

11 Einführung von Pay for Performance

11.1 P4P – Vorbemerkung

Pay for Performance spielt international eine große Rolle und ist neben Public Reporting eines der wichtigsten Instrumente zur Qualitätsverbesserung, das auf Systemebene einsetzbar ist (s. Kap. 6). Wenngleich die Evaluationsstudien nicht so klare Ergebnisse erbracht haben wie erhofft (s. Kap. 6.4), 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 (Koalitionsvertrag der Großen Koalition vom 27.11.2013, Value-Based Purchasing Programm für den Krankenhausbereich und das Physician Value-Based Payment Modifier [PVBPM] für den ambulanten Bereich in den USA [s. Kap. 6.4.1], Quality and Outcome Framework [QOF] mit jährlich neu justierten Indikatoren in Großbritannien [s. Kap. 6.4.2]).

Führend sind in den umfangreichen internationalen P4P-Programmen Prozessindikatoren, Patientensicherheitsindikatoren (PSI) und Indikatoren auf der Basis von Patientenerfahrungen (*patient-related outcome measures*, PROM), im Mittelpunkt stehen zunehmend Probleme der ambulanten Versorgung und Koordination. Ergebnisindikatoren werden risikoadjustiert eingesetzt. Enttäuschenderweise wurde kein *spill over*-Effekt zur Verbesserung von nicht angereizten Indikatoren beobachtet, dafür aber ein deutlicher *ceiling*-Effekt; die

poor performers profitierten anders als vermutet nicht stärker als die Top-Einrichtungen. Nach Beendigung des Anreizes scheint die Qualität der Versorgung wieder auf das Ausgangsniveau zurückzufallen. Die zusätzlichen (oder einbehaltenen) Zahlungen sind teilweise recht hoch.

Die Gründe für die geringe langfristige Wirkung von P4P sind darin zu suchen, dass die Indikatoren bereits langfristig bekannt waren, es zu einem *ceiling*-Effekt kam, dass durch eine mangelhafte Kopplung von Vergütung und Qualitätssystematik nicht zu einer Verbesserung der *poor performers* gekommen war (s. Kap. 8.5.2), dass es einen gleichzeitigen Einsatz von *Public Reporting* und P4P gab (*dual use*), dass die Vergütung (immer noch) zu niedrig war (vor allem unter Berücksichtigung der Opportunitäts- und Grenzkosten) und dass man die Risikoaversion der Entscheidungsträger unterschätzte, ebenso wie die Dominanz des grundlegenden Vergütungssystems (DRG).

Die Einführung von P4P stellt einen sehr differenzierten (komplexen) Eingriff in den Kontext eines Gesundheitssystems dar und macht daher ein Rahmenkonzept notwendig, das auf organisationstheoretischen Überlegungen, auf Überlegungen zur Komplexität des Gesundheitssystems, auf Annahmen zur Verhaltensänderung, auf ökonomischen Grundannahmen, auf einer Analyse von Wechselwirkungen zwischen den qualitätsorientierten Vergütungsbestandteilen von P4P und dem zugrundeliegenden Vergütungssystem (z.B. DRG) und auf politikwissenschaftlichen Annahmen (s. Kap. 9.4) beruht:



Rahmenkonzept: Zugang

- **Organisationstheorie: Expertenorganisation** (professional bureaucracy)
- **Systemtheorie: Komplexe Systeme**
- **Verhaltensänderung**
 - *lerntheoretische Modelle*
 - *soziale Wahrnehmung*
 - *organisatorischer Wandel*
 - *Lernen durch Kontext*
- **Ökonomische Grundannahmen**
 - *Opportunitätskosten und Diskontierung*
 - *Principal-Agent-Theorie*
 - *Verhaltensökonomie*
- **Kombination andere Vergütungssysteme**
- **Politikwissenschaftliche Annahmen**

Die beiden wichtigsten Konsequenzen für die praktische Umsetzung von P4P, die sich aus dem entwickelten Rahmenkonzept ableiten lassen, sind

- die aus hohen Opportunitäts- und niedrigen Grenzkosten, Diskontierung und Risikoaversion resultierende **Verstärkung des Mengenan-**

reizes fallpauschalierender Vergütungssysteme (z.B. DRG-System) bei Leistungen, bei denen eine Mengenausweitung z.B. durch Indikationsausweitung machbar ist (s. Empfehlung 42) und

- der Übergang bzw. die **Überschneidung mit der Einzelleistungsvergütung** bei Leistungen, bei denen die Informationsasymmetrie insofern aufgehoben ist, als dass sie vollständig durch Leitlinien oder EBM abgesichert und daher „allseits bekannt“ ist; eine Einzelleistungsvergütung ist in diesem Fall sinnvoller (s. Empfehlung 45).

11.2 Allgemeine Empfehlungen für die Einführung von P4P

1. Rahmenkonzept zugrunde legen: Die Einführung von P4P muss sich, wie oben ausgeführt, auf ein Rahmenkonzept beziehen, wie es in Kapitel 8 und 9 in diesem Gutachten entwickelt wurde.



Empfehlung 31:

Die Einführung von P4P basiert auf dem „Rahmenkonzept Qualitätsentwicklung“, das vom Bundesministerium für Gesundheit dem Gesetzgeber, den Verbänden (einschließlich dem Gemeinsamen Bundesausschuss und der Öffentlichkeit vorgelegt wird (s. Empfehlung 1). Dieses Rahmenkonzept beruht auf einem umfassenden Qualitätsverständnis.

2. Quantitative Erfassung und Indikatoren voneinander trennen: In Ergänzung zu den grundlegenden Aussagen zu Indikatoren in Kapitel 2.4 soll hier nur nochmals hervorgehoben werden, dass man die Vollerhebung eines qualitätsrelevanten Parameters von der Erhebung eines Indikators mit Vorhersagefunktion unterscheiden muss. Der erste Fall verstärkt die Tendenz von P4P zur Einzelleistungsvergütung (zur diesbezüglichen Bedeutung der Informationsasymmetrie s. 8.5.4.). Will man aber im Sinne des Indikatorenkonzeptes einen Indikator mit Vorhersagefunktion für einen Versorgungsbereich einführen, tritt die direkte Qualitätsfunktion des Indikators in den Hintergrund. Dies hat für die (politisch) Handelnden zwei Konsequenzen, man muss sich einerseits weniger mit der Gefahr der Mengenausweitung durch P4P selbst auseinandersetzen, andererseits muss man sich jedoch mit strategischen Überlegungen, Schwerpunktsetzungen und Schwachpunktanalysen beschäftigen. Um es kurz zu sagen: kein echter Indikator ohne Problemanalyse. In dem vorliegenden Gutachten sind die sog. „fünf Ziele“ herausgearbeitet worden, die auf die Notwendigkeiten der Gesundheitsversorgung der alternden Gesellschaft Bezug nehmen (s. Kap. 1.4, s. Empfehlung 4). Im Weiteren bleiben die Aussagen zur den Charakteristika valider Indikatoren (s. Kap. 2.4) 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 32:

Die Wahl von Indikatoren mit Vorhersagefunktion für die Einführung von P4P erfordert eine strategische Problemanalyse, die sich an den zukünftigen Herausforderungen des Gesundheitssystems orientieren („fünf Ziele“, s. Empfehlung 4).

3. Inhaltliche Auswahl von Indikatoren – dreidimensionaler Orientierungsrahmen: Inhaltlich müssen die Indikatoren für ein P4P-Programm einem umfassenden Qualitätsverständnis entsprechen, so wie es in dem vorliegenden Gutachten unter dem Begriff der 6 Qualitätsperspektiven erstellt worden ist (s. Kap. 1.3). Zusammen mit der Integrationsdimension und der Morbiditätsdimension (Berücksichtigung insbesondere von chronischen Erkrankungen und Prävention) bewegt sich die Indikatorenausstattung des P4P-Programms in einem dreidimensionalen Orientierungsrahmen (s. Kap. 1.4 und 9.3, s. Empfehlung 4). Im Speziellen stehen in den internationalen P4P-Programmen Patientensicherheits- und Indikatoren auf der Basis von Patientenerfahrungen im Vordergrund.



Empfehlung 33:

Die Indikatoren des P4P-Programmes müssen das Krankheitsspektrum der zukünftigen Versorgung der Gesellschaft (chronische Mehrfacherkrankungen) sowie die Prävention abdecken (Morbiditätsdimension) und die Integration des Gesundheitssystems fördern (Integrationsdimension). Patientensicherheitsindikatoren und Patientenerfahrungen sind zu integrieren (s. Empfehlung 4).

4. Indikatoren auf der Basis von Routinedaten sind insensitiv und können in P4P-Programmen zu Mengenausweitung führen: Wie in Empfehlung 22 bereits ausgeführt, sind administrative Daten primär für Vergütungszwecke und nicht für Qualitätsmessungen entwickelt worden und weisen in vielen Fällen deutliche Sensitivitätsdefizite auf. Es besteht die Gefahr einer Mengenausweitung, wenn insensitive administrative Indikatoren bei Leistungen mit entsprechendem Potenzial als Indikatoren in P4P-Programmen eingesetzt werden. In Einzelfällen können sie begleitend jedoch eingesetzt werden (z.B. standardisierte Mortalität), sie geben u.U. sinnvolle Hintergrundinformationen. 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 infektionsepidemiologische Falldefinitionen in nosokomialen Infektions-

erfassungen. 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 34:

Indikatoren auf der Basis von Routinedaten haben eine schlechte Sensitivität und weisen bei P4P-Programmen die Gefahr einer Mengenausweitung auf. Klinische Daten sind im Bereich der Infektionsepidemiologie (Erfassung nosokomialer Infektionen) gut als Falldefinitionen etabliert, diese Entwicklung sollte für die Qualitätsindikatoren übernommen werden (s. Empfehlung 22). Administrative Indikatoren sollten wegen ihrer mangelnden Sensitivität nur in P4P-Konzepten mit Einzelleistungsvergütungscharakter verwendet werden.

5. Prozessindikatoren führen vor Ergebnisindikatoren: Die großen laufenden P4P-Projekte in den USA (Value-Based Purchasing (VBP) und in Großbritannien (Quality and Outcome Framework [QOL]) verwenden fast ausschließlich Prozessindikatoren (s. Kap. 6.4.1 und 6.4.2), da bei Ergebnisindikatoren die Problematik der Risikoselektion zu gewichtig ist, und selbst wenn die Risikoadjustierung perfekt entwickelt wäre, kleine Einrichtungen systematisch schlechter gestellt werden. Prozessindikatoren entsprechen den zukünftigen Herausforderungen des Gesundheitswesens besser als Ergebnisindikatoren und sind gut zuzuordnen (s. Empfehlung 21). Unter den Prozessindikatoren sollten solche ausgewählt werden, bei denen die Informationsasymmetrie erhalten ist, also nicht zur Einzelleistung führende Indikatoren wie „Blutgasanalyse bei Pneumonie“, sondern eher „Arztbrief nach Entlassung innerhalb zwei Tagen“ (s. Abb. 50). Für die Zukunft sind Prozessindikatoren mit starkem Outcome-Charakter (in erster Linie Patientensicherheitsindikatoren [z.B. Katheterinfektionen]) die erste Wahl.

Indikatoren:	Informationsasymmetrie	Risikoselektion	Wertung
■ Ergebnis	+	+ ⚡	Risikoselektion ist Problem
■ Prozess (a)	–	+	Einzelleistung
■ Prozess (b)	+	–	optimal
■ Struktur	–	–	Investitionsbeihilfe

Abb. 50 Differenzierter Gebrauch von Prozessindikatoren, in Abgrenzung auch von Struktur- und Ergebnisindikatoren (aus Kap. 8.4.5)



Empfehlung 35:

Prozessindikatoren führen vor Ergebnisindikatoren (s. Empfehlung 21), am ehesten sind Prozessindikatoren aus dem Bereich der Koordination und Kommunikation geeignet, die nicht einer Mengenausweitung ausgesetzt sind.

6. Risikoselektion und exception reporting können Unterversorgung verursachen: Besonders bei zugrundeliegenden Vergütungsansätzen, die bereits ihrerseits zur Risikoselektion neigen (z.B. Pauschalen), 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. Eine weitere Möglichkeit ist die Einrichtung eines *exception reporting* (s. Kap. 6.4.2), das es unter bestimmten Bedingungen möglich macht, Patienten aus dem P4P-Programm herauszunehmen, insbesondere wenn patientenseitige Gründe für ein Nicht-Erreichen des angestrebten Ziels vorliegen. Diese Regelung soll einer Risikoselektion und Schlechterstellung (bzw. niedrigerer Attraktivität) von Praxen in sozial unterprivilegierten Wohngebieten entgegenwirken, beinhaltet aber ebenfalls die Gefahr einer daraus resultierenden Unterversorgung als schwerwiegender unerwünschter Nebeneffekt der P4P-Einführung



Empfehlung 36:

Risikoselektion und evtl. exception reporting können zur Unterversorgung führen und sollten vermieden werden.

11.3 Empfehlungen: Organisation und System beachten

1. Professionalismus-Diskussion: 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 Ebe-

ne) 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, wissenschaftsgestütztes 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 u. Mauck 2005), insbesondere Veränderungen der Rollen und der interprofessionellen Zusammenarbeit betonend, die schon die unten beschriebenen Aspekte des organisatorischen Wandels mit einbeziehen.



Empfehlung 37:

Die multiprofessionelle und multidimensionale Implementierung ist zu bevorzugen. Die monoprofessionelle Beschränkung des Professionalismuskonzeptes 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.

2. Organisatorisches Lernen fördern: Es ist unumgänglich, wie in Empfehlung 36 bereits angesprochen, das lerntheoretische Instrument P4P nicht nur

um Aspekte der sozialen Wahrnehmung (Rollen, Professionalismus), sondern auch um Elemente des organisatorischen Wandels und um kontextbezogene 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.



Empfehlung 38:

Die Organisation als Ganzes ansprechen, den organisatorischen Wandel fördern – P4P setzt die Organisation unter „internen Stress“, es müssen jetzt nämlich zusätzlich zu Kosten und mengenbezogenen Erlösen auch erlöswirksame Qualitätsinformationen intern bearbeitet werden, und zwar hinsichtlich sowohl der Abgrenzung von Verantwortung als auch der Aufgabenstellungen und in der Führungsarbeit.

3. Motivation crowding out im Blick behalten, aber nicht überschätzen:

Grundsätzlich muss bei jedem Vergütungssystem, das im beruflichen Kontext angewandt wird, mit Auswirkungen auf Motivation, Rolle und Zusammenarbeit gerechnet werden (Prendergast 1999; s.a. 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 qualitätsbezogene Vergütung gemindert wird („crowding out“; Berenson et al. 2013; Cassel u. 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 39:

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.).

4. Value (Effizienz) statt allein Qualität als Zielgröße für P4P-Programm in Betracht ziehen: 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 u. Rosenthal 2013; CMS 2011A; CMS 2011B; Ryan et al. 2012). Dieser Effizienzgedanke 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. 8.4.2) 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 40:

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

5. Mögliche Attraktoren einsetzen: Komplexe Systeme orientieren sich an sog. Attraktoren (s. Kap. 8.3.1), von außen nicht erkennbare Systemzustände, die eine relative Stabilität aufweisen, in deren Richtung sich die Systeme entwickeln. Wenn man über *framing* und *direction pointing* (s. Kap. 8.7.2) 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 u. Mauck 2005). Ein zweiter Attraktor besteht sicherlich in professioneller Autonomie und Kollegialität, ganz wie im Konzept der Expertenorganisation beschrieben (B) (Avorn u. 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 41:

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 bei der Implementierung von P4P-Programmen einzusetzen.

11.4 Empfehlungen: 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 zusätzlicher ökonomischer (und außerdem politischer) Weichenstellungen (s. Kap. 8). 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.

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 potenziell 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 Zusammen-

hang 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; wohlgemerkt: 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. 8.5). So ist es sicher nicht möglich, mit P4P ein eng sektoral begrenztes, mit hochtouriger Volumenexpansion einhergehendes Vergütungssystem wie das des klassischen DRG-Systems 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.

1. Die notwendige Höhe der P4P-Vergütung hängt stark mit dem zugrundeliegenden Vergütungssystem zusammen: es 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 Opportunitätskosten hoch bzw. 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 qualitätsbezogenen Erlösbestandteiles (Frolich et al. 2007). Die Mehrzahl der Ana-

lysen plädiert daher für höhere P4P-Vergütungen (Chung et al. 2010; Ryan u. Blustein 2011; Werner et al. 2011B; Werner u. Dudley 2012).



Empfehlung 42:

Höhe der P4P-Zahlungen muss besonders im DRG-System die Opportunitäts- und Grenzkosten berücksichtigen. Die Höhe des qualitätsbezogenen Erlösanteiles sollte den Messaufwand, die Opportunitätskosten, die Grenzkosten und die Diskontierung berücksichtigen und hängt daher stark von der Art des dominierenden Vergütungssystems ab. Bei Einzelleistungsvergütung und DRGs sind hohe Opportunitätskosten (und niedrige Grenzkosten) anzusetzen, die Höhe der P4P-Vergütung muss entsprechend hoch angesetzt werden.

2. Frequenz, Stückelung und Zahlungsmodus bedenken: Ein intakter *Feedback*-Mechanismus vorausgesetzt (s. Kap. 8.4.2), bei dem die Zahlungen zeitnahe erfolgen, kann der *Feedback* noch verbessert werden, wenn die sog. Überschätzung relativer Risiken berücksichtigt wird, ein Phänomen, das in der Verhaltensökonomie beschrieben ist (s. Kap. 8.5.5, 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 ereignisbezogen vorgenommen werden.



Empfehlung 43:

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

3. Monetäre Kopplung differenziert gestalten: Eine P4P-Vergütung besteht aus zwei Bestandteilen, einem System der Qualitätsmessung und einem System der monetären Bewertung (s. Kap. 6.2). 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 performers* 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. 39 in Kap. 8.5.2) sollten daher immer in Kombination eingesetzt werden.



Empfehlung 44:

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.

4. Informationsasymmetrie und Principal-Agent-Theorie – besonders bei Prozessindikatoren Übergang zur Einzelleistungsvergütung beachten: 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 u. 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. 6.5) 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. 6.4.2). 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 u. 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; hier besteht der zentrale Einsatzbereich von Prozessindikatoren.** Umso 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 Patientensicherheitsindikatoren (z.B. vermeidbare nosokomiale Infektionen) sind hier zu disku-

tieren (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 u. Ward 2006). Indikatoren der Strukturqualität sind meist weitgehend bekannt sind, eine Informationsasymmetrie liegt nicht vor, ein reiner Investitionszuschuss ist sinnvoller (s. Kap. 4.3).



Empfehlung 45:

Konsequenzen aus der principal-agent Theorie und damit Überschneidungen zur Einzelleistungsvergütung gehören zu den schwierigsten Fragestellungen, mit denen sich das P4P-Konzept auseinanderzusetzen 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 Patientensicherheitsindikatoren 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 mitarbeiterbezogene Händedesinfektionsmittelverbrauch oder *Readmission*-Indikatoren sind ebenfalls sinnvoll, genauso Indikatoren, die prozessuale Aspekte der Patientensicherheit beschreiben (z.B. Zahl der bearbeiteten Meldungen im CIRS).

5. Aktualität und Wechsel der Indikatoren wichtig: 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 46:

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.

6. Wettbewerbsdichte und dual use: Auf der Basis der hier diskutierten ökonomischen Konzepte muss auf die Frage der Kombination mit anderen Qualitätsinitiativen eingegangen werden, vor allem wenn sie Wettbewerbscharakter haben (z.B. *public reporting*), s. auch Empfehlung 12). 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 (sog. *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. 5), 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 u. 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:

- **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 das mögliche Verbesserungspotenzial 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 Program*, s. Ryan et al. 2011).
- **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 di-

rekte Anreizwirkung durch die qualitätsbezogenen Zahlungen (Werner et al. 2011). Der indirekte Anreiz durch *public reporting* wirkt besonders in hochkompetitiven 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 wettbewerbswirksame Instrumente nicht ausreichen, z.B. in ländlichen Gegenden. In hochkompetitiven 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). Hochkompetitive 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 47:

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 und Fachgebieten differenzierter Einsatz ist ein interessanter Ansatz, der in der Zukunft zu diskutieren wäre.

7. Verhaltensökonomie – framing beachten: Entscheidungen werden in einem referentiellen Zusammenhang getroffen, der die Interpretation erleichtert oder sogar erst ermöglicht. Dies betrifft z.B. die sog. Risikoaversion („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 erbrin-

gen 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 Risikoaversion 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 Verlustavers, d. h. ein entgangener Gewinn wird weniger hoch bewertet als ein gleich großer Verlust (Werner u. Dudley 2012). Die Komponenten eines adäquaten *framing* werden sehr gut durch das *Mindscape*-Instrument beschrieben (s. Abb. 42). 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 48:

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.

8. Verlässlichkeit und Nachvollziehbarkeit: Die Zahlung der P4P-Vergütung muss nicht nur zeitnah (s. Kap. 8.5.5), sondern wegen der Diskontierungstendenz verlässlich und im Zustandekommen nachvollziehbar erfolgen. Die Gesundheitsberufe und Einrichtungen können nur dann ein *Feedback* erkennen, wenn sie verstehen können, wie eine qualitätsbezogene Vergütung ausgelöst wird und gestaltet ist.


Empfehlung 49:

Verlässliche und nachvollziehbare Zahlungen: das Verständnis für das Zustandekommen der P4P-Zahlungen verbessert die Wirksamkeit des Feedback.

9. Integration in andere Vergütungssysteme kritisch und differenziert vornehmen: P4P wird grundsätzlich im Rahmen anderer Vergütungssysteme eingesetzt (s. Kap. 8.6). Getreu der Maßgabe „Qualität statt Menge!“ kommt P4P gerade in Gesundheitssystemen in die Diskussion, die durch mengenorientierte 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), wenn Indikatoren mit geringer Sensitivität verwendet werden, wenn die Risikoadjustierung durch *up-coding* der zur Adjustierung verwendeten Parameter (z.B. Komorbidität) zur Absenkung der „standardisierten Raten“ verwendet wird (Pseudoverbesserung), eine aktive Risikoselektion möglich erscheint und sich mit den resultierenden Qualitätsdaten dann gegenüber den Kostenträgern eine Mengenausweitung begründen lässt. 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 dringendsten 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. 36, Kap. 8.4.2) zu fördern.

**Empfehlung 50:**

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.

11.5 Empfehlungen: 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. 8.7). 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 kontextbezogenen 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 „hineg-diskutieren“. 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 Kapitel 8.7.2 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.

1. direction pointing als elementares Mittel zur Implementierung nutzen: 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 wettbewerbsbezogene Elemente geht (s. 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 Geschehenlassen reicht nicht aus (vgl. Empfehlungen 1–5, Empfehlung 14). Den professionellen Fachgesellschaften und Institutionen, die sich mit den Themen Qualität und Patientensicherheit auseinandersetzen und teilweise Berufsverbandscharakter haben, müssen öffentlichkeitswirksame Unterstützungen erfahren. Getreu dem *Governance*-Konzept ist die aktive Einbeziehung von Medien und Einrichtungen aus der Zivilgesellschaft notwendig (s. unten). In Kapitel 8.5 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 interessengeleiteten Perspektive die für die Gesamtgesellschaft optimalen Entscheidungen treffen kann.



Empfehlung 51:

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 (s. Empfehlung 14).

2. Strategische Ziele setzen – Prozessindikatoren problemorientiert auswählen: 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 (Empfehlungen in Kap. 10.2). 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, problemorientiert 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. 4.5) und in Großbritannien (s. Kap. 6.4.2) gewesen und werden natürlich umso wichtiger, je 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. Empfehlungen 12). Auch einer P4P-Implementierung, die ausschließlich oder überwiegend auf Ergebnisindikatoren setzt, wird wegen der Risikoselektion (s. Kap. 6.5) 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. 8.5.4).

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 52:

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, 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 53:

Nicht alle Prozessindikatoren sind sinnvoll bei P4P einzusetzen. 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.

3. Verschlechterung des Zugangs und Datenmanipulation 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 Kapitel 6.5 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-Pro-

grammen in jedem Fall stattfinden und frühzeitig in den Implementierungsprozess integriert werden muss.



Empfehlung 54:

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.

4. Rahmenbedingungen schaffen – eine politische Aufgabe: 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 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 u. 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 (s. Empfehlung 14). 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 (s. Empfehlung 1). Ein solches Rahmenkonzept ist handlungsleitend und für die Beteiligten letztlich handlungslegitimierend, eine wichtige Voraussetzung für die Einführung eines Instrumentes wie *Pay for Performance* – qualitätsorientierte Vergütung.



Empfehlung 55:

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.