstanden worden sein, und der derzeit sehr oft genutzte Rückgriff auf den Begriff der Organisationskultur ist - bei aller Bedeutung - wichtig, aber nicht hinreichend, denn außer der Einsicht, dass in einem Nuklearkraftwerk offensichtlich eine Organisationskultur etabliert werden konnte, die an eine Sicherheitskultur heranreicht, erklärt er nicht, warum im Gesundheitswesen bzw. Krankenhaus eine entsprechende "Kultur" bislang nicht geschaffen wurde (Chassin 2013).

Tableau 23: Vergleich

Kompliziert	Komplex
Ferrari	Urwald
Computer	Internet
Gallen-OP	Antibiotika-Resistenz
Dienstanweisung	Einführung Leitlinie
Kondensation	Wetter
Kreuzworträtsel	Schach

Hinzu kommt, dass nicht nur der Kontext, sondern auch die Interventionen, die in Ansatz gebracht werden, komplexer Natur sind. Die Situation wird also reichlich unübersichtlich: wir haben komplexe Interventionen (z.B. die Einführung von P4P, die ja nicht als Einzelintervention analog einer Medikamenteneinnahme zu betrachten ist) in einer komplexen Umgebung (z.B. einem Gesundheitssystem oder einem Krankenhaus) (Frolich et al. 2007). Diese "**doppelte Komplexität**" (zum Begriff vgl. Shojania 2013) bildet den Hintergrund für die notwendige und gewünschte wissenschaftliche Evaluation,

von der "eindeutige" Ergebnisse erwartet werden. Ohne hier *in extenso* auf die Aufgabenstellung und das methodische Konstrukt der Versorgungsforschung eingehen zu können (vgl. hierzu Pfaff und Schrappe 2011), soll nur festgestellt werden, dass die Auseinandersetzung um die Frage geht, inwieweit in nicht-linearen, hochgradig interdependenten Systemen Einzelfaktoren isoliert werden können, die getrennt zu beobachten und im klassischen Ansatz des randomisierten Versuchs zu untersuchen sind. Besonders gilt dies für den Fall, dass komplexe Interventionen in komplexen Umgebungen (z.B. in Populationen) untersucht werden (s. Abb. 7). In jedem Fall sind Konzeptbildung, Pilotierung und spätere Implementierung der Ergebnisse immanente Bestandteile der Untersuchungen (Avorn und Fisher 2010). Der Medical Research Council in Großbritannien hat hierzu in zwei Memoranden Stellung genommen (MRC 2000, 2008), sah sich jedoch harscher Kritik durch Vertreter der Komplexitätstheorie ausgesetzt, die auch diesen Ansätzen ein falsches (zu mechanistisches) Verständnis von Komplexität vorwarfen (z.B. Cohn et al. 2013, zusammenfassende Darstellung der Kontroverse s. Mühlhauser et al. 2011). Die Auseinandersetzung hat Parallelen zu derjenigen um die Bedeutung des randomisierten Versuchs bei Interventionen zur Patientensicherheit (Leape et al. 2002, Shojania et al. 2002). Ganz grundsätzlicher Art ist die Auseinandersetzung um das richtige Verständnis der Komplexitätstheorie, die sich um die o.g. Serie im BMJ (Plsek und Greenhalgh 2001, Wilson und Holt 2001, Plsek und Wilson 2001, Fraser und Greenhalgh 2001) ergab und in der den Autoren insbesondere eine Psychologisierung des Attraktoren-Konzeptes vorgeworfen wird (Paley 2010, Greenhalgh et al. 2010, Paley 2011).



Weiter: 3. Motivation ..., 3.4. Konzeptioneller Rahmen

3.4. Konzeptioneller Rahmen: komplexe professionelle Systembürokratie

Expertenorganisation und komplexes System, so wie dargestellt in den beiden vorangegangenen Kapiteln, reichen aus mehreren Gründen für sich allein nicht aus, um den konzeptionellen Rahmen - wie im vorliegenden Artikel angestrebt - für Einsatz und Evaluation eines Instrumentes wie P4P zu bilden. Das Konzept der Expertenorganisation beschreibt, ausgehend vom Konflikt zwischen dezentral organisierten professionellen Experten und zentralem Management, die Steuerungs- und Anpassungsproblematik von Einrichtungen im Gesundheitswesen hervorragend, ist aber auf Organisationen beschränkt und für sich genommen kaum umfassend auf das ganze Gesundheitssystem anwendbar, insbesondere nicht auf die Gestaltung der Finanzierung. Die Komplexitätstheorie ergibt andererseits ein gutes Modell für die Steuerungsproblematik auf Systemebene und natürlich auch in Organisationen, bleibt auf der organisatorischen Ebene aber relativ undifferenziert und könnte auch zu Entschuldigungstendenzen führen ("ist ja doch nicht zu machen"). Gesundheitsversorgung und Prävention werden nun tatsächlich aber lokal und durch Organisationen erbracht, daher bedarf es eines spezifischeren Konzeptrahmens, der die Gestaltung der Versorgung sowohl in der Peripherie als auch auf Systemebene beschreibt. Hier scheint die **Synthese der Konzepte Expertenorganisation (professional bureaucracy) und komplexes System** einen guten Ansatz zu bilden.

Tableau 24: Gemeinsamkeiten von Expertenorganisation und komplexen Systemen

- Verdeckte interne Regeln
- Teile/Experten verfügen über große Autonomie
- Neigung zur Selbstorganisation
- Innovationsparadox: hochinnovativ, Innovation aber nicht vorhersehbar
- Unsicherheit wird akzeptiert ("intrinsische Unsicherheit")

Analogien beider Konzepte (s. Tableau 24) sind bereits auf den ersten Blick erkennbar: die verborgenen Regeln der dezentral agierenden Experten, die geringe Steuerungsfähigkeit, die Schwierigkeit, von außen einzuwirken. Die weitergehende Analyse erbringt dann an mehreren Punkten systematische Parallelen, die die Verwendung des kombinierten Konzeptes sinnvoll erscheinen lassen:

1) Beide Konzepte betonen die **Unsichtbarkeit der (gleichwohl vorhandenen) internen Regeln**. In der Expertenorganisation sind diese in der dezentralen Vernetzung der Fachexperten verborgen, im komplexen System sind sie zwischen allen Teilen des Systems als veränderbare Strukturen vorhanden.

2) Die Teile des Systems und die Experten (im operativen Kern der Expertenorganisation) verfügen über ein **hohes Maß an Autonomie**, gleichzeitig sind sie lern- und anpassungsfähig. Trotz der Autonomie sind die autonomen Teile miteinander verbunden, sie neigen zur **Selbstorganisation** (in der Expertenorganisation auf fachlichen Ebene in Abgrenzung zur Managementebene). Da die Regeln nicht bekannt sind, sind unvorhergesehene und in ihrer Ausprägung stark differierende Reaktionen der Organisation bzw. des Systems die Regel.

3) Beide Konzepte können hochinnovativ sein, jedoch sind die Innovationen in Art, Menge und Zeitpunkt nicht vorhersehbar und können von außen nicht mit voraussagbarem Erfolg angestoßen werden. Der Grund für dieses **“Innovations-Paradoxon”** liegt darin, dass einerseits einengende und innovationsfeindliche Strukturen fehlen, andererseits die zur Innovation erforderliche Kooperation nur spontan erfolgen kann und daher nicht zu antizipieren ist. Jegliche Reduktion der Autonomie (z.B. im Versuch einer Komplexitätsreduktion mit dem Versuch der besseren Steuerbarkeit) als innovationsfeindlich interpretiert. Die Innovationsfähigkeit der Expertenorganisation bezieht sich primär auf Produktinnovationen, da Prozess- oder Strukturinnovationen ohne die nicht-professionellen Kräfte erschwert sind.

4) Beide Konzepte akzeptieren **Unsicherheit** - dieser Punkt erscheint als der Wichtigste. Die Diagnose- und Standardisierungs-gestützte Koordination in der Expertenorganisation ist immer nur begrenzt in der Lage, den Experten in seiner Unabhängigkeit zu leiten, er wird immer improvisieren müssen (s. das Beispiel “Lehrer vor der Klasse”). Die Komplexitätstheorie sieht Unsicherheit, Spannung und Paradoxie (im Gegensatz zu linearen Modellen) sogar als konstituierendes Bestandteil von Systemen an und erklärt eine Reduktion dieser Erscheinungen als aussichtslos oder nur in Ansätzen erreichbar. Die Medaille “Unsicherheit”, so könnte man sagen, hat hier in besonderem Maße zwei Seiten: einerseits ist die Toleranz gegenüber Unsicherheit sicherlich eine adäquate Eigenschaft, denn sie schützt vor irreführenden linearen Konzepten. Andererseits sind Unsicherheiten natürlich soweit nicht erwünscht, als dass sie mit Qualitäts- und Sicherheitsmängeln vergesellschaftet sind. Unter dem Begriff der “intrinsic Unsicherheit” wird weiter unten darauf näher eingegangen.

Natürlich haben beide Konzepte **auch maßgebliche Unterschiede**, in der ersten Linie durch ihren Horizont. Die *professional bureaucracy* ist ein Konzept für die Gestaltung einer Organisation, die Komplexitätstheorie bildet durch ihre Nähe zur Systemtheorie einen sehr viel weiteren Horizont ab. Außerdem ist die Expertenorganisation ein bürokratisches System, basiert somit auf der funktionellen (verrichtungsorientierten) Arbeitsteilung und hat mit der sog. Maschinenbürokratie deshalb ihre

Inflexibilitätsproblematik und Resistenz gegenüber Prozess- und Strukturinnovationen gemeinsam. In der System-/Komplexitätstheorie ist dagegen eine verrichtungsorientierte Arbeitsteilung nur *ein* möglicher Mechanismus der Aufgabenerledigung unter vielen, und lange nicht der typische; die Flexibilität ist hoch. Komplexe Systeme haben auch eine höhere Integrationsfähigkeit als Expertenorganisationen, letztere haben (fast) ausschließlich die Fähigkeit zur Integration auf fachlicher Ebene.

Tableau 25: Definition "Komplexe professionelle Systembürokratie"

Aus der Synthese der Konzepte Expertenorganisation (*professional bureaucracy*) und Komplexitätstheorie entwickelter Arbeitsbegriff, der auf gemeinsamen Eigenschaften beider Konzepte wie Autonomie, Tendenz zur Selbstorganisation, Toleranz von Unsicherheit ("intrinsische Unsicherheit") und "Innovationparadoxon" (Innovationen nicht planbar bei großer Innovationsnähe) beruht und zur Beschreibung sowohl der organisatorischen als auch der Systemebene verwendet wird.

Trotz dieser Unterschiede ist die Verbindung der Konzepte Expertenorganisation auf der einen Seite und Komplexitätstheorie auf der anderen Seite in der Anwendung auf das Gesundheitswesen sinnvoll. Beide Konzepte ergänzen sich in ihrem Organisations- bzw. Systemhorizont, sie haben sich überschneidende zentrale Eigenschaften (s.o.) und sie haben phänomenologische Übereinstimmungen, wenn man z.B. an die Parallelitäten zwischen den Spannungsfeldern Experte vs. Management und Standesorganisation/Selbstverwaltung vs. Politik denkt (zur Einbeziehung der politischen Ebene s. Kap. 6). Ohne dass der Begriff m.E. bislang anderenorts kodifiziert wurde, wird hier deshalb als Arbeitsbegriff für diese Arbeit der Terminus "**komplexe professionelle Systembürokratie**" vorgeschlagen (zur Definition s. Tableau 25). Der Begriffsbestandteil "Systembürokratie" gibt darin wieder, dass es sich einerseits um ein (komplexes) System handelt, andererseits aber um ein verrichtungsorientiert strukturiertes System mit hochgradiger und strukturelevanter Arbeitsteilung.

Weiter: 3. Motivation ..., 3.5. Verhaltensänderung, 3.5.1. Einführung

3.5. Verhaltensänderung in der “komplexen professionellen Systembürokratie”

3.5.1. Einführung, lerntheoretische Modelle

Bereits in der ersten Annäherung wird klar, dass auf dem Hintergrund der hier so bezeichneten “komplexen professionellen Systembürokratie” i.S. einer Synthese von Expertenorganisation und komplexem System externe Anforderungen an Qualität und Sicherheit nur schwer umsetzbar sind. In der Expertenorganisation werden diese als “*management-behavior*” abgelehnt und an die (schwache) Managementebene delegiert. Im komplexen System ist die Wahrscheinlichkeit hoch, dass solche Anforderungen “verschluckt” werden, dass sie also nur selten zu den erwarteten Auswirkungen führen. Beide Systeme können zwar einerseits mit Spannungen, Paradoxien und Unsicherheit umgehen, lassen andererseits solche aber auch zu, sie neigen zur “intrinsischen Unsicherheit” (s.o.). In dieser Situation haben Verhalten und Verhaltensänderung von Individuen und Gruppen eine besondere Bedeutung, denn es wird offensichtlich von beiden die Übernahme erheblicher Verantwortung erwartet. Dies gilt natürlich auch im Hinblick auf die Einführung bzw. Umsetzung einer Intervention wie P4P. Daher werden im nächsten Schritt die wichtigsten **Theorien zur Verhaltensänderung** dargestellt und als Hintergrund eines konzeptionellen Rahmens diskutiert.

Tableau 26: Bedürfnistheorie nach Maslow (zit. n. Staehle 1999, S. 170)

▶ Defizit-Motive

- Physiologische Bedürfnisse (z.B. Hunger)
- Sicherheitsbedürfnisse (z.B. finanzielles Auskommen)
- Soziale Bedürfnisse

▶ Wachstums-Motive

- Wertschätzung
- Selbstverwirklichung

Zur Systematik werden vier Modelle der Verhaltensänderung unterschieden (s. z.B. Anonymous 1999, Grol und Grimshaw 2003, Shojania und Grimshaw 2005):

- Lerntheoretische Modelle
- Soziale Wahrnehmung
- Organisatorischer Wandel
- Lernen durch Kontext.

▶ Die **lerntheoretischen Modelle** umfassen in der ersten Linie Motivation, Feedback und Belohnung. Motivation als Voraussetzung zielorientierten Handelns wird nach der

Bedürfnistheorie von Maslow (zit. n. Staehle 1999, S. 170, 218ff) in die sog. Defizit-Motive und die Wachstums-Motive unterteilt (s. Kasten), weiterhin differenziert man interne Motivation (z.B. Einstellungen, s. unten) und externe Motivation (z.B. Belohnungssysteme). Voraussetzung für eine internen Motivation ist die gleiche Ausrichtung von Handlungsziel und Handeln. Externe Motivation kann die interne Motivation schwächen (sog. Untergrabungswirkung; Berenson et al. 2013), dies ist aber nicht zwangsläufig so; eine externe Motivation kann durch Übung dazu führen, dass die interne Motivation überhaupt erst aktiviert wird.

<u>Präventive Maßnahmen</u>		
	Wirksamkeit	Negative Folgen
Akut therapeutische Maßnahmen	SOFORT	SPÄT
Präventive Maßnahmen	SPÄT	SOFORT

Abb. 8: Feedback am Beispiel der Prävention im Vergleich zu akutmedizinischen Maßnahmen. Die negativen Folgen als Teil des Aufwandes erscheinen sofort, der Erfolg erst in der Zukunft (n. Cook et al. 2004)

Der Feedback-Mechanismus, und bei P4P handelt es um einen solchen, ist ein zentraler Mechanismus des Lernens, seine Wirksamkeit ist jedoch von einer Reihe methodischer Voraussetzungen abhängig. So muss ein Feedback an den verantwortlichen Handelnden adressiert sein, es muss als solches interpretiert werden, es muss auf einem erreichbaren Auslöser beruhen, es darf nicht mit anderen Feedback-Systemen kollidieren, es muss zeitnah sein und der Aufwand zur Erlangung des Feedbacks darf nicht als zu hoch empfunden werden.

Am Beispiel der Prävention (und hierzu gehören ja auch Maßnahmen zur Verbesserung von Qualität und Sicherheit) lässt sich im Vergleich zu akutmedizinischen Maßnahmen (z.B. neue Operationsmethode, neues Gerät) die Bedeutung des Zeitfaktors gut verdeutlichen (Cook et al. 2004). Der Aufwand für die Präventions- bzw. akutmedizinische Maßnahme wird bei Letzterer zeitnah gemindert durch die eintretende Wirksamkeit ("Ertrag"), negative unerwünschte Folgen treten dagegen erst spät ein, können oft nicht mehr zugerechnet werden und sind u.U. sogar nur statistisch nachweisbar. Bei den präventiven Maßnahmen (z.B. perioperative Antibiotikaphylaxe) sind die evtl. auftretenden negativen Folgen (z.B. Allergie) sofort

spürbar und vergrößern den Aufwand, während die Wirksamkeit (Verhinderung von postoperativen Wundinfektionen) erst bei hoher Patientenzahl und in Zukunft auftritt, oft wegen der statistischen Effekte für den Handelnden gar nicht erfahrbar. Finanzielle Anreize wie P4P können diesen strukturellen Nachteil der Präventionsmaßnahmen eventuell ausgleichen.

Tableau 27: Umsetzung der variablen, leistungsbezogenen Vergütung bei Führungskräften (vgl. Lebrez 2013)

- ▶ Übergeordnete, allgemeine Faktoren
 - Sind die Ziele planbar und für den Mitarbeiter/die Gruppe umsetzbar?
 - Besteht ein Zusammenhang zwischen Leistung und Zielerreichung?
 - Sind die Ziele zuverlässig quantifizierbar, kein *gaming*?
 - Ist die Motivation und das Ziel adäquat?
 - Werden andere Motivatoren negativ beeinflusst?
- ▶ Mitarbeiter
 - Hat der Mitarbeiter genügend Erfahrung und Freiräume, um ein Ziel durch eigene Leistung erreichen zu können?
 - Ist eine Mehr an Selbstbestimmung vom Mitarbeiter gewollt?
 - Hat der Mitarbeiter ausreichend Kompetenzen und Ressourcen, um die Ziele zu erreichen?
 - Ist der finanzielle Anreiz für Motivationsänderung hoch genug?
- ▶ Organisation
 - Existiert eine offene Gesprächskultur, die Zielvereinbarung zulässt?
 - Ist der Mitarbeiter fähig zu einem Zielvereinbarungsgespräch?
 - Steht die Leistungssteigerung in einem vernünftigen Verhältnis zum Aufwand der Vergütung (Verhandlungen, finanzielle Mittel etc.)?

Belohnungssysteme verstärken den Feedback (so z.B. bei P4P) und können soziale (z.B. Anerkennung) sowie materielle Gratifikationen beinhalten. Die externe Belohnung muss sowohl hinsichtlich des primären Aufwandes für den Handelnden (z.B. für Verhaltensänderung, Bereitstellung der Daten) als auch gegenüber der internen Rückkopplung (z.B. Wertschätzung von Kollegen, Patienten) in einem akzeptablen Verhältnis stehen. Wenn der primäre Aufwand zu hoch ist, oder wenn durch die Verhaltensänderung zwar die externe Belohnung erlangt wird, die interne Anerkennung aber vermindert wird, ist das Belohnungssystem nicht wirksam. Veränderungen im sozialen Kontext sind notwendig (s.u).

Feedback-Systeme (mit oder ohne Belohnung) haben den Nachteil, dass sie sehr häufig auf "niedrigen" Ebenen der Maslow'schen Bedürfnispyramide ansetzen und als externe Motivatoren mit der internen Motivation im Widerspruch stehen können (einen Freundschaftsdienst soll man nicht versuchen zu bezahlen). Sie wirken daher besser, wenn zusätzliche Interventionen auf der Ebene der Wertschätzung und zur Steigerung der Selbstverwirklichung eingesetzt werden.

Im geschäftlichen Kontext gibt es mit Belohnungssystemen reichhaltige Erfahrungen durch die variable leistungsabhängige Vergütung (s. Kasten), die allerdings nicht direkt auf P4P übertragbar sind, weil sie sich auf den innerinstitutionellen Rahmen beziehen. Man erkennt jedoch hier bereits, dass es neben allgemeinen Faktoren (z.B. bezieht sich die individuelle Leistung auf das Ziel?) vor allem Eigenschaften des Mitarbeiters (z.B. ist er überhaupt in der Lage, das Ziel zu erreichen?) und der Organisation (z.B. Gesprächskultur) sind, die bedacht werden müssen, bevor man das Instrument der Zielvereinbarungen einsetzt (Lebrenz 2013, Prendergast 1999).

Weiter: 3. Motivation ..., 3.5. Verhaltensänderung, 3.5.2. Soziale Wahrnehmung

3.5.2. Soziale Wahrnehmung

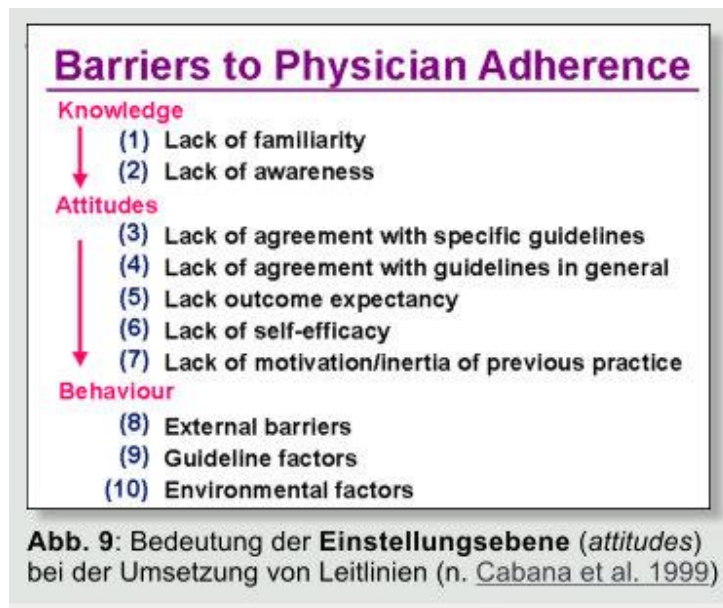
► Die Theorien der **Sozialen Wahrnehmung** stellen Einstellungen (*beliefs*), Haltungen (*attitudes*) und Absichten (*intentions*) in den Vordergrund: nicht das bewusste System, sondern das autonome System ("Bauchgefühl") steuern das Verhalten (King et al. 2013). Es gibt enge Bezüge zu den *behavioural economics* und der Kritik der bewußten, auf der Basis optimaler Information vorgenommenen Nutzenmaximierung. Wissen (*knowledge*) und Können (*skills*) sind von notwendiger Bedeutung, werden aber durch die Einstellungsebene dominiert, die ihrerseits kein individuelles, sondern ein soziales Konstrukt darstellt (s. Tableau 28). Im Ergebnis definiert sich das Individuum in seiner sozialen Rolle und entwickelt auch seine Veränderungsbereitschaft im Rahmen dieser Rolle. Rollen existieren auch im Innovationsprozess, man kann auf diese Weise Innovatoren (Erfinder), *early adopters* (frühe Anwendung), *early majority* (beginnende Umsetzung in der Breite), *late majority* (Umsetzung in der Breite) und *laggard*s (Nachzügler) unterscheiden.

Tableau 28: "If hand hygiene were a new drug it would be used by all" (Stone 2001), knapper kann man das Dilemma nicht fassen. In Deutschland könnten durch adäquate Händedesinfektion mehrere 1000 Todesfälle jährlich verhindert werden. Würden diese 1000 Todesfälle auf eine behandelbare Krankheit zurückgehen, die mit einer (nebenwirkungsarmen) Medikation zu behandeln wäre, wäre es kaum denkbar, dass Ärzte und Pflegepersonal angeben, sie hätten keine Zeit, dieses Medikament zu verabreichen (zu den Gründen der Non-Compliance s. z.B. Grol und Grimshaw 2003).

Um Verhaltensänderung zu erreichen, beschränkt man sich nicht mehr auf die Vermittlung von Wissen und auf Feedback-Mechanismen, sondern man versucht gezielt über Veränderungen der sozialen Rollen die Einstellungsebene zu beeinflussen. So wird aus *education*, der Wissensvermittlung, die Identifikation von Meinungsführern (*opinion leader*), die dann gezielt beeinflusst werden, so dass sie aufgrund ihrer Vorbildfunktion (Pittet et al. 2004) das Verhalten der breiten Masse im gewünschten Sinne beeinflussen (sog. *targeted education*). Ähnlich geht das *academic detailing* vor, bei dem Ärzten bestimmte Verhaltensweisen direkt "auseinandergesetzt" werden, wobei der wissenschaftlichen Ebene große Bedeutung zukommt, oder auch der *local consensus process*, bei denen die lokalen Akteure als Gruppe hinsichtlich der Veränderung miteinander in Bezug gesetzt werden (Bero et al. 1998).

Dem Konzept und den Methoden dieser Richtung wurde bereits früh große Bedeutung

und auch eine gewisse Akzeptanz zuteil, so z.B. in der Diskussion um die Implementierung von Leitlinien (Cabana et al. 1999) und der Umsetzung EBM-basierter Behandlungsmethoden (Greco und Eisenberg 1993, Bero et al. 1998, Grol und Grimshaw 2003, Bloom 2005). Neben den Methoden zur Förderung der Veränderung wurden auch die entgegenstehenden Barrieren genauer spezifiziert, die die Umsetzung in das entsprechende Verhalten behindern können, insbesondere wenn Wissen/Können und die Einstellungsebene nicht kongruent sind (Cabana et al. 1999).



Inhaltlich orientieren sich die traditionellen Rollenvorstellungen stark an den drei Begriffen interne Motivation (Cassel und Jain 2012), Professionalismus (mit dem Kernbereich Autonomie) und Altruismus (Patientenorientierung, Gesundheitsversorgung darf nicht "ökonomisiert" werden) (Frolich et al. 2007). Die Rollenzuschreibung des Teamplayers und der aktiven Organisationsveränderung (sowie des politischen Umfeldes) stehen nicht im Vordergrund, allerdings zeigen sich zwischen den verschiedenen Gesundheitsberufen deutliche Unterschiede. Man muss allerdings fast zwingend davon ausgehen, dass dieses Rollenverständnis in Zukunft unter einen erheblichen Veränderungsdruck gerät, der von einer mehr auf Koordination, Teamarbeit und Interprofessionalität ausgerichteten Gesundheitsversorgung ausgehen, weil die Anforderungen aus der Alterung der Gesellschaft und der Zunahme multipler, chronischer Erkrankungen anders nicht zu bewältigen sind (Ricketts und Fraher 2013). Unweigerlich wird sich daraus auch eine verstärkte Diskussion von "Management-Themen" wie Führung ("*medical leadership*") und Vorbildfunktion ergeben (Schrappe 2009), denn beide sind unabhängige Prädiktoren wichtiger Maßnahmen in der Gesundheitsversorgung (z.B. Händedesinfektion (Pittet et al. 2004)). Wenn es gut läuft,

könnte man sagen, wird zu den Rollen-Kompetenzen, die in diesem Zusammenhang vermittelt werden, auch das Umgehen mit komplexen Situationen gehören, die ja einerseits den Alltag bestimmen, dem sich die Gesundheitsberufe schon immer gegenübersehen, die andererseits ein Thema darstellen, das durchaus auch zum Motor für die Entwicklung neuer Rollenbildern taugen kann (Fraser und Greehalgh 2001).

Weiter: 3. Motivation ..., 3.5. Verhaltensänderung, 3.5.3. Organisatorischer Wandel

3.5.3. Organisatorischer Wandel

► Theorien, die auf dem **Organisatorischen Wandel** als entscheidendem Merkmal von Veränderung aufbauen, sind in der auf das Gesundheitswesen Bezug nehmenden Literatur weitaus weniger häufig anzutreffen als die lerntheoretischen oder Rollenbezogenen Zugänge. Möglicherweise steht dies mit dem (mono-)professionellen Zugang in Zusammenhang, der den meisten Betrachtungen zugrundeliegt, und der geringen Bedeutung, der in Expertenorganisationen organisatorischen Zugängen eingeräumt wird. Häufig werden nur einzelne Merkmale (z.B. Zeitmangel, juristische Risiken, finanzielle Zwänge, s. z.B. Grol und Grimshaw 2003) oder einzelne Mechanismen genannt (z.B. Computer-gestützte *reminder* (Shojania et al. 2010)), um organisatorische Bezüge bei der Verhaltensänderung herzustellen. Nur selten nähert man sich grundlegend den Konzepten organisatorischen Wandels an und diskutiert dessen Bedeutung im Hinblick auf die Außenwelt (z.B. Kommunikation und Anpassungsprozesse) bzw. bzgl. der internen Strukturen (z.B. die Regeln der Ressourcenverteilung, Organisationskultur etc.) (Frolich et al. 2007). Dieser Mangel steht in krassem Widerspruch zur betriebswirtschaftlichen oder organisationspsychologischen Literatur.

Veränderungen der IT-Ausstattung und die daraus resultierenden Instrumente wie die bereits genannten *reminder* oder *computer-assisted physician order entry*-Systeme (CPOE's) (Metzger et al. 2010) bilden vielleicht die einzige Ausnahme, denn wie jedem bewusst ist, der in der Organisation von Gesundheitseinrichtungen Erfahrungen sammeln konnte, stellt die IT einen der massivsten Eingriffe in das organisatorische Geschehen dar. Trotzdem zeigt gerade dieses Beispiel, dass ein einfaches "*If you completely rebuild it, they will come*" als Metapher für ein *systems reengineering* kaum ausreicht, wenn nicht grundsätzliche Fragen des organisatorischen Wandels mitgedacht werden (Shojania und Grimshaw 2005).

Wie wichtig allerdings eine rege Diskussion der Veränderung der Organisationen wäre, kann man an den nur verhalten geführten Diskussionen zu den folgenden Themen erkennen (die oben angeschnittenen Themenbereiche bzgl. Expertenorganisation und Komplexität werden hier ausgespart):

- Zentralität vs. Dezentralität: welche Aufgaben müssen zentral, welche können dezentral am besten erledigt werden (vgl. Schrappe 2007),
- wie weit sind die Organisationen des Gesundheitswesens von der verrichtungsorientierten Arbeitsteilung (Bürokratie) über die divisionale Organisation (Spartensystem) zur Matrixorganisation vorangeschritten bzw. welche Formen der

Aufbauorganisation und Aufgabenbewältigung erscheinen sinnvoll und angebracht,

- welcher besonderen Gestaltung bedarf das Verhältnis von erster (Geschäftsführung) und zweiter (Abteilungsleitungen) Ebene im Gesundheitswesen (s.

Expertenorganisation), einschließlich der Frage nach angemessenen Forderungen an Führung im Gesundheitswesen (Schrappe 2009), und schließlich der

- Frage, welches Grundverständnis von Veränderung im Gesundheitswesen zugrundegelegt werden sollte: Kontingenz (kontinuierliche Veränderung als Reaktion auf veränderte Umweltbedingungen) oder Konsistenz (diskontinuierlicher Übergang zwischen definierten organisatorischen Grundzuständen). Gerade dieser Punkt ist implizit in der zunehmenden Bedeutung von sog. *multifaceted* oder *bundle-interventions* enthalten (Bero et al. 1998, Robb et al. 2010), die ja gerade auch im Zusammenhang mit der Patientensicherheit eine große Rolle spielen (z.B. Jain et al. 2006).

Über diese Fragen der organisatorischen Strukturgestaltung geht jedoch das Problem hinaus, inwiefern eine Verhaltensänderung durch ein "Lernen der Organisation" unterstützt oder bewirkt werden kann. Im Rahmen der *Human Resources* Theorien gibt es dazu Vorstellungen, dass durch ein sog. organisatorisches Lernen nicht nur das Individuum verändert wird, sondern auch der Vorrat an Werten, Normen und Wissen, den die Organisation als solche vorhält (s. Staehle 1999, S. 914ff). Die Mitglieder der Organisation werden in ihrem Veränderungsprozess gefördert (oder gehemmt), wenn durch organisatorisches Lernen der Bedeutungszusammenhang verändert wird, in dem der individuelle Lernschritt (z.B. die Akzeptanz einer Patientensicherheits-relevanten Maßnahme) steht. Man spricht auch von "Deutero-Lernen", bei dem ein Rückkopplungs-*loop* auf das Individuum und der andere *loop* auf die Organisation zurückwirkt.

Weiter: 3. Motivation ..., 3.5. Verhaltensänderung, 3.5.4. Kontext, Zusammenfassung

3.5.4. Lernen im Kontext, Zusammenfassung

► Verhaltensänderung durch **Lernen im Kontext bezieht** alle drei vorgenannten Vorgehensweisen mit ein und basiert mehr noch als diese auf einer vorweg ausgearbeiteten, tragfähigen Umsetzungsstrategie (Anonymous 1999). Rollenveränderungen der Berufsgruppen (Grol und Grimshaw 2003, Timmermanns und Mauck 2005), Einbeziehung der Patienten (Bero et al. 1998, Hibbard et al. 2012) und Kampagnen der Massenmedien (Grol und Grimshaw 2003) gehören ebenso dazu wie die Evaluation der erreichten Wirksamkeit und deren öffentlichen Kommunikation. Als Beispiel können in Deutschland die Kampagnen "Gib AIDS keine Chance" der Bundeszentrale für gesundheitliche Aufklärung und die Aktion Saubere Hände (Abb. 10) des Robert-Koch-Institutes, der Gesellschaft für Qualitätsmanagement in der Gesundheitsversorgung e.V. (GQMG) und des Aktionsbündnis Patientensicherheit e.V. (APS) dienen. Diese Kampagnen weisen eine Nähe zum *social marketing* auf und müssen in mehreren Phasen geplant werden (z.B. Konzeption und Planung, Identifikation von Kommunikationskanälen, Entwicklung der Informationskanäle, Implementierung der Intervention, Evaluation, Feedback, vgl. Cook et al. 2004). Bezüglich P4P können sicherlich das VBP-Programm des CMS und das QOL-Programm des NHS als solche umfassenden Konzepte angesprochen werden.



Diese Konzepte sind selbstredend am mächtigsten, wären aber bei institutionellen

Veränderungen natürlich überdimensioniert. Für Instrumente wie P4P, die auf Systemebene eingesetzt werden, sind sie aber von großer Bedeutung. Wie im Kapitel über die Verantwortung der Politik (s. Kap. 6) weiter auszuführen sein wird, sind diese Konzepte nur dann erfolgreich, wenn alle Player hinsichtlich der entscheidenden Merkmale einer solchen Intervention kongruente Meinungen vertreten, ein nur selten zu erreichender Zustand.

Zusammenfassend kann hinsichtlich der Frage, welche Rolle den unterschiedlichen Konzepten der Verhaltensänderung in einer komplexen professionellen Systembürokratie zukommt, folgendes festgehalten werden:

- **lerntheoretische Ansätze**, insbesondere Feedback-Verfahren, sind wichtig, greifen aber zu kurz, wenn sie allein angewandt werden. Der Grund liegt in der großen Bedeutung von Einstellungen, professionellen Loyalitäten und sozialen Rollen, ganz abgesehen von ökonomischen und politischen Faktoren (s.u.). Es hat also keinen Zweck, das Feedback- und Belohnungsverfahren P4P als "Insellösung" einzusetzen, ohne dass die Einstellungs- und Rollenebene mit verändert wird.
- Konzepte der **sozialen Wahrnehmung** reichen insofern näher an den Kern des Problems heran, weil sie zusätzlich zu Feedbackverfahren auch die Einstellungsebene und das Rollenverständnis thematisieren. Insbesondere erscheint dies bezüglich des Themas Professionalismus wichtig zu sein, denn die traditionellen Rollenbilder werden zunehmend unter Druck geraten und müssen sich an veränderte Bedingungen und veränderte Aufgaben innerhalb des Gesundheitssystems anpassen.
- Verhaltensänderung durch **organisatorischen Wandel** setzt auf einen unabhängigen Wissens- und Werte"vorrat" der Organisation, der einer Veränderung durch Lernprozesse zugänglich ist. Auf den ersten Blick scheinen diese Ansätze den Anforderungen zu genügen, denn die Organisation "mit einem eigenen Leben" bildet ein Gegengewicht gegen die professionelle Gebundenheit der Experten in der professionellen Systembürokratie. Allerdings kann der organisatorische Wandel bzw. das Organisationslernen nur gelingen, wenn der beidseitige Austausch zwischen Organisation bzw. System und den Experten möglich ist, wenn also die "Kanäle offen sind". Hier sind genauso Zweifel angebracht wie in der Frage, ob das Organisationslernen es mit der hochgradigen Komplexität auf Organisations- und Systemebene aufnehmen kann, ungeachtet der offenen Problematik, ob organisationelle Konzepte auf das Gesamtsystem des Gesundheitswesens anwendbar sind.
- Die **Kontext-bezogenen Theorien** sind am weitesten gespannt und stellen daher die

Favoriten dar, insbesondere da es mit dem VBP-Programm in den USA und dem QOL-Programm in Großbritannien potente Beispiele gibt, in denen man bei der Implementierung von P4P auf diese Konzepte setzt. Zunächst darf nicht vergessen werden, dass diese Methoden des sozialen Marketings einer langfristigen, strategisch aufgebauten Planung bedürfen, die von der politischen Seite aus gesteuert werden muss und in ihrer Dauer weit über eine Legislaturperiode hinausreichen. Jenseits der Kontinuität ist eine Übereinstimmung der Akteure in den wichtigsten inhaltlichen Fragen notwendig; es sei daran erinnert - nur als Beispiel -, dass in Deutschland noch nicht einmal der differenzierte Stellenwert von Ergebnis- vs. Prozessindikatoren geklärt ist. Wie später noch auszuführen sein wird, muss das Thema Qualität und Sicherheit im Gesundheitswesen über den i.e.S. politischen Raum hinaus in den zivilgesellschaftlichen Bereich hineinwachsen, ähnlich wie es die Leapfrog-Initiative (www.leapfrog.org) in den USA getan hat.

Ohne den letztendlich in Kapitel 7 zusammenfassend dargestellten Empfehlungen vorgreifen zu wollen, kann man jedoch jetzt schon festhalten, dass das Feedback-Verfahren P4P nur im Rahmen von begleitenden Maßnahmen sinnvoll einsetzbar ist, die weitergehende Veränderungspotentiale eröffnen (Einstellungs- und Rollenveränderungen, Organisationslernen, soziales Marketing). Nur unter diesen Bedingungen können die vier Kontextdomänen, die in Kapitel 3.1. aufgeführt sind (s. Abb. 6), in einer Art beeinflusst werden, dass eine Realisierung des durch P4P bewirkten Effektes möglich erscheint. Bevor die Empfehlungen auf diesem Hintergrund konkreter ausformuliert werden können, müssen noch die ökonomischen Anreize einschließlich ihrer Grundannahmen, die "konkurrierend" wirksamen Vergütungssysteme und die Rolle der politischen Gestaltung diskutiert werden, was in den nächsten drei Kapiteln geschehen soll.

Weiter: 4. Ökonomie

4. Ökonomie

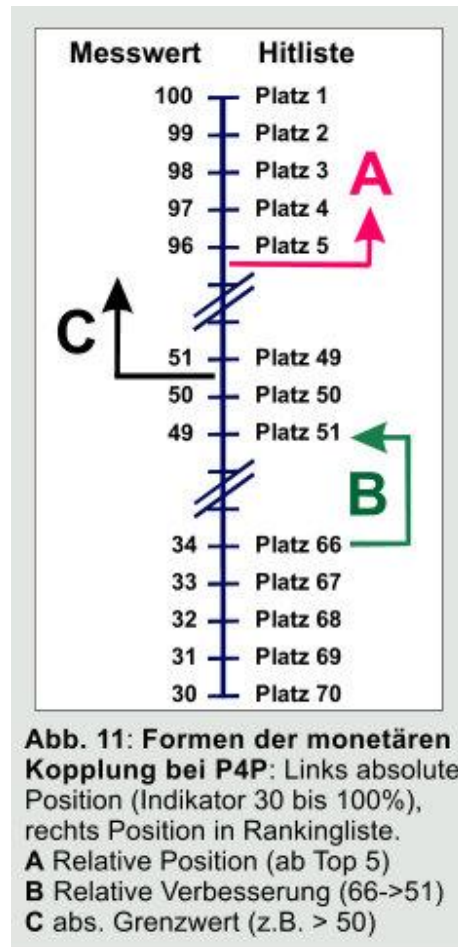
4.1. Einführung

Obwohl es sich bei P4P um ein Instrument handelt, das mit monetären Anreizen arbeitet, wird ökonomischen Betrachtungen in der Diskussion erstaunlich wenig Raum gegeben. Die Aufmerksamkeit ist meist auf die Messung der Qualität gerichtet, wenig werden die Fragen thematisiert, wie die gemessene Qualität ökonomisch zu bewerten sei, welche Höhe die finanziellen Anreize haben müssen, um optimal wirksam zu sein, und welche Grundannahmen eigentlich zu treffen sind, wenn man finanzielle Anreize in einem Zusammenhang einsetzen möchte, der alles andere als übersichtlich ist. In den vorangegangenen Kapiteln wurde daher, anschließend an die Darstellung der langfristigen Evaluationsergebnisse der großen P4P-Projekte, ausführlich auf die organisatorischen (Expertenorganisation) und Systemeigenschaften (Komplexität) sowie die zur Verfügung stehenden Konzepte der Verhaltensänderung eingegangen, wobei der Arbeitsbegriff der "komplexen professionellen Systembürokratie" zugrundegelegt wurde. Zur Verhaltensänderung scheinen in der Konsequenz nur Ansätze sinnvoll zu sein, die sämtliche in Kapitel 3.1. genannten vier Kontextfaktoren einbeziehen. P4P kann als Feedback-Verfahren nicht allein Verhaltensveränderungen im Gesundheitswesen bewirken, sondern muss in Zusammenhang mit Veränderungen der sozialen Rollen, mit organisatorischem Lernen und mit veränderten Kontextbedingungen gesehen werden. Ökonomische Fragen, so wie eingangs erwähnt, bedürfen jedoch einer gesonderten Bewertung.

Tableau 29: Definition P4P (aus Kap. 1.2.): *Pay for Performance (P4P)* im Gesundheitswesen basiert auf einer Qualitätsmessung durch definierte Indikatoren und koppelt die Qualität der Versorgung an Vergütungsbestandteile mit dem Ziel, die Qualität der Gesundheitsversorgung zu verbessern. P4P kann sich auch auf Qualität im Verhältnis zu den Kosten (Effizienz, *value*) beziehen.

Diese Überlegungen gehen von der Definition von P4P aus (s. Kap. 1.2., s. Tableau 29), die beinhaltet, dass P4P analog zum DRG-System, das ein epidemiologisches Instrument der Krankheitssystematik mit einem Vergütungssystem koppelt, ebenfalls aus zwei Systemen zusammengesetzt ist, nämlich einerseits einem System der Qualitätsmessung (Indikatoren einschl. Skalierung, Grenzwerten etc.) und andererseits einem System der monetären Bewertung. Die Situation gerät noch komplizierter, wenn man die Effizienz der Behandlung zugrundelegt, also das Verhältnis von Qualität und

Kosten, so wie es im *Value-Based Purchasing*-System des CMS der Fall ist (Damberg et al. 2009). Hier muss ein System der Effizienzmessung (Qualität plus Kosten) mit einem System der finanziellen Bewertung von Effizienz in Bezug gesetzt werden.



Auf vier Punkte soll hier eingegangen werden, die von der Frage der monetären Bewertung (Kopplung) bis zu grundlegenden ökonomischen Annahmen des Verhaltens der Akteure unter finanziellen Anreizbedingungen reichen:

- Kopplung der monetären Bewertung an die Qualitätssystematik
- Höhe der monetären Bewertung
- Informationsasymmetrie
- Ökonomische Grundannahmen.

► **Monetäre Bewertung:** Die Kopplung der Vergütungsbestandteile an die Qualitätsindikatoren muss praktikabel (administrabel), nachvollziehbar (verständlich), justizierbar und von seinen Anreizwirkungen sinnvoll gestaltet werden. Unter Praktikabilität ist hier nicht der Umgang mit den Qualitätsindikatoren (Messaufwand etc.)

gemeint, sondern allein die finanzielle Bewertung. Eine Vermischung mit einer Einzelleistungsvergütung sollte vermieden werden (s.u.), daher empfiehlt es sich, kumulative Bewertungen über einen bestimmten Zeitraum vorzunehmen. Um seine Anreizwirkung auf der Ebene der Leistungserbringer erzielen zu können, muss die Kopplung an die Vergütung in einem kurzen zeitlichen Zeitraum stattfinden, weil ein zeitnaher Feedback generell wirksamer ist und die Diskontierung der zusätzlichen Vergütung nicht so attraktiv erscheint (Damberg et al. 2007, Torchiana et al. 2013). Die Kopplung muss weiterhin nachvollziehbar sein, weil sie sonst keine Anreizwirkung erzielen kann, die handlungsleitend wäre (Ryan et al. 2011). Natürlich wird es auch zu juristischen Auseinandersetzungen kommen, wie es in Deutschland schon bei der Mindestmengen-Problematik geschehen ist, wo die Validität (also die Aussagekraft dieses Indikators hinsichtlich der Qualität der Versorgung) angegriffen wurde (s. Bundessozialgerichts-Urteil vom 18.12.2012 zur Frühgeborenenversorgung, vgl. Trefz 2013). Eine besondere Rolle spielen dabei, vor allem soweit Ergebnis-Indikatoren verwendet werden, Mängel in der Risiko-Adjustierung und eine mangelnde Aussagekraft der Indikatoren bei kleinen Einrichtungen, insbesondere bei Erkrankungen mit niedriger Inzidenz.

Am wichtigsten ist jedoch die Frage, wie Indikator und finanzielle Bewertung in Bezug gesetzt werden (**monetäre Kopplung**). Grundsätzlich sind drei verschiedene Verfahren möglich und in den bisherigen Studien angewendet worden (s. Abb. 11): Relative Position in einer Rankingtabelle (A in Abb. 11), relative Veränderung im Vergleich zur letzten Messung (B) und absoluter Grenzwert (C) sowie Kombinationen dieser Verfahren (Cannon 2006). Inhaltlich geht es darum zu vermeiden, dass sich die Verbesserung nur bei den Einrichtungen manifestiert, die schon vorher zu den Besten gehörten, während die *poor performers* nicht profitieren (Lindenauer et al. 2007). Diese Gefahr besteht vor allem bei der Verwendung von relativen Positionen ("die besten 5", verschärft durch Abschläge für die "schlechtesten 5"). Diese Gefahr wird auch nicht durch absolute Grenzwerte ausgeglichen (C), denn falls diese so hoch liegen, dass sie von den schlechteren Einrichtungen nicht erreicht werden können, unternehmen diese keine (frustranten) Anstrengungen. Es sind daher mehrere, gestaffelte Grenzwerte sinnvoll, die "weit nach unten reichen", so dass sie auch von den *poor performers* erreicht werden können. Für die Förderung der *poor performers* kann man aber insbesondere nicht auf relative Verbesserungen (B) verzichten. Darüberhinaus spielt auch die Sicherheit, mit der ein "Erfolg" vorherzusehen ist, eine Rolle: bei relativen Positionen mit Rankingtabelle ist die Situation für den Leistungserbringer zunächst unklar, denn seine Position hängt von der Leistung der Wettbewerber ab, erst am

Schluss ergibt sich die eigene Position. Falls diese "Chance" als zu gering angesehen wird, wird - insbesondere bei starker Risiko-Aversion - der Nutzen stark diskontiert (s.u.; Zusammenstellung s. Cannon 2006, zu den entsprechenden Befunden im HQIP-Projekt s. Jha et al. 2012, Ryan et al. 2012B).

Zusammenfassend kann festgehalten werden, dass relative Anreize nach Position, alleine eingesetzt, schlecht geeignet sind. Sie sollten mit der Belohnung relativer Positionsverbesserungen "auf den unteren Rängen" (B) und mit absoluten Anreizen in Form gestaffelter Grenzwerte, die auch für die "Schlechteren" erreichbar sind, kombiniert werden (Werner 2011).

Weiter: 4. Ökonomie, 4.2. Höhe der Bewertung und Informationsasymmetrie

4.2. Höhe der monetären Bewertung, Informationsasymmetrie

► **Höhe der monetären Bewertung:** Implizit wird in den meisten Überlegungen davon ausgegangen, dass die Anreizwirkung von P4P unabhängig von der Höhe der Qualitätsbezogenen Vergütung ist. In einem Review mit 8 randomisierten Studien war die Höhe des Anreizes nur in 4 Studien überhaupt genannt (Frolich et al. 2007). Über die Kosten, die den beteiligten Einrichtungen für die Erhebung der Qualitätsdaten (einschließlich Personalkosten) entstehen, hat keine Studie in diesem Review Angaben gemacht (Frolich et al. 2007). Im HQIP-Program in den USA wurden im ersten Jahr 850 Mill \$. bei mehr als 3000 teilnehmenden Krankenhäusern bewegt, danach durchschnittlich 35.000\$ pro Jahr pro Haus (Werner et al. 2011B). Nimmt man für ein Krankenhaus einen zusätzlichen Vergütungsbestandteil in dieser Größenordnung an, dann ist es durchaus denkbar, dass der Messaufwand (Erhebung der Daten, Durchführung von zusätzlichen Untersuchungen, Bereitstellung entsprechender Strukturen und von Personal etc.) weitaus höhere Kosten verursacht als später als *incentive* gezahlt werden - wohlgermerkt, die Kosten fallen sofort an, die Zahlungen (eventuell) in der Zukunft.

Für eine Beurteilung der notwendigen Höhe der Anreizzahlungen müssen außerdem die **Opportunitätskosten** mit einbezogen werden, also die Kosten, die dadurch entstehen, dass eine andere Option des Einsatzes dieser finanziellen Aufwendung nicht realisiert wird (entgangener Erlös) (Frolich et al. 2007). Die Opportunitätskosten sind im Gesundheitswesen sehr für das zugrundeliegende Vergütungssystem, in das P4P "eingebettet" ist, empfindlich. besonders dann, wenn von ihm ein starker Mengenanreiz ausgeht (z.B. Einzelleistungsvergütung, DRG). Es ist in diesem Fall äußerst fraglich, ob eine P4P-Vergütung auf institutioneller Ebene eher realisiert wird als eine zusätzliche Erlöseinheit (zusätzlicher Fall), die evtl. nur mit variablen und evtl. ohne zusätzliche sprungfixe Kosten zu realisieren ist: "*chase the utilization dollars rather than the quality dollars*" (s. Damberg et al. 2009). Erschwerend kommt hinzu, dass die in Frage stehenden Erlöse (Qualitäts- bzw. Leistungs-bezogen) unterschiedlich diskontiert werden können (**Diskontierung** betrifft die Abwertung von Erlösen, die erst in Zukunft anfallen, "lieber 10€ heute als 20€ in einem Jahr" (Damberg et al. 2007)). Wird der Qualitäts-bezogene Erlösbestandteil z.B. erst im nächsten oder übernächsten Jahr gezahlt (auf der Basis der im laufenden Jahr durchgeführten Erhebung - ein sehr realistisches Szenario), der durch eine zusätzliche Leistung ausgelöste Erlös dagegen aber sofort, dann wird man sich selbst dann für die zeitnah zu realisierende Option der Fallsteigerung entscheiden, wenn der Qualitäts-bezogene Erlösbestandteil etwas höher liegt (Frolich et al. 2007). Die Gesamtsicht der Überlegungen macht es verständlich, dass einige mehr analytisch unterfütterte Untersuchungen für höhere P4P-Vergütungen

plädieren (Chung et al. 2010, Ryan und Blustein 2011, Werner et al. 2011B, Werner und Dudley 2012).

Um die Höhe und damit potentielle Anreizwirkung des Qualitäts-bezogenen Erlösanteiles beurteilen zu können, muss man den Messaufwand berücksichtigen, die Opportunitätskosten einbeziehen und einer eventuellen Diskontierung Rechnung tragen.

An dieser Stelle sei noch angefügt, dass auch die Finanzierungsseite Trade-offs zwischen Risiko und *incentives* zu beachten hat, denn die Leistungsanbieter werden ihrerseits höhere Preise verlangen, um das Risiko durch einen schwer einzuschätzenden Vergütungsbestandteil auszugleichen. So ist auch für die Finanzierungsseite die Betrachtung der Frage alternativer Verwendungszwecke der P4P-Mittel durchaus von Bedeutung (Prendergast 1999).

► **Informationsasymmetrie und die Alternative der Einzelleistungsvergütung:**

Ohne den Überlegungen zur Integration von P4P in übergeordnete Vergütungssysteme vorzugreifen (s. unten), soll hier kurz auf Aspekte der *principal-agent*-Theorie eingegangen werden, die auf der Informationsasymmetrie zwischen Finanzierer (z.B. Versicherung) und Leistungserbringer beruht (der Einfachheit halber wird es bei einem bipolaren System belassen, andere Akteure wie Staat, ständische Organisationen etc. werden hier nicht berücksichtigt) (Nicholson et al. 2008, Frolich et al. 2007). Hätten beide Seiten den gleichen Informationsstand über die Realisierung einer qualitativ hochstehenden Behandlung, wäre es für den Finanzierer die beste Option, jeden der Bestandteile dieser Behandlung per Einzelleistungsvergütung "zu kaufen". Dieses Informationsgleichgewicht liegt jedoch nicht vor, denn die Angehörigen der Gesundheitsberufe und die Organisationen des Gesundheitswesens verfügen über einen erheblichen Informationsvorsprung vor der Finanzierungsseite, wenngleich etwa durch *Evidence-Based Medicine* und durch daraus abgeleitete Leitlinien ein erheblicher Teil des Wissens "säkularisiert" worden ist, also der Allgemeinheit und somit auch den Finanzierungsseite zur Verfügung steht. In diesem Fall der Informationsasymmetrie muss sich die Versicherung (*principal*) darauf verlassen, dass sich der *agent*, also die Leistungserbringer, in ihrem Sinne verhält. Erschwerend kommt hinzu, dass der *principal* dies nicht oder nur sehr schwierig kontrollieren kann.

Anreize wie P4P sind in dieser Situation grundsätzlich durchaus geeignet, den Informationsnachteil *ex post* auszugleichen, den die Finanzierungsseite *ex ante* gegenüber dem Leistungserbringer hat. Wenn sich jedoch die Informationsasymmetrie abschwächt, verliert P4P den Charakter eines Anreizes und wird zu einer (evtl. zusätzlichen) Zahlung im Sinne einer leistungsbezogenen (Einzelleistungs-)Vergütung

(Werner und Dudley 2012, Wodchis et al. 2007). In diesem Fall wäre zu erwarten, dass die angereizte Leistung nach Beendigung des P4P-Programmes wieder zurückgeht; genau dies ist in dem mit viel Geld ausgestatteten, auf distinkte Leistungen zugeschnittenen QOL-Programm in Großbritannien zu beobachten gewesen (Lester et al. 2010).

Folgende Fälle sind hier zu unterscheiden, die in der gegenwärtigen Diskussion wenig Beachtung finden, obwohl sie als relativ wirkmächtig angesehen werden müssen:

- die Qualitätsindikatoren, die Gegenstand der P4P sind, wurden bereits **vorher durch andere Anreizsysteme** (z.B. *public reporting*) adressiert. In diesem Fall ist es den Partnern auf beiden Seiten bereits bekannt, „wo der Hund begraben liegt“, und eine eigentliche (zusätzliche) Anreizwirkung kann nicht mehr erzeugt werden, ganz abgesehen davon, dass durch den *ceiling*-Effekt das mögliche Verbesserungspotential schon ausgeschöpft ist (Damberg et al. 2009). Gleiches gilt für den Fall, dass Indikatoren sehr lange in Gebrauch sind, man muss sich daher an die alte Forderung halten, dass Indikatoren regelmäßig zu wechseln sind.
- bei den Qualitätsindikatoren handelt es sich um Indikatoren der **Strukturqualität**, die weitgehend bekannt sind - hier wird aus P4P ein reiner Investitionszuschuss.
- bei den Qualitätsindikatoren handelt es sich um **Prozessindikatoren**. Ganz gegen den üblichen Trend ist festzuhalten: um so mehr EBM, um so weniger sinnvoll ist P4P (Nicholson et al. 2008), denn wenn Prozessindikatoren durch EBM-gestützte Leitlinien derart vorgegeben sind, dass sie die Qualität der Leistung weitgehend abbilden (aufgehobene Informationsasymmetrie), ist eine Einzelleistungsvergütung sinnvoller und von der Anreizwirkung her auch effektiver (Damberg et al. 2007). Um ein Beispiel zu nennen: die Blutkultur vor erster Antibiotika-Gabe bei der ambulant erworbenen Pneumonie kann man auch direkt vergüten. Solche Indikatoren mit Nähe zur Einzelleistungsvergütung zeigen eine starke Tendenz zur Vernachlässigung anderer, nicht angereizter Leistungen (*teaching to the test*) und dazu, nach Beendigung des P4P-Programms auf das Ursprungsniveau zurückzufallen (Lester 2010). Anders ist es in Situationen, in denen die Informationsasymmetrie nicht aufgehoben ist, z.B. bei der Behandlung chronischer Erkrankungen. Prozessindikatoren z.B. zur Koordination wie Vollständigkeit der Informationsweitergabe sind sinnvoll, denn das Geschehen „dahinter“ ist für die Finanzierungsseite nicht zu durchblicken, und sie dienen im Übrigen strategischen Zielen (s.u.). Ähnlich verhält es sich mit einem Indikator wie der Zahl der bearbeiteten Meldungen eines CIRS (*Critical Incident Reporting System*), der Benefit ist (wenn man den Indikator stichprobenartig kontrolliert) gegeben, die Informationsasymmetrie ist gewahrt.

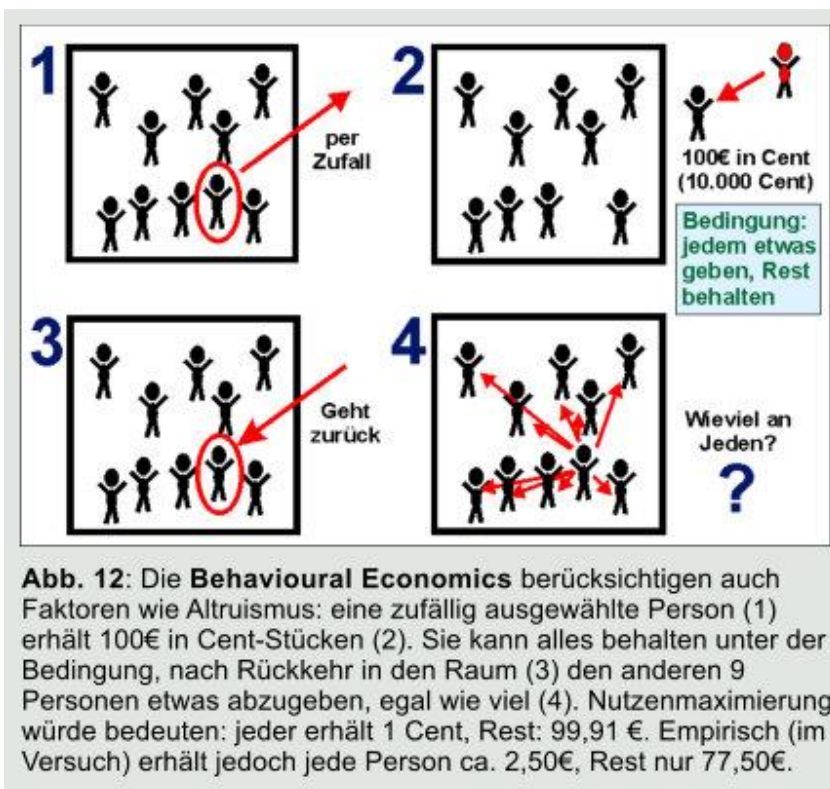
- bei den Qualitätsindikatoren handelt es sich um **Ergebnisindikatoren**. Hier ist es kaum möglich, die Informationsasymmetrie aufzuheben, selbst bei einfachen Elektivleistungen ist es der Finanzierungsseite nicht bekannt, wie das optimale Ergebnis zustandekommt. Das Problem liegt hier jedoch beim starken Anreiz zu einer Risikoselektion, die einerseits schwer durch Risikoadjustierung zu neutralisieren ist, und andererseits kleine Einrichtungen systematisch benachteiligt (Nicholson et al. 2008).

In der Konsequenz: P4P nur mit "frischen" Prozessindikatoren in komplexen Situationen mit aufrechterhaltener Informationsasymmetrie sinnvoll, die regelmäßig geändert werden.

Weiter: 4. Ökonomie, 4.3. *Behavioural Economics*

4.3. Ökonomische Grundannahmen: *Behavioural Economics*

► **Ökonomische Grundannahmen:** Spätestens seit der Finanzkrise 2007 werden in der ökonomischen Diskussion und in der Öffentlichkeit die Grundannahmen des *rational choice* als Basis für ökonomische Entscheidungen zunehmend in Frage gestellt. Es gibt eine große Zahl empirischer Untersuchungen z.B. aus dem Bereich der Verhaltensökonomie (*behavioural economics*, sog. *Prospect Theory*), die nachweisen, dass ökonomische Entscheidungen nicht zuverlässig und ausschließlich unter dem Aspekt der individuellen Nutzenoptimierung zu beschreiben sind (Kahnemann und Tversky 1979). Auch Im Gesundheitswesen existieren entsprechende Hinweise, dass Patienten und Ärzte nicht als *"perfectly rational decision makers"* (Loewenstein et al. 2013) anzusehen sind, auch wenn eine Zeitlang die Theorie des health care consumerism hoch im Kurs stand (Angell und Kassirer 1996). Eine große Rolle spielt diese Thematik besonders beim sog. Qualitätswettbewerb (s.o.), einem Konstrukt, das Qualitätsinformationen neben Menge und Preis als wettbewerbswirksam ansieht, und dem P4P und *Public Reporting* zugerechnet werden.



Bei *Public reporting*, der Veröffentlichung von Leistungserbringer-bezogenen Qualitätsindikatoren ohne finanzielle Anreize, ist man lange davon ausgegangen, dass diese Informationen Patienten und Zuweiser dazu veranlassen würden, die Anbieter mit

besserer Qualität zu bevorzugen und diesen dadurch eine bessere Marktchance zu verschaffen. Solche Daten werden in größerem Umfang in Deutschland im Rahmen der Qualitätssicherung im Krankenhausbereich nach §137 SGB V (AQUA 2013) veröffentlicht. In den USA werden Indikatoren-Sets zu insgesamt neun Bereichen der Gesundheitsversorgung publiziert, u.a. aus der stationären Krankenhausversorgung (*Hospital Inpatient Quality Reporting Program* (Hospital-IQR)), der ambulanten Krankenhausversorgung, aus Arztpraxen, der stationären und häuslichen Pflege und den Krankenkassen bzw. *Managed Care* Organisationen (Medicare C und D, s. CMS 2012A). Einer der Unterschiede zum deutschen System besteht darin, dass Patientenbefragungen eine große Bedeutung zukommt (z.B. 11 Indikatoren beim Hospital-IQR). 90% der anderen Indikatoren sind Prozessindikatoren und werden nicht risikoadjustiert.



Die verfügbaren empirischen Studien konnten den Effekt einer “informierten Wahl” des Patienten jedoch nie belegen (Werner und Ash 2005). Vielmehr war es so, dass die Anbieter, insbesondere die Krankenhäuser, auf der “Hitliste”, die auf Basis der Indikatoren gebildet werden, nicht an hinterer Stelle stehen wollten, und daher von sich aus Qualitätsverbesserungsmaßnahmen einleiteten, ohne durch Patienten dazu veranlasst worden zu sein (zusammenfassende Darstellung der Studien und Systematischer Review s. SVR 2008, Nr. 685ff, aktueller Berenson et al. 2013). Die Patienten fallen als “Nutzenmaximierer” sozusagen aus. Hinzukommt, dass die Rolle der anderen “Player”, also z.B. der Zuweiser, der Versicherer, der öffentlichen Meinung, die

Haltung der Patienten hinsichtlich Zugehörigkeit zum Stadtviertel, zurückzulegende Entfernungen, die Bedeutung anderer Eigenschaften der betreffenden Anbieter, die von den veröffentlichten Qualitätsindikatoren nicht beschrieben werden, bis hin zur Frage, ob die für die Patienten relevanten Informationen überhaupt und wenn dann verständlich dargestellt wurden, dass alle diese Fragen letztendlich unklar bleiben. Bei P4P und *Value-based Purchasing* wird die Situation noch unübersichtlicher.

Auch Ärzte sind nicht als reine Nutzenmaximierer anzusehen. Zwar gibt es Studien, die zeigen, dass Ärzte auf finanzielle Anreize reagieren (Hickson et al. 1987), aber es geschieht nicht regelmäßig (Rosenthal und Frank 2006) und der Effekt ist insbesondere bei bestimmten Themen wie Prävention nur gering (Hillman et al. 1998, Hillman 1999, Cook et al. 2004). Die Verhaltensökonomie hat hier aber entscheidende neue Impulse gegeben, die letztendlich auf eine Integration von psychologischen, soziologischen sowie später auch neurobiologischen Erkenntnissen in das Modell des Verhaltens unter ökonomischen Anreizen hinausläuft (Kahnemann und Tversky 1979, aktuell Verma et al. 2014). Diesen Ansätzen ist gemeinsam, dass Entscheidungen nicht allein auf der Basis einer monetären oder sekundär in monetären Einheiten zu beschreibenden Abwägung erfolgen, sondern maßgeblich unter dem Einfluss des autonomen Systems ("Bauchgefühl") (s. Kap. 3.5.2.) und sozialer Faktoren. Dabei treten einige mittlerweile gut beschriebene Effekte auf (Damberg et al. 2007, Verma et al. 2014, zu den entsprechenden Befunden im HQIP-Projekt s. Jha et al. 2012, Ryan et al. 2012B):

- **Framing:** Entscheidungen werden in einem referentiellen Zusammenhang getroffen, der die Interpretation erleichtert oder sogar erst ermöglicht. Dies betrifft z.B. die sog. **Risiko-Aversion** ("der Spatz in der Hand ist besser als die Taube auf dem Dach"); befragte Personen akzeptieren lieber eine Woche sicher garantierten Urlaub als die 50%ige Chance auf einen dreiwöchigen Urlaub (Damberg et al. 2007). Hiermit hängt die Diskontierung zusammen: weil die Gefahr gesehen wird, dass ein Ereignis in der Zukunft nicht eintritt, obwohl es mehr Nutzen erbringen könnte, wenn es tatsächlich eintritt, wird lieber ein geringerer Nutzen in der Gegenwart realisiert (lieber 10€ heute als 20€ in einem Jahr). Genauso sind Unterschiede zu beobachten, je nachdem ob eine Entscheidung als möglicher Verlust (*loss-frame*) oder als möglicher Gewinn (*gain-frame*) interpretiert wird. So wird ein leistungsabhängiger Gehaltsanteil gleicher Größe höher eingeschätzt, wenn das Gesamtgehalt am Jahresende ausgezahlt und der variable Anteil bei Nicht-Erreichen der Ziele am Jahresende abgezogen wird, als wenn er am Jahresende erst gewährt wird. Personen sind meist Verlust-avers, d.h. ein entgangener Gewinn wird weniger hoch bewertet als ein gleich großer Verlust (Werner und Dudley 2012).

- Es wird versucht, die Entscheidung mit einer überschlägigen **globalen Situationseinschätzung** zu bewältigen, die sich meist auf einen gewählten Fixpunkt bezieht. Wenn z.B. ein Ziel zu weit weg ist, als dass es erreicht werden kann, wird auch keine Anstrengung mehr unternommen (*goal gradient*), insbesondere wenn das Ziel auch nicht im Bereich der Handlungsoptionen liegt (z.B. Ergebnisindikatoren bei niedriger Inzidenz, oder zu hoch liegender absoluter Grenzwert (s. Abb. 11, Fall C)). Es können auf diese Weise Fehleinschätzungen groben Ausmaßes vorkommen, Personen fahren z.B. durch die ganze Stadt, um 10€ beim Kauf eines Radioweckers zu sparen, würden aber nie ähnliche Anstrengungen unternehmen, um 10€ bei einem LED-Fernseher zu sparen (Damberg et al. 2007). Vor die Wahl gestellt, ob ein Todesfall schlimmer wäre als kein Todesfall (Reduktion um 100%), oder 1000 Todesfälle schlimmer als 999 (Reduktion um 0,1%), entscheiden sich die meisten Personen für den ersten Fall, obwohl absolut die Differenz jedes Mal ein Todesfall beträgt (Überschätzung relativer Risiken (Verma et al. 2014)).

- **Isolation effect:** zur Beurteilung einer Entscheidungssituation wird eine der möglichen Dimensionen selektiert und die anderen werden unterdrückt. Werden z.B. mehrere Interventionen gleichzeitig implementiert, kann die Vielzahl nicht erfasst werden, nur einzelne werden herausgegriffen, es kann aber nicht vorhergesagt werden welche dieser Interventionen.

Eine gut handhabbare und etwas stärker praxis-orientierte Zusammenfassung haben diese entscheidungsrelevanten Elemente im sog. Mindspace-Konzept gefunden (King et al. 2013). Wenn z.B. Patienten dazu veranlasst (*to nudge them* - "schupsen") werden sollen, an Präventionsmaßnahmen teilzunehmen, ist es oft weniger wichtig, dieses mit Überlegungen zum absoluten Nutzen z.B. in Form von gewonnenen Lebensjahren zu tun (Rauchen aufhören!), sondern sich zu überlegen, wer die Botschaft überbringt (*messenger*), wie die Umwelt sich verhält (*norms, priming*) und was die *default*-Einstellung ist, also was als Normalfall angesehen wird - die Abweichung muss vom Individuum aktiv betrieben und begründet werden. Das Mindscape-Konzept ist sehr auf das *framing* fokussiert.

Zusammenfassend können aus den Kapiteln 4.1. bis 4.3. unter besonderer Berücksichtigung der *Principal Agent*- und der *Prospect Theory* (Verhaltensökonomie) folgende Grundsätze abgeleitet werden:

- Der alleinige Einsatz relativer Anreize über Rankinglisten ("die besten 5") ist kritisch zu sehen, da die *poor performers*, evtl. sogar die Mehrheit der teilnehmenden Institutionen verloren gehen, weil sie kaum Chancen sehen (*goal gradient*). Zu empfehlen ist eine

Kombination mit relativer Verbesserung und gestaffelten absoluten Grenzwerten.

- Die Zahlung der Anreize muss zeitnah, verlässlich und nachvollziehbar erfolgen, weil sonst Diskontierung und Risiko-Aversion ein zu großes Gewicht erhalten.
- Insbesondere bei Ergebnisindikatoren muss die Risikoadjustierung optimal gestaltet werden, weil sich ansonsten bei Erkrankungen mit niedriger Inzidenz vor allem kleinere Einrichtungen wegen des Morbiditäts- und Komorbiditätsrisikos nicht an dem P4P-Programm beteiligen bzw. alternativ Risikoselektion betreiben. Daher sind Prozessindikatoren, die keiner Risikoadjustierung bedürfen, zu präferieren, evtl. ergänzt durch einige Ergebnis- (z.B. adjustierte Mortalität) und Strukturindikatoren.
- Kleinere, häufigere Zahlungen, die aber relativ hoch erscheinen, sind wirkungsvoller als größere aber nur selten gezahlte Summen, die z.B. im Rahmen des Gesamterlöses eines Krankenhauses kaum auffallen (Überschätzung relativer Gewinne).
- Grundsätzlich ist die Einbehaltung von Vergütungsbestandteilen der stärkere Anreiz als die zusätzliche Vergütung (Verlust-Aversion), ein solches Vorgehen wird jedoch evtl. als unfair und motivationsmindernd erlebt.
- Die Indikatoren sollten “unverbraucht” sein und nicht schon vorher Gegenstand anderer Systeme (z.B. *Public Reporting*) gewesen sein, weil dann kein weiterer Effekt mehr auftritt (*ceiling*).
- Ein P4P-Programm sollte nicht gleichzeitig mit mehreren anderen Regelungen eingesetzt werden (*isolation effect*).
- Um so mehr ein Behandlungsablauf durch Leitlinien beschrieben und durch EBM “durchgedrungen” ist, um so mehr gewinnt eine Einzelfallvergütung gegenüber P4P an Vorteil, deren Anreizwirkung in jedem Fall überlegen ist. Ist die Informationsasymmetrie aber erhalten, ist P4P sinnvoll einzusetzen.
- Außerordentlich wichtig ist ein adäquates *framing* (Messenger, Normen, Default-Einstellungen, Kontext-Bedingungen).

Weiter: 5. Integriertes Vergütungssystem

5. Integration in bestehende Vergütungssysteme

5.1. Dominierende Vergütungsanreize im deutschen Gesundheitssystem

Das Rahmenkonzept, das für die vorliegende Analyse unter dem Arbeitsbegriff "komplexe professionelle Systembürokratie" Verwendung findet, versucht möglichst umfassend, die Kontextbedingungen zu identifizieren und einzubeziehen, die den Einsatz von P4P beeinflussen, sowohl hinsichtlich der Modelle zur Verhaltensänderung als auch bezüglich der ökonomischen Entscheidungsmodelle. Auf keinen Fall darf dabei natürlich das Gesamt-Vergütungssystem fehlen, in das P4P als Vergütungsbestandteil eingebettet ist (Mayes 2011). So ist z.B. die Frage zu stellen, ob in einem DRG-System oder in einem Vergütungssystem, das auf Einzelleistungen beruht, der immanente Mengenanreiz durch P4P wirklich aufgehoben werden kann, denn "... *the toxicity of the existing payment system may prevent weakly powered incentives from overcoming the Sisyphian task of QI*" (Damberg et al. 2009). Ob es nun notwendig ist, das System ganz zu wechseln, wie es gelegentlich in den USA formuliert wird (Corrigan und McNeill 2009), oder ob es möglich ist, mittels des Instrumentes P4P Schwächen des Systems an einzelnen Stellen zu heilen, soll anhand einer groben Charakterisierung der Verfassung von Gesundheitssystemen im folgenden Absatz kurz diskutiert werden. Dabei wird davon ausgegangen, dass das deutsche Gesundheitssystem wie die meisten anderen westlichen Gesundheitssysteme dadurch charakterisiert ist, dass

- eher die Leistungsmenge vergütet wird als die Qualität der Leistung, dass
- es sehr stark sektoral gegliedert ist, dass
- Erkrankungen und damit verbundene Untersuchungs- und Behandlungsmethoden und nicht die Krankheitsvermeidung (Prävention) vergütet wird, und dass es
- auf Akuterkrankungen und nicht auf die kontinuierliche Betreuung chronisch und mehrfach Erkrankter ausgerichtet ist.

In mehreren Übersichtsarbeiten sind die die Zusammenhänge zwischen P4P und anderen Vergütungsinstrumenten ausführlich dargestellt worden (Miller 2009, Rosenthal 2008). Die **Mengenorientierung** ist dabei in zweifacher Weise von Bedeutung, zum einen stellt sie traditionell das Hauptargument für die Einführung von P4P dar ("wir vergüten Menge statt Qualität, daher muss P4P kommen"), ist aber gleichzeitig eines der Haupthindernisse für dessen Wirksamkeit, weil in einem Mengen-orientierten System die Opportunitätskosten besonders hoch sind - als Alternative zur Investition in Qualitätsverbesserung und Dokumentation gilt immer: "ein Fall geht noch" (ausführliche Darstellung einschl. Risiko-Aversion und Diskontierung s.o.). Aber auch die **sektorale**

Gliederung schränkt den Horizont und damit den Effekt einer P4P-Vergütung stark ein, denn Anbieter werden entsprechende Maßnahmen in erster Linie dann in Betracht ziehen, wenn der daraus resultierende Nutzen in ihrem Sektor und nicht im Nachbarsektor anfällt. Generell kann man nur betonen, dass eine ausgeprägte Sektorierung für alle Qualitätsverbesserungsmaßnahmen ein großes Hemmnis ist, weil immer auch Nachbarsektoren profitieren, die sich selbst nicht engagiert haben. Der **Erkrankungs-Bezug** der Finanzierung verweist auf eine weitere grundsätzliche Weichenstellung in unseren Gesundheitssystemen, denn trotz aller Sonntagsreden finanzieren wir Akuterkrankungen als **“misslungene Prävention”**. Wenn hier also kein Paradigmenwechsel stattfindet, wird auch P4P dort eingesetzt werden, wo das Gesundheitssystem seinen Schwerpunkt hat und - das darf man nicht vergessen - wo Daten verfügbar sind: in der (prozedural orientierten, operativ dominierten) Akutmedizin und nicht in der präventiv ausgerichteten Gesundheitsversorgung. Für die **Orientierung an Akuterkrankungen** gilt Ähnliches, P4P verstärkt sogar die Bindung an Akuterkrankungen (z.B. Myokardinfarkt, ambulant erworbene Pneumonie), sowohl gegenüber Prävention als auch gegenüber den chronischen Erkrankungen, indem sie die Vergütung dieser Erkrankungen erhöht - es sei denn, man steuert aktiv um und nimmt mittels des Instrumentes P4P gezielt und strategisch überdacht die chronischen und Mehrfacherkrankungen in den Blick. Das QOL des NHS, das sich primär an chronische Erkrankungen richtet, hat hier eine wichtige Vorreiterfunktion.

Weiter: 5. Integriertes Vergütungssystem, 5.2. Einzelleistungsvergütung und sektorale Pauschalen

5.2. Einzelleistungsvergütung und sektorale Pauschalen

Geht man von den vier vorgenannten Systemeigenschaften aus (s. Tableau 30), kann man Vergütungssysteme daraufhin untersuchen, ob sie sich hinsichtlich Mengenanreiz, Prävention und Behandlung chronischer Erkrankungen günstig mit P4P-Vergütungselementen kombinieren lassen (die Sektorbezogenheit ist selbst Strukturmerkmal der Vergütungssysteme). Als weiterer Parameter für diese Überlegungen strategischer Natur ist die Tendenz zur Risikoselektion durch P4P hinzuzuziehen (s. Tableau 30). Zur Systematik der unterschiedlichen Vergütungssysteme kann man nach Gegenstand folgende Einteilung vornehmen:

- Einzelleistungsvergütung
- Vergütung nach Zeiteinheit
- Sektorale Pauschalen
- Transsektorale Pauschalen für (Akut-)Erkrankungen und Erkrankungsepisoden
- Erkrankungspauschalen
- Sektorale Populationspauschalen
- Populationspauschalen

Die resultierenden Kosten entstehen nach Zahl der Erkrankungen, den Behandlungsepisoden pro Erkrankung, der Zahl und dem Typ der Leistungen in jeder Episode, der Zahl und der Art der individuellen Prozesse, Medizinprodukte und Medikamente in jeder Behandlung und letztendlich der Preise jedes Prozesses, Produktes bzw. Medikamentes (Miller 2009).

Tableau 30: Eigenschaften der Vergütungssysteme: Kriterien für eine Kombination mit P4P

Verringerung des Mengenanreizes
Förderung von Prävention
Förderung der Behandlung chronischer Erkrankungen
Hemmung der Tendenz zur Risikoselektion

► Bei **Einzelleistungsvergütung** besteht ähnlich wie bei sektoralen Pauschalen ein starker Mengenanreiz, der Eingriffe und andere abrechenbare Leistungen fördert und Krankheitsvermeidung sowie die “begleitende Behandlung” chronischer Erkrankungen unattraktiv macht. Obwohl Einzelleistungsvergütung immer noch eine große Rolle spielt (z.B. Privatversicherte in Deutschland, *fee for service* bei Medicare in den USA, vgl. hier Chen und Ackerly 2014), wird deren Einsatz in einem Gesundheitssystem der Zukunft (Alterung, Chronizität, Multimorbidität) auf Bereiche mit vermuteter Unterversorgung

beschränkt werden. Im breiten Einsatz ist wenigstens nicht zu erwarten, dass P4P den Mengenanreiz austarieren kann. Soweit sich P4P auf distinkte Prozessindikatoren bezieht, besteht sogar die Gefahr, dass der Mengenanreiz zunimmt (z.B. durch zusätzliche diagnostischen Maßnahmen, die P4P-angereizt sind; s. hierzu auch Kap. 4.2.). Man kann zwar, wie es in vielen Projekten der Fall ist, einfache Präventionsmaßnahmen mit P4P belohnen (z.B. Impfprogramme), aber schon bei den chronischen Erkrankungen wird das schwieriger, ganz abgesehen davon, dass bei Aufhebung der Informationsasymmetrie die direkte Kontrahierung der betreffenden Leistungen sinnvoller ist. Die Einzelleistungsvergütung ist eine der Situationen, wo Ergebnisindikatoren größere Bedeutung zukommen kann. Allerdings ist darauf hinzuweisen, dass es in diesem Fall die Gefahr einer Risikoselektion durch P4P besteht, obwohl eine Einzelleistungsvergütung eigentlich keinen großen Anreiz hierzu bietet - man würde also einen unerwünschten Effekt durch P4P (soweit Ergebnisindikatoren genutzt werden) "importieren". Ein schlagendes Argument ist aber, dass die Opportunitätskosten so hoch sind, da die alternative Erlössteigerung durch die Erbringung einer zusätzlichen Leistung leichter möglich ist als komplexe Maßnahmen zur Qualitätsverbesserung bzw. -dokumentation. Vordergründig erscheint also P4P durch Ergebnisindikatoren unter Einzelleistungsbedingungen plausibel, in Gesamtsicht der Argumente hat P4P bei diesem Vergütungssystem aber keine Zukunft.

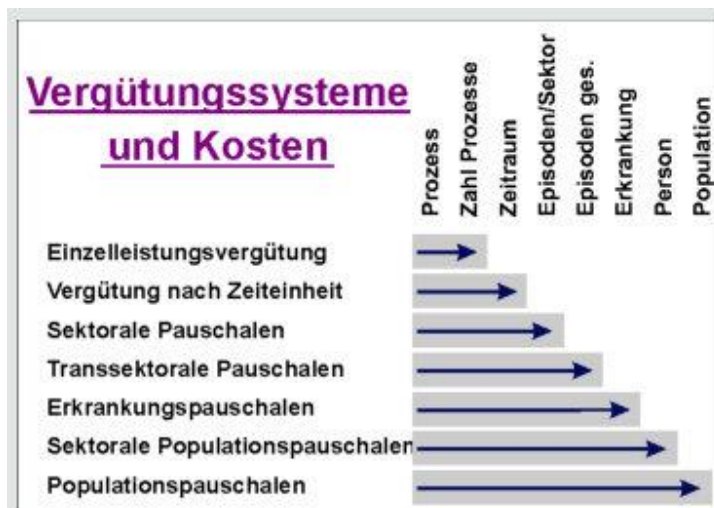


Abb. 14: Die Vergütungssysteme erfassen je nach Sektorzugehörigkeit Kosten, die sich aus Prozessen (Leistungen), Krankheitsepisoden (oder Akuterkrankungen), Erkrankungen in ihrer Gesamtheit und den Kosten, die für die Behandlung einer Person bzw. einer Population entstehen, zusammensetzen (modif. n. Miller 2009)

► **Vergütung über Zeiteinheit:** Ob man es nun bedauert oder nicht, dieses Vergütungssystem (z.B. tagesgleiche Pflegesätze) ist ein Auslaufmodell. Der Mengenanreiz ist auf die Zahl der abgerechneten Pflage-tage beschränkt. Prävention und Behandlung von Chronizität wird nicht angereizt, ist aber nicht ausgeschlossen - hier wäre durchaus ein Einsatzgebiet von P4P zu sehen. Auch könnte man einzelne Qualitäts-relevante Indikatoren sinnvoll einsetzen, u.U. kombiniert mit Kostenaspekten, um der Verlängerung der Verweildauern etwas entgegenzusetzen. Eine Gefahr der Risikoselektion besteht kaum.

► **Vergütung durch sektorale Pauschalen:** Die wichtigsten sektoralen Pauschalen sind die DRG's in der stationären Versorgung. Ihre Funktion besteht in der Transparenzsteigerung in der Krankenhausversorgung, die wichtigste Voraussetzung für eine stärkere Integration des ambulanten und stationären Sektors. Solange die sektorale Abgrenzung jedoch aufrecht erhalten wird, geht von den DRG's ein starker Mengenanreiz aus, zusammen mit einer deutlichen Akzentuierung operativer, Prozeduren-gebundener, operativer Akutleistungen zulasten der Versorgung chronisch-mehrfach Erkrankter und zulasten der Prävention. Wenngleich es sich nur um eine sektorale Pauschale handelt, besteht in diesem Bereich ein Anreiz zur Risikoselektion, dem man durch fortlaufende Anpassung des DRG-Systemes zuvorzukommen versucht, die letztendlich nichts anderes ist als die laufende Integration von Einzelleistungskomponenten. Auf den ersten Blick erscheint eine Kombination von DRG-Vergütung und P4P sinnvoll, wobei in erster Linie auf die Nutzung der im DRG-System vermehrt vorliegenden administrativen Abrechnungsdaten Bezug genommen wird. Dieses kann allerdings nur dann gelten, wenn es sich um ein Qualitäts-adjustiertes DRG-System handelt, was - das muss immer wieder betont werden - beim deutschen DRG-System nicht der Fall ist. Die Abrechnungsdaten weisen insbesondere ein Defizit in ihrer zentralen Eigenschaft auf, nämlich der Sensitivität (s. oben), mit anderen Worten, man übersieht die Qualitätsprobleme, um die es geht (s. z.B. Azaouagh und Stausberg 2008, Calderwood et al. 2014, Miller et al. 2001, Pawlson et al. 2007, Powell et al. 2001, s. auch SVR 2008 Nr.654ff).

Ganz abgesehen von dieser Problematik ist es vor allem aber fraglich, ob P4P mit dem immanenten Mengenanreiz des DRG-Systemes kompatibel ist. Im besten Falle würde P4P diesen Anreiz dämpfen können, aber dies ist nicht anzunehmen. Wie in Kap. 4.2. ausgeführt (*principal agent*-Problematik), sprechen sowohl

- die hohen Opportunitätskosten: eine zusätzlich erbrachte DRG-Leistung wird immer einer eventuellen Qualitäts-bezogenen Vergütung vorgezogen, als auch

- die Diskontierung: die Qualitäts-bezogene Vergütung wird im Allgemeinen später als die Leistungsvergütung realisiert, und

- die Risikoaversion: die zusätzliche DRG-Leistung ist "sicher", die zusätzliche Vergütung durch Qualität hängt z.B. vom Ranking ab,

eindeutig dagegen (Frolich et al. 2007). Dies gilt zumindestens so lange, bis der Mengenanreiz durch Veränderungen im Mehrleistungsausgleich abgemildert wird (s. Koalitionsvertrag 27.11.2013). Dabei ist darauf hinzuweisen, dass aufgrund der vorgenannten Ausführungen der Mehrleistungsausgleich die Vergütung der zusätzlichen Leistung *unter* die variablen Kosten drücken müsste, um wirksam zu werden. Die Problematik besteht vor allem bei Verwendung administrativ erhobener Ergebnisindikatoren, denn bei einem immanenten *underreporting* dieser Indikatoren (mangelnde Sensitivität administrativer Daten) bei gleichzeitiger Fähigkeit zur Risikoselektion in großen Verbänden lässt sich eine Mengensteigerung mit dem Argument der "guten Qualität" hervorragend durchsetzen. Die einzige Möglichkeit besteht in der Verwendung von Prozessindikatoren ohne Mengenanreiz, idealerweise solcher Indikatoren, die die transsektorale Integration widerspiegeln (z.B. Vereinbarung mit dem ambulanten oder Reha-Bereich zum MRSA-Screening).

Weiter: 5. Integriertes Vergütungssystem, 5.3. Von der transsektoralen zur Populationspauschale

5.3. Von der transsektoralen zur Populationspauschale

► **Transsektorale Pauschalen für (Akut-)Erkrankungen und**

Erkrankungsepisoden: Diese Vergütungsform wird in Deutschland in erster Linie durch die Integrierte Versorgung nach §140a ff SGB V repräsentiert, in den USA entspricht dieser Vergütungsform das *bundle payment*. Es werden alle Leistungen, die mit einer Erkrankung zusammenhängen, unabhängig von ihrer sektoralen Erbringung erfasst. Der Mengenanreiz ist daher geringer, die Krankheitshäufigkeit lässt sich nicht ohne Ende steigern. Allerdings ist die Zentrierung auf Akuterkrankungen unter gleichzeitiger Vernachlässigung chronischer Erkrankungen schon allein wegen der besseren Abgrenzbarkeit ungebrochen, auch gibt es keinen Anreiz zur Prävention, denn man generiert Erlöse für den behandelten Krankheitsfall und nicht für die verhinderte Erkrankung bzw. Krankheitsepisode. Wie bei allen (höherrangigen) Pauschalen besteht ein Anreiz zur Risikoselektion, denn der Deckungsbeitrag der Pauschale steigt mit sinkender Komorbidität bzw. Komplikationshäufigkeit. P4P hat in diesem Zusammenhang trotzdem bessere Einsatzmöglichkeiten als in den vorgenannten Vergütungssystemen. Zum Beispiel können sehr gut Prozessindikatoren wie die Zeit bis zur stationären Aufnahme oder betreffend die Behandlung am Wochenende eingesetzt werden, selbst die stationäre Verweildauer kann hier diskutiert werden. Auch Ergebnisindikatoren können ihren Einsatz finden, soweit die Fallzahl nicht zu gering ist. Zusätzlich zu den besseren Umsetzungsbedingungen von P4P kann dieses Instrument hier auch strategischer eingesetzt werden, denn die Integrationsfähigkeit des Systems kann bei gezieltem Einsatz gestärkt werden.

Tableau 31: Vergütungssysteme (aus 5.2.):

- Einzelleistungsvergütung
- Vergütung nach Zeiteinheit
- Sektorale Pauschalen
- Transsektorale Pauschalen für (Akut-) Erkrankungen und Erkrankungsepisoden
- Erkrankungspauschalen
- Sektorale Populationspauschalen
- Populationspauschalen

► **Erkrankungspauschale:** Bei der pauschalierten Vergütung von Patienten mit einer chronischen bzw. komplexen Erkrankung, wie sie z.B. in den deutschen strukturierten Behandlungsverfahren (DMP) nach § 137f,g angedacht (aber noch nicht umgesetzt) ist, kann der gesamte Krankheitsverlauf einschließlich aller anfallenden Behandlungen für einen gewissen Zeitraum (z.B. 1 Jahr) erfasst werden (s. *Comprehensive Care Payment* in den USA (Miller 2009)). Der Unterschied zur Populations-bezogenen Pauschalierung

(*Managed Care*) besteht darin, dass das Morbiditätsrisiko auf Populationsniveau beim Versicherer verbleibt. Trotzdem brauchen P4P-Programme es nicht mehr mit einem starken Mengenreiz "aufzunehmen", stattdessen können hier Prozess- und Ergebnisindikatoren Anwendung finden, eine Risikoselektion ist allerdings möglich. Typischerweise sind chronische Erkrankungen Gegenstand von *Disease Management* oder *Comprehensive Care* Programmen, auch zumindest sekundärpräventive Maßnahmen finden ihren Sinn. Eine Primärprävention wird jedoch nicht gefördert.

Vergütung und Anreizwirkung				
	Menge	Prävention	Chron. Erkr.	Risiko-selektion
• Einzelleistungs-Vergütung	+			
• Zeiteinheit	+			
• Pauschale /sektoral	+			+
• Pauschale /transsekt.				+
• Pauschale, Erkrankung			+	+
• Population /sektoral			+	+
• Pauschale, Population		+	+	+

Abb. 15: Schematische und stark vereinfachende Darstellung der **Vergütungssysteme**, für die hier eine Kombination mit P4P diskutiert wird. Als Kriterien werden der Mengenreiz, die Förderung der Prävention sowie der Behandlung chronischer Mehrfacherkrankungen und die Risikoselektion in Betracht gezogen.

► **Sektorale Populationspauschalen:** Sektorale Populationspauschalen sind eine Besonderheit in Deutschland und wurden als Hausarzt-zentrierte Verträge nach §73b SGB V eingeführt. Zunächst hat die Sektorierung des deutschen Gesundheitssystems dadurch zugenommen, weil der ambulante Bereich jetzt aus einem primär- und einem sekundärfachärztlichen Bereich besteht, die mit gesonderten Budgets und nach dem Willen der jetzigen Regierung (s. Koalitionsvertrag) auch mit getrennten körperschaftsrechtlichen Strukturen ausgestattet werden sollen. Es wird sich zeigen, ob sich durch diese Entwicklung der Druck auf den ambulanten und den stationären sekundärfachärztlichen Bereich in Richtung auf eine bessere und engere Kooperation erhöht und dies zu einer nachweisbar besseren Integration führt (s. Ambulante Spezialfachärztliche Versorgung nach §116b SGB V). In der HZV besteht kaum ein sinnvoller Anreiz zur Mengenausweitung, allerdings ein Anreiz zur Risikoselektion gegenüber den Nachbarsektoren. Chronische Erkrankungen sind Gegenstand der HZV,

Prävention kann durch P4P angereizt werden, wenn das Instrument entsprechend eingesetzt wird.

► **Populationspauschalen** werden in *Managed Care*-Systemen verwendet. "Als Managed Care wird ein Versorgungssystem bezeichnet, das die Leistungserbringung und Finanzierung in unterschiedlichem Ausmaß zusammenfasst. Dabei sieht es ein prospektiv pauschaliertes Finanzierungssystem vor. Managed Care verfolgt die Ziele, Sektoren und Leistungserbringer im Sinne einer regionalen, Outcome-orientierten Gesundheitsversorgung zu integrieren, sowie deren Effizienz u. a. durch Zielgruppenorientierung und Prävention sowie Generationenbezug zu verbessern. In der Umsetzung werden selektives Kontrahieren und die Etablierung von Managementstrukturen bei Kostenträgern und Leistungserbringern eingesetzt. Die Arztwahl kann in unterschiedlichem Maße eingeschränkt sein, Instrumente wie Gatekeeping, Disease Management, Case Management, Qualitätsmanagement, Leitlinien und *utilization review* kommen in verschiedenen Kombinationen und wechselndem Umfang zur Anwendung." (SVR 2009, Nr. 988). Die Leistungsanbieter, das ist das entscheidende Charakteristikum, übernehmen das Morbiditätsrisiko auf Populationsebene, die sog. *provider-payer-split* ist aufgehoben. Allerdings besteht kein Mengenanreiz mehr, dagegen jedoch durchaus ein großer Anreiz zur Risikoselektion und in gewisser Weise auch zur Vorenthaltung von Leistungen. Der Anreiz für Prävention ist gegeben, er hängt stark von der Bindungszeit der eingeschriebenen Patienten ab. So werden Maßnahmen mit kurzem Zeithorizont wie die Grippeimpfung gefördert, langfristig wirksame Leistungen wie die Zervix-Karzinom-Prophylaxe weniger, weil das Ereignis evtl. bereits in die nächste Versicherungsperiode fällt (Frolich et al. 2007). P4P wird innerhalb des Managed Care Organisation zur internen Steuerung eingesetzt, kann aber auch für die gesamte MCO zur Qualitätsverbesserung eingesetzt werden. Zu berücksichtigen ist, dass MCOs primär keinen großen Anreiz zur Offenlegung ihrer Qualitätsdaten haben. Nicht zuletzt deswegen ist es fraglich, ob P4P in der externen Steuerung von Managed Care Organisationen eine wichtige Rolle spielen kann.

Zusammenfassend können folgende Konsequenzen abgeleitet werden (zur stark schematisierenden Illustration s. Abb. 15):

- Auf den ersten Blick und getreu der Losung "**Qualität statt Menge!**" drängt sich P4P in der Kombination mit Einzelleistungsvergütung, der Vergütung von Zeiteinheiten und den sektoralen Pauschalen (z.B. DRG) gewissermaßen auf. Wie allerdings oben bereits ausgeführt, kann P4P aus mehreren Gründen den Mengenanreiz nicht ausgleichen. Ausgenommen können Erkrankungen werden, für die eine Mengenausweitung nicht

möglich ist (z.B. Geburten, Polytraumata); hier besteht für P4P ein sinnvolles Einsatzgebiet, wenngleich auch bei diesen Erkrankungen die Opportunitätskosten (s.o.) ein Ausweichen auf andere Erkrankungen bewirken können.

- Es besteht sogar die Gefahr, dass bei Einzelleistungsvergütung und sektoralen Pauschalen der **Mengenreiz** durch P4P **verstärkt** wird, und zwar wenn folgende Bedingungen zusammentreffen: (1) es handelt sich um Leistungen, bei denen die Möglichkeit zur Mengenausweitung besteht (z.B. Endoprothetik), (2) es werden Indikatoren mit geringer Sensitivität verwendet, wie es bei Indikatoren auf der Basis administrativer Daten der Fall ist, (3) mit den resultierenden Qualitätsdaten wird eine Mengenausweitung begründet und in den Verhandlungen mit den Kostenträgern durchgesetzt, und eventuell wird zusätzlich (4) eine aktive Risikoselektion betrieben, weil die Einrichtung aufgrund ihrer Größe mit eigenen Daten eine Risikobewertung ihrer Patienten betreiben kann.

- Sieht man allerdings vom Mengenreiz ab, können bei Einzelleistungsvergütung und sektoraler Pauschalierung durchaus interessante Einsatzmöglichkeiten für P4P darin bestehen, die **Koordination** der Behandlung und überhaupt die Behandlung von **chronischen Erkrankungen** zu verbessern. Man kann den Nachteil chronischer Erkrankungen aus dem nicht-operativen Bereich, der dadurch entsteht, dass der primäre ökonomische Anreiz bei diesen Vergütungsformen eher auf operativ zu behandelnde Akuterkrankungen gerichtet ist, versuchen auszugleichen und hätte damit eines der dringenden Qualitätsprobleme in Deutschland aufgegriffen. Gleiches gilt für auch für andere Themen wie Patientensicherheit (z.B. Indikatoren zur Einführung und sinnvollen Handhabung von Instrumenten wie CIRS). In den genannten Fällen sind Prozessindikatoren sinnvoll einzusetzen, die keiner Risikoselektion bedürfen.

- Bei höhergradigen Pauschalen (integrierte transsektorale Versorgung, Erkrankungspauschalen, *Managed Care*) tritt die Koordination der Behandlung und die Versorgung von Patienten mit **chronischen, multiplen Erkrankungen** mehr in den Vordergrund. Da hier der konkurrierende Mengenreiz weniger stark ausgeprägt ist, kann man mit einem gezielten Einsatz von P4P sinnvoll eingreifen und Schwerpunkte in der Qualitätsentwicklung setzen (s. QOF-Projekt in Großbritannien). Auf eine Risikoselektion ist Rücksicht zu nehmen (s.u.).

- Der Anreiz zur **Risikoselektion** verhält sich (*cum grano salis*) spiegelbildlich zum Mengenreiz. Um so höhergradig die Pauschalierung ist (bis hin zum Übergang der Versicherungsfunktion auf die Leistungserbringer bei *Managed Care*), um so mehr profitieren die Anbieter von der Vermeidung von Patienten mit teuren Erkrankungen

bzw. hoher Komorbidität. Man kann durch die Wahl von Prozessindikatoren und/oder eine adäquate Risikoadjustierung von Ergebnisindikatoren gegensteuern. Die Risikoadjustierung findet jedoch ihre Grenzen bei Erkrankungen mit geringerer Fallzahl, bei denen es zu einer Benachteiligung von kleineren Einrichtungen kommt (s. Kap. 1.3.). Es ist auch darauf hinzuweisen, dass nicht alle Prozessindikatoren von der Risikoselektion ausgenommen sind (vgl. Mehta et al. 2008). Allerdings sind Erkrankungen und Behandlungen ohne die Möglichkeit zur Risikoselektion (Unfälle oder Geburten bei Aufnahmehzwang, Impfungen) nicht davon betroffen.

- Analog kann man zum Thema **Prävention** vorgehen. Einen direkten Anreiz für Prävention bietet, abgesehen von Managed Care, keines der hier diskutierten Vergütungssysteme. Durch einen adäquaten Einsatz von P4P ist es aber grundsätzlich möglich, den vor allem durch Diskontierungsaspekte gehemmten Einsatz der Prävention (s. Abb. 8, Kap. 3.5.1.) zu fördern, ohne dabei den Grundanreiz der Vergütung zu verändern.

Weiter: 6. Politische Verantwortung

6. Politische Verantwortung

6.1. Governance im Gesundheitssystem

Wie in den vorangegangenen Kapiteln bei der Skizzierung eines Rahmenkonzeptes für die Einführung, Funktion und Evaluation von P4P dargelegt, werden die Konzepte von Expertenorganisation und Systemkomplexität (zusammengefasst als komplexe professionelle Systembürokratie) stark durch die Annahmen, die man zu grundsätzlichen ökonomischen Fragen trifft (z.B. *principal agent* Konzept und Risikoaversion), und durch die Auswirkungen der Vergütungssysteme geprägt, in die P4P im Einzelfall "eingebettet" ist. Zur Komplettierung eines Rahmenkonzeptes verbleibt nun noch ein weiteres und letztes Element, nämlich die politische Ebene. Auch hier spielt es eine große Rolle, von welchen Grundannahmen zur Funktion "der Politik" man dabei ausgeht.

Tableau 32a: Begriff **Governance** in der Darstellung n. Mayntz (2010): "Heute wird der Begriff Governance im Kontext nationaler ebenso wie internationaler Politik benutzt, um die Gesamtheit der in einer politischen Ordnung mit- und nebeneinander bestehenden Formen der kollektiven Regelung gesellschaftlicher Sachverhalte zu bezeichnen. Dabei liegt der Akzent auf den verschiedenen Formen zivilgesellschaftlicher Beteiligung an Prozessen politischer Regelung und Problemlösung auf allen Ebenen des politischen Systems, von der lokalen bis zur nationalen."

Sehr häufig geht man bei der Betrachtung der politischen Ebene von einem unhinterfragten, hierarchischen Verständnis aus: die Politik strukturiert, verordnet, beschließt, und die Peripherie folgt. Je nach Standpunkt wird die Professionalisierung der Politik akzeptiert, die Intransparenz beklagt und das "Versagen der Politik" irritiert zur Kenntnis genommen, wenn sich Einzelinteressen (wieder einmal) durchgesetzt haben. Dass es so einfach nicht ist, dass das hierarchische Modell zumindestens überdacht werden muss, liegt andererseits fast auf der Hand und ist dem Betrachter direkt zugänglich, denn ganz offensichtlich ist die politische Funktion, besonders im Gesundheitswesen, selbst Bestandteil der ausgeprägten Komplexität, die einen solchen gesellschaftlichen Bereich kennzeichnet. Aber ebensowenig, wie Expertenautonomie und Komplexität die Existenz von Regeln und eine externe Einflussnahme kategorisch ausschließen (sondern sie lediglich als verdeckt und wenig vorhersehbar beschreiben), so wenig kann sich die Politik untätig hinter der Komplexität des Systems verstecken, die Verantwortung ablehnen und das Gesundheitswesen als Regelwerk zur Verausgabung von jährlich 200 Mrd. € ohne aktiven Gestaltungsauftrag abtun.

Es stellt sich natürlich allerdings die Frage, ob es ein eigenständiges Politikfeld "Gesundheitspolitik" überhaupt gibt, neben den klassischen Politikfeldern wie Außen-, Innen- und Finanz- bzw. Wirtschaftspolitik. Wenn man Gesundheitspolitik nicht einfach über den Gegenstand "Gesundheitsversorgung und Prävention" definieren will, fällt zunächst die staatliche Gewährleistung für die Gesundheitsversorgung der Bevölkerung ins Auge, die zumindestens im Bereich der Krankenhausversorgung (Art. 74 GG) und für eine Mindestversorgung besteht (zur Diskussion s. Landau 2009). Die weitgehend staatlich organisierte gesetzliche Krankenversicherung für den größten Teil der Bevölkerung über steuer-analoge Beiträge auf die Arbeitseinkommen und die konkurrierende Zuständigkeit von Bund und Ländern z.B. für die Finanzierung oder die Investitionen im Krankenhausbereich stellen ein weiteres Kennzeichen dar, ganz abgesehen von den Zuschüssen an die gesetzlichen Krankenversicherungen aus dem Bundeshaushalt. Auch erscheint die dichte Verknüpfung von wirtschaftlichen und sozialen Interessen und die umfangreiche korporatistische Verfasstheit der Akteure von Bedeutung, die an der Steuerung und Umsetzung beteiligt sind - wie im Begriff der "komplexen professionellen Systembürokratie" bereits antizipiert (zu diesem Arbeitsbegriff s. Kap. 3.4.).

Falls man also aufgrund dieser Merkmale von der Existenz einer eigenständigen Gesundheitspolitik ausgehen möchte, stellt sich in zweiter Linie die Frage, ob sich das Politikmanagement im Gesundheitswesen, also die Prozesse und Inhalte, von anderen Politikfeldern abgrenzen lassen. In der Darstellung der drei "Arenen" nach Korte (Korte 2012) umfassend

- **Parteiendemokratie:** öffentliche Verhandlungsprozesse, Mehrheitsentscheidungen, Repräsentation der Parteien und Regierungsvertreter
- **Verhandlungsdemokratie:** Netzwerke, nicht-öffentliche Abstimmungen, Konsensbildung, Paketlösungen
- **Mediendemokratie:** Entscheidungen werden durch Stimmungen vorbereitet, Bürger treten in Kommunikation, Aufmerksamkeit als Machtprämie

ist augenfällig, dass im Gesundheitswesen die mittlere Arena der Verhandlungsdemokratie mit ihren korporatistischen Strukturen und nichtöffentlichen Konsensprozessen gegenüber der öffentlichen Parteiendemokratie enorm an Bedeutung gewonnen hat. In erster Linie sind hier die Verbände zu nennen, teilweise mit staatlichen Aufgaben betraut (z.B. Sicherstellung der ambulanten Versorgung) und in korporatistisch zusammengesetzten Institutionen eingebunden, die zentrale Steuerungsfunktionen direkt unter und in Überlappung mit der gesetzgeberischen Ebene wahrnehmen (z.B. Gemeinsamer Bundesausschuss), gemeinhin

zusammenfassend als Selbstverwaltung bezeichnet. Die Aushandlungsprozesse, die hier ablaufen, haben teilweise direkte Auswirkungen auf die Versorgung der Patienten und die Finanzierung im Gesundheitswesen (z.B. Nutzenbewertung und Finanzierungsentscheidungen zu Medikamenten). Diese enge Verknüpfung von Politik und Wirtschafts- bzw. zivilgesellschaftlicher Ebene prägt heute die Gesundheitspolitik ganz erheblich - ganz abgesehen von der rasanten Bedeutungszunahme der Mediendemokratie im Zeitalter von Facebook und *shitstorm*. Ausgedrückt in der klassischen politikwissenschaftlichen Terminologie *polity* - *politics* - *policy* (s. Tableau 32) werden Themenstellung und Ergebnisse (*policy*) nicht nur durch die normative Ebene (*polity*) beeinflusst, sondern in starkem Maße auch durch Interessen, Konflikte und informelle Konsensprozesse (*politics*) (Schubert und Bandelow 2008).

Tableau 32a: Begriff **Governance** in der Darstellung n. Mayntz (2010): "Heute wird der Begriff Governance im Kontext nationaler ebenso wie internationaler Politik benutzt, um die Gesamtheit der in einer politischen Ordnung mit- und nebeneinander bestehenden Formen der kollektiven Regelung gesellschaftlicher Sachverhalte zu bezeichnen. Dabei liegt der Akzent auf den verschiedenen Formen zivilgesellschaftlicher Beteiligung an Prozessen politischer Regelung und Problemlösung auf allen Ebenen des politischen Systems, von der lokalen bis zur nationalen."

Folgt man diesem Verständnis einer Gesundheitspolitik, steht deren Existenz natürlich sogleich wieder in Frage, um so mehr als nach dem in diesem Artikel verwendeten system- bzw. komplexitätstheoretischen Ansatz (Rahmenkonzept, s. Kap. 3.3) weder Regeln erkennbar noch Interventionen mit vorhersehbarer Wirkung möglich wären. Hier ist die Diskussion jedoch nicht stehengeblieben, so hat die Neokorporatismus-Debatte mit ihrer Betonung institutioneller Elemente, die zwischen politischer Entscheidungsebene und Wirtschaft existieren oder aufgebaut werden, klargestellt, dass ein erweitertes Verständnis von "Politik", das sich nicht nur auf das hierarchische Top-Down-Verständnis beruft und die letztlich dann auch mit Umsetzungsfragen betrauten Institutionen einbezieht, für die Steuerung eines solchen hochkomplexen Systems wie der Gesundheitsversorgung durchaus funktional ist. Gemeinsam als wichtig erkannte Probleme können unter Nutzung der jeweiligen Kernkompetenzen sinnvoll angegangen werden, auch wenn teilweise informelle Kommunikationskanäle genutzt werden müssen (Lauth und Thiery 2012). Der aktuelle Begriff, unter dem diese Diskussion derzeit geführt wird, die sog. **Governance** (Benz und Dose 2010, Mayntz 2010, s. Tableau 32a), stellt ein Politikverständnis dar, das davon Abstand nimmt, dass allein die politische Hierarchie oder der Markt die Koordination übernimmt, sondern dass das

Zusammenwirken staatlicher, privater und zivilgesellschaftlicher Strukturen für diese Koordination notwendig ist (Lauth und Thiery 2012). Es ist durchaus bemerkenswert, wie offen und explizit die "politische Ebene" dieses Konzept für das Gesundheitswesen in den Vordergrund stellt (Bundeszentrale für politische Bildung 2014).

Natürlich verbergen sich hinter diesen Konzepten erhebliche Gefahren, insbesondere bezüglich Intransparenz und politischer Einflussnahme. Einerseits kann, wie bereits in Kap. 3.4 angedeutet, das Konzept der *Governance* sehr gut in das Rahmenkonzept der komplexen professionellen Systembürokratie integriert werden, es repräsentiert sozusagen Komplexität und Einbeziehung der "Experten" auf der Makroebene und betont die Kompetenz zur spontanen Selbstorganisation, die ja nicht zu übersehen ist. Andererseits darf durch diese Konzepte die Funktion und Verantwortung der drei Gewalten, vor allem der Legislative, nicht eingeschränkt werden, Interessenlagen müssen erkennbar bleiben, Konsensergebnisse transparent (Mayntz 2010). Am wichtigsten erscheint jedoch, dass sie für die politische Ebene nicht als Entschuldigung für Tatenlosigkeit, Konzeptarmut und Verantwortungslosigkeit dienen darf.

Weiter: 6. Politische Verantwortung, 6.2. Handlungsfelder der Politik

6.2. P4P: Handlungsfelder der Politik (1)

Die genannte Kritik an modernen neokorporatistischen, "neoinstitutionalistischen" Politikkonzepten wie *Governance* ist nur allzu berechtigt, denn so wichtig es ist, die Akteure mit in Themensetzung und Entscheidungsbildung einzubeziehen, um eine bessere Umsetzbarkeit zu gewährleisten, so wichtig ist es auch, dass die Politik i.S. der legislativen Entscheidungen, der exekutiven Umsetzung und der juriskativen Kontrolle sich nicht selbst entmachtet (vgl. Mayntz 2010). Eine solche vorausseilende Selbstentmachtung verbietet sich nicht nur wegen der Gewährleistung der Gesundheitsversorgung (s.o.), sondern weil - ohne politischen Druck bewegt sich nichts (Moran und Scanlon 2013). Auch wenn Veränderungen auf der Ebene der Individuen und Institutionen vor Ort ablaufen, das "System" ist deswegen nicht aus der Verantwortung. Die politische Sphäre hat ihre spezifischen Aufgaben zu erfüllen, muss hierzu ihre spezifischen Methoden anwenden und hat dazu ihre spezifischen Strukturen zu nutzen.

Der Rückgriff auf die politische Theorie findet nun im Zusammenhang mit P4P seine Begründung darin, dass für die Einführung und die erfolgreiche Umsetzung eines Instrumentes wie P4P auch auf politischer Ebene Weichenstellungen notwendig sind, die die vorangehenden Überlegungen zur Rolle der Organisation, der individuellen Verhaltensänderung, der Komplexität des Systems und der ökonomischen Grundannahmen ergänzen. Immerhin, das darf nicht vergessen werden, handelt es sich bei P4P um einen massiven Eingriff in das Gesundheitswesen: der Fokus wird von der Menge (*volume*) auf die Qualität der Gesundheitsversorgung gelegt, ein - falls erfolgreich - geradezu paradigmatischer Wandel. Im Einzelnen müssen auf der politischen Ebene folgende vier Bedingungen thematisiert und implementiert werden:

- *Direction pointing*
- Strategische Ziele setzen
- Negative Auswirkungen kontrollieren
- Rahmenbedingungen schaffen

► ***Direction pointing***: Dieser Begriff stammt aus dem IOM-Werk "*Crossing the Quality Chasm*" von 2001 und bringt es auf den Punkt: die Richtung vorgeben, neben "*prohibitions*" und "*resource of permission providing*" eines der drei *simple rules*, die für eine Intervention in komplexen Systemen geeignet sind (Plsek 2003). Die system- oder komplexitätstheoretische Provenienz dieses Begriffes ist offensichtlich: da die internen Regeln nicht bekannt sind, die Auswirkungen von Interventionen nicht vorhersehbar sind und das System darüberhinaus zur spontanen Selbstorganisation neigt, wäre alles

sinnlos, was über eine Richtungsanzeige hinausgeht. Diese Richtungsanzeige kann sich auch auf die in komplexen Systemen vorhandenen Attraktoren beziehen, die als End- oder Zwischenzustände dienen und - obwohl von außen nicht direkt sichtbar - eine relative Stabilität bieten (s. Kap. 3.3.1.).

Man muss dem Konzept nicht einmal folgen, allerdings gibt es empirische Daten, die belegen, dass die Akteure im Gesundheitswesen eine solche "Richtung" durchaus wünschen und unterstützen. In der Befragung der *early adopters* von P4P in den USA gaben diese an, dass sie P4P trotz des Fehlens überzeugender Wirksamkeitsnachweise weiter verfolgen wollen, weil es besser sei, für Qualität bezahlt zu werden als für die Menge, und weil sie sich davon langfristig Transparenz für die Patienten versprechen (Rosenthal 2007). Die Richtungsangabe hat in diesem Sinne eine wichtige Kohärenzverstärkende Wirkung, ähnlich wie die Organisationskultur in Organisationen: das Große Ganze muss sichtbar sein. Im Rahmen des *Value-Based Purchasing* (VBP)-Programms in den USA wird eine solche Vorgabe beispielhaft gemacht: "*to foster joint clinical and financial accountability*", die Verantwortung auf klinischer und finanzieller Ebene stärken (CMS 2011B).

Natürlich muss das *direction pointing* in einer glaubwürdigen Form geschehen. Dies ist leichter gesagt als getan, denn die Angehörigen der Gesundheitsberufe haben ein feines Gespür, sie fassen solche Vorgaben durch die Politik schnell als "Manager-Sprech" auf. Logik und Sprache auf der Ebene von Politik (und Management) sind anders als am Krankenbett, gerade wenn es - wie bei P4P - um wettbewerbs-bezogene Elemente geht (vgl. Kap. 3.2.2., auch Edwards 2005). Insofern ist die immer wiederkehrende Bezugnahme auf mangelnde Qualität dann schwierig, wenn andere, gleichzeitig wirksame Anreize (z.B. Mengenanreiz) dominant sind und die Qualitätserwartung konterkarieren. Wie weiter unten (s. Kap 7.5.) ausführlicher dargestellt, ist daher z.B. zu diskutieren, ob - allen Ökonomisierungs-Debatten zum Trotz - in diesem Kontext das Konzept des *value*, also das Verhältnis von Qualität und Kosten, als Zielgröße für P4P am Ende nicht glaubwürdiger ist als die Zielgröße Qualität. In den USA wird es im *Value-Based-Purchasing*-Programm (VBP) so praktiziert. Diese Annahme erscheint auf den ersten Blick paradox, aber von außen auf Geld angesprochen zu werden wird von den Angehörigen der Gesundheitsberufe unter Umständen weniger befremdlich erlebt als auf die Qualität ihrer Arbeit.

In jedem Fall ist die Diskussion zu führen: wie kann ein authentisches *direction pointing* durch die politische Ebene aussehen? Eine "stumme Gesundheitspolitik" ist auf jeden Fall nicht die Lösung. Als Beispiel sei hier auf eine bislang erstaunlicherweise wenig

beachtete Möglichkeit zur Förderung und Verstärkung des Qualitätsgedankens hingewiesen, nämlich die verstärkte Förderung der Ausbildung in den Bereichen Ethik, Ökonomie und Qualität/Patientensicherheit für die Angehörigen der Gesundheitsberufe. Man kann sich auf den Standpunkt stellen, dass der Fortschritt, den das deutsche Gesundheitswesen in punkto Qualität und Patientensicherheit zu einem definierten Zeitpunkt in der Zukunft erreicht hat, gut durch die Integration dieser Aspekte in die Ausbildung beschrieben werden kann (Schrappe 2013). Nebenbei: damit hätte die Integration dieser Themen in die Ausbildung die Funktion eines klassischen Indikators.

Weiter: 6. Politische Verantwortung, 6.3. Handlungsfelder der Politik (2)

6.3. P4P: Handlungsfelder der Politik (2)

► **Strategische Ziele setzen:** Wie aus den vorangehenden Kapiteln bereits deutlich geworden ist, kann P4P keine Wunder vollbringen, insbesondere kann es dominante Anreize wie z.B. den Mengenanreiz durch das DRG-System nicht ausgleichen. Ein Instrument wie P4P bietet aber bei gezieltem Einsatz die Gelegenheit, zentrale Problemfelder eines Gesundheitssystems anzugehen - jedes Gesundheitssystem hat seine charakteristischen Stärken und Schwächen. Im deutschen Gesundheitssystem wären hier neben der Orientierung an Mengen und (operativ zu behandelnden) Akuterkrankungen vor allem die Koordinationsdefizite durch die ausgeprägte Sektorierung zu nennen (s.o., z.B. Schoen et al. 2011). Die Voraussetzung ist jedoch, dass man sich vor (!) Implementierung von P4P nach strategischen Gesichtspunkten ein Bild davon macht, welche dieser grundlegenden Probleme in welcher Reihenfolge anzugehen sind (Berenson et al. 2013). Diese strategischen Überlegungen können dann Indikatoren, meist Prozessindikatoren ohne Risikoselektion und ohne Aufhebung der Informationsasymmetrie (s.o.), identifizieren, die dann gerade für die Versorgung chronisch und mehrfach Erkrankter sowie die Versorgung älterer Patienten eine große Bedeutung haben. Andere Indikatoren können sich z.B. mit der Unterversorgungsproblematik in ländlichen Regionen beschäftigen, man muss hier jedoch diskutieren, ob die angestrebten Ziele nicht eher durch direkte Unterstützung und Einzelleistungsvergütung anzugehen wären. Diese strategischen Ziele zu setzen, ist eine genuin politische, gestaltende Aufgabe, die zwar die Unterstützung der Selbstverwaltungspartner bedarf, in ihrer Ausgestaltung jedoch in den Händen des Gesetzgebers liegen muss, nicht zuletzt deswegen, weil evtl. nicht unerhebliche Steuermittel aufgebracht werden müssen. Auch in den USA werden solche strategische Ziele derzeit intensiv diskutiert, z.B. die bessere Koordination über die Sektorgrenzen (Daughtridge et al. 2014) oder die optimale Versorgung der Bevölkerung, insbesondere der *underserved populations* (Rosenthal et al. 2007, Brown et al. 2014).

► **Negative Auswirkungen kontrollieren:** Gerade wenn ein neues gesundheitspolitisches Instrument eingesetzt wird, besteht die Verantwortung des Gesetzgebers in der Vermeidung bzw. Kontrolle etwaiger unerwünschter Auswirkungen. Diese Kontrolle kann in Zusammenarbeit mit Verbänden und anderen Partnern des Gesundheitswesens zusammen erfolgen, die Letztverantwortung liegt jedoch auf der Ebene der Politik. Die möglichen unerwünschten Auswirkungen der Einführung von P4P wurden zusammenfassend in Kap. 2.6. aufgeführt und diskutiert. Ganz im Vordergrund stehen Probleme des Zugangs (sozial benachteiligte Bevölkerungsgruppen, Kontinuität der Versorgung, Risikoselektion, vgl. z.B. Ryan 2013) und der Verzerrung durch

Datenmanipulation.

► **Rahmenbedingungen schaffen:** Die Rahmenbedingungen einer P4P-Einführung sind Gegenstand dieses Artikels (Rahmenkonzept). Es handelt sich dabei um Systemfaktoren und Umsetzungsparameter, die die möglichen Auswirkungen von P4P direkt oder indirekt fördern oder konterkarieren, und sie liegen in der Verantwortung der gesetzgeberischen Ebene, sie sind, um diesen Ausdruck im alten Sinne zu gebrauchen, ein politisches Problem (Miller 2009). Es handelt sich dabei nicht nur um reine Motivations- oder ökonomische Faktoren, wie z.B. konkurrierende ökonomische Anreize, die den Anreizen der P4P-Vergütung entgegengesetzt sind (z.B. Einzelleistungsvergütung). Dies ist ein ganz entscheidendes Problem, ohne Frage, aber negative Rahmenbedingungen können auch darin bestehen, dass bei Umsetzung und Spezifikation eines Instrumentes wie P4P Fehler gemacht werden. Da es hier im wahrsten Sinne des Wortes um Geld geht, kann man z.B. nicht erwarten, dass auf der Ebene der "Systembürokratie", als der Selbstverwaltung, eine zwanglose Einigung zu den anzuwendenden Indikatoren zustande kommt, eher ist eine langdauernde Blockade mit langfristiger Ergebnislosigkeit zu erwarten. Die *Governance* findet also dort ihre Grenzen, wo Interessen zurückgeschnitten werden müssen, Beispiele für ein Versagen der "Verhandlungsdemokratie" gibt es zu Genüge (z.B. Mindestmengenindikatoren im Krankenhausbereich). Zu den Rahmenbedingungen gehört es auch, die Komplexität des Geschehens nicht aus den Augen zu verlieren. Es reicht nicht, an einer Schraube zu drehen (s. Kap. 1.5.). Die Implementierung eines Instrumentes wie P4P muss immer im Kontext anderer Interventionen gesehen werden, nicht nur der Rahmen, der durch die zugrundeliegenden Vergütungssysteme (z.B. DRG) gegeben wird, sondern auch die wirtschaftliche Situation der Krankenhäuser, die Veränderungen in der Trägerschaft, die Sektorierung des Gesundheitswesens, die Morbidität etc.

Natürlich ist die Ausgestaltung der Institutionen im Bereich der Selbstverwaltung (z.B. Gemeinsamer Bundesausschuss) rational, weil nur so eine bessere Partizipation erreicht werden kann, aber es besteht die Gefahr der Blockade und Suboptimierung (Mayntz 2010). Hinzu kommt die Tendenz jeder Institution, Funktion mit Selbsterhaltung und Selbstbeschäftigung zu verwechseln und letztendlich notwendige Entwicklungen zu blockieren. Es muss daher verstärkt darüber nachgedacht werden, wie die "Zivilgesellschaft", damit ist hier gemeint: die nicht mit dem Gesundheitswesen verbundenen Gesellschaftsgruppen, Berufe, Verbände, Unternehmen, in das Thema Qualität und Sicherheit eingebunden wird, ein Ansatz, der in den USA z.B. mit der *Leapfrog*-Initiative erfolgreich beschritten wird (Galvin et al. 2005, Moran und Scanlon 2013).

Weiter: 7. Empfehlungen für die zukünftige Nutzung von P4P, 7.1. Allgemeines

7. Empfehlungen für die zukünftige Nutzung von P4P

7.1. Allgemeines

Es betrifft den institutionellen Rahmen, und es bezieht sich ausschließlich auf die ärztliche Vergütung, aber dieses Wortspiel bringt es trotzdem auf den Punkt: es geht nicht so sehr um die Frage "Was ist die richtige Vergütung ärztlicher Leistungen?", sondern um das Problem "Wie vergütet man ärztliche Leistungen richtig?" (Molzberger 2011). Welches Vergütungssystem, welche Kombination von Vergütungselementen wirkt so, dass die Gesundheitsversorgung optimal gestaltet ist? Am Anfang steht folglich eine strategische, weitreichende Überlegung, die den Bedarf an Gesundheitsversorgung und Prävention reflektiert, und zwar unter den Bedingungen einer immer älter werdenden, in einigen Regionen des Landes auch geringer werdenden Bevölkerung, die vorhersehbar in erster Linie an chronischen, multiplen Erkrankungen leiden wird (SVR 2009). Die Versorgung von Akuterkrankungen muss bestehen bleiben, aber die Zukunft wird nicht ohne eine verbesserte Koordination und Integration der Versorgung zu gestalten sein, die für die Behandlung chronischer Erkrankungen unerlässlich ist - und nicht ohne eine gestärkte Prävention, damit das Auftreten von chronischen Erkrankungen verhindert oder zumindest hinausgeschoben wird. Defizite in Koordination (Schoen et al. 2005, zuletzt Schoen et al. 2011) und Prävention betreffen zwei der größten Qualitätsprobleme des deutschen Gesundheitssystems, ein weiteres die augenscheinlichen Defizite im Bereich der Patientensicherheit (Behnke et al. 2013, Ott et al. 2013, Schrappe et al. 2008).

Tableau 33: Definition von P4P (s. Kap. 1.2.): "Pay for Performance (P4P) im Gesundheitswesen basiert auf einer Qualitätsmessung durch definierte Indikatoren und koppelt die Qualität der Versorgung an Vergütungsbestandteile mit dem Ziel, die Qualität der Gesundheitsversorgung zu verbessern. P4P kann sich auch auf Qualität im Verhältnis zu den Kosten (Effizienz, *value*) beziehen."

Es ist also Zeit zu realisieren, dass die im deutschen Gesundheitssystem angewendeten Vergütungssysteme weiterentwickelt werden müssen, und dies insbesondere zugunsten einer Qualitäts-Orientierung statt einer Orientierung auf die Menge der erbrachten Leistung und die Prozeduren-lastige Akutversorgung. Qualität statt Menge, das ist die *headline*, unter der international die Diskussion um *Pay for Performance* (P4P) vor ca. 15 Jahren aufgenommen wurde und aktuell weitergeführt wird. Der politische Druck zur Implementierung von P4P ist erheblich, denn so unterschiedliche Gesundheitssysteme wie das der USA und von Großbritannien haben sich langfristig auf die Verwendung

Qualitäts-bezogener Vergütungselementen festgelegt. Im Einzelnen handelt es sich um

- das umfassende und verpflichtende *Value-based Purchasing* Programm von Medicare in den USA (CMS 2011A, Belmont et al. 2011, Ryan et al. 2012, Chien und Rosenthal 2013) und
- das *Quality and Outcome Framework* mit jährlich neu justierten Indikatoren des *National Health Service* in Großbritannien (NHS 2013B).

Auch in Deutschland lässt der Koalitionsvertrag der Großen Koalition vom 27.11.2013 den festen Willen erkennen, Qualität (z.B über den Mehrerlösausgleich) bei der Vergütung zu berücksichtigen.

In der vorliegenden Arbeit werden die bisherigen Ergebnisse zu P4P auf diesem Hintergrund dargestellt und in ihrer Bedeutung diskutiert. Obwohl das Konzept einer qualitätsorientierten Vergütung außerordentlich plausibel ist, haben die wissenschaftlichen Evaluationen der ersten groß angelegten Programme keinen durchschlagenden Erfolg gezeigt (s. Kap. 2). Aus diesem Grund ist es unumgänglich, sich ausführlicher mit den impliziten Annahmen, den Kontextfaktoren bei der Implementierung und der Konzeption der Evaluation von P4P zu beschäftigen. Ein solches *conceptual framework* wurde auch in anderen Gesundheitssystemen als Basis für die P4P-Einführung entwickelt (*Crossing the Quality Chasm* des IOM (2001) bzw. *The NHS Outcome Framework 2013/2014* (NHS 2013B)). Ein solches **Rahmenkonzept** muss in der Lage sein,

- die Vielfalt der Faktoren, die auf das Verhalten der Leistungsanbieter einwirken und den Effekt von P4P mit beeinflussen, zu berücksichtigen,
- das Gesundheitssystem als Ganzes in seiner Reaktion auf P4P mit einzubeziehen,
- die Situationen, in denen P4P als wirksames Instrument sinnvoll einzusetzen ist, und die methodischen Anforderungen an die Gestaltung des Instrumentes zu charakterisieren,
- die eventuell auftretenden negativen Konsequenzen zu antizipieren, damit hier Gegenmaßnahmen getroffen werden können,
- P4P in das Zusammenspiel mit anderen Vergütungssystemen und anderen Qualitätsrelevanten Entwicklungen im Gesundheitswesen (z.B. Leitlinienentwicklung, *Evidence-based Medicine* und *Infection Control*) einzufügen,
- realistische Zielerwartungen für die Evaluation zu formulieren, und
- auf dieser Basis die politischen Entscheider realistisch zu informieren.

Um einige konkrete Beispiele zu nennen, P4P dürfte nicht in der Lage sein, bei Einsatz parallel zu einem klassischen DRG-System dessen unerwünschten Wirkungen (z.B.

Mengenanreiz, Förderung akutmedizinischer, operativ orientierter Leistungen) komplett auszugleichen und Qualität gegen diese Anreize zu verbessern (es kann aber sehr wohl bei gezieltem Einsatz einige Schwerpunkte anders setzen, z.B. Förderung der Behandlung chronisch erkrankter Patienten). Bereits an dieser Stelle kann darauf hingewiesen werden, dass der geringe Effekt von P4P im *Premier Hospital Quality Incentive Demonstration Project* (HQIP) - Programm (s. Kap. 2.2.) zumindest teilweise auf der zu wenig beachteten Wechselwirkung mit dem DRG-System und dem für die Mehrheit der Medicare-Patienten relevanten *Fee for Service* (Einzelfall)-Vergütung beruht. P4P kann weiterhin keine Qualitätsverbesserung erreichen, wenn die eingesetzten ökonomischen Anreize nicht ausreichen, z.B. weil sie quantitativ zu niedrig sind, ihre Zahlung unsicher erscheint oder zu weit in der Zukunft liegt (Diskontierung). P4P kann außerdem keinen Erfolg haben, wenn ein authentisches *direction pointing* durch die Politik fehlt, keine Richtung vorgegeben wird, oder die professionellen Werte verletzt werden.

Zusammenfassung kann daher als erste Empfehlung formuliert werden:

Empfehlung 1: Umfassendes Rahmenkonzept

Pay for Performance (P4P) ist kein isoliert einzusetzendes Instrument, sondern nur im Zusammenhang mit zahlreichen Kontextfaktoren (z.B. anderen Vergütungselementen) wirksam. Die Implementierung, Weiterentwicklung und Evaluation von P4P kann folglich nur vor dem Hintergrund eines umfassenden Rahmenkonzeptes erfolgen, das sowohl die professionellen als auch die organisatorischen, ökonomischen und Systemfaktoren umfasst. Eine erkennbare Wirksamkeit von P4P ist nur dann zu erwarten, wenn durch eine vorangegangene strategische Analyse die Ansatzpunkte im Gesundheitssystem identifiziert werden konnten, die von P4P beeinflussbar sind.

Das Kapitel 7 dieser Arbeit zum Thema "Empfehlungen" ist in sechs weitere Abschnitte aufgeteilt, die Empfehlungen zu den folgenden Themen enthalten:

- ▶ Hat P4P als Vergütungselement eine Zukunft? (Kap. 7.2.)
- ▶ Etablierung eines Rahmenkonzeptes (Kap. 7.3.)
- ▶ Welche Voraussetzungen sind für P4P als Feedback-Instrument notwendig? (Kap. 7.4.)
- ▶ Organisation und System (Kap. 7.5.)
- ▶ Ökonomische Fragestellungen und Annahmen (Kap. 7.6.)
- ▶ Politische Verantwortung (Kap. 7.7.)

Weiter: 7. Empfehlungen für die zukünftige Nutzung von P4P, 7.2. "P4P Is Here To Stay!"

7.2. "P4P Is Here To Stay!"

"Like it or not, Hospital Pay for Performance is here to stay" schließen Andrew Ryan und Jan Blustein ihren Artikel im NEJM (Ryan und Blustein 2012). Im Jahr 2017 werden allein im *Value-Based Purchasing* Programm (VBP) des Krankenhausbereichs knapp 2 Mrd. \$ Qualitäts-bezogen umverteilt, derzeit handelt es sich noch um die Hälfte dieser Summe - jährlich. P4P, da gibt es keinen Zweifel, stellt im Gesundheitswesen international eine der wichtigsten Systeminnovationen der letzten Jahre dar. Warum ist das so, und dauert dieser Trend an?

Tableau 34: Qualitätsdefizite im deutschen Gesundheitssystem (s. Kap. 1.1.):

- Kerndaten (z.B. OECD-Datensatz)
- *Infection Control* und Patientensicherheit (Behnke et al. 2013, Gastmeier und Geffers 2008, Ott et al. 2013, Schrappe et al. 2008)
- Varianz der Ergebnisse im regionalen Vergleich (Aqua-Institut 2013)
- Prozessuale Parameter wie z.B. Koordination der Behandlung (Schoen et al. 2005, zuletzt Schoen et al. 2011).

Der **erste Grund** ist offensichtlicher Natur: das Niveau von Qualität und Patienten entspricht nicht den Erwartungen (ausführliche Quellen-Angaben in Kap. 1.1., Tableau 34). Man hat die Hoffnung, mittels P4P einen Anreiz für "mehr Qualität" zu bieten. Wie die vorliegende Arbeit zeigt, kein Selbstläufer, aber grundsätzlich - bei richtiger Umsetzung - ein guter Gedanke.

Der **zweite Grund** liegt in der Sektorierung des Gesundheitssystems und der rein sektoralen Perspektive, unter der die Akteure Kosten- und Qualitätsaspekte handhaben. Die MRSA-Problematik in Deutschland kann gut und gerne als "*sectoral disease*" gelten; warum denn vor der Einweisung screenen, "wenn ich das nicht bezahlt bekomme"? Paradigmatisch wird die Sektorierung durch das DRG-System repräsentiert: eine sektorale Pauschalierung mit ihren imposanten Auswirkungen auf Menge und Optimierung zulasten der Nachbarsektoren. Dem System ist es jedoch nicht möglich, das DRG-System zu verlassen, denn nur durch DRGs kann die notwendige Transparenz im stationären Sektor gewährleistet werden, die für die (spätere) weitergehende Integration der Sektoren insbesondere in Abgleich mit dem ambulanten Sektor notwendig ist. Solange die Integration und damit die Fortentwicklung des Gesundheitssystems nicht weitergeführt wird, bleibt keine andere Möglichkeit, als allzu öffentliche Defizite des DRG-Systemes durch Vergütungselemente wie P4P "zu

heilen“ zu versuchen - ein schwieriger Weg, denn die ökonomischen Anreize des DRG-Systemes sind außerordentlich mächtig, viel wirkstärker als die einer Standard-P4P-Lösung. Es kommt in dieser Situation darauf an, die optimalen Einsatzbedingungen von P4P zu wählen und besonders solche Indikatoren zu identifizieren, die Koordination und Integration sowie die Behandlung chronischer Erkrankungen und die Prävention betonen.

Der **dritte Grund** betrifft die gestiegene informationelle Selbstbestimmung der (zukünftigen) Patienten. Qualitätsdaten (z.B. §137 SGB V) sind in großen Mengen vorhanden, *Public Reporting* ist Usus geworden, also warum nicht danach bezahlen? Dieser nur zu plausible Gedanke ist nicht von der Hand zu weisen und wird mit jeder Berichterstattung und Diskussion über Sicherheits- und Qualitätsmängel stärker. Hinzu kommt die Möglichkeit, in P4P-Instrumente *Patient-Reported-Outcomes* und Indikatoren, die auf Patientenbeobachtungen beruhen, zu integrieren, so wie es international in den großen P4P Proframmen bereits der Fall ist.

Es lohnt sich an dieser Stelle nochmals ein Blick auf die beiden Programme in den USA und in Großbritannien. So ist das VBP-Programm in den USA in mehrfacher Hinsicht bemerkenswert: (1) es basiert weitgehend auf Prozessindikatoren, für die eine Risikoadjustierung nicht notwendig ist, (2) 60% der Indikatoren beziehen sich auf Patientensicherheit (einschließlich nosokomialer Komplikationen und Sterblichkeit), (3) rund ein Viertel der Zahlungen gehen auf Indikatoren aus Patientenbefragungen bzw. *Patient-Reported Outcomes* (PRO) zurück, (4) das Programm richtet sich primär auf Effizienz (Qualität zu Kosten) und ist (5) als komplexes System eingebettet in ein umfassendes systemtheoretisches Konzept des amerikanischen Gesundheitswesens. Auch das QOL-Programm in Großbritannien basiert auf Prozessindikatoren und bezieht Patientenerfahrungen mit ein; es ist vor allem auf chronische Erkrankungen ausgerichtet, hat einen massiven Einstieg der Vergütung im ambulanten Bereich erbracht und weist einen breiten *Public Health*-Bezug auf.

Empfehlung 2: P4P ist als langfristige Entwicklung zu etablieren

Es gibt keine Alternative zum Einsatz von P4P, allerdings muss es kritisch, differenziert und integriert in andere Entwicklungen im Gesundheitswesen implementiert werden, denn in der derzeitigen Situation kann auf kein Instrument verzichtet werden, das die Qualitäts- bzw. Sicherheitsprobleme günstig zu beeinflussen verspricht. Aber P4P ist keine *magic bullet*, mit dem alle Defizite des Gesundheitssystemes geheilt werden können, es ersetzt keine genuine Weiterentwicklungsperspektive auf Systemebene (Beispiel DRG-System und Sektorproblematik).

Weiter: 7. Empfehlungen für die zukünftige Nutzung von P4P, 7.3. *Conceptual Framework*

7.3. *Conceptual Framework* - Rahmenkonzept als notwendige Bedingung

Bei der Entwicklung eines Rahmenkonzeptes, dessen Anforderungen in Kap. 7.1. formuliert wurden (s. Tableau 35) kann auf andere Gesundheitssysteme Bezug genommen werden (*Crossing the Quality Chasm* des IOM (2001) bzw. *The NHS Outcome Framework 2013/2014* (NHS 2013B)), trotzdem müssen die Eigenschaften des deutschen Gesundheitssystems im Vordergrund stehen (vgl. Kap. 5.1.):

- Orientierung an der Leistungsmenge,
- starke sektorale Gliederung,
- Erkrankungs- statt Präventions-Orientierung,
- Ausrichtung auf Akuterkrankungen statt auf chronische und Mehrfacherkrankungen.

Tableau 35: Rahmenkonzept - Anforderungen (Kap. 7.1.)

- die Vielfalt der Faktoren, die das Verhalten der Leistungsanbieter und den Effekt von P4P beeinflussen, berücksichtigen,
- das Gesundheitssystem als Ganzes in seiner Reaktion auf P4P mit einbeziehen,
- die Situationen, in denen P4P als wirksames Instrument sinnvoll einzusetzen ist, und die methodischen Anforderungen an die Gestaltung des Instrumentes charakterisieren,
- die eventuellen negativen Konsequenzen antizipieren, damit Gegenmaßnahmen getroffen werden können,
- P4P in das Zusammenspiel mit anderen Vergütungssystemen und anderen Qualitäts-relevanten Entwicklungen im Gesundheitswesen (z.B. Leitlinienentwicklung, *Evidence-based Medicine* und *Infection Control*) einfügen,
- realistische Zielerwartungen für die Evaluation formulieren,
- auf dieser Basis die politischen Entscheider realistisch informieren.

Die Faktoren, die in das Rahmenkonzept einfließen müssen, sind ausserordentlich zahlreich und in ihren Wechselwirkungen vielfältig (s. Tableau 36). Das in dieser Arbeit entwickelte Konzept stützt sich zunächst auf das aus der Organisationstheorie stammende Konzept der Expertenorganisation (*professional bureaucracy*) und auf die aus der Systemtheorie stammende Komplexitätstheorie. Diese Konzepte weisen in bestimmten Bereichen Parallelen auf und werden hier unter dem Arbeitsbegriff der **komplexen professionellen Systembürokratie** zusammengefasst (Definition s. Tableau 25a):

- beide Konzepte betonen die **Unsichtbarkeit der** (gleichwohl vorhandenen) **internen**

Regeln, in der Expertenorganisation in Form der dezentralen Vernetzung der Fachexperten, im komplexen System grundsätzlich zwischen allen Teilen des Systems;

- die Teile des Systems und die Experten im operativen Kern der Expertenorganisation verfügen über ein **hohes Maß an Autonomie**, gleichzeitig sind sie lern- und anpassungsfähig und neigen zur **Selbstorganisation** (in der Expertenorganisation auf fachlichen Ebene in Abgrenzung zur Managementebene), unvorhergesehene und in ihrer Ausprägung stark differierende Reaktionen der Organisation bzw. des Systems sind die Regel;

Tableau 36: Zugänge für die Etablierung eines P4P-Rahmenkonzeptes

- Organisationstheorie: Expertenorganisation (*professional bureaucracy*) (s. Kap. 3.2.)
- Systemtheorie: Komplexe Systeme (s. Kap. 3.3.)
- Verhaltensänderung im Gesundheitswesen (Kap. 3.5.)
 - lerntheoretische Modelle (Kap. 3.5.1.)
 - soziale Wahrnehmung (Kap. 3.5.2.)
 - organisatorischer Wandel (Kap. 3.5.3.)
 - Lernen durch Kontext (Kap. 3.5.4.)
- Ökonomische Grundannahmen
 - Opportunitätskosten und Diskontierung (Kap. 4.1./2.)
 - *Principal Agent* Theorie (Kap. 4.2.)
 - Verhaltensökonomie (Kap. 4.3.)
- Kombination mit anderen Vergütungssystemen (Kap. 5.)
- Politikwissenschaftliche Annahmen (s. Kap. 6)

- beide Konzepte können hochinnovativ sein, jedoch sind die Innovationen in Art, Menge und Zeitpunkt nicht vorhersehbar und können von außen nicht mit voraussagbarem Erfolg angestoßen werden, insbesondere wenn es sich um Prozessinnovationen handelt ("**Innovations-Paradoxon**"). Jegliche Versuche, hier Einfluss zu nehmen, wird als Reduktion der Autonomie und als innovationsfeindlich interpretiert;

- beide Konzepte akzeptieren **Unsicherheit**: der Experte in der Expertenorganisation wird immer improvisieren müssen ("Lehrer vor der Klasse"), die Komplexitätstheorie sieht Unsicherheit, Spannung und Paradoxie (im Gegensatz zu linearen Modellen) sogar als konstituierendes Bestandteil von Systemen an und erklärt eine Reduktion dieser Erscheinungen als aussichtslos oder nur in Ansätzen erreichbar. Diese "**intrinsische Unsicherheit**", die einerseits als Toleranz gegenüber Unsicherheit sicherlich eine adäquate Eigenschaft darstellt, denn sie schützt vor irreführenden linearen Konzepten, ist andererseits gerade beim Thema Qualität und Patientensicherheit ein kritischer

Punkt.

Hinzu kommt, dass nicht nur der durch die professionelle komplexe Systembürokratie beschriebene Kontext, sondern auch die Intervention P4P selbst komplexer Natur ist (sog. "**doppelte Komplexität**", zum Begriff vgl. Shojanian 2013). Unter dem Strich kann man von

drei Basis-Merkmalen sprechen, die bei der Implementierung von P4P in einem entwickelten Gesundheitssystem eine Rolle spielen: **Innovations-Paradoxon** und **intrinsische Unsicherheit** des Systems sowie **doppelte Komplexität** des Instrumentes.

Vor diesem Hintergrund erscheint es notwendig, sich ergänzend der Frage zuzuwenden, welche Rolle den unterschiedlichen **Konzepten der Verhaltensänderung** in einer komplexen professionellen Systembürokratie zukommt:

- **lerntheoretische Ansätze** umfassen insbesondere Feedback-Verfahren mit Belohnung, zu denen auch P4P gehört. Sie sind wichtig, greifen aber zu kurz, da Einstellungen, professionellen Loyalitäten und sozialen Rollen, ganz abgesehen von ökonomischen und politischen Faktoren (s.u.), eine sehr große, konkurrierende Rolle zukommt. Es hat also keinen Zweck, das Feedback- und Belohnungsverfahren P4P als "Insellösung" einzusetzen, ohne dass die Einstellungs- und Rollenebene mit verändert wird.

- Konzepte der **sozialen Wahrnehmung** thematisieren zusätzlich die Einstellungsebene und das Rollenverständnis, wichtig vor allem bezüglich des Themas Professionalismus, denn die traditionellen Rollenbilder geraten unter Druck und müssen sich an veränderte Bedingungen und veränderte Aufgaben innerhalb des Gesundheitssystems anpassen.

- Verhaltensänderung durch **organisatorischen Wandel** setzt auf einen unabhängigen Wissens- und Werte"vorrat" der Organisation, der einer Veränderung durch Lernprozesse zugänglich ist. Auf den ersten Blick scheint dieser Ansatz den Anforderungen zu genügen, denn die Organisation "mit einem eigenen Leben" bildet ein Gegengewicht gegen die professionelle Gebundenheit der Experten. Allerdings kann der organisatorische Wandel bzw. das Organisationslernen nur gelingen, wenn der beidseitige Austausch zwischen Organisation/System und den Experten möglich ist, wenn also die "Kanäle offen sind". Hier sind genauso Zweifel angebracht wie in der Frage, ob das Organisationslernen es mit der hochgradigen Komplexität auf Organisations- und Systemebene aufnehmen kann.

- Die **Kontext-bezogenen Theorien** sind am weitesten gespannt und stellen daher die Favoriten dar, insbesondere da es mit dem VBP-Programm in den USA und dem QOL-Programm in Großbritannien potente Beispiele gibt, in denen man bei der Implementierung von P4P auf diese Konzepte setzt. Allerdings bedürfen diese Methoden des sozialen Marketings einer langfristigen, strategisch aufgebauten Planung, die von der politischen Seite aus gesteuert werden muss und in ihrer Dauer weit über eine Legislaturperiode hinausreichen. Jenseits der Kontinuität ist eine Übereinstimmung der Akteure in den wichtigsten inhaltlichen Fragen notwendig, weiterhin sollte der zivilgesellschaftliche Bereich integriert werden.

In der Vergangenheit wurde bei den Diskussionen um P4P ebenso wie um Leitlinien-Implementierung oder *Evidence-Based Medicine* sehr stark auf lerntheoretische und Modelle der sozialen Wahrnehmung (einschließlich Professionalismus) gesetzt, Konzepte des organisatorischen Wandels und Kontext-bezogene Modelle standen und stehen nicht im Vordergrund.

Es ist daher äußerst fraglich, ob P4P, selbst ein komplexes Instrument, unter diesen Voraussetzungen in der Lage ist, die hohen Erwartungen in einem System wie dem des Gesundheitswesens zu erfüllen, ohne dass noch weitere Voraussetzungen erfüllt sind. In erster Linie wurden in der vorliegenden Arbeit **ökonomische und Vergütungsfragen** (s. Kap. 4.), in der Folge auch **politikwissenschaftliche Fragen** (s. Kap. 5.) diskutiert:

- Die **Höhe der Qualitäts-bezogenen Vergütung** und das System der **monetären Kopplung** an Qualitätsindikatoren und müssen kritisch gewertet werden, insbesondere unter Berücksichtigung der Opportunitäts- und Grenzkosten.
- Zusätzlich ist die **Diskontierung** der P4P-Vergütung zu berücksichtigen, falls die Zahlungen nicht zeitnah, sondern mit Zeitverzug erfolgen, u.U. sogar später als die Vergütung eventueller zusätzlicher Fallerlöse.
- Aus den Argumentationszusammenhängen der **behavioural economics** ist zusätzlich zu den Opportunitätskosten und der Diskontierung eine mögliche **Risikoaversion** einzubeziehen, wenn für den Leistungserbringer nicht sicher ist, ob er eine Qualitäts-bezogene Vergütung erhält (und er sich dann für die "sichere" zusätzliche Fallerbringung entscheidet). Dies kann z.B. dann der Fall sein, wenn bei einer relativ sortierten Rankingliste die relative Position genauso von der Performance der Mitbewerber wie von der eigenen Position abhängt. Genauso sind ein adäquates **framing** und die Berücksichtigung der Tendenz zur Überschätzung zeitnaher, kleiner Zahlungen von Wichtigkeit.
- Von großer Bedeutung ist die Tatsache, dass P4P nicht als alleiniges

Vergütungssystem eingeführt werden kann, sondern immer **in ein allgemeines Vergütungssystem eingebettet** ist. Dieses zugrundeliegende Vergütungssystem weist immer Anreizfunktionen auf, die unter Umständen der Qualitäts-bezogenen Vergütung durch P4P zuwiderlaufen können (z.B. DRG-System und starker Mengenanreiz). Man muss die Elemente einer P4P-Vergütung folglich sehr vorsichtig und gezielt in die allgemeine "Vergütungslandschaft" integrieren.

Aus *politikwissenschaftlicher Sicht* sind bei der heute vorherrschenden Tendenz zu *Governance*-Konzepten (s. Kap. 6.1.) vier Aspekte von Bedeutung:

- Es muss ein Rahmen erkennbar sein, der für die angestrebten Veränderungen einen Interpretations- und Sinnzusammenhang darstellt (***direction pointing***).
- Ein sinnvoller Einsatz von P4P ist ohne **strategische Überlegungen** nicht möglich, damit die optimalen und entscheidenden Ansatzpunkte für eine Weiterentwicklung des Gesundheitssystems identifiziert und angegangen werden können.
- Von politischer Seite muss erkennbar Verantwortung und Vorsorge getragen werden für etwaige **negative Effekte** der Qualitäts-orientierten Vergütung.
- Es müssen günstige **Rahmenbedingungen** geschaffen werden, z.B. durch Vorgaben zur Gestaltung der institutionellen Kooperationen, die für die Umsetzung verantwortlich sind. Die stärkere Einbeziehung der zivilgesellschaftlichen Ebene, die mit dem Gesundheitssystem primär nicht verbunden, jedoch an Qualität und Sicherheit genuin interessiert ist, ist zu diskutieren.

Empfehlung 3: Rahmenkonzept für die Implementierung und die Umsetzung von P4P

P4P sollte nicht als rein lerntheoretisch begründetes, einfaches Feedback-Instrument mit Belohnungskomponente verstanden und eingesetzt werden, dies wäre "zu kurz gesprungen". Stattdessen müssen sowohl Konzepte des organisatorischen Wandels und Kontext-bezogene Veränderungsstrategien als auch ökonomische, vergütungstechnische und politische Weichenstellungen mit einbezogen werden, erst dann kann von einer realistischen Erfolgsschance ausgegangen werden. Aus ökonomischer Sicht sind vor allem Opportunitätskosten, Diskontierung bzw. Zeitachse und Aspekte der Risikoaversion zu bewerten, außerdem ein optimales *framing*. Auf der *policy*-Ebene geht es in erster Linie um das *direction pointing* (Interpretationsebene für die anstehenden Entwicklungen) und um strategische Überlegungen bzgl. der optimalen Einsatzgebiete der P4P-Vergütung. Weiterhin müssen günstige Rahmenbedingungen geschaffen und eventuelle negative Nebeneffekte kontrolliert werden.

Weiter: 7. Empfehlungen für die zukünftige Nutzung von P4P, 7.4. P4P als Feedback-Instrument

7.4. Gestaltung des *Feedback*

P4P als lerntheoretisch begründetes Feedback-System bedarf, wie im vorangehenden Abschnitt ausgeführt, der Erweiterung um Instrumente der Verhaltensänderung aus dem Bereich des organisatorischen Wandels und Kontext-Lernens, um gut bedachte und fein justierte ökonomische Kontextfaktoren sowie um politische Weichenstellungen, die sich gut mit dem Begriff *direction pointing* beschreiben lassen. Allerdings nützt dies alles nichts, wenn der eigentliche Feedbackmechanismus selbst nicht intakt ist. In diesem Abschnitt werden daher kurz die technischen und methodischen Aspekte, die die Voraussetzung für das Funktionieren des Feedback-Mechanismus bilden, dargestellt und unter Einbeziehung des Rahmenkonzeptes (s. Kap. 7.3.) diskutiert. Weitergehende, die technische Ebene überschreitende Aspekte werden in den nächsten Kapiteln behandelt. Die Darstellung hat folgende Gliederung:

- Indikatoren-Konzept
- Inhaltliche Auswahl der Indikatoren
- Administrative vs. klinische Daten
- Prozess- vs. Ergebnisindikatoren I
- Risikoadjustierung
- Gaming
- Adressaten des Feedback
- Wirkung des Feedback: Zeitnähe
- *Ceiling*

► **Indikatoren-Konzept:** In Ergänzung zu den grundlegenden Aussagen zu Indikatoren in Kap. 1.4. sollen hier nur zwei Dinge hervorgehoben werden, die bei der Arbeit mit Indikatoren immer wieder zu Verwirrung führen. Zunächst muss man klären, ob man eine Vollerhebung eines Qualitäts-relevanten Parameters anstrebt (z.B. rechtzeitige Blutkultur bei Pneumonie), oder ob man einen Indikator mit Vorhersagefunktion implementieren will. Der erste Fall verstärkt die Tendenz von P4P zur Einzelleistungsvergütung (zur diesbezüglichen Bedeutung der Informationsasymmetrie s. 7.6. Empfehlung 21). Will man im Sinne des Indikatorenkonzeptes einen Indikator mit Vorhersagefunktion z.B. für einen Versorgungsbereich einführen und nutzen, tritt die direkte Qualitätsfunktion des Indikatores in den Hintergrund (z.B. Desinfektionsmittelverbrauch - *eo ipso* kein Qualitätsaspekt, wohl aber ein gutes Surrogat für nosokomiale Infektionen); neben die Selektivität des Indikators tritt die Abstraktion (s. Schrappe 2010). Dies hat für die (politisch) Handelnden zwei Konsequenzen, man muss sich nämlich weniger mit der Gefahr der Mengenausweitung durch P4P selbst (u.U. innerhalb eines stark Volumen-orientierten Systems wie DRG,

dessen Mengenanreiz somit noch weiter verstärkt würde, s. hierzu 5.3.) auseinandersetzen, und zweitens - vielleicht lästige Pflicht - können und müssen politische, aufgrund strategischer Überlegungen für notwendig erachtete Schwerpunktsetzungen und Schwachpunktanalysen greifen. Um es kurz zu sagen: kein echter Indikator ohne Problemanalyse. So kann man z.B. die in Kap. 5.1. bzw. 7.3. genannten grundsätzlichen Charakteristika des deutschen Gesundheitssystems heranziehen und entsprechende Indikatoren z.B. zur Stärkung der Versorgung chronisch Kranker oder der Prävention auswählen, und zwar mit dem Ziel, einen ganzen Versorgungsbereich zu verbessern (und nicht nur eine einzelne Leistung einzukaufen). Im weiteren bleiben die Aussagen zur den Charakteristika valider Indikatoren (s. Kap. 1.4.) bestehen, insbesondere sollte ein Indikator zuverlässig zu messen sein, valide (sensitiv) das betreffende Qualitätsproblem vorhersagen (die Spezifität steht nicht im Vordergrund), und das Qualitätsproblem muss veränderbar sein.

Empfehlung 4: Die Wahl von Indikatoren mit Vorhersagefunktion erfordert eine strategische Problemanalyse

Am Anfang der Implementierung eines P4P-Programmes steht die Entscheidung, ob einzelne (Qualitäts-relevante) Leistungen eingekauft werden sollen oder ob Indikatoren mit Vorhersagefunktion für die Qualitätsprobleme eines Versorgungsbereiches genutzt werden sollen. Im ersten Fall etabliert man eine Einzelleistungsvergütung, im zweiten Fall kann man von P4P im eigentlichen Sinne sprechen, weil die Problemlösungskompetenz der Leistungsanbieter vor Ort anerkannt wird. Eine vorherige Problemanalyse ist jedoch unverzichtbar.

► **Inhaltliche Auswahl von Indikatoren:** Eng mit der vorgeschalteten Problemanalyse verknüpft ist die Auswahl des inhaltlichen Problemfeldes, für das Indikatoren eingesetzt werden. Grundsätzlich werden Indikatoren wahlweise nach epidemiologischen Kriterien, nach Inhalt, Qualitäts-Systematik, Erhebungsmethodik und eingenommener Perspektive eingeteilt. Die Systematisierung nach inhaltlichen Kriterien umfasst wiederum Fach- und Diagnose-bezogene Indikatoren, organisatorische Indikatoren, Patientensicherheitsindikatoren und Indikatoren auf der Basis von Patientenerfahrungen. Es ist sehr interessant, wie stark die großen aktuellen P4P-Projekte in den USA (*Value-Based Purchasing* (VBP), s. Ryan et al. 2012C) und in Großbritannien (*Quality and Outcome Framework* (QOL), s. NHS 2013A) ihren Schwerpunkt auf **Patientensicherheits-Indikatoren** (s. Kap. 2.4.) und auf **Patienten-Erfahrungen** bzw. *patient-reported outcomes* (s. Kap. 2.4., 2.5.) legen. Im VBP-Programm machen diese beiden Gruppen 60% der Indikatoren und der Zahlungen (!)

aus (Ryan et al. 2012C).

Empfehlung 5: Patientensicherheits-Indikatoren und Patientenerfahrungen integrieren

Im Vordergrund der internationalen P4P-Programme stehen Patientensicherheits- und Indikatoren auf der Basis von Patientenerfahrungen. Neben strategisch positionierten Indikatoren zu verstärkter Integration, Behandlung chronischer Erkrankungen und zur Stärkung der Prävention sollten diese auch in Deutschland besonders beachtet werden.

► **Administrative vs. klinische Daten:** Ein weiterer "Evergreen" ist die Frage der Verwendung **administrativer oder klinischer Daten** (zur umfangreichen Literatur s. Kap. 1.3.). Administrative Daten sind primär für Vergütungszwecke und nicht für Qualitätsmessungen entwickelt worden (so auch das G-DRG-System) und weisen in vielen Fällen deutliche Sensitivitätsdefizite auf, erfüllen also hiermit das wichtigste Validitätskriterium eines Indikators nicht. Es besteht die Gefahr einer Mengenausweitung, wenn insensitive administrative Indikatoren bei Leistungen mit entsprechendem Potential eingesetzt werden (s. Kap. 5.3.). In Einzelfällen können sie begleitend jedoch eingesetzt werden (z.B. standardisierte Mortalität), sie geben u.U. sinnvolle Hintergrundinformationen. Allerdings verstärken sie in der Expertenorganisation die Spaltung zwischen Management und *operating core*, also den Experten, die diese Indikatoren - bei aller Erleichterung, dass sie mit der Datenerhebung nichts zu tun haben - als *management-business* abtun. Klinische Daten haben ihrerseits den klaren Nachteil, dass sie aufwendiger und nicht rein administrativ, sondern nur unter Beteiligung von Experten zu erheben sind. In Teilbereichen haben sie sich jedoch längst durchgesetzt, z.B. als epidemiologische Falldefinitionen in nosokomialen Infektionserfassungen (vgl. hierzu Talbot et al. 2013). Streng ist hierbei darauf zu achten, dass epidemiologische Falldefinitionen keine klinischen Diagnosen darstellen, selbst wenn diese Falldefinitionen sich klinischer Parameter bedienen, sondern lediglich einer standardisierten Erfassung dienen.

Empfehlung 6: Administrative Indikatoren haben schlechte Sensitivität , klinische Indikatoren sind keine Diagnosen

Administrative Indikatoren sollten wegen ihrer mangelnden Sensitivität nur in P4P-Konzepten mit Einzelleistungsvergütungs-Charakter verwendet werden (s. Empfehlung 4). Wenn man auf administrative Indikatoren setzt, läuft man die Gefahr einer verstärkten Mengenausweitung, außerdem wird der Qualitätsgedanke geschwächt, weil diese Indikatoren nicht sensitiv genug sind, also die Ereignisse, die abgebildet werden

sollen, nur unvollständig wiedergeben. In der Expertenorganisation wird Qualitätserfassung als *management-business* angesehen. Bei der Verwendung von klinischen Indikatoren ist der Erhebungsaufwand einzuplanen (und in der Höhe der Anreize zu berücksichtigen), weiterhin ist auf den Unterschied von epidemiologischen Falldefinitionen (s. nosokomiale Infektionen) zu klinischen Diagnosen hinzuweisen.

► **Prozess- vs. Ergebnisindikatoren:** Zur Abwägung zwischen Prozess- bzw. Ergebnisindikatoren ist in dieser Arbeit bereits mehrfach Stellung genommen worden (Kap. 1.3., 2.2., 2.4., 2.5., 4.2. bzw. 1.3., 2.4., 4.2.). Die großen laufenden P4P-Projekte in den USA (*Value-Based Purchasing* (VBP), s. Ryan et al. 2012C) und in Großbritannien (*Quality and Outcome Framework* (QOL), s. NHS 2013A) verwenden fast ausschließlich Prozessindikatoren (s. Kap. 2.4., 2.5.). Natürlich ist z.B. auch in den USA der Ruf nach Ergebnisindikatoren deutlich vernehmbar, da ihre Relevanz hoch ist (Rosenthal et al. 2007), aber die Problematik ihres Einsatzes (in erster Linie Risikoadjustierung und strukturelle Benachteiligung kleiner Häuser) ist bekannt (Cannon 2006, Davidson et al. 2007, Nicholson et al. 2008, Scott und Ward 2006). Die Leistungsanbieter können Prozess und evtl. auch Struktur beeinflussen, wegen der Risikoverteilung aber weniger die Ergebnisse, daher reagieren sie weniger auf Ergebnisindikatoren (Frolich et al. 2007). Ungeachtet der entsprechenden Beiträge der *principal-agent* Theorie (s. Empfehlung 21), wo die Thematik wieder aufgenommen wird, sollte der Schwerpunkt bei den Prozessindikatoren liegen, häufig wechselnd und koordinative Prozesse betonend (z.B. in der Behandlung von chronischen Erkrankungen, transsektorale Koordination). Besonders ist hier auf die Notwendigkeit vorangehender strategischer Überlegungen zu verweisen, da diese Indikatoren sinnvoll positioniert werden müssen (vgl. Handlungsfeld Politik). Für die weitere Entwicklung ist der Einsatz von Ergebnisindikatoren nicht ausgeschlossen, wenn sich die Praxis von P4P eingespielt hat und die methodischen Probleme erfolgreich bearbeitet wurden; besonders Outcome-relevante Patientensicherheits-Indikatoren (z.B. vermeidbare nosokomiale Infektionen) sind hier zu diskutieren (s. Entwicklung in Maryland, Calikoclu et al. 2012). Die Validität des Indikators Mortalität im Krankenhaus in seinen unterschiedlichen Spielarten (30-Tage, standardisiert) ist auf Systemebene sinnvoll (Heller 2010), auf der Ebene der einzelnen Einrichtungen jedoch schlecht, insbesondere kleine Häuser sind benachteiligt (Davidson et al. 2007, Scott und Ward 2006). Indikatoren der Strukturqualität sind meist weitgehend bekannt, ein reiner Investitionszuschuss wäre sinnvoller (s. Kap. 4.2.).

Empfehlung 7: Prozessindikatoren sind entscheidend, Ergebnisindikatoren können ergänzen

Im Mittelpunkt sollten wie in den großen internationalen Referenzprojekten (USA, UK) Prozessindikatoren stehen, da bei Ergebnisindikatoren die Problematik der Risikoselektion zu gewichtig ist, und selbst wenn die Risikoadjustierung perfekt entwickelt wäre, kleine Einrichtungen systematisch schlechter gestellt werden. Für die Zukunft sind jedoch Ergebnisindikatoren nicht völlig ausgeschlossen (in erster Linie Outcome-relevante Patientensicherheits-Indikatoren (z.B. Katheterinfektionen).

► **Risikoadjustierung:** Die Leistungsanbieter stehen unter dem Eindruck, dass die natürliche Varianz höher ist als der Einfluss einer eventuellen Qualitätsverbesserung (Pronovost und Kazandjian 1999). Aus dieser Sicht ist es nicht unverständlich, dass eine Risikoselektion zu den Verhaltensoptionen unter einer Qualitäts-orientierten Vergütung gehört. Es besteht dabei die Gefahr, dass der Anreiz zur Risikoselektion verstärkt wird, der bereits vom zugrundeliegenden Vergütungssystem ausgeht, in das P4P integriert ist. Um so weitergehend die Pauschalierung ist (bis hin zum Übergang der Versicherungsfunktion auf die Leistungserbringer bei *Managed Care*), um so mehr profitieren die Anbieter von der Vermeidung von Patienten mit teuren Erkrankungen bzw. hoher Komorbidität (vgl. Kap. 5.3.). Man kann durch die Wahl von Prozessindikatoren und/oder eine adäquate Risikoadjustierung von Ergebnisindikatoren gegensteuern. Die Risikoadjustierung findet jedoch ihre Grenzen bei Erkrankungen mit geringerer Fallzahl, bei denen es zu einer Benachteiligung von kleineren Einrichtungen kommt (s. Kap. 1.3.). Erkrankungen und Behandlungen ohne die Möglichkeit zur Risikoselektion (Unfälle oder Geburten bei Aufnahmezwang, Impfungen) sind davon ausgenommen. Eine weitere Möglichkeit ist die Einrichtung eines *exception reporting* (s. Kap. 2.5.), das es unter bestimmten Bedingungen möglich macht, Patienten aus dem P4P-Programm herauszunehmen, insbesondere wenn Patienten-seitige Gründe für ein Nicht-Erreichen des angestrebten Ziels vorliegen (NHS 2013A). Diese Regelung soll einer Risikoselektion und Schlechterstellung (bzw. niedrigerer Attraktivität) von Praxen in sozial unterprivilegierten Wohngebieten entgegenwirken. Wenn keine oder eine ungenügende Risikoadjustierung durchgeführt wird, besteht die Gefahr einer daraus resultierenden Unterversorgung als schwerwiegender unerwünschter Nebeneffekt der P4P-Einführung (s. Kap. 2.6.).

Empfehlung 8: Risikoselektion bedeutet Unterversorgung

Besonders bei zugrundeliegenden Vergütungsansätzen (z.B. Pauschalen), die bereits

ihrerseits zur Risikoselektion neigen, ist eine zusätzliche P4P-bedingte Risikoselektion geeignet, eine relevante Unterversorgung zu verursachen. Zumindest soweit Ergebnisindikatoren eingesetzt werden, ist daher eine möglichst optimale Risikoadjustierung notwendig. Prozessindikatoren bedürfen im allgemeinen keiner Risikoadjustierung.

► **Gaming:** Das *gaming* durch Mängel der Dokumentation und Fälschung der Daten ist von anderen Formen der mangelnden Reliabilität abzugrenzen (s. Kap. 2.6.2.) und betrifft die Organisationskultur bzw. die Fähigkeit zum organisatorischen Lernen (Maisey et al. 2008). Regelmäßige, nicht angemeldete Stichproben und andere Reliabilitätsüberprüfungen sind nicht verzichtbar.

Empfehlung 9: Sobald Indikatoren finanziell relevant werden, sind effektive Stichproben notwendig

Eine P4P-Einführung ohne die gleichzeitige Implementierung von Stichproben zur Kontrolle des *gaming* ist nicht sinnvoll.

► **Adressaten des Feedback:** Zwei Komplexe sind hier zu diskutieren. Für ein funktionierendes Feedback ist es einerseits von größter Bedeutung, die Adressaten der Belohnungskomponente richtig auszuwählen, mit anderen Worten Verantwortung und Anreiz zur Deckung zu bringen (vgl. Kap. 3.5.1.). Die entsprechende Literatur scheint sich auf den ersten Blick mehrheitlich dafür auszusprechen, die für die Behandlung verantwortlichen Ärzte direkt anzusprechen (Petersen et al. 2013, Torchiana et al. 2013), die nähere Analyse der Arbeiten zeigt aber, dass sich diese Empfehlungen auf die ambulante Versorgung beziehen, im Krankenhaus dagegen eher die Gesamtorganisation als Adressat präferiert wird (Berenson 2013, Rolnick et al. 2002). Bei persönlichem Anreiz wird eine Verstärkung der Risikoselektion befürchtet (Werner et al. 2011). Der zweite Komplex geht jedoch dahin, ob es in einem als desintegriert und stark sektoriert bezeichneten Gesundheitssystem grundsätzlich überhaupt möglich ist, einzelne Ärzte bzw. Abteilungen zusätzlich Qualitäts-bezogen zu vergüten, weil die Abgrenzung des Verantwortungsbereichs zwischen den einzelnen Leistungserbringern kaum machbar erscheint (Cannon 2006). Diese Schwierigkeit wird bei Verwendung von Summenscores noch verstärkt (Tompkins et al. 2009). Dieser Punkt ist insofern sehr wichtig, als dass die in dieser Arbeit präferierten Prozessindikatoren mit Schwerpunkt Koordination und Behandlung chronischer Erkrankungen naturgemäß mehrere Versorger und Versorgungssektoren betreffen. Es muss bei der Wahl der Indikatoren also darauf hingewirkt werden, dass die die finanziellen Anreize auch die notwendigen

organisatorischen Voraussetzungen betreffen, angefangen von der Schaffung gemeinsamer Dokumentation (IT) bis hin zu gemeinsamer Führung der Patienten. In den USA werden diesbezüglich *bundled payment* und die *accordable care organizations* diskutiert (Ryan et al. 2012C). Dass dies die schon genannten strategischen Vorüberlegungen voraussetzt, ist offensichtlich.

Empfehlung 10: Die organisatorischen Voraussetzungen einer verbesserten Integration der Versorgung fördern

Da in einem stark sektorierten Gesundheitssystem Verantwortungsbereich und Anreiz nur schwer in Deckung zu bringen sind, müssen die Indikatoren (Prozessindikatoren) auch organisatorische Aspekte einer verstärkten Integration (IT, gemeinsame Verantwortung für die Patienten) mit einbeziehen.

► **Wirkung des Feedback - Zeitnähe:** Es ist keine Frage, zu den Grundregeln eines funktionierenden Feedback gehört die Zeitnähe der Rückkopplung, in unserem Beispiel P4P also die Zeitnähe der Zahlung des P4P-Bestandteiles der Vergütung. In einer Studie werden z.B. mit sehr guten Ergebnissen 6 Monate gewählt (Torchiana et al. 2013). Einer der Gründe bezieht sich auf die Diskontierung, die nicht nur um so höher ausfällt, je länger der Zeitraum ist, über den diskontiert wird, sondern auch je höher die Unsicherheit der entgeltigen Zahlung ist (zur Risikoaversion s. Kap. 7.6., s. auch Damberg et al. 2007, Torchiana et al. 2013). Ein weiterer Aspekt betrifft die Überschätzung relativer Risiken: wenn die qualitäts-orientierte Vergütung z.B. im Gesamterlös untergeht, übt er eine sehr viel geringere Anreizwirkung aus als wenn es sich um eine eigenständige Zahlung handelt (s. Kap. 7.6.).

Empfehlung 11: Feedback nur zeitnah gestalten!

Der Feedback muss zeitnahe gestaltet sein, am besten außerhalb der normalen Erlösvergütung, so dass die P4P-bezogene Vergütung klar als Anreiz erkennbar ist und nicht von Zeitferne und Unsicherheit geschmälert wird.

► **Ceiling:** Sowohl im HQIP-Programm in den USA (Ryan et al. 2011) als auch im QOF-Pilotprojekt in Großbritannien gab es Hinweise auf ausbleibende Wirkungen von P4P durch Erreichen der maximal erreichbaren Performance. In der Evaluation von Campbell et al. (2007) war in einer *interrupted time-series*-Analyse die Qualitätsverbesserung insbesondere bei den Krankheitsbildern Asthma und Typ 2-Diabetes von dem langjährigen säkularen Trend einer langsamen auch ohne Intervention sichtbaren Verbesserung deutlich abzugrenzen, dieser Effekt war aber in

den Jahren danach nicht mehr nachweisbar (Campbell et al. 2009). Dieser Effekt werden als *ceiling*-Effekt interpretiert und z.B. eine Anhebung der Grenzwerte empfohlen (Doran und Roland 2010). Verstärkt wird dieser Effekt durch nicht aktuelle Indikatoren (s. Kap. 7.6.) und Aufhebung der Informationsasymmetrie, weil dann ein zusätzlicher Mengenanreiz auftritt (s. Kap. 7.6.).

Empfehlung 12: Ein *ceiling*-Effekt muss bedacht und verhindert werden

Ein *ceiling*-Effekt tritt vor allem bei Übergang in Einzelleistungsvergütung und bei nicht aktuellen Indikatoren auf.

Weiter: 7. Empfehlungen für die zukünftige Nutzung von P4P, 7.5. Organisation und System beachten!

7.5. Organisation und System beachten!

P4P ist ein komplexes Feedback-Instrument mit lerntheoretischem Hintergrund, das im Gesundheitssystem eingesetzt wird, das einerseits als Organisationsform besonderen Zuschnitts (Expertenorganisation), andererseits als umfassendes System mit hochgradiger Komplexität zu beschreiben ist. Für die Kombination dieser beiden Konzepte wird in der vorliegenden Arbeit der Begriff der "komplexen professionellen Systembürokratie" verwendet (s. Kap. 3.4., zur Definition s. Tableau 25a). Neben einer Unsichtbarkeit der Regeln, der ausgeprägten Autonomie und der Tendenz zur Selbstorganisation ist dieses System besonders durch das Innovations-Paradoxon und die intrinsische Unsicherheit gekennzeichnet. Das Innovations-Paradoxon beschreibt die Gleichzeitigkeit von Innovationsresistenz gegenüber äußeren Anforderungen und Innovationsnähe im *operating core* der Experten. Unter intrinsischer Unsicherheit wird in dieser Arbeit die Toleranz von Unsicherheit sowohl in Expertenorganisationen als auch in komplexen Systemen verstanden. Aus diesem Blickwinkel betrachtet, erstaunt es kaum, dass ein Feedback-Instrument wie P4P, wenngleich Belohnungs-bewehrt, als isolierte Intervention im Gesundheitswesen kaum eine Chance hat, die angestrebte Qualitätsverbesserung zu erreichen - die Veränderungen werden vom System quasi "verschluckt", zumindestens in dem Fall, dass notwendige flankierende Maßnahmen ausbleiben. Aus diesem Grund - nämlich weitere flankierende Weichenstellungen zu identifizieren - ist es sinnvoll, sich mit einigen Aspekten eingehender zu beschäftigen:

- Professionalismus als Chance?
- Organisatorisches Lernen fördern
- *Motivation crowding out*
- Value statt Qualität
- Attraktoren

► **Professionalismus:** Die Veränderungsresistenz des Gesundheitswesens hat in den letzten 20 Jahren innerhalb und außerhalb des Gesundheitswesens breite Aufmerksamkeit erhalten. Die bisherigen Ansätze zur Beförderung von notwendigen Veränderungen bedienten sich einer Erweiterung der verhaltenstheoretischen Konzepte um Elemente der sozialen Wahrnehmung betreffend Einstellungen, Haltungen, Rollenverständnis und Professionalismus (z.B. Greco und Eisenberg 1993, Bero et al. 1998, Grimshaw et al. 2004). So stellen Interventionen wie Leitlinien und *Evidence-Based Medicine* als externe Wissensbasen neue Anforderungen an die sozialen Rollen in den Mittelpunkt, der Typus des impliziten Heilers wird abgelöst durch den expliziten Experten, Rollenveränderungen zwischen den Berufsgruppen im Gesundheitswesen haben Hochkonjunktur (vgl. SVR 2008). Die Thematisierung ökonomischer

Gesichtspunkte durch externe Akteure (z.B. die politische Ebene) wird forciert, hat letztendlich aber nur zum Vorwurf der "Ökonomisierung" geführt, als ob die Berufsgruppen im Gesundheitswesen nicht auch schon vorher ausgeprägte Erfahrungen mit Situationen der Knappheit gemacht hätten. Nun wird auch bei der Thematik P4P an den "Professionalismus" appelliert, insbesondere der Ärzteschaft.

Wenn man sich diese Situation vor dem Hintergrund der organisatorischen und Systemeigenschaften betrachtet, die hier ausgeführt wurden, erscheint es außerordentlich unwahrscheinlich, dass der Rekurs auf Ansätze der sozialen Wahrnehmung wie Rollenveränderungen und Professionalismus Erfolg haben werden, obwohl sie durchaus als Voraussetzungen einer Veränderung angesehen werden müssen. Diese Aussage gilt trotz der unzweifelhaften Vorteile gerade für die Berufsgruppen im Gesundheitswesen, die sich aus P4P ergeben, denn was wäre sinnvoller, nach Qualität statt nach Stückzahl bezahlt zu werden? Es muss eine gewisse Analogie zur Diskussion um EBM konstatiert werden, wo die Differenzen zwischen den beiden Positionen Handwerk vs. Wissenschaft (s. Kap. 3.2.2.) auf der professionellen Ebene nicht gelöst werden konnte, obwohl sich speziell für die Ärzteschaft die einmalige Chance ergeben hätte, Einfluss auf ein wichtiges, Wissenschafts-gestütztes Instrument der Allokation von Mitteln zu erhalten. Allerdings kommt es zu Verschiebungen in den beruflichen Rollen, insbesondere ändert sich im Rahmen einer P4P-Einführung die Arbeitsteilung zwischen Ärzten und Pflege. Bei der Einführung von P4P kommt es nach den empirischen Daten durch die Übernahme vorher ärztlicher Tätigkeiten durch die Pflege dort zu zusätzlicher Belastung und Unzufriedenheit (Kurtzman et al. 2011, Maisey et al. 2008, McDonald et al. 2007). Eventuell wird die Teamarbeit erschwert (Maisey et al. 2008), wenngleich im Allgemeinen eine Verbesserung berichtet wird. Aus diesem Grund werden dezidiert multiprofessionelle und multidimensionale Implementierungsstrategien für P4P empfohlen (Timmermanns und Mauck 2005), insbesondere Veränderungen der Rollen und der interprofessionellen Zusammenarbeit betonend, die schon die unten beschriebenen Aspekte des organisatorischen Wandels mit einbeziehen.

Empfehlung 13: Multiprofessionelle und multidimensionale Implementierung ist zu bevorzugen

Die monoprofessionelle Beschränkung des Professionalismus-Konzeptes auf interne Motivation, Altruismus und Autonomie hat zu keiner tragfähigen Strategie geführt. Die Implementierung von P4P sollte einem multiprofessionellen Ansatz folgen und primär die Rollenverständnisse der beteiligten Berufsgruppen weiterzuentwickeln versuchen, ohne die Miteinbeziehung von Konzepten des organisatorischen Wandels und des

Kontextlernens wird jedoch kein Erfolg zu erzielen sein.

► **Organisatorisches Lernen fördern:** Es ist unumgänglich, wie in Empfehlung 3 bereits angesprochen, das lerntheoretische Instrument P4P nicht nur um Aspekte der sozialen Wahrnehmung (Rollen, Professionalismus), sondern auch um Elemente des organisatorischen Wandels und um Kontext-bezogene Modelle zu ergänzen. Als Beispiel aus der Vergangenheit kann die Bildung von Zentren herangezogen werden, wo - bei aller Kritik - doch im Krankenhausbereich und darüber hinaus Prozesse in Gang kamen, die einem organisatorischen Wandel mit Veränderungen der Verantwortung und Wertevorstellungen nahekommen. Allerdings werden organisatorische Ansätze im Gesundheitswesen nur sehr zurückhaltend eingesetzt, wahrscheinlich um der Angst vor Einschränkungen der professionellen Autonomie keine Nahrung zu geben. Gleichwohl ist dieser Aspekt des organisatorischen Lernens ein wichtiges Argument dafür, die Organisation als Ganzes zum Adressaten für die P4P-Zahlungen zu machen, und nicht einzelne Personen (s. Empfehlung 10).

Empfehlung 14: Die Organisation als Ganzes ansprechen, den organisatorischen Wandel fördern

P4P setzt die Organisation unter "internen Stress", es müssen jetzt nämlich zusätzlich zu Kosten und Mengen-bezogenen Erlösen auch Erlös-wirksame Qualitätsinformationen intern bearbeitet werden, und zwar sowohl hinsichtlich der Abgrenzung von Verantwortung als auch von Aufgabenstellungen und in der Führungsarbeit.

► **Motivation crowding out:** Grundsätzlich muss bei jedem Vergütungssystem, das im beruflichen Kontext angewandt wird, mit Auswirkungen auf Motivation, Rolle und Zusammenarbeit gerechnet werden (Prendergast 1999, s. auch Kap. 3.5.1.). Zur Motivation der Ärzte besteht die wichtigste Befürchtung darin, dass die interne Motivation (professionelle Einstellung), durch die externe Motivation i.S. einer Qualitäts-bezogene Vergütung gemindert wird ("*crowding out*"; Berenson et al. 2013, Cassel und Jain 2012). Empirisch ist dies jedoch nicht zu belegen, wie Untersuchungen z.B. im QOL-Projekt in Großbritannien zeigen (McDonald et al. 2007), und auch von theoretischer Seite gibt es Gegenargumente (Prendergast 1999, Staehle 1999, S. 242). Dieses häufig geäußerte Argument ist also nicht so durchschlagend, wie man es auf den ersten Blick meint. In diesem Zusammenhang ist die Diskussion um den Zielpunkt von P4P interessant, u.U. ist es sinnvoller, die Effizienz (value) statt Qualität allein anzureizen.

Empfehlung 15: *Motivation Crowding Out* - im Blick behalten, aber nicht überschätzen

Es ist nach den empirischen Ergebnissen und aus theoretischer Sicht unklar, ob sich bei P4P externe und interne Motivation gegenseitig schwächen. Das Problem sollte begleitend untersucht werden (Befragungen etc.).

► **Value statt Qualität:** Gegenwärtig wird in Deutschland davon ausgegangen, dass ein P4P-Programm im Gesundheitswesen allein der Qualität der Versorgung Rechnung tragen muss ("*quality-based purchasing*", Cannon 2006). In den USA wurde jedoch durch das VBP-Programm, das im Jahr 2010 im Rahmen des *Patient Protection and Affordable Care Act* beschlossen wurde und seit Ende 2012 implementiert ist (seit 1.10.2013 sind die finanziellen Regelungen im Krankenhausbereich aktiv), ein Anreizsystem geschaffen, das nicht nur die Qualität, sondern die Effizienz der Versorgung zum Gegenstand hat (Brown et al. 2014, Chien und Rosenthal 2013, CMS 2011A, CMS 2011B, Ryan et al. 2012). Dieser Effizienz-Gedanke wird als Value ("*quality in relation to the cost of care*", Tompkins et al. 2009, s. auch Damberg et al. 2009) bezeichnet und wird bereits seit Anfang der 90er Jahre in den USA intensiv diskutiert (Wenzel 1992). Gegenwärtig ist Value auch der Kernbegriff der gegen die Überversorgung mit nutzlosen bzw. schädlichen Verfahren gerichteten *Choosing Wisely Initiative* der *National Physicians Alliance* in den USA (Morden et al. 2014).

Es wäre auch in Deutschland eine wichtige Diskussion, welche Vorteile es haben könnte, auch in Deutschland primär diesen Ansatz zu wählen. Dagegen spricht wohl in erster Linie die Annahme, insbesondere bei Ärzten würde eine externe finanzielle Motivation die *a priori* gegebene interne Motivation (zu den Begriffen s. Kap. 3.5.1.) schwächen (zum *motivation crowding out* s.o.). Abgesehen davon, dass es dafür kaum empirische und theoretische Argumente gibt, wäre auch zu überlegen, ob es nicht gerade die ausschließliche Orientierung solcher externen Anreize auf die Qualität ist, die Widerstand erzeugt. Bei aller Bezugnahme auf die professionellen und altruistischen Motivationsebenen bleibt ja eines klar: Ärzte und die anderen Berufsgruppen sehen es schon als richtig an, sie adäquat zu bezahlen. Von daher ist es durchaus wahrscheinlich, dass die Kopplung von P4P an Parameter nach dem Muster "Kosten für Qualität" eine gute Akzeptanz erfährt.

Empfehlung 16: *Value* (Effizienz) statt allein Qualität als Zielgröße für P4P-Programm in Betracht ziehen

Das VBP-Programm in den USA setzt auf *value* (Effizienz) als Zielgröße für P4P und

damit evtl. eine bessere Akzeptanz erreichen, denn die Kosten der Versorgung sind ein etabliertes Kriterium für die Außenbeziehungen der Gesundheitseinrichtungen.

► **Attraktoren:** Komplexe Systeme orientieren sich an sog. Attraktoren (s. Kap. 3.3.1. und 6.2.), von außen nicht erkennbare Systemzustände, die eine relative Stabilität aufweisen, in deren Richtung sich die Systeme entwickeln. Wenn man über *framing* (s. Kap. 7.6.) und *direction pointing* (s. Kap. 7.7.) nachdenkt, erscheint es sinnvoll, über Attraktoren in der komplexen professionellen Systembürokratie zumindest Hypothesen aufzustellen. Wichtig ist der Altruismus der Berufsgruppen im Gesundheitswesen, dem Menschen zu helfen, ist der Antrieb vieler dort Tätigen (**A**) (Timmermanns und Mauck 2005). Ein zweiter Attraktor besteht sicherlich in professioneller Autonomie und Kollegialität, ganz wie im Konzept der Expertenorganisation beschrieben (**B**) (Avorn und Fischer 2010, Cook et al. 2004). Als weitere Attraktoren kommen Wissenschaftsorientierung (**C**) und die finanzielle Honorierung im Sinne der Wertschätzung in Frage (**D**) (Frolich et al. 2007). Und letztendlich: es soll organisiert sein, die Supportfunktionen müssen stimmen (**E**).

Diese hypothetische Aufzählung hat ihre Bedeutung, solange system- und komplexitätstheoretische Konzepte weiter verfolgt werden, um Entwicklungen im Gesundheitswesen zu erklären bzw. vorherzusagen. Nach den Botschaften der Komplexitätstheorie, sind die Effekte oft überraschend. Als Beispiel mag die große Akzeptanz der Patientensicherheitsthematik in den Gesundheitsberufen gelten, die nur über die Attraktoren Altruismus und Steigerung der professionellen Autonomie zu erklären ist. Wenn man über die Implementierung von P4P-Programmen nachdenkt, ist die finanzielle Honorierung sicher ein wichtiges Argument.

Empfehlung 17: Attraktoren des komplexen Systems diskutieren, Attraktoren nutzen

Es ist lohnenswert, sich getreu der Konzepte der System- bzw. Komplexitätstheorie Gedanken über mögliche Attraktoren zu machen und sie bei der Implementierung von P4P-Programmen zu nutzen.

Weiter: 7. Empfehlungen für die zukünftige Nutzung von P4P, 7.6. Ökonomie und Vergütungssystem

7.6. Ökonomie und Vergütungssystem

Selbst wenn P4P als perfektes, "erweitertes" Feedback verstanden wird, so wie in den vorangehenden Kapiteln dargestellt, kann dieses Instrument im isolierten Einsatz keine grundlegenden Umkehrungen bestehender Anreize im Gesundheitswesen bewirken, sondern bedarf, so wie in Kap. 4 ausführlich dargestellt, zusätzlicher ökonomischer (und außerdem politischer) Weichenstellungen. Diese Weichenstellungen betreffen neben einfachen ökonomischen Überlegungen (z.B. Opportunitätskosten) vor allem Konsequenzen aus der *principal agent* Theorie (Informationsasymmetrie) und der Verhaltensökonomie (z.B. Risikoaversion) sowie Wechselwirkungen von P4P mit dem dominierenden Vergütungssystem, denn man muss grundsätzlich davon ausgehen, dass P4P-Elemente in ein anderes Vergütungssystem (z.B. DRG) integriert werden (s. Tableau 37).

- Tableau 37: Ökonomische Ausgangsüberlegungen** (s. Kap. 4)
- Grundlegende Überlegungen (vgl. Kap. 4.2.)
 - Opportunitätskosten
 - Diskontierung
 - *Principal agent* Theorie (vgl. Kap. 4.2.)
 - Informationsasymmetrie
 - Verhaltensökonomie (*behavioural economics*) (vgl. Kap. 4.3.)
 - *framing*, insbes. Risikoaversion
 - globale Einschätzung
 - *isolation effect*
 - Kombination mit dominierendem Vergütungssystem (vgl. Kap. 5.), insbes.
 - Einzelleistungsvergütung
 - Sektorale Pauschalen
 - Transssektorale Pauschalen
 - Populationspauschalen

Ganz im Vordergrund stehen Fragen der **Informationsasymmetrie**, die für das Konzept von P4P als durchaus kritisch anzusehen sind und die Berechtigung des P4P-Ansatzes in Abgrenzung von der Einzelleistungsvergütung potentiell in Frage stellen: "*Is P4P really Fee for Service?*" (Wodchis et al. 2007). Behandlungsverfahren und Qualitätsaspekte, die sich allseitiger Kenntnis erfreuen und (z.B. durch *Evidence-Based Medicine*) abgesichert sind, bedürfen keiner *Pay for Performance*, sondern können direkt durch die Finanzierungsseite vergütet werden. Besondere Bedeutung kommt in diesem Zusammenhang der Abwägung zwischen Prozess- und Ergebnisindikatoren sowie, weitergehend, der Auswahl derjenigen Prozessindikatoren zu, die der strategischen Zielsetzung des P4P-Programmes adäquat sind.

Im weiteren zeigt die **Verhaltensökonomie**, dass es ohne zutreffende Annahmen zur Risiko- oder Verlustaversion kaum möglich ist, durch P4P eine sinnvolle Anreizbildung zu erzeugen; wohlgedenkt: einen Anreiz erzielt man immer, aber es ist in diesem Zusammenhang ja entscheidend, dass dieser in der Kombination mit den anderen Anreizen, die im System wirksam sind, in vernünftiger Form die Schwächen des Systems mindert und die Stärken stärkt. Dies führt zur Frage der **Integration von P4P in das zugrundeliegende Vergütungssystem** und zu den Wechselwirkungen mit dem dominierenden Vergütungssystemen, denn P4P wird nie der alleinige, dominierende Vergütungsansatz sein (s. Kap. 5). So ist es sicher nicht möglich, mit P4P ein eng sektoral begrenztes, mit hochtouriger Volumenexpansion einhergehendes Vergütungssystem wie das des klassischen DRG-Systemes zu beherrschen - hier ist eine grundsätzliche Weiterentwicklung notwendig. Aber vielleicht kann P4P den Fokus leicht verschieben, z.B. weg von den Akuterkrankungen und operativen Behandlungsmethoden hin zur kontinuierlichen Therapie der schon jetzt und in Zukunft noch mehr dominierenden chronischen, multiplen Erkrankungen einer alternden Bevölkerung - und zur Prävention.

Zu folgenden Aspekten wird auf dieser Basis hier Stellung genommen:

- Höhe der P4P-Vergütung einschl. Opportunitätskosten und Diskontierung (vgl. Kap. 4.2.)
- Frequenz, Stückelung und Zahlungsmodus (vgl. Kap. 4.2.)
- Monetäre Kopplung (vgl. Kap. 4.2.)
- Informationsasymmetrie und *principal-agent* Theorie (vgl. Kap. 4.2.)
- Aktualität und Wechsel der Indikatoren
- Wettbewerbsdichte und *dual use* (vgl. Kap. 4.3.)
- Verhaltensökonomie - *framing* (vgl. Kap. 4.3.)
- Verlässlichkeit und Nachvollziehbarkeit
- Integration in Vergütungssysteme (vgl. Kap. 5.)

► Zur **Höhe der P4P-Vergütung** fällt schnell ins Auge, dass in den bislang evaluierten Programmen besonders im Krankenhaus (HQIP in den USA) die zusätzlich ausgeschütteten Mittel mit weniger als 50.000.- € pro Jahr eher im niedrigen Bereich liegen, wenn man sie mit den Gesamtvergütungen vergleicht (Werner et al. 2011B). Im QOF-Programm in Großbritannien handelt es sich um die gleiche Größenordnung, allerdings auf der Ebene einer Arztpraxis (Roland 2004, Gillem et al. 2012). Die Höhe der P4P-Vergütung stellt bei Einführung von P4P eine der meist diskutierten Fragen dar und muss daher schlüssig beantwortet werden. Man darf sich dabei nicht nur auf die

Aufwendungen bzgl. des Messaufwandes (z.B. Erhebung der Daten, Bereitstellung von Personal), die der dokumentierenden Einrichtung entstehen, beschränken, sondern man muss die Opportunitätskosten mit einbeziehen, die die entgangenen Erlösmöglichkeiten bei alternativer Verwendung der Mittel wiedergeben. Die Opportunitätskosten sind im Gesundheitswesen sehr für das zugrundeliegende Vergütungssystem, in das P4P "eingebettet" ist, empfindlich, besonders wenn von diesem ein starker Mengenanreiz ausgeht (z.B. Einzelleistungsvergütung, DRG) und die Grenzkosten niedrig sind. Es ist in diesem Fall wahrscheinlich, dass eine zusätzliche Erlöseinheit (zusätzlicher Fall), die evtl. nur mit variablen und ohne zusätzliche sprungfixe Kosten zu realisieren ist, einer P4P-Vergütung vorgezogen wird (Damberg et al. 2009). Hinzu kommt die Diskontierung des Qualitäts-bezogenen Erlösbestandteiles (Frolich et al. 2007). Die Mehrzahl der Analysen plädiert daher für höhere P4P-Vergütungen (Chung et al. 2010, Ryan und Blustein 2011, Werner et al. 2011B, Werner und Dudley 2012).

Empfehlung 18: Höhe der P4P-Zahlungen muss besonders im DRG-System Opportunitätskosten berücksichtigen

Die Höhe des Qualitäts-bezogenen Erlösanteiles muss den Messaufwand, die Opportunitätskosten und die Diskontierung berücksichtigen und hängt daher stark von der Art des dominierenden Vergütungssystems ab. Bei Einzelleistungsvergütung und DRGs sind hohe Opportunitätskosten anzusetzen, die Höhe der P4P-Vergütung muss entsprechend hoch angesetzt werden.

► **Frequenz, Stückelung und Zahlungsmodus:** Ein intakter Feedback-Mechanismus vorausgesetzt (s. Kap. 7.4.), bei dem die Zahlungen zeitnahe erfolgen (s. Empfehlung 11), kann der Feedback noch verbessert werden, wenn die sog. Überschätzung relativer Risiken berücksichtigt wird, ein Phänomen, das in der Verhaltensökonomie beschrieben ist (s. Kap. 4.3., s. Damberg et al. 2007 und Verma et al. 2014). Häufigere und dafür kleinere Zahlungen mit on/off-Charakteristik sind wirksamer als Zahlungen, die integriert in andere, umfangreichere Zahlungen (z.B. Gesamterlös) erfolgen, insbesondere wenn sie Ereignis-bezogen vorgenommen werden.

Empfehlung 19: Kleine und häufigere Zahlungen sind zu präferieren

Kleine, häufigere und Ereignis-bezogene P4P-Zahlungen mit on/off-Charakteristik sind größeren integrierten Zahlungen vorzuziehen.

► **Monetäre Kopplung:** Eine P4P-Vergütung besteht aus zwei Bestandteilen, einem System der Qualitätsmessung und einem System der monetären Bewertung (s. Kap.

4.1.). Die Kopplung der Vergütungsbestandteile an die Qualitätsindikatoren muss praktikabel (administrabel), nachvollziehbar (verständlich), justiziabel und von seinen Anreizwirkungen sinnvoll gestaltet werden. Das inhaltliche Hauptproblem der monetären Kopplung besteht aber darin, dass sich die Verbesserung nicht nur bei den Einrichtungen manifestieren sollte, die schon vorher zu den Besten gehörten, während sich der Abstand zu den *poor performer* weiter vergrößert (s. z.B. HQIP-Programm, Lindenauer et al. 2007). Diese Gefahr besteht vor allem bei der Verwendung von relativen Positionen auf Rankinglisten ("die besten 5"). Die sog. *poor performer*, evtl. sogar die Mehrheit der teilnehmenden Institutionen, können verloren gehen, weil sie kaum Chancen sehen, das Ziel zu erreichen (sog. *goal gradient*). Absolute Grenzwerte können Abhilfe schaffen, aber auch nur dann, wenn sie gestaffelt eingesetzt werden, so dass die Grenzwerte auch durch die *poor performer* erreicht werden können. Keinesfalls kann aber auf die Honorierung relativer Verbesserungen ("von Position A auf Position B um X Plätze") verzichtet werden. Die drei Kopplungsmechanismen (s. Abb. 11) sollten daher immer in Kombination eingesetzt werden.

Empfehlung 20: Relative Position, relative Verbesserungen und absolute Grenzwerte kombinieren

Die monetäre Kopplung der P4P-Vergütung kann nicht alleine aufgrund der relativen Position auf einer Rankingliste vorgenommen werden, sondern muss ebenso gestaffelte Grenzwerte und relative Positionsverbesserungen mit einbeziehen, damit auch die *poor performer* einen realistischen Anreiz zur Qualitätsverbesserung haben.

► **Informationsasymmetrie und *principal-agent* Theorie:** Obwohl Anreize wie P4P grundsätzlich dazu geeignet sind, den Informationsnachteil *ex post* auszugleichen, den die Finanzierungsseite *ex ante* gegenüber dem Leistungserbringer hat, verliert P4P den Charakter eines Anreizes, wenn sich bei gut abgesicherten Leistungen die Informationsasymmetrie abschwächt, und mutiert damit zu einer zusätzlichen Zahlung im Sinne der Einzelleistungsvergütung (Werner und Dudley 2012, Wodchis et al. 2007). Dies hat zwei Konsequenzen: einerseits die Gefahr der Vernachlässigung anderer, nicht angereizter Leistungen (*teaching to the test*, s. Kap. 2.6.3.) und andererseits der Rückgang der Qualität nach Beendigung des P4P-Programmes, weil der Qualitätsrelevante Leistungsbestandteil ja nicht mehr vergütet wird - so beobachtet im QOL-Programm in Großbritannien (Lester et al. 2010, s. Kap 4.2.). Dieser Umstand ist weniger bei **Ergebnisindikatoren** möglich, ungeachtet der Problematik Risikoadjustierung und strukturelle Benachteiligung kleiner Häuser (Cannon 2006, Davidson et al. 2007, Nicholson et al. 2008, Scott und Ward 2006). Der hier diskutierte

Effekt tritt in erster Linie bei **Prozessindikatoren** auf, vor allem wenn über ihren Einsatz allgemeine Einigkeit besteht und sie z.B. durch EBM gut abgesichert sind (Nicholson et al. 2008). Anders ist es in Situationen, in denen die Informationsasymmetrie nicht aufgehoben ist, z.B. bei der Behandlung chronischer Erkrankungen, der Prävention von nosokomialen Komplikationen (z.B. Händedesinfektionsmittelverbrauch) oder im Bereich Patientensicherheit. Um so mehr ist auf die Wichtigkeit vorangehender strategischer Überlegungen zu verweisen, da diese Indikatoren sinnvoll positioniert werden müssen (vgl. Handlungsfeld Politik). Für die weitere Entwicklung ist der Einsatz von Ergebnisindikatoren nicht ausgeschlossen, wenn die Praxis von P4P sich eingespielt hat und die methodischen Probleme erfolgreich bearbeitet wurden; besonders Outcome-relevante Patientensicherheits-Indikatoren (z.B. vermeidbare nosokomiale Infektionen) sind hier zu diskutieren (s. Entwicklung in Maryland, Calikoclu et al. 2012). Die Validität des Indikators Mortalität im Krankenhaus in seinen unterschiedlichen Spielarten (30-Tage, standardisiert) ist auf Systemebene vielleicht sinnvoll (Heller 2010), auf der Ebene der einzelnen Einrichtungen jedoch schlecht, insbesondere kleine Häuser sind benachteiligt (Davidson et al. 2007, Scott und Ward 2006). Indikatoren der **Strukturqualität** sind meist weitgehend bekannt sind, eine Informationsasymmetrie liegt nicht vor, ein reiner Investitionszuschuss ist sinnvoller (s. Kap. 4.2.).

Empfehlung 21: Besonders bei Prozessindikatoren Übergang zur Einzelleistungsvergütung beachten!

Konsequenzen aus der *principal-agent* Theorie und damit Überschneidungen zur Einzelleistungsvergütung gehören zu den schwierigsten Fragestellungen, mit denen sich das P4P-Konzept auseinandersetzen hat. Die wichtigste Einsicht besteht darin, dass die Einzelleistungsvergütung überlegen ist, wenn es sich um gut bekannte, wissenschaftlich abgesicherte Prozessindikatoren handelt. P4P ist nur sinnvoll, wenn eine Informationsasymmetrie besteht, d.h. das eigentliche Wissen um die Gestaltung der qualitativ angestrebten Leistung bei den Leistungserbringern liegt. Dies ist z.B. bei Indikatoren aus dem Grenzbereich zwischen Ergebnis- und Prozessindikatoren (z.B. zur Prävention vermeidbarer Komplikationen), bei der Behandlung chronischer oder Mehrfacherkrankungen oder bei prozessualen Patientensicherheits-Indikatoren der Fall. Bei Ergebnisindikatoren ist die Informationsasymmetrie meist kein Problem, hier stellt sich allerdings mit Macht die Problematik der Risikoselektion. Strukturindikatoren sind als Investitionszuschuss anzusehen.

Konkrete Beispiele: Ein Prozessindikator wie die Abnahme einer Blutkultur vor Antibiotikagabe bei der ambulant erworbenen Pneumonie ist für P4P nur schlecht

geeignet, eine direkte Vergütung ist sinnvoller. Prozessindikatoren zur Koordination z.B. wie Vollständigkeit der Informationsweitergabe sind sinnvoll, denn das Geschehen "dahinter" ist für die Finanzierungsseite nicht zu durchblicken, und sie dienen im Übrigen strategischen Zielen (s.o.). Outcome-nahe Prozessindikatoren wie der Mitarbeiter-bezogene Händedesinfektionsmittelverbrauch oder Readmission-Indikatoren sind ebenfalls sinnvoll, genauso Indikatoren, die prozessuale Aspekte der Patientensicherheit beschreiben (z.B. Zahl der bearbeiteten Meldungen im CIRS).

► **Aktualität und Wechsel der Indikatoren:** Eine weitere Konsequenz aus den Überlegungen der *principal-agent* Theorie ist die Notwendigkeit, Indikatoren hoher **Aktualität** einzusetzen (Sorbero et al. 2006, Torchiana et al. 2013) bzw. die Indikatoren regelmäßig zu wechseln. Hat man sich erst einmal an die Indikatoren "gewöhnt", geht von ihnen keine Anreizwirkung mehr aus, daher ist es sinnvoll, diese spätestens alle drei Jahre neu festzulegen (Werner et al. 2011, Jha et al. 2012, Ryan et al. 2012). Zur Gewährleistung von Aktualität und regelmäßigem Wechsel der Indikatoren empfiehlt sich die Delegation der Indikatorenentwicklung an eine spezialisierte Institution, wie es in Großbritannien mit dem NICE geschehen ist (Gillem et al. 2012), und wie es in Deutschland mit dem neuen Qualitätsinstitut geplant ist.

Empfehlung 22: Indikatoren müssen aktuell sein und regelmäßig gewechselt werden

Um Aktualität und Wechsel zu gewährleisten, ist die Delegation an ein Institut wie dem geplanten Qualitätsinstitut sinnvoll.

► **Wettbewerbsdichte** und **dual use**: Auf der Basis der hier diskutierten ökonomischen Konzepte (s. Tableau 37) muss auf die Frage der Kombination mit anderen Qualitätsinitiativen eingegangen werden, vor allem wenn sie Wettbewerbscharakter haben (z.B. *public reporting*). Ganz allgemein ist davor zu warnen, eine allzu hohe Vielzahl von parallelen, um die Aufmerksamkeit der Leistungsanbieter konkurrierenden Qualitätsinitiativen zu initiieren, weil dies zu einem Wirkungsverlust führen kann (*overwhelming*, s. Ryan et al. 2011). Spezifischere Fragen ergeben sich jedoch, wenn die Qualitätsindikatoren, die Gegenstand der P4P sind, vorher oder parallel bereits durch andere Anreizsysteme adressiert werden. In der ersten Linie ist hier das *public reporting* von Bedeutung (s. Kap. 4.3.), die Kombination mit P4P wird als *dual use* bezeichnet (zu den entsprechenden Erfahrungen im HQIP- und im VBP-Programm in den USA s. Werner et al. 2011 und Ryan et al. 2011). Auch der Sachverständigenrat hatte die Empfehlung abgegeben, P4P und *public reporting* zu kombinieren (SVR 2008, Nr. 748 Ziffer 4). Diese Empfehlung bietet sich auf den ersten

Blick ja auch unbedingt an: warum Indikatoren, die sich bei *public reporting* schon bewährt haben, nicht auch bei P4P weiter nutzen? Trotzdem kann diese Empfehlung aus heutiger Sicht aus Sicht des Autors der vorliegenden Arbeit nicht (mehr) gehalten werden, und zwar aus zwei Gründen:

1. *public reporting* hebt die **Informationsasymmetrie** auf, wenn es der P4P-Einführung vorgelagert ist, beiden Seiten ist bereits bekannt, "wo der Hund begraben liegt", und eine eigentliche (zusätzliche) Anreizwirkung kann nicht mehr erzeugt werden - ganz abgesehen davon, dass bereits durch einen *ceiling*-Effekt (s. Empfehlung 12) das mögliche Verbesserungspotential schon ausgeschöpft sein kann (Damberg et al. 2009). Diese Situation war in Massachusetts zu beobachten, wo *public reporting* seit 2004 implementiert war und P4P 2009 zusätzlich eingeführt wurde; es war kein zusätzlicher Effekt durch P4P zu beobachten (*Massachusetts Medicaid's Hospital-Based P4P Programm*, s. Ryan et al. 2011).

2. beide Systeme haben eine unterschiedliche Wirkung je nach **Wettbewerbsdichte**. In beiden Fällen handelt es sich um ein finanzielles Anreizsystem (vgl. Frolich et al. 2007), im Fall von *public reporting* durch indirekte Anreize (z.B. Einweiserverhalten), im Fall von P4P durch eine direkte Anreizwirkung durch die Qualitäts-bezogenen Zahlungen (Werner et al. 2011). Der indirekte Anreiz durch *public reporting* wirkt besonders in hoch-kompetitiven Märkten (z.B. im städtischen Umfeld), weniger jedoch in nicht-kompetitiven Märkten (z.B. ländliche Regionen), wo man als Einrichtung nicht befürchten muss, durch Qualitätsinformationen Nachteile zu erfahren. Im ersteren Fall kann der direkte Anreiz durch P4P kaum noch weitere Wirkung entfalten, in wettbewerbsschwachen Märkten stellt P4P jedoch das wirkstärkere Prinzip dar, da der direkte finanzielle Stimulus auch unabhängig von der Wettbewerbsdichte wirkt.

Diese Überlegungen führen folglich nicht nur zu einer kritischen Haltung gegenüber dem gleichzeitigen Einsatz der beiden Instrumente des Qualitätswettbewerbes, sondern weisen dem Instrument P4P auch präferentielle Einsatzgebiete entsprechend der Ausprägung des Wettbewerbs in Teilmärkten zu: P4P ist besonders wirksam, wenn in nicht-kompetitiven Regionen andere, indirekt Wettbewerbs-wirksame Instrumente nicht ausreichen, z.B. in ländlichen Gegenden. In hoch-kompetitiven Regionen ist *public reporting* u.U. das wirksamere Instrument, zumindest erzeugt P4P keine zusätzliche Qualitätsverbesserung. Diese Argumentation kann evtl. auch auf die fachliche Ebene ausgeweitet werden, so wird vereinzelt diskutiert, ob man P4P nicht differenziert nach Krankheits- bzw. Fallgruppen einsetzen kann (Werner et al. 2011). Hoch-kompetitive Fallgruppen wie Herzkatheterinterventionen oder operative Elektivtherapien wären aus dieser Sicht nicht in dem Maße ein sinnvolles Einsatzgebiet für P4P wie weniger

kompetitiv ausgerichtete Fachgebiete, z.B. die Behandlung von chronisch Kranken.

Empfehlung 23: Keine unhinterfragte Kombination mit *public reporting*, beide Instrumente sind nicht synergistisch

Entgegen der bisherigen Ansicht ist die Kombination mit *public reporting (dual use)* äußerst kritisch zu sehen (*ceiling*, Aufhebung der Informationsasymmetrie, komplementäre Wirkung hinsichtlich Wettbewerbsdichte), eher ist ein differenzierter Einsatz beider Instrumente sinnvoll. P4P ist im Vergleich zu *public reporting* besonders wirksam in Bereichen (Regionen, Krankheitsgruppen), in denen die Wettbewerbsdichte gering ist. Ein nach Wettbewerbsdichte differenzierter Einsatz nach Fachgebieten ist ein interessanter Ansatz, der in der Zukunft zu diskutieren wäre.

► **Verhaltensökonomie - *framing***: Entscheidungen werden in einem referentiellen Zusammenhang getroffen, der die Interpretation erleichtert oder sogar erst ermöglicht. Dies betrifft z.B. die sog. Risiko-Aversion ("der Spatz in der Hand ist besser als die Taube auf dem Dach", s. Damberg et al. 2007), die zu einer höheren Diskontierung führt: weil die Gefahr gesehen wird, dass ein Ereignis in der Zukunft nicht eintritt, obwohl es mehr Nutzen erbringen könnte, wenn es tatsächlich eintritt, wird lieber ein geringerer Nutzen in der Gegenwart realisiert. Gerade bei der Kopplung der P4P-Vergütung an die Qualitätsindikatoren spielt das eine große Rolle, wenn mit relativen Positionen ("die besten 10") gearbeitet wird, denn in diesem Fall hängt die P4P-Vergütung nicht nur von der eigenen Leistung, sondern ebenso von der Leistung der Wettbewerber ab, erst am Schluss ergibt sich die eigene Position. Falls diese Sicherheit zu gering ist, wird insbesondere bei starker Risiko-Aversion der Nutzen zu stark diskontiert (Zusammenstellung s. Cannon 2006, zu den entsprechenden Befunden im HQIP-Projekt s. Jha et al. 2012, Ryan et al. 2012B). Dies gilt auch für den Fall, dass die (technische) Berechnung der Anreize nicht klar ist (Ryan et al. 2011), wenn man sozusagen "nicht weiß, woran es liegt". Genauso sind Unterschiede zu beobachten, je nachdem ob eine Entscheidung als möglicher Verlust (*loss-frame*) oder als möglicher Gewinn (*gain-frame*) interpretiert wird. Ein leistungsabhängiger Gehaltsanteil gleicher Größe wird höher eingeschätzt, wenn das Gesamtgehalt am Jahresende ausgezahlt und der variable Anteil bei Nicht-Erreichen der Ziele am Jahresende abgezogen wird, als wenn er am Jahresende erst gewährt wird (Frolich et al. 2007). Personen sind meist Verlust-avers, d.h. ein entgangener Gewinn wird weniger hoch bewertet als ein gleich großer Verlust (Werner und Dudley 2012). Die Komponenten eines adäquaten *framing* werden sehr gut durch das Mindscape-Instrument beschrieben (s. Abb. 13). Für die Einführung eines kritischen Instrumentes wie P4P sind also authentische *messenger (opinion leader)*

genauso notwendig wie erfolgreiche Pilotprojekte (s. HQIP in den USA), Wertschätzung der Akteure bzw. Gesundheitsberufe statt “*bad apples*” und vor allem gesellschaftlich Rahmenbedingungen, die Normen und Kontext der Einführung dieses neuen Instrumentes wiedergeben (s. Kap. 7.7.).

Empfehlung 24: Berücksichtigung von Risiko- und Verlustaversion sowie optimales *framing* erforderlich

Die Einführung von P4P muss auf die Risiko- und Verlustaversion der Einrichtungen Bezug nehmen, der Diskontierung entgegenzutreten und die positiven Aspekte (Qualitätsverbesserung, professionelle Autonomie) in den Vordergrund stellen. Das *framing* kann nicht ohne gesellschaftliche bzw. politische Rahmenbedingungen geschehen, die die Notwendigkeit und die Perspektiven eines Wechsels der Vergütungslogik als allgemeinen Konsens wiedergeben.

► **Verlässlichkeit und Nachvollziehbarkeit:** Die Zahlung der P4P-Vergütung muss nicht nur zeitnah (s. Kap. 7.4.), sondern wegen der Diskontierungstendenz verlässlich und im Zustandekommen nachvollziehbar erfolgen. Die Gesundheitsberufe und einrichtungen können nur dann einen sinnvoll Feedback erkennen, wenn sie verstehen können, wie eine Qualitäts-bezogene Vergütung ausgelöst wird und gestaltet ist.

Empfehlung 25: Verlässliche und nachvollziehbare Zahlungen

Das Verständnis für das Zustandekommen der P4P-Zahlungen verbessert die Wirksamkeit des Feedback.

► **Integration in Vergütungssysteme:** P4P wird grundsätzlich im Rahmen anderer Vergütungssysteme eingesetzt (s. Kap. 6). Getreu der Maßgabe “Qualität statt Menge!” kommt P4P gerade in Gesundheitssystemen in die Diskussion, die durch Mengen-orientierte Vergütungssysteme wie DRGs oder Einzelleistungsvergütung geprägt sind. Allerdings kann P4P aus mehreren Gründen (z.B. hohe Opportunitätskosten) den Mengenanreiz nicht ausgleichen. Es besteht sogar die Gefahr, dass bei Einzelleistungsvergütung und sektoralen Pauschalen der Mengenanreiz durch P4P verstärkt wird, und zwar wenn bei Leistungen, bei denen die Möglichkeit zur Mengenausweitung besteht (z.B. Endoprothetik) Indikatoren mit geringer Sensitivität verwendet werden, sich mit den resultierenden Qualitätsdaten eine Mengenausweitung begründen lässt und eine aktive Risikoselektion möglich erscheint. Sieht man vom Mengenanreiz ab, können bei Einzelleistungsvergütung und sektoraler Pauschalierung jedoch interessante Einsatzmöglichkeiten für P4P darin bestehen, die Koordination der

Behandlung und überhaupt die Behandlung von chronischen Erkrankungen zu verbessern. Man gleicht damit den Nachteil chronischer, konservativer Erkrankungen aus, der dadurch entsteht, dass die ökonomische Anreize bei diesen Vergütungsformen eher auf operativ zu behandelnde Akuterkrankungen gerichtet sind, und greift damit eines der dringenden Qualitätsprobleme auf. Gleiches gilt für auch für andere Themen wie Patientensicherheit (z.B. Indikatoren zur Einführung und sinnvollen Handhabung von Instrumenten wie CIRS). Bei der integrierten transsektoralen Versorgung, Erkrankungspauschalen oder *Managed Care* tritt der Mengenanreiz zurück und die Koordination der Behandlung und die Versorgung von Patienten mit chronischen, multiplen Erkrankungen mehr in den Vordergrund, so dass man mit einem gezielten Einsatz von P4P sinnvoll eingreifen und Schwerpunkte in der Qualitätsentwicklung setzen kann (s. QOF-Projekt in Großbritannien). Analog lässt sich zum Thema Prävention vorgehen, hier ist es grundsätzlich möglich, den vor allem durch Diskontierungsaspekte gehemmten Einsatz der Prävention (s. Abb. 8, Kap. 3.5.1.) zu fördern.

Empfehlung 26: Integration in andere Vergütungssysteme kritisch und differenziert vornehmen

P4P kann grundsätzlich nicht die dominanten Anreize der Vergütungssysteme ausgleichen, in die es "eingebettet" wird. Dies gilt in Deutschland vor allem für das DRG-System im Krankenhausbereich. Allerdings ist es möglich, eine vorgeschaltete strategische Analyse vorausgesetzt, durch einen differenzierten Einsatz von P4P Bereiche zu fördern, die vom DRG-System vernachlässigt werden. Hier ist insbesondere an die Versorgung chronisch, mehrfach erkrankter Patienten zu denken, weiterhin an die Koordination über die Sektorgrenzen hinweg und die Prävention.

Weiter: 7. Empfehlungen für die zukünftige Nutzung von P4P, 7.7. Politische Verantwortung

7.7. Politische Verantwortung

Die korporatistischen Strukturen der Selbstorganisation, in Deutschland unter dem Begriff der Selbstverwaltung zusammengefasst, sind in den letzten beiden Jahrzehnten immer wichtiger geworden, parallel zum Wandel des politischen Grundverständnisses weg von einem hierarchischen Modell zu Konzepten wie dem *Governance*-Konzept (s. Kap. 6.1.). Dieser Wandel ist funktional, entspricht er doch auf der organisatorischen Ebene der Expertenautonomie, auf der Systemebene der Komplexität des Gesundheitssystems, auf der Ebene der (notwendigen) Verhaltensänderung den Kontext-bezogenen Theorien und hinsichtlich der ökonomischen Grundannahmen den verhaltensökonomischen Erkenntnissen.

Durch diese Verbreiterung der Regelungskompetenz in der Peripherie (und die Minderung des Gegensatzes Zentralität vs. Dezentralität) lässt sich jedoch nicht die Notwendigkeit politisch gesetzter Rahmenbedingungen "hinwegdiskutieren". Expertenorganisation und Komplexitätstheorie, integriert im verwendeten Rahmenkonzept, weisen den Institutionen und den professionellen Strukturen der Selbstorganisation eine wichtige Rolle zu, aber zu den Koordinationsmechanismen einer Gesellschaft gehören neben Markt, Organisationen und Gemeinschaften eben auch die Hierarchie (Mayntz 2010). In Kap. 6 sind auf diesem Hintergrund vier Grundkompetenzen und -aufgaben der politischen Ebene identifiziert worden, die nur von dieser Ebene wahrgenommen werden können. Neben dem *direction pointing* spielt hier die strategische Analyse und das daraus resultierende Setzen strategischer Ziele der Gesundheitspolitik eine große Rolle, weiterhin die Kontrolle eventueller negativer Auswirkungen einer Innovation wie P4P, und viertens die Etablierung von Rahmenbedingungen, die einer Innovation erst die "Luft zum Atmen" geben. Für diese vier Bereiche werden hier zum Abschluss dieser Arbeit konkrete Empfehlungen ausgesprochen:

- *Direction pointing* (s. Kap. 6.2.)
- Strategische Ziele setzen (s. Kap. 6.3.)
- Negative Auswirkungen kontrollieren (s. Kap. 6.3.)
- Rahmenbedingungen schaffen (s. Kap. 6.3.).

► ***Direction pointing*** gehört zur kleinen Zahl von Regelungsmechanismen, die in komplexen Systemen sinnvoll anzuwenden sind (s. "*Crossing the Quality Chasm*" des IOM (2001), Plsek 2003). Empirische Daten zeigen, dass die Akteure im Gesundheitswesen eine solche "Richtungsweisung" als Instrument der Koordination durchaus schätzen (Rosenthal 2007), wenn sie in einer glaubwürdigen Form geschieht.

Gerade wenn es - wie bei P4P - um wettbewerbs-bezogene Elemente geht (vgl. Kap. 3.2.2., auch Edwards 2005), kann die Bezugnahme auf mangelnde Qualität irritieren, wenn andere, gleichzeitig wirksame Anreize (z.B. Mengenanreiz) dominant sind und die Qualitätserwartung konterkarieren. Wie groß gerade unter den fortbestehenden DRG-Bedingungen die Möglichkeiten sind, ein solches Instrument zu umgehen, ist in dieser Arbeit in zahlreicher Hinsicht beschrieben worden (Opportunität, Diskontierung etc.). Daher steht fest, dass im politischen Kontext die Zielsetzung, mehr für die Qualität der Gesundheitsversorgung der Bevölkerung tun zu wollen, ganz unübersehbar im Mittelpunkt stehen muss, anders ist eine derart eingreifende Umstellung im Gesundheitswesen nicht zu realisieren.

Eine "stumme" Gesundheitspolitik ist also keine Option. Die führenden Gesundheitspolitiker müssen sich persönlich und aktiv für "mehr Qualität" einsetzen, es müssen dafür Foren geschaffen werden, ein passives Geschehen-lassen reicht nicht aus. Den professionellen Fachgesellschaften und Institutionen, die sich mit den Themen Qualität und Patientensicherheit auseinandersetzen und teilweise Berufsverbands-Charakter haben, müssen öffentlichkeitswirksame Unterstützung erfahren. Getreu dem *Governance*-Konzept ist die aktive Einbeziehung von Medien und Einrichtungen aus der Zivilgesellschaft notwendig (s. unten). In Kap. 6.2. wurde auf die verstärkte Förderung der Ausbildung in den Bereichen Ethik, Ökonomie und Qualität/Patientensicherheit für die Angehörigen der Gesundheitsberufe als Möglichkeit hingewiesen, eine authentische Politik "pro Qualität" zu unterlegen. Wie es in Deutschland schon fast vorbildlich geschehen ist, müssen die Strukturen der Selbstverwaltung aktiv eingebunden und beauftragt werden. Allerdings darf das nicht dazu führen, dass die politische Ebene sich aus der Verantwortung zurückzieht. Gerade bei der strategischen Zielsetzung und auch bei finanziell wirksamen Beschlüssen (z.B. Identifizierung von Indikatoren für P4P) ist nicht zu erwarten, dass die Selbstverwaltung aus ihrer Interessen-geleiteten Perspektive die für die Gesamtgesellschaft optimalen Entscheidungen treffen kann.

Empfehlung 27: *direction pointing* als elementares Mittel zur Implementierung nutzen

Direction pointing weist wichtige koordinative und unterstützende Eigenschaften auf, ohne die eine Innovation wie ein P4P-Programm nicht erfolgreich eingeführt und umgesetzt werden kann. *Direction pointing* setzt einen aktiven und sichtbaren Einsatz der politischen Ebene voraus, der Qualität als Richtschnur der gesundheitspolitischen Agenda betont, der alle Ebenen des Gesundheitssystems einbezieht, öffentliche Foren schafft, die mit der Thematik beschäftigten Fachgesellschaften fördert, die Ausbildung entsprechend differenziert und Kontakte zu Einrichtungen der Zivilgesellschaft schafft.

► **Strategische Ziele setzen:** Die Notwendigkeit, vor Einführung eines P4P-Systems sich aus strategischer Perspektive über die Ziele und Schwerpunkte einer solchen Initiative klar zu werden, ist in der vorliegenden Arbeit immer wieder betont worden (z.B. Kap. 4.2. und Empfehlung 1). P4P ist ein Instrument, das Feedback mit finanziellen Anreizen verbindet, es ist aufwendig und kann durchaus auch negative Auswirkungen haben - also sollte man es gezielt, Problem-orientiert und nur in Bereichen einsetzen, an denen eine Wirksamkeit auch wirklich zu erwarten ist. Solche strategischen Überlegungen sind Bestandteil der Vorbereitungen der P4P-Einführung in den USA (s. gesundheitspolitischer Kontext des VBP-Programms Kap. 2.3.) und in Großbritannien (s. Kap. 2.5.) gewesen und werden natürlich um so wichtiger, um so differenzierter man das Instrument P4P einsetzen will (z.B. bessere Koordination über die Sektorgrenzen (Daughtridge et al. 2014) oder die optimale Versorgung der Bevölkerung, insbesondere der *underserved populations* (Rosenthal et al. 2007, Brown et al. 2014)). Bei einer "Weiter so"-Strategie, bei der man z.B. einfach die Indikatoren aus der Qualitätssicherung nach §137 SGB V fortschreibt und finanziell bewertet, sind weitergehende strategische Überlegungen nicht notwendig, schon allein deswegen, weil eine derartige P4P-Initiative aller Voraussicht nach vollständig wirkungslos bleiben wird (s. in erster Linie Empfehlungen 21, 22 und 23). Auch einer P4P-Implementierung, die ausschließlich oder überwiegend auf Ergebnisindikatoren setzt, wird wegen der Risikoselektion (s. Kap. 2.6.1.) und der strukturellen Benachteiligung kleiner Einrichtungen kein Erfolg beschieden sein. Die Verwendung von Prozessindikatoren bietet sich an, allerdings nur in definierter Form, die strategische Überlegungen dringend notwendig macht (Berenson et al. 2013). Wie an anderer Stelle hervorgehoben, besteht hier einer der entscheidenden Problemstellungen eines aktuellen P4P-Konzeptes, denn wegen der Gefahr der Aufhebung der Informationsasymmetrie sind auch nicht alle Prozessindikatoren für P4P geeignet, insbesondere nicht diejenigen, die Gefahr laufen, als Einzelleistungsvergütung zu fungieren (s. Kap. 4.2., Kap. 5.2. und Kap. 7.6.). Um also den im deutschen Gesundheitssystem herrschende Orientierung an Mengen und (operativ zu behandelnden) Akuterkrankungen zu begegnen und die Koordinationsdefizite durch die ausgeprägte Sektorierung zu vermindern (s.o., z.B. Schoen et al. 2011), muss man Prozessindikatoren ohne Risikoselektion und ohne Aufhebung der Informationsasymmetrie (s.o.), identifizieren, die dann gerade für die Versorgung chronisch und mehrfach Erkrankter sowie die Versorgung älterer Patienten ihre Bedeutung haben. Andere Indikatoren können sich z.B. mit der Unterversorgungsproblematik in ländlichen Regionen beschäftigen, man muss hier jedoch diskutieren, ob die angestrebten Ziele nicht eher durch direkte Unterstützung und Einzelleistungsvergütung anzugehen wären. Diese strategischen Ziele zu setzen, ist

eine genuin politische, gestaltende Aufgabe, die zwar die Unterstützung der Selbstverwaltungspartner bedarf, in ihrer Ausgestaltung jedoch in den Händen des Gesetzgebers liegen muss, nicht zuletzt deswegen, weil evtl. nicht unerhebliche Steuermittel aufgebracht werden müssen.

Empfehlung 28: Strategische Ziele setzen I - kein “weiter so”

Aufgrund der Analysen dieser Arbeit kann eine Weiternutzung von langjährig verwendeten Indikatoren nicht empfohlen werden, genauso wenig wie ein *dual use* mit Indikatoren aus dem *public reporting* (s. Empfehlung 23), weil P4P in dieser Konstellation nach den vorliegenden Daten keine Wirksamkeit aufweist. Da der Einsatz von Ergebnisindikatoren wegen der Risikoselektion und der strukturellen Benachteiligung kleinerer Einrichtungen nicht unproblematisch ist, bleiben Prozessindikatoren, die nach einer vorangegangenen Analyse des Handlungsbedarfs ausgewählt werden.

Empfehlung 29: Strategische Ziele setzen II - Prozessindikatoren Problemorientiert auswählen

Nicht alle Prozessindikatoren sind sinnvoll bei P4P einzusetzen (s. Empfehlung 21). Es müssen diejenigen Prozessindikatoren identifiziert werden, bei denen die Informationsasymmetrie nicht aufgehoben ist (s.o.) und die übergeordneten Zielen in der strukturellen Weiterentwicklung des deutschen Gesundheitssystems dienen: Verbesserung der Versorgung von chronisch und mehrfach erkrankten Patienten hohen Alters, Abbau der Mengenorientierung, Verbesserung der Koordination, Förderung der Prävention und Integration der Sektoren. Diese Zielsetzung kann nur auf der politischen Ebene unter “hierarchischer Verantwortung” erfolgen.

► **Negative Auswirkungen kontrollieren:** Gerade wenn ein neues gesundheitspolitisches Instrument eingesetzt wird, besteht die Verantwortung der politischen Ebene in der Vermeidung bzw. Kontrolle etwaiger unerwünschter Auswirkungen. Diese Kontrolle kann in Zusammenarbeit mit Verbänden und anderen Partnern des Gesundheitswesens zusammen erfolgen, die Letztverantwortung liegt jedoch auf der Ebene der Politik. Die möglichen unerwünschten Auswirkungen der Einführung von P4P wurden zusammenfassend in Kap. 2.6. aufgeführt und diskutiert. Ganz im Vordergrund stehen Probleme des Zugangs (sozial benachteiligte Bevölkerungsgruppen, Kontinuität der Versorgung, Risikoselektion, vgl. z.B. Ryan 2013) und der Verzerrung durch Datenmanipulation. Die rechtzeitige Erkennung und Analyse von unerwünschten Auswirkungen ist eines der Hauptgründe für die

Evaluationsforschung, die bei der Einführung von P4P-Programmen in jedem Fall stattfinden und frühzeitig in den Implementierungsprozess integriert werden muss.

Empfehlung 30: Verschlechterung des Zugangs und Datenmanipulation kontrollieren

Die Politik verantwortet die *surveillance* und Kontrolle etwaiger negativer Auswirkungen einer P4P-Einführung wie Zugangsprobleme (sozial benachbarte Gruppen, Kontinuität, Risikoselektion) und Verzerrung durch Datenmanipulation. Frühzeitig im Implementierungsprozess müssen Evaluationen geplant und angestoßen werden.

► **Rahmenbedingungen schaffen:** Die Rahmenbedingungen bestehen aus Systemfaktoren und Umsetzungsparametern, die die möglichen Auswirkungen von P4P direkt oder indirekt fördern oder hemmen, und sie liegen in der Verantwortung der gesetzgeberischen Ebene, sie sind, um diesen Ausdruck im alten Sinne zu gebrauchen, ein "politisches Problem" (Miller 2009). Es handelt sich dabei nicht nur um reine Motivations- oder ökonomische Faktoren, wie z.B. konkurrierende ökonomische Anreize, die den Anreizen der P4P-Vergütung entgegengesetzt sind (z.B. Einzelleistungsvergütung). Ohne Frage sind dies ganz entscheidende Probleme, aber negative Rahmenbedingungen können auch darin bestehen, dass bei Umsetzung und Spezifikation eines Instrumentes wie P4P Fehler gemacht werden. Da es hier im wahrsten Sinne des Wortes um Geld geht, kann man z.B. nicht erwarten, dass auf der Ebene der "Systembürokratie", also der Selbstverwaltung, eine zwanglose Einigung zu den anzuwendenden Indikatoren zustande kommt, eher ist eine langdauernde Blockade mit langfristiger Ergebnislosigkeit zu erwarten. Die *Governance* findet also dort ihre Grenzen, wo Interessen zurückgeschnitten werden müssen, Beispiele für ein Versagen der "Verhandlungsdemokratie" gibt es zu Genüge (z.B. Mindestmengenindikatoren im Krankenhausbereich). Zu den Rahmenbedingungen gehört es auch, die Komplexität des Geschehens nicht aus den Augen zu verlieren. Es reicht nicht, an einer Schraube zu drehen (s. Kap. 1.5.). Die Implementierung eines Instrumentes wie P4P muss immer im Kontext anderer Interventionen gesehen werden, nicht nur der Rahmen, der durch die zugrundeliegenden Vergütungssysteme (z.B. DRG) gegeben wird, sondern auch die wirtschaftliche Situation der Krankenhäuser, die Veränderungen in der Trägerschaft, die Sektorisierung des Gesundheitswesens, die Morbidität etc.

Natürlich ist die Ausgestaltung der Institutionen im Bereich der Selbstverwaltung (z.B. Gemeinsamer Bundesausschuss) rational, weil nur so eine bessere Partizipation erreicht werden kann, aber es besteht die Gefahr der Blockade und Suboptimierung (Mayntz 2010). Hinzu kommt die Tendenz jeder Institution, Funktion mit Selbsterhaltung

und Selbstbeschäftigung zu verwechseln und letztendlich notwendige Entwicklungen zu blockieren. Es muss daher verstärkt darüber nachgedacht werden, wie die "Zivilgesellschaft", damit ist hier gemeint: die nicht mit dem Gesundheitswesen verbundenen Gesellschaftsgruppen, Berufe, Verbände, Unternehmen, in das Thema Qualität und Sicherheit eingebunden wird, ein Ansatz, der in den USA z.B. mit der *Leapfrog*-Initiative erfolgreich beschrritten wird (Galvin et al. 2005, Moran und Scanlon 2013). Das Thema Qualität und Sicherheit im Gesundheitswesen muss über den i.e.S. politischen Raum hinaus in den zivilgesellschaftlichen Bereich hineinwachsen, nur dann das *framing* funktionieren, nur dann hat das *direction pointing* Aussicht auf Erfolg.

In aller erster Linie bedeutet das Setzen von Rahmenbedingungen aber die Etablierung eines Handlungsrahmens, eines *conceptual frameworks*, wie es in den USA durch "*Crossing the Quality Chasm*" des IOM (2001) bzw. in Großbritannien durch *The NHS Outcome Framework 2013/2014* (NHS 2013B) etabliert werden konnte. Ein solches Rahmenkonzept ist handlungsleitend und für die Beteiligten letztlich handlungslegitimierend, eine wichtige Voraussetzung für die Einführung eines Instrumentes wie *Pay for Performance* - Qualitäts-orientierte Vergütung.

Empfehlung 31: Rahmenbedingungen als politische Aufgabe

Explizite Rahmenbedingungen bilden den Hintergrund für die Initiierung des politischen Prozesses, für die Implementierung, die Umsetzung und die Evaluation. Die Rahmenbedingungen betreffen die Koordination der Zuständigkeiten im Implementierungsprozess genauso wie das Setzen von Meilensteinen für die Akteure, die an der Einführung von P4P beteiligt sind, die Antizipation von Fehlentwicklung und hemmenden Faktoren, und die Integration in die anderen Entwicklungen des Gesundheitssystemes, mit denen P4P in Wechselwirkung tritt.

Literatur und Quellen

- Albrecht M (2013), Loos S, Otten M: Sektorenübergreifende Qualitätssicherung in der ambulanten Versorgung. *Z. Evid. Fortbild. Qual. Gesundh. wesen* 107, 2013, 528-33
- Alshamsan R (2010), Majeed A, Ashworth M, Car J, Millett C: Impact of Pay for Performance on Inequalities in Health Care: Systematic Review. *J. Health Serv. Res. Policy* 15, 2010, 178-84
- Angell M (1996), Kassirer JP: Quality and the Medical Marketplace - Following Elephants. *N. Engl. J. Med.* 335, 1996, 883-5
- Anonymous (1999): Getting Evidence into Practice. *Effective Health Care* 5 (1), 1999
- Anonymous (2011): Die Kunst der Qualitätsmessung. Titel des Schwerpunktheftes der Zeitschrift für Evidenz, Fortbildung und Qualität, 105, 2011, Heft 1
- Anonymous (2013): Deutschlands Zukunft gestalten. Koalitionsvertrag zwischen CDU, CSU und SPD, 18. Legislaturperiode, 27.11.2013
- Anonymous (2014): Fragwürdige Zahlen des Krankenhaus-Report 2014 - Krankenhäuser erwarten Entschuldigung der AOK. *Das Krankenhaus* 2/2014 S. 101ff
- AQUA-Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH (2013): Qualitätsreport 2012: Länderauswertung mit Geodarstellung für veröffentlichungspflichtige Indikatoren. Göttingen, 2013
- Arte-Sendung "Heilen wie am Fließband" v. 17.9.2013, hier z.B. Statements von G. Maio, Medizinerethiker aus Freiburg: "Was die evidenzbasierte Medizin letztendlich auf den Weg bringt, ist eine Standardisierung der Entscheidungen, so dass quasi der einzelne Mensch durch das Raster der Standardisierung fällt. Je mehr Sie sich orientieren an der EBM, desto mehr versuchen Sie, Schablonen herzustellen, und Sie passen dann nicht so sehr die Verfahren an den Patienten an, sondern Sie kategorisieren die Patienten, ..." und "Von den Ärzten wird im Grunde eine Vorhersagbarkeit verlangt, und da wird die EBM als Instrument benutzt, um quasi nur solche Entscheidungen zuzulassen, von denen man annehmen kann, dass sie gut eingetaktet werden können und in ein solches auf Effizienz orientiertes System eben gut hineinpassen."
- Azaouagh A (2008), Stausberg J: Häufigkeit der nosokomialen Pneumonie im Krankenhaus – Vergleich von rechnergestützter Basisdokumentation und papierbasierter Krankenakte. *Pneumologie* 62, 2008, 273-8
- Avorn J (2010), Fischer M: 'Bench to Behavior': Translating Comparative Effectiveness Research into Improved Clinical Practice. *Health Aff.* 29, 2010, 1891-1900
- Baker GR (2001): Healthcare Managers in the Complex World of Healthcare. *Frontiers of Health Services Management* 18, 2001, 23-32
- Bardach NS (2013), Wang JJ, De Leon SF, Shih SC, Boscardin WJ, Goldman LE, Dudley RA: Effect of Pay-for-Performance Incentives on Quality of Care in Small Practices With Electronic Health Records. A Randomized Trial. *Jama* 310, 2013, 1051-9
- Bhattacharyya T (2009), Freiberg AA, Mehta P, Katz JN, Ferris T: Measuring The Report Card: The Validity Of Pay-For-Performance Metrics In Orthopedic Surgery. *Health Aff.* 28, 2009, 526–532

- Behnke M (2013), Hansen S, Leistner R, Diaz LAP, Gropmann A, Sohr D, Gastmeier P, Piening B: Nosokomiale Infektionen und Antibiotika-Anwendung: Zweite nationale Prävalenzstudie in Deutschland. *Dtsch. Aerztebl.* 110, 2013, 627-33
- Belmont E (2011), Haltom CC, Hastings DA, Homchick RG, Morris L, Taitsman J, Peters BM, Nagele RL, Schermer B, Peisert KC: A New Quality Compass: Hospital Boards' Increased Role Under The Affordable Care Act. *Health Aff.* 30, 2011, 1282-9
- Benz A (2010), Dose N: Governance - Modebegriff oder nützliches sozialwissenschaftliches Konzept? In: Benz A, Dose N (Hrsg.): *Governance – Regieren in komplexen Regelsystemen*, 2. Aufl., Wiesbaden 2010, S. 13-36
- Berenson RA (2013), Pronovost PJ, Krumholz HM: Achieving the potential of health care performance measures: timely analysis of immediate health policy issues. Princeton, NJ: Robert Wood Johnson Foundation, May 2013 (<http://www.rwjf.org/en/research-publications/find-rwjf-research/2013/05/achieving-the-potential-of-health-care-performancemeasures.html>, access 24.2.2014)
- Bero LA (1998), Grilli R, Grimshaw JW, Harvey E, Oxman AD, Thomson MA, on behalf of the Cochrane Effective Practice and Organisation of Care Review Group. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *BMJ* 317, 1998, 465-468
- Bloom BS (2005): Effects of Continuing Medical Education on Improving Physician Clinical Care and Patient Health: A Review of Systematic Reviews. *Int. J. Techn. Assessm. Health Care* 21, 2005, 380-5
- Blustein J (2011), Weissman JS, Ryan AM, Doran T, Hasnain-Wynia R: Analysis Raises Question on Whether Pay-For-Performance In Medicaid Can Efficiently Reduce Racial and Ethnic Disparities. *Health Aff.* 30, 2011, 1165-75
- Broom A (2009), Adams J, Tovey P: Evidence-Based Health Care in Practice: A Study of Clinician Resistance, Professional De-Skilling, and Inter-Specialty Differentiation in Oncology. *Social Science and Medicine* 68, 2009, 182-200
- Brown JR (2014), Sox HC, Goodman DC: Financial Incentives to Improve Quality. Skating to the Puck or Avoiding the Penalty Box? *JAMA* 311, 2014, 1009-10
- Bundeszentrale für politische Bildung (2014): Dossier Gesundheitspolitik. Die wichtigsten Akteure im deutschen Gesundheitswesen, Teil 1 bis 3. Version 26.2.2014. <http://www.bpb.de/politik/innenpolitik/gesundheitspolitik/72565/staat-und-politik?p=all>, Zugriff 6.3.2014
- Cabana MD (1999), Rand CS, Powe NR, Wu AW, Wilson MH, Abboud P-AC, Rubin HR: Why Don't Physicians Follow Clinical Practice Guidelines? A Framework for Improvement. *JAMA* 282, 1999, 1458-65
- Calderwood M S (2014), Kleinman K, Soumerai SB, Jin R, Gay C, Platt R, Kassler W, Goldmann DA, Jha AK, Lee GM: Impact of Medicare's Payment Policy on Mediastinitis Following Coronary Artery Bypass Graft Surgery in US Hospitals. 35, 2014, 144-151
- Calikoglu S (2012), Murray R, Feeney D: Hospital Pay-For-Performance Programs in Maryland Produced Strong Results, Including Hospital-Acquired Conditions. *Health Aff.* 31, 2012, 2649-58

- Campbell S (2007), Reeves D, Kontopantelis E, Middleton E, Sibbald B, Roland M: Quality of Primary Care in England with the Introduction of Pay for Performance. *N. Engl. J. Med.* 357, 2010, 181-190
- Campbell SM (2009), Reeves D, Kontopantelis E, Sibbald B, Roland M: Effects of Pay for Performance on the Quality of Primary Care in England. *N. Engl. J. Med.* 361, 2009, 368-78
- Cannon MF (2006): Pay-For-Performance: Is Medicare a Good Candidate? *Yale Journal of Health Policy, Law, and Ethics.* 7, 2006, 1-38
- Casale AS (2007), Paulus RA, Selna MJ, Doll MC, Bothe AE, McKinley KE, Berry SA, Davis DE, Gilfillan RJ, Hamory BH, Steele GD: "ProvenCare": A Provider-Driven Pay-for-Performance Program for Acute Episodic Cardiac Surgical Care. *Ann. Surg.* 246, 2007, 613-23
- Casalino LP (2007), Elster A: Will Pay-For Performance and Quality Reporting Affect Health Care Disparities? *Health Aff.* 26, 2007, w405-14
- Cassell CK (2012), Jain SH: Assessing individual physician performance: does measurement suppress motivation? *JAMA* 2012;307:2595-6
- Centers for Medicare and Medicaid Services (2011A): Medicare program; hospital inpatient value-based purchasing program. Final rule. *Fed Regist.* 2011;76(88):26490–547
- Centers for Medicare and Medicaid Services (2011B): Roadmap for Implementing Value Driven Healthcare in the Traditional Medicare Fee-for-Service Program. https://www.cms.gov/QualityInitiativesGenInfo/downloads/VBPRoadmap_OEA_1-16_508.pdf (Zugriff 3.12.2013)
- Centers for Medicare and Medicaid Services, Health Services Advisory Group (2012): National impact assessment of Medicare quality measures. Baltimore: Centers for Medicare & Medicaid Services, March 2012 (<http://www.cms.hhs.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityMeasures/Downloads/NationalImpactAssessmentofQualityMeasuresFINAL.pdf>)
- Chandler C (1999): The Role and Education of Doctors in the Delivery of Health Care. *Lancet.* 353, 1999, 1178-81
- Chassin ME (2013): Improving the Quality of Health Care: What's Taking so Long? *Health Aff.* 32, 2013, 1761-65
- Chen C (2014), Ackerly DC: Beyond ACOs and Bundled Payments. Medicare's Shift Toward Accountability in Fee-for-Service. *JAMA* 311. 2014, 673-4
- Chew-Graham CA (2013), Hunter C, Langer S, Stenhoff A, Drinkwater J, Guthrie EA, Salmon P: How QOF is shaping primary care review consultations: a longitudinal qualitative study. *BMC Family Practice* 14, 2013, 103-10
- Chien AT (2013), Rosenthal MB: Medicare's Physician Value-Based Payment Modifier – Will the Tectonic Shift Create Waves? *New Engl. J. Med.* 269, 2013, 2076-8
- Chung S (2010), Palaniappan L, Wong E, Rubin H, Luft H: Does the Frequency of Pay for Performance Payment Matter? Experience from a Randomized Trial. *Health Serv. Res.* 45, 2010, 553-64
- Classen DC (2011), Resar R, Griffin F, Federico F, Frankel T, Kimmel N, Whittington JC,

- Frankel A, Seger A, James BC: 'Global Trigger Tool' Shows That Adverse Events in Hospitals May Be Ten Times Greater Than Previously Measured. *Health Aff.* 30, 2011, 581-9
- Cohn S (2013), Clinch M, Bunn C, Stronge P: Entangled Complexity: Why Complex Interventions Are Just Not Complicated Enough. *JHSRP* 18, 2013, 40-43
- Cook DJ (2004), Montori VM, McMullin JP, Finfer SR, Rocker GM: Improving patients' safety locally: changing clinician behaviour. *Lancet* 2004; 363: 1224-30
- Corrigan J (2009), McNeill D: Building Organizational Capacity: A Cornerstone of Health System Reform. *Health Aff.* 28, 2009, w205-15
- Damberg CL (2007), Sorbero M, Mehrotra A, Teleki S, Lovejoy S, Bradley L. An environmental scan of pay for performance in the hospital setting: final report [Internet]. Santa Monica (CA): RAND; 2007 Nov [cited 2014 Feb 21]. Available from: <http://aspe.hhs.gov/health/reports/08/payperform/PayPerform07.html>
- Damberg CL (2009), Raube K, Teleki SS, dela Cruz E: Taking Stock of Pay-For-Performance: A Candid Assessment From the Front Lines. *Health Aff.* 28, 2009, 517-25
- Davidson G (2007), Moscovice I, Remus D: Hospital Size, Uncertainty, and Pay-For-Performance. Upper Midwest Rural Health Research Center, Minnesota, Febr. 2007
- Daughtridge GW (2014), Archibald T, Conway PH: Quality Improvements of Care Transitions and the Trend of Composite Hospital Care. *JAMA* 311, 2014, 1013-4
- Dijkstra R (2006), Wensing M, Thomas R, Akkermans R, Braspenning J, Grimshaw J, Grol R: The Relationship Between Organizational Characteristics and the Effect of Clinical Guidelines on Medical Performance in Hospitals, a Meta-Analysis. *Health Serv. Res.* 6, 2006, 53-62
- Doran T (2006), Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hiroeh U, Roland M: Pay-for-Performance Programs in Family Practices in the United Kingdom. *N Engl J Med* 2006;355:375-84
- Doran T (2010), Roland M: Lessons from Major Initiatives to Improve Primary Care in the United Kingdom. *Health Aff.* 29, 2010, 1023-9
- Dudley RA (2006), Rosenthal MB: Pay for Performance: A Decision Guide for Purchasers. AHRQ Publication No. 06-0047, Rockville 2006
- Edwards N (2003): Doctors and Managers: Poor Relationships May Be Damaging Patients – What Can Be Done? *Qual. Saf. Health Care* 12 (Suppl), 2003, i21-i24
- Edwards N (2005): Doctors and Managers: Building a New Relationship. *Clin. Med.* 5, 2005, 577-9
- Eijkenaar F (2013), Emmert M, Scheppach M, Schöffski O: Effects of Pay for Performance in Health Care: A Systematic Review of Systematic Reviews. *Health Pol.* 110, 2013, 115-30
- Erasmus V (2010), Daha TJ, Brug H, Richardus JH, Behrendt MD, Vos MC, van Beeck EF: Systematic Review of Studies on Compliance with Hand Hygiene Guidelines in Hospital Care. *Infect. Control Hosp. Epidemiol.* 31, 2010, 283-94
- Fraser SW (2001), Greenhalgh T: Coping with Complexity: Educating for Capability. *Brit.*

- Med. J. 323, 2001, 799-803
- Freidson E (2001): Professionalism. The Third Logic. The University of Chicago Press, Chicago 2001
- Frolich A (2007), Talavera JA, Broadhead P, Dudley RA: A Behaviorable Model of Clinician Responses to incentives to Improve Quality. Health Policy 80, 2007, 179-93
- Galvin RS (2005), Delbanco S, Milstein A, Belden G: Has The Leapfrog Group Had An Impact On The Health Care Market? Health Aff. 24, 2005, 28-33
- Gastmeier P (2008), Geffers C: Nosokomiale Infektionen in Deutschland: Wieviele gibt es wirklich. Eine Schätzung für das Jahr 2006. Dtsch. Med. Wochenschr 2008, 133, 111-15
- Gillam SJ (2012), Siriwardena AN, Stelle N: Pay-for-Performance in the United Kingdom: Impact of the Quality and Outcomes Framework – A Systematic Review. Ann. Fam. Med. 10, 2012, 461-8
- Glickman SW (2007), Ou F-S, DeLong ER, Roe MT, Lytle BL, Mulgund J, Rumsfeld JS, Gibler WB, Ohman EM, Schulman KA, Peterson ED: Pay for Performance, Quality of Care, and Outcomes in Acute Myocardial Infarction. JAMA 297, 2007, 2373-80
- Graves N (2008), McGowan JE: Nosocomial Infection, the Deficit Reduction Act, and Incentives for Hospitals. JAMA 300, 2008, 1577-79
- Greco PJ, Eisenberg JM: Changing Physicians' Practices. N. Engl. J. Med. 329, 1993, 1271-4
- Greenhalgh T (2010), Plsek P, Wilson T, Fraser S, Holt T: Response to „Appropriation of Complexity Theory in Health Care. J. Health Serv. Res. Pol. 15, 2010, 115-7
- Grimshaw JM (2004), 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 Techn. Assessm. 8, 2004, Nr. 6
- Groene O: Does Quality Improvement Face a Legitimacy Crisis? Poor Quality Studies, Small Effects. J. Health Serv. Res. Pol. 16, 2011, 131-2
- Grol R (2003), Grimshaw J: From Best Evidence to Best practice: Effective Implementation of Change in Patients' Care. Lancet 362, 2003, 1225-30
- Grossbart SR (2006): What's the Return. What is the Effect of "Pay for Performance" Initiatives on the Quality of Care Delivery. Med. Care Res. Rev. 63, 2006, 29S-48S
- Heller G (2010) für das ATRÄK-Projektteam: Der Einfluss von strukturellen klinikspezifischen Variablen auf die Qualität der Krankenversorgung. In: KLAUBER, J., Geraedts, M., Friedrich, J. (Hrsg.): Krankenhaus-Report 2010. Schwerpunkt: Krankenhausversorgung in der Krise? Schattauer Verlag, Stuttgart, 2010, S. 255-74
- Hibbard JH (2012), Greene J, Sofaer S, Friminger K, Hirsh J: An Experiment Shows That A Well-Designed Report On Costs And Quality Can Help Consumers Choose High-Value Health Care. Health Aff. 31, 2012, 560-8
- Hickson GB (1987), Altemeier WA, Perrin JM: Physician Reimbursement by Salary or Fee-for-Service: Effect on Physician Practice Behavior in a Randomized Propective Study. Pediatrics 80, 1987, 344-50

- Hillman AL (1998), Ripley K, Goldfarb N, Nuamah I, Weiner J, Lusk E: Physicians Financial Incentives and Feedback: Failure To Increase Cancer Screening in Medicaid Managed Care. *Am. J. Publ. Health* 88, 1998, 1699-1701
- Hillman AL (1999), Ripley K, Goldfarb N, Weiner J, Nuamah I, Lusk E: The Use of Physician Financial Incentives and Feedback to Improve Pediatric Preventive Care in Medicaid Managed Care. *Pediatrics* 104, 1999, 931-4
- Institute of Medicine (2001): *Crossing the Quality Chasm: A New Health System for 21st Century*. National Academy Press, Washington, 2001
- Institute of Medicine (2007): *Rewarding Provider Performance: Aligning Incentives in Medicare*. National Academy Press, Washington, 2007
- Jain M (2006), Miller L, Belt D, King D, Berwick DM: Decline in ICU adverse events, nosocomial infections and cost through a quality improvement initiative focusing on teamwork and culture change. *Qual Saf Health Care* 2006;15:235–239
- JCAHO (1991): *Primer on Indicator Development and Application*, Joint Commission on Accreditation of Healthcare Organizations. One Renaissance Blvd, Oakbrook Terrace, Illinois 60181, 1991
- Jha AK (2010), Orav EJ, Epstein AM: The effect of financial incentives on hospitals that serve poor patients. *Ann Intern Med* 2010;153:299- 306.
- Jha AK (2012), Joynt KE, Orav EJ, Epstein AM: The Long-Term Effect of Premier Pay for Performance on Patient Outcomes. *N. Engl. J. Med.* 366, 2012, 1606-15
- Kahnemann D (1979), Tversky A: Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 4, 1979, 263-91
- Kannampalli T (2011), Schauer GF, Cohen T, Patel VL: Considering Complexity in Health Care Systems. *J. Biomed. Inform.* 44, 2011, 943-7
- Kazandjian VA (1995), Wood P, Lawthers J: Balancing Science and Practice in Indicator Development: The Maryland Hospital Association Quality Indicator (QI) Project. *Intern. J. Qual. Health Care*, 7, 1995, 39-46
- King D (2013), Greaves F, Vlaev I, Darzi A: Approaches Based on Behavioral Economics Could Help Nudge Patients and Providers Towards Lower Health Spending Growth. *Health Aff.* 32, 2013, 661-8
- Kluge (2002), *Etymologisches Wörterbuch der deutschen Sprache*, 24. Auflage, Berlin/New York 2002
- Kohn LT (1999), Corrigan JM, Donaldson MS (eds.): *To Err Is Human. Building a Safer Health System*. Committee on Quality of Health Care in America, Institute of Medicine, Washington 1999
- Korte KR (2012): *Das politische System der Bundesrepublik Deutschland*. In: Lauth HJ, Wagner C: *Politikwissenschaft: Eine Einführung*, S. 63-97, 7. überarb. Auflage, Verlag Ferdinand Schöningh, Paderborn 2012
- Kurtzman ET (2011), O'Leary D, Sheingold BH, Devers KJ, Dawson EM, Johnson JE: Performance-Based Payment Incentives Increase Burden and Blame for Hospital Nurses. *Health Aff.* 30, 2011, 211-8
- Landau H (2009): *Gesundheit als Staatsziel? Verfassungsrecht und Staatsaufgaben*. In:

- Konrad-Adenauer-Stiftung e.V. (Hrsg.): Volkskrankheiten. Gesundheitliche Herausforderungen in der Wohlstandsgesellschaft. Herder Verlag, 2009, S. 589-602
- Langdown C (2013), Peckham S: The use of financial incentives to help improve health outcomes: is the quality and outcomes framework fit for purpose? A systematic review. *J. Publ. Health* doi:10.1093/pubmed/fdt077
- Lauth HJ (2012), Thiery, P.: Politikfeldanalyse. In: Lauth HJ, Wagner C: Politikwissenschaft: Eine Einführung, S. 263-293, 7. überarb. Auflage, Verlag Ferdinand Schöningh, Paderborn 2012
- Leape LL (2002), Berwick DM, Bates DW: What practices will most improve safety? Evidence-based medicine meets patient safety. *JAMA* 288, 2002, 501–7
- Lebrenz C (2013): Leistungsbezogenes Gehalt lohnt – selten. *Frankfurter Allg. Sonntagszeitung* 23.12.2013
- Lega F (2005), DePietro C: Converging Patterns in Hospital Organization: Beyond the Professional Bureaucracy. *Health Policy* 74, 2005, 261-81
- Lester H (2010), Schmittziel J, Selby J, Fireman B, Campbell S, Lee J, Whippy A, Madvig P: The Impact of Removing Financial Incentives from Clinical Quality Indicators: Longitudinal Analysis of Four Kaiser Permanente Indicators. *Brit. Med. J.* 340, 2010, c1898
- Lindenauer PK (2007), Remus D, Roman S, Rothberg MB, Benjamin EM, Ma A, Bratzler DW: Public Reporting and Pay for Performance in Hospital Quality Improvement. *N. Engl. J. Med.* 356, 2007, 486-96
- Loewenstein G (2013), Asch DA, Volpp KG: Behavioral Economics Holds Potential To Deliver Better Results for Patients, Insurers, and Employers. *Health Aff.* 32, 2013, 1244-50
- Luft HS (2012): Advancing Public Reporting Through a New ‘Aggregator’ To Standardize Data Collection on Provider’s Cost and Quality. *Health Aff.* 31, 2012, 619-26
- Maisey S (2008), Steel N, Marsh R, Gillam S, Fleetcroft R, Howe A: Effects of payment for performance in primary care: qualitative interview study. *J. Health Serv. Res. Pol.* 13, 2008, 133–139
- Marshall MN (2010), Heath I, Sweeney K: Clinical Practice: When Things Go Wrong. *Lancet* 375, 2010, 1491-3
- Matern U (2006), Koneczny S, Scherrer M, Gerlings T: Arbeitsbedingungen und Sicherheit am Arbeitsplatz OP. *Dtsch. Ärztebl.* 103, 2006, A3187-92
- Mayes R (2011): Moving (Realistically) from Volume-Based to Value-Based Health Care Payment in the USA: Starting with Medicare Payment Policy. *JHSRP* 16, 2011, 249-51
- Mayntz R (2010): Governance im modernen Staat. In: Benz A, Dose N (Hrsg.): Governance – Regieren in komplexen Regelsystemen, 2. Aufl., Wiesbaden 2010, S. 37-48
- McCulloch P (2006): Surgical Professionalism in the 21st Century. *Lancet* 367, 2006, 177-81

- McDonald R (2007), Harrison S, Checkland K, Campbell SM, Roland M: Impact of Financial Incentives on Clinical Autonomy and Internal Motivation in Primary Care: Ethnographic Study. *Brit. Med. J.* 334, 2007, 1357
- McLean P (2006), Sutton M, Guthrie B: Deprivation and quality of primary care services: evidence for persistence of the inverse care law from the UK Quality and Outcomes Framework. *J. Epidemiol. Community Health* 60, 2006, 917–922
- Mears A (2010), Webley P: Gaming of Performance Measurement in Health Care: Parallels with Tax Compliance. *J. Health Serv. Res. Pol.* 15, 2010, 236-42
- Medical Research Council (2000). A Framework for the Development and Evaluation of RCTs for Complex Interventions to Improve Health 2000
- Medical Research Council (2008). Developing and evaluating complex interventions: new guidance 2008
- Mehrotra A (2009), Damberg C, Sorbero MES, Teleki SS: Pay for Performance in the Hospital Setting: What Is the State of the Evidence? *Am. J. Med. Qual.* 24, 2009, 19-28
- Mehta RH (2008), Liang L, Karve AM, Hernandez AF, Rumsfeld JS, Fonarow GC, Peterson ED: Association of Patient Case-Mix Adjustment, Hospital Process Performance Rankings, and Eligibility for Financial Incentives. *JAMA* 300, 2008, 1897-1903
- Metzger J (2010), Welebob E, Bates DW, Lipsitz S, Classen DC: Mixed Results in the Safety Performance of Computerized Physician Order Entry. *Health Aff.* 29, 2010, 655-63
- Miller HD (2009): From Volume to Value: Better Ways to Pay for Health Care. *Health Aff.* 28, 2009, 1418-28
- Miller M (2001), Elixhauser A, Zhan C, Meyer G: Patient Safety Indicators: using administrative data to identify potential patient safety concerns. *Health Serv. Res.* 36, 2001, 110-132
- Millett C (2007), Gray J, Saxena S, Netuveli G, Majeed A: Impact of a pay-for-performance incentive on support for smoking cessation and on smoking prevalence among people with diabetes. *CMAJ* 176, 2007, 1705-10
- Mintzberg H (1979): The Professional Bureaucracy, in: Mintzberg, H.: *The Structuring of Organisations*. Prentice-Hall, Upper Saddle River, NJ 07458, p. 348-79
- Mintzberg H (1997): Toward Healthier Hospitals: *Health Care Man. Rev.* 22, 1997, 9-18
- Molzberger B (2011), Marienhaus GmbH Waldbreitbach, fuw 3/2011, S. 255
- Moran J (2013), Scanlon D: Slow Progress on Meeting Hospital Safety Standards: Learning from the Leapfrog Group's Efforts. *Health Aff.* 32, 2013, 27-32
- Morden NE (2014), Colla CH, Sequist TD, Rosenthal MB: Choosing Wisely – The Politics and Economics of Labeling Low-Value Services. *N. Engl. J. Med.* 370, 2014, 589-92
- Mühlhauser I (2011), Lenz M, Meyer G: Entwicklung, Bewertung und Synthese von komplexen Interventionen – eine methodische Herausforderung. *Z. Evid. Forttbild. Qual. Gesundheth. wesen* 105, 2011, 751-61

- National Health Service (2013A): 2013/14 General Medical Services (GMS) Quality and Outcomes Framework (QOF). Guidance for GMS-Contract 2013/14. Department of Health 2013
- National Health Service (2013B): The NHS Outcome Framework 2013/14. Department of Health 2013
- Nicholson S (2008), Pauly MV, Wu AYJ, Murray JF, Teutsch SM, Berger ML: Getting Real Performance out of Pay-for-Performance. *Milbank Quart.* 86, 2008, 435-57
- Organisation for Economic Co-Operation and Development (2014), OECD.StatExtracts, http://stats.oecd.org/Index.aspx?DatasetCode=HEALTH_STAT, Zugriff 8.1.2014
- Ott E (2013), Saathoff S, Graf K, Schwab F, Chaberny IF: Prävalenz von nosokomialen und mitgebrachten Infektionen in einer Universitätsklinik. *Dtsch. Ärztebl.* 110, 2013, 533-40
- Paley J (2010): The Appropriation of Complexity Theory in Health Care. *J. Health Serv. Res. Pol.* 15, 2010, 59-61
- Paley J (2011): Complexity in Health Care: A Rejoinder. *J. Health Serv. Res. Pol.* 16, 2011, 44-45
- Pawlson LG (2007), Scholle SH, Powers A: Comparison of Administrative-Only Versus Administrative Plus Chart Review Data for Reporting HEDIS Hybrid Measures. *Am. J. Manag. Care* 13, 2007, 553-8
- Pearson SD (2007), Schneider EC, Kleinman KP, Coltin KL, Singer JA: The Impact of Pay-For-Performance on Health Care Quality in Massachusetts, 2001-2003. *Health Aff.* 27, 2007, 1167-76
- Petersen LA (2013), Simpson K, Pietz K, Urech TH, Hysong SJ, Profit J, Conrad DA, Dudley R, Woodard LD: Practice-Level Financial Incentives on Hypertension Care. A Randomized Trial. *Jama* 310, 2013, 1042-50
- Petzold T (2013), Steinwitz A, Schmitt J, Eberlein-Gonska M: Evaluation der Externen Qualitätssicherung gemäß §137 SGB V am Universitätsklinikum Carl Gustav Carus Dresden. *Z. Evid. Fortbild. Qual. Gesundh. wesen* 107, 2013, 541-7
- Pfaff H (2011), Schrappe M: Einführung in die Versorgungsforschung. In: H. Pfaff, G. Glaeske, E. Neugebauer, M. Schrappe (eds.): *Lehrbuch Versorgungsforschung*. Schattauer Verlag, Stuttgart 2011, S. 2-41
- Pham HH (2007), Schrag D, O'Malley AS, Wu B, Bach PB: Care Patterns in Medicare and Their Implications für Pay for Performance. *N. Engl. J. Med.* 356, 2007, 1130-9
- Pittet D (2004), Simon A, Hugonnet S, Pessoa-Silva CL, Sauvan V, Perneger TV: Hand Hygiene among Physicians: Performance, Beliefs, and Perceptions. *Ann Intern Med.* 141, 2004, 1-8
- Plsek PE (2001A), Greenhalgh T: The Challenge of Complexity in Health Care. *Brit. Med. J.* 323, 2001, 625-8
- Plsek PE (2001B), Wilson T: Complexity, Leadership, and Management in Healthcare Organisations. *Brit. Med. J.* 323, 2001, 746-9
- Plsek P (2001C): Redesigning Health Care with Insights from the Science of Complex Adaptive Systems. In: *Institute fo Medicine: Crossing the Quality Chasm: A New*

- Health System for 21st Century. National Academy Press, Washington, 2001, p.-309-22
- Plsek P (2003): Complexity and Adoption of Innovation in Health Care. Vortrag auf der Conference "Accelerating Quality Improvement in Health Care - Strategies to Speed the Diffusion of Evidence-Based Innovations, Washington, D.C., January 27-28, 2003, <http://www.nihcm.org/pdf/Plsek.pdf> (08.01.2014)
- Powell H (2001), Lim LL-Y, Heller RF: Accuracy of administrative data to assess comorbidity in patients with heart disease: an Australian perspective. *J. Clin. Epidemiol.* 54, 2001, 687-93
- Prendergast C (1999): The Provision of Incentives in Firms. *J. Econ. Lit.* 37, 1999, 7-63
- Pronovost PJ (1999), Kazandjian VA: A New Learning Environment: Combining Clinical Research with Quality Improvement. *J. Eval. Clin. Pract.* 5, 1999, 33-40
- Pronovost JP (2008), Goeschel CA, Wachter RM: The Wisdom and Justice of Not Paying for "Preventable Complications". *JAMA* 299, 2008, 2197-2199
- Reichert S (2000), Simon T, Halm, EA: Physicians' Attitudes About Prescribing and Knowledge of the Costs of Common Medications *Arch Intern Med.* 2000;160:2799-2803
- Relman AS (2007): Medical Professionalism in a Commercialized Health Care Market. *JAMA* 298, 2007, 2668-70
- Richardson KA (2008): Managing Complex Organizations: Complexity Thinking and the Science and Art of Management. *E:Co* 10, 2008, 13-26
- Ricketts TC (2013), Fraher EP: Reconfiguring Health Workforce Policy So That Education, Training, And Actual Delivery of Care are Closely Connected. *Health Aff.* 32, 2013, 1874-80
- Robb E (2010), Jarman B, Sunharalingam B, Higgins C, Tennant R, Elcock K: Using Care-Bundles to Reduce In-Hospital Mortality: Quantitative Survey. *Brit. Med. J.* 340, 2010, c1234
- Roland M (2004): Linking Physicians' Pay to the Quality of Care — A Major Experiment in the United Kingdom. *N. Engl. J. Med.* 251, 2004, 1448-1454
- Rolnick SJ (2002), Margolis KL, Fortman KK, Maciosek MV, Grimm RH: How acceptable are financial incentives and written feedback for improving hypertension control? Perspectives from physicians, clinic administrators, and patients. *Am. J. Man. Care* 8, 2002, 441-47
- Rosenthal MB (2006), Frank RG: What is the Empirical Basis for Paying for Quality in Health Care. *Med. Care Res. Rev.* 63, 2006, 135-57
- Rosenthal M (2007), Landon BE, Howitt K, Song HS R, Epstein AM: Climbing Up the Pay for Performance Learning Curve: Where Are the Early Adopters Now? *Health Aff.* 26, 2007, 1674-82
- Rosenthal MB (2007), Dudley, R.A.: Pay-For-Performance. Will the Latest Payment Trend Improve Care? *JAMA* 297, 2007, 740-7
- Rosenthal MB (2008): Beyond Pay for Performance – Emerging Models of Provider-Payment Reform. *N. Engl. J. Med.* 359, 2008, 1197-2000

- Ryan AM (2010): Has Pay-for-Performance Decreased Access for Minority Patients? Health Serv. Res. 45, 2010, 6-23
- Ryan AM (2011), Blustein J: The Effect of the MassHealth Hospital Pay-for-Performance Programm on Quality. Health Serv. Res. 46, 2011, 712-28
- Ryan AM (2012A), Nallamotheu BK, Dimick JB: Medicare's Public Reporting Initiative on Hospital Quality Had Modest or No Impact on Mortality from Three Key Conditions. Health Aff. 31, 2012, 585-92
- Ryan AM (2012B), Blustein J, Casalino LP: Medicare's Flagship Test of Pay-For-Performance Did Not Spur More Rapid Quality Improvement Among Low-Performing Hospitals. Health Aff. 31, 2012, 797-805
- Ryan AM (2012C), Blustein J: Making the Best of Hospital Pay for Performance. N. Engl. J. Med. 366, 2012, 1557-9
- Ryan AM (2013): Will Value-Based Purchasing Increase Disparities in Care? N. Engl. J. Med. 269, 2013, 2472-4
- Sachverständigenrat für die Konzertierte Aktion im Gesundheitswesen (1997): Gesundheitswesen in Deutschland. Kostenfaktor und Zukunftsbranche. Kurzfassung, Nomos 1998
- Sachverständigenrat für die Begutachtung der Entwicklung im Gesundheitswesen (2008): Kooperation und Verantwortung. Voraussetzungen für eine zielorientierte Gesundheitspolitik. Gutachten 2007, Band I und II, Nomos, Baden-Baden 2008
- Sachverständigenrat für die Begutachtung der Entwicklung im Gesundheitswesen (2009): Koordination und Integration – Gesundheitsversorgung in einer Gesellschaft des längeren Lebens. Sondergutachten 2009 des Sachverständigenrates für die Begutachtung der Entwicklung im Gesundheitswesen, www.svr-gesundheit.de
- Schmitt J (2013), Petzold T, Eberlein-Gonska M, Neugebauer EAM: Anforderungsprofil an Qualitätsindikatoren. Relevanz aktueller Entwicklungen der Outcomes Forschung für das Qualitätsmanagement. Z. Evid. Fortbild. Qual. Gesundh. wesen 107, 2013, 516-2
- Schoen C (2005), Osborn R, Huynh PT, Doty M, Zapert K, Peugh J, Davis K: Taking The Pulse Of Health Care Systems: Experiences Of Patients With Health Problems In Six Countries. Health Aff. 2005, W5-509-25, DOI 10.1377/htlaff.W5.509
- Schoen C (2011), Osborn R, Squires D, Doty M, Pierson R, Applebaum S: New 2011 Survey of Patients with Complex Care Needs in Eleven Countries Finds That Care is Often Poorly Coordinated. Health Aff. 20, 2011, 2437-42
- Schrappe M (2006), Lessing C, Jonitz G, Grandt D, Conen D, Gerlach F, Hart D, Lauterberg J, Loskill H, Rothmund M: Agenda Patientensicherheit 2006, Witten, 2006, www.aps-ev.de/fileadmin/fuerRedakteur/PDFs/Agenda_Patientensicherheit/Agenda_Patientensicherheit_2006.pdf
- Schrappe M (2007): Medizinische Zentren – Systematik und Nutzen. Z. ärztl. Fortbild. Qual. Gesundh.wes. 101, 2007, 141-6
- Schrappe M (2008), Lessing C, Conen D, Hart D, Hoppe-Tichy T, Jonitz G, Lauterberg J, Leppin G, Lichte T, Loskill H: Agenda Patientensicherheit 2008, Witten, 2008,

- http://www.aps-ev.de/fileadmin/fuerRedakteur/PDFs/Agenda_Patientensicherheit/Agenda_2008.pdf
- Schrappe M (2009): Führung im Krankenhaus – Clinical Corporate Governance. Z. ärztl. Fortbild. Qual. Gesundh.wes 103, 2009, 198-204
- Schrappe M (2010): Indikatoren. In: K.W. Lauterbach, M. Lungen, M. Schrappe (Hrsg.): Gesundheitsökonomie; Management und Evidence-based Medicine. Handbuch für Praxis, Politik und Studium. 3., völlig neu bearbeitete und erweiterte Auflage, Schattauer, Stuttgart, 2010, S. 329-348
- Schrappe M (2013): Patientensicherheit in Deutschland: Wo stehen wir im Jahr 2020? Präv. Gesundheitsf. 8, 2013, 49-55
- Schubert K (2008), Bandelow NC: Politikfeldanalyse: Dimensionen und Fragestellungen, in: Schubert, K., Bandelow, N.C. (Hrsg.): Lehrbuch der Politikfeldanalyse 2.0, München, S. 1-24
- Scott A (2011), Sivey P, Ait Ouakrim D, Willenberg L, Naccarella L, Furler J, Young D: The effect of financial incentives on the quality of health care provided by primary care physicians. Cochrane Database Syst Rev. 2011;9:CD008451
- Scott IA (2006), Ward M: Public reporting of hospital outcomes based on administrative data: risks and opportunities. Med J Aust, 184: 571-575
- Sens B (2003), Fischer B: Begriffe und Konzepte des Qualitätsmanagements. Inform. Biom. Epidemiol. Med. Biol. 34, 2003, 1-64
- Share DA (2012), Masen M: Michigan's Physician Group Incentive Program Offers a Regional Model for Incremental 'Fee for Value' Payment Reform. Health Aff. 31, 2012, 1993-2001
- Shen Y (2003): Selection Incentives in a Performance-Based Contracting System. Health Serv. Res. 38, 2003, 535-52
- Shojania KG (2002), Duncan BW, McDonald KM, Wachter RM: Safe But Sound. Patient Safety Meets Evidence-Based Medicine. JAMA 288, 2002, 508-13
- Shojania KG (2005), Grimshaw JM: Evidence-Based Quality Improvement: The State Of The Science. Health Affairs 24, 2005, 138-150
- Shojania KG (2010), Jennings A, Mayhew A, Ramsay C, Eccles M, Grimshaw J: Effect of Point-of-care Computer Reminders on Physician Behaviour: A Systematic Review. CMAJ 182, 2010, E216-25
- Shojania KG (2013): Conventional Evaluations of Improvement Interventions: More Trials or Just More Tribulations? BMJ Qual. Saf. 22, 2013, 881-4
- Snowden DJ (2007), Boone ME: Entscheiden in chaotischen Zeiten. Harvard Business Manager Dez. 2007, 27-42
- Sorbero MES (2006), Damberg CL, Shaw R, Teleki S, Lovejoy S, Decristofaro A, Dembosky J, Schuster C: Assessment of Pay-for-Performance Options for Medicare Physician Services: Final Report. Executive Summary. Rand Cooperation, May 2006, prepared for the Assistant Secretary for Planning and Evaluation, US Department for Health and Human Services
- Staehele WH (1999): Management: eine verhaltenswissenschaftliche Perspektive.

- Vahlens Handbücher der Wirtschafts- und Sozialwissenschaften, 8. Auflage, München 1999
- Stone S (2001): Hand Hygiene: The Case for Evidence-Based Education. *J. R. Soc. Med.* 94, 2001, 278-81
- Talbot TR (2013), Bratzler DW, Carrico RM, Diekema DJ, Hayden MK, Huang SS, Yokoe DS, Fishman NO, for the Healthcare Infection Control Practices Advisory Committee: Public Reporting of Health Care-Associated Surveillance Data: Recommendations From the Healthcare Infection Control Practices Advisory Committee. *Ann. Intern. Med.* 159, 2013, 631-5
- Timmermanns S (2005), Mauck A: the Promises and Pitfalls of Evidence-based Medicine. *Health Aff.* 24, 2005, 18-28
- Tompkins CP (2009), Higgins AR, Ritter GA: Measuring Outcomes and Efficiency in Medicare Value-Based Purchasing. *Health Aff.* 28, 2009, w251-61
- Torchiana DF (2013), Colton DG, Rao SK, Lenz SK, Meyer GS, Ferris TG: Massachusetts General Physicians Organization's Quality Incentive Program Produces Encouraging Results. *Health Aff.* 32, 2013, 1748-56
- Trefz U (2013): Rechtlich problematische Mindestmenge: Struktur- statt Mengenvorgaben? *F&W* 4/2013
- Van Herck P (2010), De Smedt D, Annemans L, Remmen R, Rosenthal MB, Sermeus W: Systematic Review: Effect, Design Choices, and Context of Pay-For-Performance in Health Care. *BMC Health Services Res.* 10, 2010, 247-60
- VanLare JM (2012), Conay PH: Value-Based Purchasing – National Programs to Move from Volume to Value. *N. Engl. J. Med.* 367, 2012, 292-5
- Veit C (2012), 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. Ein Gutachten im Auftrag des Bundesministeriums für Gesundheit, 2012
- Verma AA (2014), Razak F, Detsky AL: Understanding choice. Why Physicians Should Learn Prospect Theories. *JAMA* 311, 2014, 571-2
- Wachter RM (2010): Patient Safety at Ten: Unmistakable Progress, Troubling Gaps. *Health Aff.* 29, 2010, 165-73
- Wenzel RP (1992): Historical Perspectives. In: Wenzel, R.P. (Ed.): *Assessing Quality Health Care. Perspectives for Clinicians.* Williams and Wilkins, Baltimore 1992, p.3-16
- Werner RM (2005), Asch DA: The Unintended Consequences of Publicly Reported Quality Information. *JAMA* 293, 2005, 1239-44
- Werner RM (2011), Kolstad JT, Stuart EA, Polsky D: The Effect of Pay-For-Performance in Hospitals: Lessons for Quality Improvement. *Health Aff.* 30, 2011, 690-8
- Werner RM (2012), Dudley RA: Medicare's New Hospital Value-Based Purchasing Program Is Likely to Have Only a Small Impact on Hospital Payments. *Health Aff.* 31, 2012, 1932-40
- Willms G (2013), Bramesfeld A, Pottkämper K, Broge B, Szecsenyi J: Aktuelle Herausforderungen der externen Qualitätssicherung im deutschen

Gesundheitswesen. Z. Evid. Fortbild. Qual. Gesundh. wesen 107, 2013, 523-7
Wilson T (2001), Holt T: Complexity and Clinical Care. Brit. Med. J. 323, 2001, 685-8
Wodchis WP (2007), Ross JS, Detsky AS: Is P4P Really FFS? JAMA 298, 2007, 1797-9
Young GJ (2007), Meterko M, Beckman H, Baker E, White B, Sautter KM, Greene R,
Curtin K, Bokhour BG, Berlowitz D, Burgess JF: Effects of Paying Physicians Based
on Their Relative Performance for Quality. J. Gen Intern. Med. 22, 2007, 872-6

Abkürzungen

AHRQ: Agency for Healthcare Research and Quality

BMG: Bundesministerium für Gesundheit

BQS: vorm. Bundesgeschäftsstelle Qualitätssicherung, jetzt BQS Institut für Qualität und Patientensicherheit GmbH

CMS: Centers of Medicare and Medicaid Services, hervorgegangen aus der HCFA

CPOE: Computer-assisted Physician Order Entry Systeme

DRG: Diagnosis-Related Groups, Komorbiditäts- und Komplikations-adjustiertes Vergütungssystem in der stationären Versorgung

EBM: Evidence-Based Medicine, in Deutschland als Evidenz-basierte Medizin bezeichnet

GOÄ: Gebührenordnung Ärzte

HCFA: Health Care Financing Administration, Vorläufer der CMS

HEDIS: Health Effectiveness Data and Information Set

HMO: Health Maintenance Organization

HQIP: Premier Hospital Quality Incentive Demonstration Project (HQIP)

IOM: Institute of Medicine, USA

JCAHO: Joint Commission on Accreditation of Healthcare Organizations

MRSA: Methicillin-resistenter Staphylococcus aureus

NHS: National Health Service in Großbritannien

NICE: National Institute for Health and Clinical Excellence

NQF: National Quality Forum

P4P: Pay for Performance

QOF: Quality and Outcome Framework (des NHS)

SGB V: Sozialgesetzbuch V

SVR: Sachverständigenrat Gesundheit, genauer: Sachverständigenrat für die Begutachtung der Entwicklung im Gesundheitswesen (vormals: für die Konzertierte Aktion im Gesundheitswesen) nach §142 SGB V

TEP: Totalendoprothese (Hüfte, Knie)

VBP-Programm: Hospital Inpatient Value-Based Purchasing Programm von Medicare, in Kraft seit 10/2013

Stichwortverzeichnis

A

Accountable Care Organizations		2.3.
Administrative Daten als Indikatoren	1.3., 1.3.	
Aktion Saubere Hände		3.5.4.
Akuterkrankungen, Betonung der		5.1.
Altruismus		3.5.2.
Anreizwirkung Vergütungssysteme		5.3.
Attraktoren	3.3.1., 6.2., 7.5.	

B

Balanced Budget Act		2.3.
Behavioural Economics		3.5.2., 4.3., 7.6.
Framing		4.3., 7.6.
Globale Situationseinschätzung		4.3.
Isolation Effect		4.3.
Bedürfnistheorie n. Maslow		3.5.1.
Belohnung		3.5.1.
BQS Institut für Qualität und Patientensicherheit GmbH	2.1.	
Bundled Payment		2.3.

C

Ceiling-Effekt	1.4., 2.6.3., 4.2., 7.4.	
Chronische Erkrankungen		5.1., 5.3.
Comprehensive Care Management		5.3.
Conceptual framework für P4P		3.1.
Crossing the Quality Chasm		2.3.

D

Deficit Reduction Act		2.3.
Deutsches Gesundheitssystem		5.1., 7.1.
Betonung Akuterkrankungen		5.1., 7.3.
Dominierende Vergütungsanreize		5.1.
Erkrankungsbezug		5.1., 7.3.
Mengenorientierung		5.1., 7.3.
Prävention, Benachteiligung der	5.1., 7.3.	
Qualitätsdefizite		7.2., 7.3.
Sektorale Gliederung		5.1., 7.2.
Direction pointing		6.2., 7.7.
Diskontierung		4.2., 5.2.
Doppelte Komplexität		3.3.2., 7.3.
DRG		5.2.

E

Einstellungen		3.5.2.
Einzelleistungsvergütung		4.2., 5.2.
Ergebnisindikator	1.3., 2.4., 4.2., 7.4.	
Ethnische Gruppen		2.6.1.
Evidence-Based Medicine	3.1., 3.2.2., 4.2.	
Exception Rules (QOF)		2.5.
Expertenorganisation s. Profess. Bureaucracy	3.2.1., 3.4.,	

F

Feedback		3.5.1., 3.5.4., 7.4.
----------	--	----------------------

Adressaten	7.4.
Zeitnähe	7.4.
Framing	4.3., 7.6.

G

Gaming	2.6.2., 7.4.
Gender	2.6.1.
Gesundheitspolitik	6.1.
Direction pointing	6.2.,
Handlungsfelder	6.2., 6.3.
Negative Auswirkungen kontrollieren	6.3.
Rahmenbedingungen	6.3.
Strategische Ziele	6.3.
USA	2.3.
Gesundheitssystem, Deutschland	5.1.
Betonung Akuterkrankungen	5.1., 7.3.
Dominierende Vergütungsanreize	5.1.
Erkrankungsbezug	5.1., 7.3.
Mengenorientierung	5.1., 7.3.
Prävention, Benachteiligung der	5.1., 7.3.
Qualitätsdefizite	7.2., 7.3.
Sektorale Gliederung	5.1., 7.2.
Governance	6.1.
Definition	6.1.
Großbritannien	
QOF-Programm	2.5.

H

Händehygiene	3.5.2.
Haltungen	3.5.2.
Hausarzt-Zentrierte Verträge	5.3.
Höhe der Mittel	4.2., 7.6.
QOF	2.5., 4.1.
HQIP Premier Hospital Quality Incentive Demonstration	2.2.

I

Indikator	1.3.,
administrativ	1.3., 1.3., 7.4.
Aktualität	7.6.
Auswahl	7.4.
Differentialdiagnostik	1.4.
Ergebnisindikator s. dort	
inhaltliche Auswahl	7.4.
klinisch	1.3., 7.4.
Messung	1.4.
Patient Experience	2.4., 2.5.
Patientensicherheitsindikatoren s. dort	
Prozessindikator s. dort	
Relevanz	1.4.
Reliabilität	1.4.
Umsetzbarkeit	1.4.
Validität	1.4.
Wechsel	7.6.
Infection Control	3.1.
Informationsasymmetrie	4.2., 7.6.

Innovation im Gesundheitswesen		2.1.
Innovations-Paradoxon	3.4., 3.4., 7.3.	
Integrierte Versorgung		5.3.
Intrinsische Unsicherheit	3.4., 7.3.	
Isolation Effect		4.3.

J

K

Koalitionsvertrag CDU/CSU und SPD 27.11.13		1.1., 7.1.
Komplexe professionelle Systembürokratie	3.4., 7.3., 7.5.,	
Definition		3.4.
Verhaltensänderung in		3.5.1.
Zugänge		7.3.
Komplexität		2.3., 3.3.1.
doppelte		3.3.2., 7.3.
Eigenschaften komplexer Systeme	3.3.1.	
Universität als komplexe Organisation		3.3.1.
Komplexität stheorie		3.3.1., 3.4.
Kompliziertheit	3.3.1., 3.3.2.	
Kontext-bezogene Konzepte	3.5.4., 3.5.4.	
Kontinuität der Versorgung		2.6.1.
Koordination		5.3.
Kosteneffektivität		2.6.3.

L

Leapfrog-Initiative		6.3.
Leistungsbezogene Vergütung	3.5.1.	
Leitlinien		3.1.
Barrieren		3.5.2.
Lernen im Kontext	3.5.4., 3.5.4.	
Lerntheoretische Modelle der Verhaltensänderung	3.5.1., 3.5.4.	

M

Managed Care		5.3.
Maslow, Bedürfnistheorie n.		3.5.1.
Medicare Prescription Drug, Improvement and Modernization Act (MMA)		2.3.
Mediendemokratie		6.1.
Mengenreiz, -orientierung	5.1., 5.3., 7.3.	
Mindscape Framework		4.3.
Monetäre Bewertung		4.1.
Motivation		3.5.1., 3.5.2.
interne		3.5.2.
Motivation crowding out		7.5.

N

Never Events		2.3.
Non-Payment for Non-Performance		2.3.
Nosokomiale Infektionen, Erhebung		1.3.

O

Ökonomie		4.1., 7.6.
Opportunitätskosten	4.2.,	5.2.
Organisationslernen	3.5.3., 3.5.4.,	7.5.
Organisatorischer Wandel	3.5.3., 3.5.4.,	7.5.

P

Parteiendemokratie	6.1.
Patient Protection and Affordable Care Act	2.3.
Patient-reported outcomes	2.4., 2.5.
Patientensicherheit	3.1.
Patientensicherheitsindikatoren	2.4.
Pauschalen	
Erkrankungs-	5.3.
Populations-	5.3.
sektoral	5.2.
Sektorale Populations-	5.3.
transsektoral	5.3.
Pay for Reporting	2.3.
Pay for Performance (P4P)	
ambulant	2.5.
Definition	1.2., 4.1.
Doppelte Komplexität	3.3.2.
Frequenz der Zahlungen	7.6.
und Gesundheitssystem	1.5.
Höhe der Mittel	4.2., 7.6.
QOF	2.5., 4.1.
Kontextdomänen	3.1.
monetäre Bewertung	4.1., 4.2., 7.6.
monetäre Kopplung	4.1., 7.6.
Nachvollziehbarkeit	7.6.
Negative Auswirkungen	2.6., 7.7.
Perspektive	7.2.
Rahmenbedingungen	7.7.
Strategische Ziele	7.7.
Technische Umsetzung	1.3.
Umsetzung	1.5.
Unerwünschte Nebeneffekte von P4P	2.6.
Datenqualität	2.6.2.
Institutionelle Ebene	2.6.2.
Kosteneffektivität	2.6.3.
Professionelle Ebene	2.6.2.
Qualität System	2.6.3.
Teaching to the Test	2.6.3.
Zugang	2.6.1.
Ethnische Gruppen	2.6.1.
Kontinuität der Versorgung	2.6.1.
Risikoselektion	2.6.1.
Sozial benachteiligte Gruppen	2.6.1.
Weibliche Patientinnen	2.6.1.
Verlässlichkeit	7.6.
Polity - Politics - Policy	6.1., 7.7.
Prävention	3.5.1., 5.1., 5.3., 7.3.,
Premier Hospital Quality Incentive Demonstration (HQIP)	2.2.
Principal Agent Theorie	1.2., 4.2., 7.6.
Professional Bureaucracy	3.2.1., 3.4.
Autonomie	3.2.1., 3.4.
Eigenschaften	3.2.1.
Funktionale Diagnose	3.2.1.
Innovations-Paradoxon	3.4., 3.4.

Interne Regeln, Unsichtbarkeit	3.4.,	
Horizontale Vernetzung	3.2.1.	
Konflikt Experten - Management	3.3.2.	
Koordination		3.2.1.
Pigeonholing		3.2.1.
Selbstorganisation	3.4.	
Unsicherheit	3.2.2., 3.4.	
Veränderungsresistenz	3.2.2.	
Professionalismus	3.2.1., 3.2.2., 3.5.2., 7.5.	
Definition	3.2.2.,	
Prospect Theory		4.3.
Prozessindikator	1.3., 2.2., 2.4., 2.5., 4.2., 7.4.	
Public Reporting	2.3., 4.2., 4.3.	

Q

Qualität		
als Systemeigenschaft	3.3.1.	
und Kunst	1.2.	
Qualitätsmanagement	3.1.	
Qualitätsmessung		
Methodik	1.3.	
Qualitätswettbewerb	1.1.	
Quality and Outcome Framework (QOF)	2.5., 7.1.	
Indikatoren	2.5.	

R

Rahmenkonzept für P4P	3.1., 7.1., 7.3.	
Rational Choice		4.3.
Risikoadjustierung	1.3., 4.3., 7.4.	
Risikoaversion	4.3., 5.2., 7.6.	
Risikoselektion	2.6.1., 5.3., 7.4.	
Rollenverständnis, berufliches	3.5.2.	

S

Sachverständigenrat Gesundheit (SVR)	1.1.	
Systematischer Review zu P4P	2.1.	
Sektorale Gliederung	5.1., 7.2., 7.3.	
Selbstbestimmung der Patienten	7.2.	
Sozial benachteiligte Gruppen	2.6.1.	
Soziale Wahrnehmung, Theorie der	3.5.2., 3.5.4.	
Strukturindikator		4.2.
Systemtheorie	2.3., 3.3.1.,	

T

Tax Relief and Health Care Act	2.3.	
Teaching to the Test	2.6.3.	
To Err Is Human		2.3.

U

Unsicherheit	3.4.	
intrinsische	3.4.	
USA		
Gesundheitspolitischer Kontext von VBP	2.3.	

V

Value	1.2., 7.5.,	
-------	-------------	--

Value-Based Purchasing (VBP)	2.3., 2.4., 6.2., 7.1., 7.2.	
Indikatoren des VBP		2.4.
Vergütung		
Leistungsbezogene		3.5.1.
Vergütungsanreize, dominierende		5.1.
Vergütungssysteme		5.1.
Anreizwirkung		5.3.
Integration von P4P in	5.1., 7.6.,	
Kriterien für Kombination mit P4P	5.1.	
Verhaltensänderung	3.5.1., 7.3.	
Lernen im Kontext		3.5.4.
Lerntheoretische Modelle		3.5.1.
Organisatorischer Wandel		3.5.3.
Soziale Wahrnehmung		3.5.2.
Verhaltensökonomie	4.3., 7.6.	
Framing	4.3., 7.6.	
Globale Situationseinschätzung	4.3.	
Isolation Effect		4.3.
Verhandlungsdemokratie		6.1.
Verlustaversion		7.6.
Versorgungsforschung		3.3.2.

W

XY

Z

Zivilgesellschaft		6.3.
-------------------	--	------

Version (in chronologischer Reihenfolge)

Version 1.0.0.

31.3.2014

Erste Veröffentlichung

Impressum

Verantwortlich für den Inhalt von <http://www.schrappe.com> sind Dr. med. Gabriele Schrappe und Prof. Dr. med. Matthias Schrappe, Venloer Str. 30, 50672 Köln, Deutschland, Tel. +49 221 245749, email matthias@schrappe.com.

Haftungsausschluss / Urheber- Kennzeichenrecht / Datenschutz

1. Haftung für den Inhalt der Homepage

Wir, Dr. G. Schrappe und Prof. Dr. M. Schrappe, übernehmen keinerlei Gewähr für die Aktualität, Korrektheit, Vollständigkeit oder Qualität der bereitgestellten Informationen. Haftungsansprüche gegen uns, welche sich auf Schäden materieller oder ideeller Art beziehen, die durch die Nutzung oder Nichtnutzung der dargebotenen Informationen bzw. durch die Nutzung fehlerhafter und unvollständiger Informationen verursacht wurden, sind grundsätzlich ausgeschlossen, sofern unsererseits kein nachweislich vorsätzliches oder grob fahrlässiges Verschulden vorliegt. Wir behalten es uns ausdrücklich vor, Teile der Seiten oder die gesamte Webseite ohne gesonderte Ankündigung zu verändern, zu ergänzen, zu löschen oder die Veröffentlichung zeitweise oder endgültig einzustellen.

2. Haftung für Verweise und Links

Bei direkten oder indirekten Verweisen auf fremde Internetseiten ("Hyperlinks"), die ausserhalb meines Verantwortungsbereiches liegen, würde eine Haftungsverpflichtung ausschliesslich in dem Fall in Kraft treten, in dem wir von den Inhalten Kenntnis haben und es uns technisch möglich und zumutbar wäre, die Nutzung im Falle rechtswidriger Inhalte zu verhindern. Wir erklären daher ausdrücklich, dass zum Zeitpunkt der Linksetzung die entsprechenden verlinkten Seiten frei von illegalen Inhalten waren. Wir haben keinerlei Einfluss auf die aktuelle und zukünftige Gestaltung und auf die Inhalte der gelinkten/verknüpften Seiten. Deshalb distanzieren wir uns hiermit ausdrücklich von allen Inhalten aller gelinkten / verknüpften Seiten, die nach der Linksetzung verändert wurden. Für illegale, fehlerhafte oder unvollständige Inhalte und insbesondere für Schäden, die aus der Nutzung oder Nichtnutzung solcherart dargebotener Informationen entstehen, haftet allein der Anbieter der Seite, auf welche verwiesen wurde, nicht derjenige, der über Hyperlinks auf die jeweilige Veröffentlichung lediglich verweist.

3. Urheber- und Kennzeichenrecht

Wir sind bestrebt, in allen Publikationen die Urheberrechte der verwendeten Grafiken, Tondokumente, Videosequenzen und Texte zu beachten, von ihr selbst erstellte Grafiken, Tondokumente, Videosequenzen und Texte zu nutzen oder auf lizenzfreie Grafiken, Tondokumente, Videosequenzen und Texte zurückzugreifen. Alle innerhalb des Internetangebotes genannten und ggf. durch Dritte geschützten Marken- und Warenzeichen unterliegen uneingeschränkt den Bestimmungen des jeweils gültigen Kennzeichenrechts und den Besitzrechten der jeweiligen eingetragenen Eigentümer. Allein aufgrund der blossen Nennung ist nicht der Schluss zu ziehen, dass Markenzeichen nicht durch Rechte Dritter geschützt sind.

4. Datenschutz

Sofern innerhalb des Internetangebotes die Möglichkeit zur Eingabe persönlicher oder geschäftlicher Daten (E-Mailadressen, Namen, Anschriften) besteht, so erfolgt die Preisgabe dieser Daten seitens des Nutzers auf ausdrücklich freiwilliger Basis.

5. Rechtswirksamkeit der Regelungen

Der Haftungsausschluss sowie die Bestimmungen zum Urheber- und Kennzeichenrecht sowie dem Datenschutz sind als Teil des Internetangebotes zu betrachten, von dem aus auf diese Seite verwiesen wurde. Sofern Teile oder einzelne Formulierungen der oben näher bezeichneten Regelungen der geltenden Rechtslage nicht, nicht mehr oder nicht vollständig entsprechen sollten, bleiben die übrigen Bestimmungen in ihrem Inhalt und ihrer Gültigkeit davon unberührt.