



Hauptverband der
österreichischen
Sozialversicherungsträger

Pay-for-Performance im Gesundheitswesen

Wissenschaftliche Evidenz zur Wirksamkeit

Dossier

Version 3

16.11.2016

Aline Dragosits MA

Hauptverband der österreichischen Sozialversicherungsträger
Evidenzbasierte Wirtschaftliche Gesundheitsversorgung
Gesundheitsökonomie

a) Abstract

Hintergrund

Defizite in der Versorgungsqualität können auf unterschiedliche Faktoren zurückgeführt werden. P4P wurde als qualitätsorientiertes Steuerungsinstrument entwickelt, welches die ärztliche Leistung an ein erbrachtes und durch Kennzahlen dargestelltes Leistungsniveau koppelt und darauf abzielt die Qualität und Effizienz der Versorgung zu optimieren.

Methode

Das Ziel dieses Dossiers ist es, die wissenschaftliche Evidenz der Wirksamkeit von P4P im Gesundheitswesen zu erörtern. Um dieses zu erreichen, erfolgt eine grobe Sichtung von systematischen Übersichtsarbeiten, welche sich mit diesem Thema beschäftigen.

Ergebnisse

Qualitätsmessung und Anreizgestaltung nehmen eine Schlüsselrolle in P4P-Programmen ein. Die Studienergebnisse haben gezeigt, dass der langfristige Erfolg von P4P neben der Auswahl geeigneter Indikatoren zur Qualitätsmessung auch von der Akzeptanz durch die LeistungserbringerInnen abhängig ist. In diesem Zusammenhang muss erwähnt werden, dass es keine Art „Kochrezept“ für die Ausgestaltung von P4P-Programmen gibt. Viel eher hängt die Wahl des Anreizmodells und der Bemessungsgrundlagen davon ab, welche Ziele mit dem Programm verfolgt werden. P4P-Programme können dazu beitragen die Behandlungsqualität zu optimieren, die Transparenz im Gesundheitswesen zu erhöhen und die Kosten zu reduzieren. Selektionseffekte, Beeinflussung der medizinischen Praxis und Akzeptanz der ÄrztInnen werden als Herausforderungen erachtet, welche bei der Konzipierung und Implementierung berücksichtigt werden sollten. Im Hinblick auf die wissenschaftliche Evidenz der Wirksamkeit von P4P zeigt sich, dass die Evaluationen zu widersprüchlichen Ergebnissen kommen. Einige Analysen deuten auf eine mögliche (Kosten) Effektivität hin. Auf Ebene einzelner Qualitätsindikatoren konnten positive Ergebnisse erzielt werden, es fehlt jedoch an einem zweifelsfreien Nachweis. Dies ist vor allem auf die Heterogenität der bis dato durchgeführten P4P-Programme und der eingesetzten Evaluationsinstrumente zurückzuführen.

Fazit

Die eruierten Ergebnisse haben gezeigt, dass ein eindeutiger Beleg zur Wirksamkeit von P4P bis dato nicht gegeben ist. Bevor P4P als qualitätsorientiertes Steuerungsinstrument eingesetzt und mit der Implementierung von P4P-Programmen begonnen werden kann, werden noch weitere Forschungsarbeiten, welche die lessons learned internationaler Programme und die Erfordernisse und Umgebungsfaktoren des nationalen Gesundheitswesens berücksichtigen, empfohlen. In diesem Zusammenhang sei erwähnt, dass P4P nur eine von vielen Möglichkeiten zur Optimierung der Qualität im Gesundheitswesen darstellt.

b) Inhalt

1	Einführung.....	1
1.1	Einleitung	1
1.2	Hintergrund	1
1.3	Zielsetzung.....	3
2	Methode	4
3	Wesentliche Elemente und Limitationen von P4P.....	5
3.1	Organisationsformen	5
3.2	Qualitätsmessung	5
3.3	Anreizgestaltung	7
3.4	Chancen und Risiken von P4P-Programmen	8
4	Aktuelle Erkenntnisse – wissenschaftlicher Status quo.....	11
4.1	Beispiel eines P4P-Programms.....	11
4.1.1	Wesentliche Merkmale – Quality and Outcomes Framework.....	11
4.1.2	Evaluationsergebnisse – Quality and Outcomes Framework.....	12
4.2	Wissenschaftliche Evidenz zur Wirksamkeit von P4P-Programmen	12
4.2.1	Erläuterung und Darstellung der wissenschaftlichen Evidenz.....	13
4.2.2	Wissenschaftliche Erkenntnisse zu P4P in Österreich.....	14
5	Best Practice bei der Implementierung von P4P-Programmen	16
5.1	Qualitätsmessung	16
5.2	Anreize und Motivation	17
5.3	Umsetzung	17
6	Zusammenfassung und Fazit	19
	Literaturverzeichnis	21
	Anhang.....	22

c) Abbildungs- und Tabellenverzeichnis

Abbildungen

Abbildung 1: Qualitätsindikatoren des QOF	12
---	----

Tabellen

Tabelle 1: Wesentliche Merkmale der Qualitätsdimensionen.....	6
Tabelle 2: Chancen von P4P-Programmen	9
Tabelle 3: Risiken von P4P-Programmen.....	9
Tabelle 4: Checkliste für die Ausgestaltung und Implementierung eines P4P-Programms ...	22

1 Einführung

1.1 Einleitung

Seit längerem ist Pay for Performance (P4P) sowie dessen Effektivität Gegenstand internationaler gesundheitspolitischer Diskussionen. Das qualitätsorientierte Steuerungsinstrument wird vor allem im Zusammenhang mit der Forderung nach einer optimalen Ressourcenallokation und Versorgungsqualität oft diskutiert. Auch im Zuge der Reformbestrebungen im österreichischen Gesundheitswesen wird das Thema vermehrt in den Fokus des nationalen Interesses treten.

Laut Emmert (2013) können Unterschiede in der Versorgungsqualität auf Fehlanreize der bisherigen Vergütungsformen und mangelnde Transparenz zurückgeführt werden. Durch die Koppelung der Vergütung der ärztlichen Leistung an das erbrachte und durch Kennzahlen dargestellte Leistungsniveau zielt das P4P - Konzept deshalb darauf ab, die Qualität und die Effizienz der Versorgung zu optimieren (Emmert 2013).

Vor allem im angelsächsischen Raum steigt die Anzahl an P4P - Programmen seit einigen Jahren. P4P stellt grob genommen eine Honorarform dar, bei welcher die Vergütung anhand der erbrachten und durch Messung von Qualitätsindikatoren nachgewiesenen ärztlichen Leistung erfolgt (Emmert 2013, Veit et al. 2012).

1.2 Hintergrund

Vergangene Reformbestrebungen in unterschiedlichen Ländern orientierten sich vorwiegend an Kostensenkungsmaßnahmen. Die Tatsache, dass in vielen Gesundheitssystemen Defizite in der Versorgungsqualität vorherrschen, kann u.a. auf die bestehende Vergütungssystematik zurückgeführt werden, da von dieser nur wenige Anreize zur Optimierung der Behandlungsqualität ausgehen. Gegenwärtig rücken deshalb Qualitätsverbesserungsinitiativen immer mehr in den Fokus. Neben der Optimierung der Behandlungsqualität erhofft man sich auch positive Auswirkungen auf künftig anfallende Behandlungskosten (Emmert 2008).

Vor dem Hintergrund steigender Kosten der Gesundheitsversorgung und gleichzeitig nachweisbaren Qualitätsmängeln, nimmt die Forderung nach Qualitätsverbesserungsprogrammen in einigen Gesundheitssystemen zu. Als Steuerungsinstrument wird P4P seit zwei Jahrzehnten vor allem im englischsprachigen Raum eingesetzt. Das Konzept hat seine Ursprünge in den 1990iger Jahren in den USA. Bereits im Jahr 1993 haben einige Health Maintenance Organizations (HMOs)¹ einzelne qualitätsbezogene Komponenten in ihre Vergütungsverträge aufgenommen. In Europa nimmt das britische Konzept „Quality and Outcomes Framework“ (QOF) aus dem Jahr 2004 eine Vorreiterrolle ein (Scheppach et al. 2011, Veit et al. 2012).

¹ Krankenversicherungs- und Versorgungsmodell; vgl. <https://www.medicare.gov/find-a-plan/staticpages/glossary/planfinder-glossary.aspx?let=H>

Viele AutorInnen haben das Konzept P4P beschrieben und definiert. Scheppach et al. (2011) haben eruiert, dass nahezu alle Definitionen des Konzepts, die Leistungsvergütung in Relation mit der Qualität und Effektivität bzw. Effizienz der Behandlung setzen (Scheppach et al. 2011). In den Fokus rücken demnach Qualitätsindikatoren und die Anreizgestaltung. Im Folgenden wird versucht P4P in Anlehnung an Veit et al. (2012) genauer zu definieren:

- Beim P4P handelt es sich um eine spezielle strategische Form der Vergütung. Das gesundheitspolitische Ziel, welches durch die Einführung derartiger Projekte verfolgt wird, ist die gleichzeitige Steuerung von Qualität und Effizienz in der Gesundheitsversorgung. Das von den GesundheitsversorgerInnen erbrachte Leistungsniveau wird dabei durch Kennzahlen dargestellt und an die Vergütung gekoppelt. Dadurch soll die Motivation gesteigert werden, die Versorgung nachhaltig und umfassend zu optimieren und weiterzuentwickeln (Veit et al. 2012).

Im Zuge der Bestrebung, die Qualität und/oder Effizienz in der Gesundheitsversorgung zu erhöhen, haben sich noch weitere Formen von Vergütungen entwickelt. Diese werden ohne Anspruch auf Vollständigkeit kurz skizziert:

- Bei **Pay for Competence** handelt es sich um eine Vergütungsform, bei welcher nicht die erreichte Qualität sondern Voraussetzungen - in Form von geschaffenen Strukturen – unter denen eine bessere Versorgungsqualität geleistet werden kann, durch Zahlungen gefördert werden (Veit et al. 2012).
- Wenn LeistungserbringerInnen ihre Versorgungsqualität messen und gegenüber Dritten darlegen, spricht man von **Pay-for-Transparency**. Bei dieser Form der Vergütung wird lediglich die Teilnahme und nicht - wie bei P4P - die Erreichung eines Zielwertes vergütet (Veit et al. 2012).
- Die **prospektive Leistungsvergütung** von Einzelleistungen oder Leistungspaketen.

Laut Veit et al. (2012) lassen sich die Ziele, welche durch die Implementierung von P4P-Programmen verfolgt werden, wie folgt kategorisieren:

- Förderung und Weiterentwicklung hervorragender Versorgungsqualität
- Nachhaltige Verbesserung von Versorgungsqualität in einem mittleren Leistungsbereich
- Stringente Verbesserung oder Verhinderung von Versorgungsqualität, welche unterhalb eines akzeptierten Niveaus liegt

Jede dieser Kategorien stellt eine andere Anforderung an die Qualitätsindikatoren und Anreize dar. Deshalb ist eine eindeutige Zielformulierung bei der Implementierung eines P4P-Projektes essentiell, damit Ziele und Instrumente zueinander passen (Veit et al. 2012). Um das übergeordnete Ziel der Qualitätsverbesserung durch die Umsetzung eines P4P-Projektes zu erreichen, setzen die Anreizprogramme überwiegend bei den LeistungserbringerInnen an, welche zu einer qualitativ hochwertigen Behandlung angespornt werden sollen (Emmert 2008).²

Im Hinblick auf die Ausgestaltung können P4P-Programme unterschiedlich organisiert sein. So können beispielsweise Netzwerke ihren angestellten ÄrztInnen Bonuszahlungen für verbesserte Qualitätskriterien gewähren, aber auch einzelne Krankenversicherungen oder

² Vgl. Kapitel 3.2 und 3.3

der Staat können P4P-Programme ausgestalten (Emmert 2008). P4P-Programme werden überwiegend im hausärztlichen Setting umgesetzt, während im Krankenhausbereich vereinzelt Projekte umgesetzt werden (Veit et al. 2012). In Kapitel 4.1 erfolgt exemplarisch die Darstellung des Referenzprojektes „Quality and Outcomes Framework“ (QOF) im hausärztlichen Setting.³

1.3 Zielsetzung

Das Ziel dieses Dossiers ist die sachliche Aufbereitung der wissenschaftlichen Evidenz zur Wirksamkeit von P4P-Programmen sowie zur Realisierung von eben diesen Programmen. Mit der Aufbereitung der vorhandenen Literatur soll vor allem folgende Frage beantwortet werden:

- *Welche Aussagen zur Evidenz im Hinblick auf die Wirksamkeit von P4P-Projekten können mithilfe von aktuellen wissenschaftlichen Daten getroffen werden?*

Dabei soll mithilfe von Erfahrungswerten neben der Effektivität auch auf die Effizienz von P4P-Projekten eingegangen werden.

Aufgrund der Tatsache, dass die Wirksamkeit von P4P-Projekten möglicherweise durch Umsetzungs- und Umgebungsfaktoren beeinflusst wird, wird in diesem Dossier auch auf evidente Empfehlungen zur Konzipierung und Implementierung eingegangen.

³ Vgl. Kapitel 3.1

2 Methode

Um die Frage nach der Evidenz der Wirksamkeit von P4P-Programmen beantworten zu können, folgt dieses Dossier einem theoretischen Ansatz. Dabei werden zum Großteil die Erkenntnisse des vorliegenden Gutachtens des Instituts für Qualität und Patientensicherheit in die Analyse einbezogen ebenso wie ein systematischer Review von Scheppach et al. (2011) und eine Übersichtsarbeit von Emmert (2008). Schließlich werden die Ergebnisse des systematischen Reviews von Ogundeji et al. (2016) zur Effektivität von P4P berücksichtigt. Um eventuell relevante Literatur zu berücksichtigen, welche zeitlich nach den eben erwähnten publiziert wurde, wurde zusätzlich eine grobe Internetrecherche mit Fokus auf systematischen Reviews und gesundheitspolitischen Stellungnahmen durchgeführt.

Bevor in Kapitel 4 damit begonnen wird, die Ergebnisse im Hinblick auf die wissenschaftliche Evidenz zur Wirksamkeit von P4P darzulegen, erfolgt zunächst in Kapitel 3 eine Darstellung der Grundlagen zu den wesentlichen Elementen von P4P (Ausgestaltungsmöglichkeiten, Qualitätsbegriff und Anreizgestaltung). Darüber hinaus werden Limitationen, in Form von Chancen und Risiken, welche im Zuge der Konzipierung und Implementierung von P4P-Programmen berücksichtigt werden sollten, aufgezeigt. Die Ergebnisse dieser Recherchen regen dazu an, Best practice Aspekte, welche bei der Konzipierung und Implementierung von P4P-Programmen Beachtung finden sollten, anzuführen. In diesem Zusammenhang erfolgt in Kapitel 5 eine Darstellung der lessons learned.

3 Wesentliche Elemente und Limitationen von P4P

3.1 Organisationsformen

Bei der Analyse bereits eingeführter P4P-Programme wird deutlich, dass diese oftmals sehr unterschiedliche Vorgehensweisen aufweisen. Diese lassen sich grob unterscheiden in:

- **klassische Bonus-Projekte** - ein Teil der Vergütung wird an die Erfüllung bestimmter Ziel- oder Schwellenwerte, welche mittels etablierten und bestehenden Qualitätsindikatorensets gemessen werden, gekoppelt
- **Zielprämien (Target Payments)** - zusätzliche Einmalzahlungen für bestimmte Leistungen
- **Non-Pay-for-Non-Performance** - Zahlungsstopp bei schwerwiegenden Versorgungsmängeln
- **Provider-driven P4P** – Initiierung erfolgt durch LeistungserbringerInnen und koppelt die Vergütung an die Optimierung einer bestimmten Versorgungssituation, wenn diese als unbefriedigend definiert wird
- **Beteiligungen und Einsparungen (Shared-Savings-Ansätze)** – Ziel ist die Erreichung einer Kosteneffizienz durch die Verbesserung von Abläufen
- **Übertragung Budget- und Qualitätsverantwortung von Versorgergruppen (Accountable Care Organizations)**
- Ferner kommen bei vielen Projekten neben den monetären auch nicht-monetäre Anreize zur Zielerreichung zum Einsatz⁴ (Veit et al. 2012).

Der überwiegende Anteil der P4P-Programme findet im hausärztlichen Setting statt⁵. Im Krankenhausbereich werden ebenfalls bereits einige Projekte umgesetzt, deren Anzahl jedoch nicht mit jenen des niedergelassenen Bereichs vergleichbar ist. Alle Programme haben die Verwendung von etablierten Indikatorensets gemeinsam (vorwiegend Struktur- und Prozessindikatoren, wenige Ergebnisindikatoren).⁶ Betrachtet man die Vergütung, so handelt es sich nicht um eine Umgestaltung bestehender Anreizsysteme. Überwiegend wird - neben der etablierten Basisvergütung (z.B. Einzelleistungsvergütung, Pauschalen) - auf zusätzliche finanzielle Anreize gesetzt um eine Optimierung der Qualität zu erreichen⁷ (Veit et al. 2012).

3.2 Qualitätsmessung

Der Qualitätsbegriff nimmt im Zusammenhang mit der Optimierung und Steuerung von Qualität im P4P eine zentrale Rolle ein. Eine valide, mit vertretbarem Aufwand realisierbare,

⁴ vgl. Kapitel 3.3

⁵ vgl. beispielsweise QOF

⁶ Vgl. Kapitel 3.2

⁷ vgl. Kapitel 3.2 und 3.3

Qualitätsmessung stellt eine der größten Herausforderungen für P4P-Programme dar (Veit et al. 2012).

Eine einheitliche Definition von Qualität im Gesundheitswesen ist herausfordernd und von der jeweiligen Perspektive⁸ abhängig. Nach Scheppach et al. (2011) in Anlehnung an Fleßa (2007) ist den Qualitätssichtweisen von PatientInnen, ÄrztInnen und der Krankenkasse der Behandlungserfolg/das Behandlungsergebnis gemein. Um dieses Ziel zu erreichen, stehen je nach Perspektive unterschiedliche Aspekte im Fokus des Interesses. Die PatientInnen definieren Qualität nach einer subjektiv als gut eingeschätzten Versorgung, der Serviceleistung, der Zuwendung und Atmosphäre. ÄrztInnen und die Krankenkassen definieren Qualität anhand des Einsatzes von Behandlungs- und Diagnostikmethoden nach dem aktuellen Stand der Technik und Wissenschaft, einer niedrigen Komplikationsrate und niedrigen Kosten (Scheppach et al. 2011).

Im Hinblick auf die Definition des Qualitätsbegriffs kommen Emmert (2008) in seiner Übersichtsarbeit ebenso wie Veit et al. (2012) in ihrem Gutachten zu dem Schluss, dass sich die Mehrheit der AutorInnen dabei einig ist, dass die von Donabedian eingeführte Dreiteilung in die Dimensionen Struktur-, Prozess- und Ergebnisqualität eine zentrale Rolle einnimmt und bei der Entwicklung eines Qualitätsmessinstruments nicht unberücksichtigt bleiben darf (Emmert 2008, Veit et al. 2012).

Auf Basis der theoretischen Ausführungen wird in Tabelle 1 der Versuch unternommen, wesentliche Merkmale der Qualitätsdimensionen nach Donabedian gegliedert darzustellen.

Tabelle 1: Wesentliche Merkmale der Qualitätsdimensionen

	Struktur	Prozess	Ergebnis
Definition	Gibt an, welche Voraussetzungen für die Patientenbehandlung zur Verfügung stehen (<i>personelle, technische und finanzielle Ressourcen</i>)	Gibt an, ob das „Richtige, richtig getan wird“. Alle für eine Behandlung erforderlichen Prozessschritte werden analysiert und optimiert	Gibt an, inwiefern sich der Gesundheitszustand der PatientInnen verändert hat (misst den Zielerreichungsgrad einer therapeutischen Maßnahme)
Beispiele	Qualifikation der ÄrztInnen Erreichbarkeit der Praxis	Anamnese- und Untersuchungstechnik Zusammenarbeit mit KollegInnen	Besserung oder Heilung von Erkrankungen Patientenzufriedenheit
Messbarkeit	Leicht messbar	Messung ist aufwendig und kostspielig	Wird beeinflusst durch Herausforderungen
Herausforderung	Reine Orientierung an dieser Dimension sagt nichts über die Qualität des Behandlungsprozesses aus	Prozesskriterien zeigen nicht auf, ob sich die Gesundheit der PatientInnen tatsächlich verbessert hat	Behandlungsergebnisse werden von mehreren Faktoren beeinflusst: Zuordnung zu einem Versorger aufgrund längeren Verläufen oder komplexeren Therapien Compliance der PatientInnen Langfristige Behandlungsergebnisse Chronische Erkrankungen Etc.

⁸ PatientInnen, ÄrztInnen, Krankenkasse

Eignung als P4P-Grundlage	Struktur korreliert sehr schwach mit Ergebnis und kann Performance ermöglichen (eher mit Pay-for-Competence vergleichbar)	Indikatoren dienen der Absicherung einer funktionalen Qualität, sind jedoch noch kein Garant für gute Versorgungsqualität. Prozessindikatoren sind geeignet, da sie für die Anbieter direkter zu beeinflussen. Darüber hinaus werden sie eher weniger von externen Faktoren beeinflusst.	Vor allem die Ergebnisqualitätsindikatoren nehmen eine zentrale Rolle in P4P-Programmen ein, da sie die Qualität der Versorgung messen Ergebnisindikatoren müssen methodisch sehr qualitativ sein, damit sie für P4P-Programme verwendet werden können
---------------------------	---	--	---

Quelle: Eigene Darstellung in Anlehnung an Emmert 2008, Scheppach et al. 2011, Ogundeji et al. 2016, und Veit et al. 2012

Die einzelnen Dimensionen beeinflussen sich in einem gewissen Grad. Gute Struktur- und Prozessqualität können beispielsweise eine gute Ergebnisqualität allenfalls fördern, jedoch nicht garantieren (Emmert 2008).

Je nach Dimension müssen für die Messung der Qualität geeignete Messgrößen, so genannte Indikatoren, definiert werden. Dadurch kann die, durch die LeistungserbringerInnen erbrachte Versorgung, qualitativ bewertet werden (Scheppach 2011).

Die Ergebnisindikatoren sind für die Messung der Versorgungsqualität methodisch qualitativ hochwertig sein, da die Messung komplex und von vielen externen Faktoren abhängig ist. Prozessindikatoren sind vom Anbieter beeinflussbarer und weniger von externen Faktoren (Emmert 2008, Ogundeji et al. 2016, Scheppach et al. 2011).

3.3 Anreizgestaltung

Neben der Auswahl geeigneter Indikatoren zur Messung des Leistungsniveaus, hängt der langfristige Erfolg von P4P auch von der Akzeptanz der Programme durch die LeistungserbringerInnen und deren Motivation zu einer Verhaltensänderung ab. In diesem Zusammenhang ist die Anreizgestaltung sowohl finanzieller als auch nicht-finanzieller Art ein entscheidender Aspekt bei der Implementierung von P4P-Programmen (Emmert 2008, Scheppach et al. 2011).

Veit et al. (2012) verweisen darauf, dass es sich bei den eigentlichen P4P-Anreizen um monetäre Anzelemente (Bonuszahlungen, Abschläge oder Varianten davon) handelt (Veit et al. 2012). Prinzipiell bestehen mehrere Möglichkeiten, die finanziellen Anreize an die LeistungserbringerInnen zu verteilen:

- Vergütung der absoluten Zielerreichung – der Bonus wird dann gewährt, wenn ein vorher definierter Zielwert erreicht wurde
- Vergütung der relativen Zielerreichung – der Bonus wird dann gewährt, wenn die LeistungserbringerInnen beispielsweise zu den zehn Prozent der Besten gezählt werden können
- Veränderung zum Vorjahr – hier wird die Leistungsverbesserung im Vergleich zu Vorjahres- oder Vergangenhheitswerten als Bemessungsgrundlage herangezogen
- Abzüge für LeistungserbringerInnen, welche einen bestimmten Schwellwert unterschreiten (Scheppach et al. 2011, Veit et al. 2012)

Die Wahl des Anreizmodells hängt davon ab, welche Ziele mit dem Programm verfolgt werden. Auch Kombinationen der Modelle sind möglich.

Neben den Verteilungsvarianten gilt es laut Scheppach et al. (2011) folgende Punkte bei der Gestaltung der Anreize zu berücksichtigen bzw. festzulegen:

- **Empfängerebene:** Anreizzahlung erfolgt entweder individuell oder auf Gruppenebene
- **Bemessungsgrundlage:** Bonuszahlung erfolgt entweder bei absoluter oder relativer Zielerreichung oder bei einer Veränderung zum Vorjahr
- **Anreizhöhe:** bisher noch keine einheitliche Meinung über die optimale Höhe des Anreizes
- **Finanzierung der Anreize:** neben den Einrichtungskosten fallen hohe Ausgaben für die Anreizzahlungen an, ebenso wie Aufwendungen für die Datenanalyse und die Verwaltung des Belohnungssystems. Privatwirtschaftliche Arbeitgeber, Krankenkassen oder öffentliche Einrichtungen können Sponsoren der P4P Programme sein.
- **Dauer der Anreizgewährung:** eine angemessene Projektlaufzeit sollte gewählt werden, um eine langfristige Verhaltensänderung bei den LeistungserbringerInnen zu bewirken

Nicht-monetäre Anreize werden in vielen Programmen in Kombination mit monetären Anreizen als Steuerungsinstrumente verwendet. Als nicht-finanzielle Qualitätsverbesserungsmaßnahmen werden u.a.

- die Veröffentlichung von ärztlichen Leistungsdaten (Public Reporting)
- interne Qualitätsmessung mit Feedback
- moderne Technologie-Ausstattung
- Schulungsmaßnahmen oder
- verpflichtende Leitlinien genannt (Emmert 2008, Scheppach et al. 2011, Veit et al. 2012).

Die Kombination von monetären und nicht-monetären Anreizen kann sich einerseits positiv auf die Qualität auswirken, andererseits wird die Evaluierung der Effektivität einzelner Steuerungsinstrumente dadurch beeinflusst und erschwert (Scheppach et al. 2011, Veit et al. 2012)

3.4 Chancen und Risiken von P4P-Programmen

Internationale Erfahrungen haben gezeigt, dass die qualitätsorientierte Vergütung im Rahmen der P4P-Programme sowohl mit Stärken als auch Schwächen behaftet ist. Basierend auf der verwendeten Literatur wird im Folgenden zunächst der Versuch unternommen, die wesentlichen Vorteile von P4P-Programmen als Chancen zu kategorisieren (vgl. Tabelle 2).

Tabelle 2: Chancen von P4P-Programmen

Chancen	Vorteile
Optimierung der Behandlungsqualität	<ul style="list-style-type: none"> ✓ Anreiz für LeistungserbringerInnen eine qualitativ hochwertige Behandlung durchzuführen (v.a. im Bereich chronischer Erkrankungen und Früherkennung/Prävention) ✓ Qualitätsverbesserung der Versorgung ✓ Reduktion von Behandlungs- als auch Medikationsfehler durch den Einsatz evidenzbasierter Medizin
Mehr Transparenz im Gesundheitswesen	<ul style="list-style-type: none"> ✓ Langfristige Messung der Qualitätsindikatoren ✓ P4P häufig in Kombination mit Veröffentlichung der Leistungsdaten
Kostenoptimierung	<ul style="list-style-type: none"> ✓ Senkung der Wahrscheinlichkeit für Folgeerkrankungen und Komplikationen

Quellen: Emmert 2008, Scheppach et al. 2011, Veit et al. 2012

Gefahren, welche mit der Implementierung verbunden sein können, werden in Tabelle 3 in Form von Risiken dargestellt.

Tabelle 3: Risiken von P4P-Programmen

Risiken	Gefahren
Selektionseffekte	<ul style="list-style-type: none"> ✓ Cherry Picking (Selektion der PatientInnen) ✓ Selektion von Maßnahmen (evtl. falsche Anreize durch vorgegebene Prozessschritte)
Beeinflussung der medizinischen Praxis	<ul style="list-style-type: none"> ✓ Gaming (verändertes Kodierverhalten um Bonuszahlungen zu erhöhen) ✓ Indikatorenbereiche, welche mit einer Bonuszahlung verbunden sind werden bevorzugt, während andere nicht belohnte Tätigkeiten evtl. vernachlässigt werden
Aussagekraft der Messergebnisse	<ul style="list-style-type: none"> ✓ Einhaltung von Prozessvorgaben garantiert nicht automatisch eine hervorragende Behandlungsqualität (z.B. bei multi-morbiden PatientInnen) ✓ Komplexität der Gesundheit
Akzeptanz der ÄrztInnen	<ul style="list-style-type: none"> ✓ Bürokratischer Mehraufwand je nach Ausgestaltung der Programme ✓ Individuelle Therapiefreiheit wird evtl. eingeschränkt ✓ P4P-Programme ✓ Senkung der intrinsischen Motivation

Quellen: Emmert 2008, Scheppach et al. 2011, Veit et al. 2012

Veit et al. (2012) weisen in ihrem Gutachten darauf hin, dass diese Risiken bei der Konstruktion der Indikatoren, der Dokumentation und dem Controlling berücksichtigt werden sollen und dabei auf Erfahrungen von bereits implementierten Programmen zurückgegriffen werden kann⁹ (Veit et al. 2012).

Folgende Maßnahmen können getroffen werden um Risiken vorzubeugen:

- Wirksame Dokumentationskontrollen durch statistische Plausibilitätsprüfungen und Stichprobenverfahren um Fehlcodierung zu vermeiden
- Präzise Indikatoren gegen unsachgemäße Leistungsausweitung wegen finanzieller Anreize
- Prüfung von Qualitätsindikatoren auf ihr Potential, Fehlanreize zu setzen

⁹ Vgl. Kapitel 5

- Risikoadjustierung zur Minderung des Anreizes zum Therapieausschluss von Hochrisikopatienten
- Umfassendes Indikatorenset zum Ausschluss von Vernachlässigungen durch Fokussierung auf P4P-Kennzahlen (Veit et al. 2012).

4 Aktuelle Erkenntnisse – wissenschaftlicher Status quo

4.1 Beispiel eines P4P-Programms

Seit einigen Jahren steigt die Anzahl der Länder, welche P4P als Steuerungsinstrument im Gesundheitswesen implementieren. Diese sind in ihrem Aufbau sehr heterogen und unterscheiden sich hinsichtlich der Anreize, Indikatoren, etc. Das Quality and Outcomes Framework (QOF) aus Großbritannien startete 2004 und nimmt in Europa eine Vorreiterrolle im Hinblick auf P4P-Programme ein (Scheppach et al. 2011).

4.1.1 Wesentliche Merkmale – Quality and Outcomes Framework

Im Folgenden wird der Versuch unternommen, die wesentlichen Merkmale eines P4P-Programms exemplarisch anhand der Inhalte des QOF darzustellen.

Das QOF wurde vom Nationalen Gesundheitsdienst (NHS) mit dem Ziel implementiert, die hausärztlich erbrachte Qualität mittels finanzieller Anreize zu verbessern. Die Teilnahme an dieser Form der qualitätsorientierten Vergütung erfolgt freiwillig. In etwa 25 Prozent des jährlichen Einkommens der HausärztInnen ist abhängig vom Erreichen von ca. 146 evidenzbasierten Qualitätsindikatoren. Gemessen wird die Qualität hauptsächlich in den Sektoren chronische Krankheiten, Prävention und Praxisversorgung. Wie Abbildung 1 verdeutlicht, wurde versucht die hausärztliche Tätigkeit in allen Facetten abzubilden, weshalb die Messung der Qualität in den vier verschiedenen Bereichen Patientenversorgung, Patientenorganisation, Patientenzufriedenheit und zusätzlichen Leistungen erfolgt (Emmert 2008, Scheppach et al. 2011, Veit et al. 2012).

Im Zuge des P4P-Programms besteht die Möglichkeit des *Exception Reportings*. In diesem Fall können LeistungserbringerInnen, PatientInnen unter bestimmten und definierten Voraussetzungen für die Messung eines Qualitätsindikators als nicht geeignet einstufen und somit von der Performance Messung ausschließen (Emmert 2008).

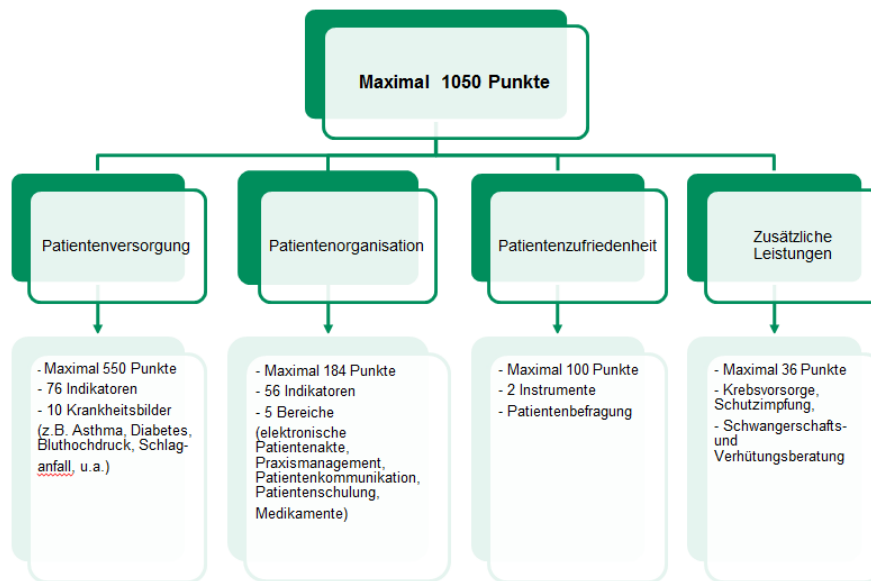


Abbildung 1: Qualitätsindikatoren des QOF¹⁰

4.1.2 Evaluationsergebnisse – Quality and Outcomes Framework

Veit et al. (2012) kommen in ihrem Gutachten zu dem Schluss, dass die Ergebnisse der Evaluationen des QOF ein inkonsistentes Bild zeigen. In den meisten Fällen ist dies auf die gewählte Untersuchungsmethodik zurückzuführen, denn einige Evaluationen beispielsweise betrachten lediglich einzelne Aspekte - z.B. Verbesserung bestimmter Qualitätsindikatoren, welche überwiegend Prozess- und Strukturindikatoren sind. Auf Ebene der Outcomes fehlt nach wie vor ein Nachweis eines Gesamtnutzens. Darüber hinaus handelt es sich in den meisten Fällen um Beobachtungsstudien, kontrollierte Interventionsstudien fehlen (Veit et al. 2012).

Ferner zeigen die Ergebnisse einer Kosteneffektivitätsanalyse zu QOF von Walker et al. (2010), dass das Programm „wahrscheinlich“ kosteneffektiv ist. Gemessen wurde der Zugewinn an sogenannten „Quality Adjusted Live Years“ (QALYs). Als Limitation ist festzuhalten, dass bei der Analyse die ausgewählten Indikatoren nur ein kleines Subset aller darstellen und die administrativen Kosten nicht berücksichtigt wurden (Veit et al. 2012).

Betrachtet man die unterschiedlichen Evaluationsergebnisse des QOF, so kann festgehalten werden, dass die Evidenz der Wirksamkeit in diesem P4P Programm derzeit nicht zweifelsfrei nachgewiesen werden kann.

4.2 Wissenschaftliche Evidenz zur Wirksamkeit von P4P-Programmen

In einigen Ländern werden P4P-Programme bereits als Steuerungsinstrumente im Gesundheitswesen eingesetzt. Um die Evidenz hinsichtlich Wirksamkeit dieser Maßnahmen zu eruieren, liegen Evaluierungsberichte vor, welche bereits von einigen AutorInnen in Form

¹⁰ in Anlehnung an Scheppach et al. (2011) nach Mc. Donald et al. (2009), Meddings u. McMahon (2008)

von Reviews und Metaanalysen zusammenfassend betrachtet wurden. Ziel dieses Kapitels ist es, den Status quo dieser Arbeiten basierend auf den verwendeten Literaturquellen darzustellen.

4.2.1 Erläuterung und Darstellung der wissenschaftlichen Evidenz

Bevor mit der Erläuterung der Evaluationsergebnisse begonnen wird, erfolgt der Verweis darauf, dass die Projekte und Programme im Zusammenhang mit P4P methodisch heterogen sind und eine generelle Beurteilung folglich als schwierig erachtet wird. Darüber hinaus betrachten die Reviewer die Ergebnisse in den meisten eingeschlossenen Studien aufgrund von Limitationen¹¹ in den Studiendesigns mit Vorsicht (Veit et al. 2012).

Veit et al. (2012) haben in ihrem Gutachten u.a. eine systematische Literaturrecherche nach Primär- und Sekundärliteratur, welche sich mit der Evidenz zur Wirksamkeit von P4P befassen, durchgeführt. Untersuchungen, welche die Langzeiterfahrungen mit P4P-Programmen eruieren, haben gezeigt, dass zwar anfänglich Qualitätsverbesserungen erzielt werden konnten, langfristig gesehen jedoch keine weiteren Verbesserungen eingetreten sind. Ferner konnte festgestellt werden, dass sich die Ergebnisse bei Wegnahme der finanziellen Anreize wieder verschlechterten. Es kann demnach nicht von einer dauerhaften Etablierung von Prozess- und Verhaltensänderungen durch P4P ausgegangen werden. Durch die, in vereinzelt Projekten bzw. Programmen vorgenommene Koppelung unterschiedlicher Interventionen ist die Wirksamkeit der zusätzlichen finanziellen Anreize nicht sauber von der Wirkung anderer Maßnahmen (z.B. Feedback) abgrenzbar. Die Evaluationen großer und bedeutsamer P4P-Programme haben gezeigt, dass die Evidenz für die Wirksamkeit finanzieller Anreize sowohl im stationären wie auch im ambulanten Bereich schwach und oftmals nur auf Ebene einzelner Qualitätsindikatoren nachweisbar ist (Veit et al. 2012).

Im Zusammenhang mit möglichen Einsparungen muss beachtet werden, dass Kosten auf unterschiedlichen Ebenen entstehen können und Einsparungen nicht unbedingt auf derselben Ebene und zeitnah erfolgen. Vor allem bei präventiven Maßnahmen oder langfristigem Outcome manifestieren sich Einsparungen erst viel später und auf Population-Health-Ebene (Veit et al. 2012).

Zusammenfassend betrachtet, kommen Veit et al. (2012) in ihrem Gutachten zu dem Fazit, dass trotz langjähriger Praxis nach wie vor eine verlässliche Evidenzbasis für die Wirksamkeit von P4P-Programmen fehlt. Um diese auf einem hohen Evidenzlevel zu eruieren, empfehlen die AutorInnen prospektive kontrollierte Interventionsstudien (Veit et al. 2012).

In einem systematischen Review von systematischen Reviews analysieren Eijkenaar et al. (2013) unter anderem die wissenschaftliche Evidenz zur Effektivität von P4P. Generell können zur Effektivität von P4P folgende Aussagen getroffen werden: auf eine mögliche (Kosten-) Effektivität wird hingedeutet. Die Ergebnisse sind jedoch widersprüchlich und die Evidenz konnte bis dato noch nicht zweifelsfrei nachgewiesen werden. Die AutorInnen führen dies – vor allem bei den Reviews zum Nachweis der Kosten-Effektivität – auf

¹¹ Fehlende Vergleichsgruppe; keine Berücksichtigung zeitlicher Trends, keine klare Abgrenzung der Auswirkungen des finanziellen Incentives von den Auswirkungen anderer qualitätsfördernder Maßnahmen, fehlende Datenvalidierung, sehr kleiner beforschter Versorgungsausschnitt, überwiegend Prozess- und Strukturindikatoren

Limitationen und der Heterogenität der Studiendesigns zurück. Zur Erueierung der Wirksamkeit werden Evaluationsverfahren empfohlen, welche vermehrt die Wirksamkeit im Hinblick auf den Outcome eruieren (Eijkenaar et al. 2013).

Die Widersprüchlichkeit der Ergebnisse zeigt sich beispielsweise bei den Evaluationsergebnissen zur Optimierung der Behandlungsqualität von chronischen Erkrankungen. Studienergebnisse haben unter anderem einen geringen positiven Impact bei der Versorgung von Diabetes und Asthma gezeigt, jedoch nicht bei kardiovaskulären Erkrankungen. Zusammenfassend kann festgehalten werden, dass die Wirksamkeit der P4P-Programme von Umsetzungs- und Umgebungsfaktoren abhängig ist (vgl. Kapitel 5) (Eijkenaar et al. 2013).

Scheppach et al. (2011) kommen zu dem Resümee, dass die Beurteilung der Effektivität von P4P mit Schwierigkeiten verbunden ist, da die monetären Anreize oft in Kombination mit anderen, teilweise nicht-monetären Qualitätsverbesserungsinitiativen gekoppelt werden. Die AutorInnen sehen in der kontinuierlichen Überwachung und Vermeidung von unerwünschten Nebenwirkungen einen entscheidenden Aspekt für den langfristigen Erfolg eines P4P-Programms (Scheppach et al. 2011).

Die Ergebnisse eines weiteren systematischen Review von Van Herck et al. (2010) haben gezeigt, dass P4P Projekte zwar überwiegend positive Auswirkung auf klinische Endpunkte haben können, im Hinblick auf komplexe Endpunkte (z.B. Koordination, Kontinuität, Kosteneffektivität) jedoch widersprüchliche Ergebnisse aufweisen (Van Herck et al. 2010 in Schrappe 2013).

Ogundeji et al. (2016) untersuchen in ihrem systematischen Review wie P4P die Effektivität hinsichtlich Verbesserung der Gesamtperformance beeinflussen können. Auch für diese AutorInnen ist eine eindeutige Empfehlung nicht möglich, da diese Effektivität innerhalb der analysierten Studien variiert. Dies ist vor allem auch auf die unterschiedliche Ausgestaltung dieser Honorierungsform in den heterogenen Systemen zurückzuführen (Ogundeji et al. 2016).

Werden starke Versorgungsdefizite in bestimmten Bereichen (z.B. Behandlung von ChronikerInnen) eruiert bzw. erfolgt bei gewissen Krankheitsbildern noch keine hochwertige Versorgung, so kann P4P laut Emmert (2013) als Steuerungsinstrument eingesetzt werden um die Versorgungsqualität auf ein höheres Level zu bringen. Zusammenfassend sollte P4P nicht als Instrument gesehen werden, welches das Gesundheitssystem revolutioniert, sondern als ein Element eines umfassenden Qualitätsverbesserungsansatzes (Emmert 2013).

Betrachtet man die Studienergebnisse so liegt ein eindeutiger Beleg zur Wirksamkeit von P4P bis dato nicht vor. Darüber hinaus muss bedacht werden, dass das qualitätsorientierte Steuerungsinstrument P4P nur eine der Möglichkeiten zur Optimierung der Qualität im Gesundheitswesen darstellt.

4.2.2 Wissenschaftliche Erkenntnisse zu P4P in Österreich

In seinem Referat zu P4P im Gesundheitswesen verweist Emmert (2013) auf einen von ihm durchgeführten „kleinen“ systematischen Review zu den Erkenntnissen zu P4P in Österreich. Zusammenfassend kann festgehalten werden, dass national bereits theoretische

Abhandlungen vorhanden sind, jedoch keine empirischen Erkenntnisse zur Wirksamkeit von P4P in Österreich vorliegen (Emmert 2013).

5 Best Practice bei der Implementierung von P4P-Programmen

Ziel dieses Kapitels ist es, die in den Studien als wesentlich für die erfolgreiche Konzipierung und Implementierung erachteten Faktoren anzuführen. Die folgenden Ausführungen dienen lediglich als evidenzbasierte Grundlage und sollen nicht als Konzept interpretiert werden.

Generell kann festgehalten werden, dass eine Übertragung bisher erfolgreicher Projekte auf nationale Gegebenheiten schier unmöglich ist. Eine Anpassung an die Erfordernisse des jeweiligen Gesundheitssystems sowie der Umweltbedingungen ist für eine erfolgreiche Implementierung essentiell (Scheppach et al. 2011, Veit et al. 2012). Strukturelle Differenzen, welche eine Übertragung bisheriger Projekte erschweren, liegen beispielsweise in der Primärversorgung vor (vgl. HausärztInnen als Gatekeeper im UK versus freie ÄrztInnenwahl im nationalen System).

Veit et al. (2012) empfehlen einen abgestuften Einstieg in die Programme. Durch Feedback, Edukation und Public Reporting können bereits viele Qualitätsprobleme reduziert werden. Die AutorInnen sehen deshalb P4P als Mittel der Wahl, wenn die Möglichkeiten der anderen Instrumente ausgeschöpft sind bzw. können zusätzliche monetäre Anreize die Wirksamkeit eines Programmes noch erhöhen. Ferner muss beachtet werden, dass beim Einsatz von P4P als gesundheitspolitisches Steuerungsinstrument, die Komplexität des Systems und des Instruments berücksichtigt werden müssen. Bei allen geplanten Projekten muss das Verhältnis von Machbarkeit, Sinnhaftigkeit und Angemessenheit gewahrt bleiben. Qualitätsaussagen sollen sich auf das beschränken, was die Messinstrumente leisten können (Veit et al. 2012).

In Anlehnung an Veit et al. (2012) werden im Folgenden die wesentliche Erfolgsfaktoren, untergliedert in Qualitätsmessung, Anreize und Motivation und Umsetzung, erörtert.

5.1 Qualitätsmessung

Im Hinblick auf die Qualitätsmessung zeigen die Ergebnisse der eruierten Erfahrungswerte durch Emmert (2008), dass die Messbedingungen getestet werden müssen. Demnach sollten die dokumentierten Daten valide und wissenschaftlich fundiert, die Qualitätsziele klinisch relevant und von den LeistungserbringerInnen anerkannt sein und die Datenauswertung zeitnah erfolgen, um eine Beeinflussbarkeit der Performance gewährleisten zu können. Die Datensammlung sollte unabhängig erfolgen. Ebenso wird ein regelmäßiges Feedback bezüglich der Performance-Messkriterien empfohlen (Emmert 2008).

Auch im Zusammenhang mit der Versorgungsart gilt es, darüber zu entscheiden, ob die Qualitätsmessung im chronischen, präventiven, akuten oder rehabilitativen Bereich erfolgt. Vor allem in den ersten beiden Bereichen sind bereits einige Indikatoren vorhanden. Darüber hinaus wurden in diesen zum größten Teil positive Ergebnisse in den Studien erzielt (Scheppach et al. 2011).

Ärztliche Leistung kann entweder durch Struktur-, Prozess- oder Ergebnisindikatoren erfasst werden (für eine detaillierte Darstellung der jeweiligen Dimensionen vgl. Kapitel 3.2).

Geeignete Indikatoren auszuwählen stellt eine große Herausforderung dar. Je nach Dimension sollten die jeweiligen Indikatoren kombiniert zur Leistungsmessung herangezogen werden. Ein guter Indikator sollte klinisch relevant, verständlich, leicht messbar und valide sein (Scheppach et al. 2011).

5.2 Anreize und Motivation

Bei der Zielformulierung muss die Frage geklärt werden, in welchen Situationen P4P bisherige Steuerungsinstrumente sinnvoll und angemessen ergänzen kann. Darüber hinaus müssen die Anreize und verwendeten Qualitätsindikatoren kompatibel sein. Neben der Empfängerebene muss auch die Dauer der Anreizgewährung, die mögliche Integration einer Strafzahlung, die Anreizquelle sowie die richtige Anreizhöhe festgelegt werden. In diesem Zusammenhang werden finanzielle Anreize zwischen zwei und vier Prozent der Normalvergütung im Krankenhausbereich bereits als wirksam angesehen. Im niedergelassenen Bereich jedoch werden Anreize in der Höhe von 10-15 Prozent als notwendig erachtet, um Effekte zu zeigen (Scheppach et al. 2011, Veit et al. 2012).

Einen essentiellen Stellenwert nehmen darüber hinaus die Gestaltungsmöglichkeiten der Bemessungsgrundlage ein. Es muss definiert werden, ob die Bonuszahlungen an Top-Performer, relative Delta-Performer oder relative Top-Performer gerichtet ist¹² (Scheppach et al. 2011).

Emmert (2008) verweist darauf, dass sowohl monetäre als auch nicht-monetäre Anreize Verwendung finden sollten, wobei die monetären Anreize hoch genug sein müssen, um eine Änderung herbeiführen zu können (diskutiert werden fünf bis zehn Prozent des jährlichen Einkommens der ÄrztInnen). Im Hinblick auf nicht-monetäre Anreize erwies sich das Public Reporting als effektiv. Dieses kann sogar einen stärkeren Anreiz darstellen als die Bonuszahlungen (Emmert 2008).

5.3 Umsetzung

Die Einbeziehung der VersorgerInnen bei der Projekterstellung trägt ebenso wie die Transparenz des Verfahrens, die Qualität der Indikatoren, die Angemessenheit der Datengrundlage, der akzeptabel niedrige Zusatzaufwand, Unterstützung in der Umsetzung und Attraktivität der Anreize zur Akzeptanz der Programme bei. Darüber hinaus nehmen die datenschutzkonforme und anonyme Speicherung von PatientInnendaten zur Qualitätsmessung ebenso eine wichtigen Stellenwert ein, wie eine standardisierte Dokumentation des Programms (Veit et al. 2012).

Emmert empfiehlt den Einsatz neutraler und zuverlässiger ModeratorInnen, welche die Konzeption begleiten und die wichtigsten Interessenvertretungen zusammenrufen. Darüber hinaus müssen die unterschiedlichen Perspektiven und Philosophien abgestimmt werden. Dazu gehören beispielsweise Fragestellungen, wie am schnellsten Verbesserungen der Behandlungsqualität erreicht werden könnten. Erfahrungen bei der Umsetzung haben

¹² Vgl. Kapitel 3.3

ebenfalls gezeigt, dass es förderlich ist Differenzierung zuzulassen (z.B. individuelle Bonusprogramme bei HMOs).

Als Zusammenfassung der Erörterungen dieses Kapitels und der Ausführungen in den vorangegangenen Kapiteln und als Orientierungshilfe wurde eine, von Scheppach et al. (2011) erarbeitete, Checkliste für die Implementierung eines P4P-Programms dem Dossier angehängt. Diese fasst die 18 wichtigsten Fragen zusammen, welche es in den verschiedenen Phasen der Implementierung eines neuen P4P-Programms zu klären gilt (Scheppach et al. 2011).

6 Zusammenfassung und Fazit

Mit diesem Dossier wird der Versuch unternommen einen Überblick zur wissenschaftlichen Evidenz der Wirksamkeit von P4P-Programmen zu geben. Basierend auf den Ergebnissen einer groben Literatursichtung und der Verwendung von systematischen Übersichtsarbeiten (vgl. Kapitel 2) wird versucht, neben der Effektivität auch auf die Effizienz von P4P-Programmen einzugehen.

In P4P-Programmen wird die Vergütung der ärztlichen Leistung an das erbrachte und durch Kennzahlen dargestellte Leistungsniveau gekoppelt. P4P als qualitätsorientiertes Steuerungsinstrument wurde - ausgehend vom englischsprachigen Raum – als eine Möglichkeit zur Optimierung der Qualität und Effizienz im Gesundheitswesen entwickelt und implementiert.

Bevor mit der Darstellung der Evidenz in diesem Dossier begonnen wurde, wurden wesentliche Elemente erarbeitet, welche die AkteurInnen bei der Entscheidung ob und wie P4P-Programme konzipiert und implementiert werden sollen, unterstützen können. Zunächst muss eine, den nationalen Gegebenheiten entsprechende, Vorgehensweise definiert werden (z.B. ob die Vergütung durch Boni oder Zielprämien erfolgt). Darüber hinaus gilt es zu bestimmen, in welches Setting P4P-Programme implementiert werden. Die Messung der Qualität und die Anreizgestaltung nehmen eine Schlüsselrolle in P4P-Programmen ein. Gemessen werden kann die Struktur-, Prozess- und Ergebnisqualität. Die einzelnen Dimensionen beeinflussen sich in einem gewissen Grad und unterscheiden sich hinsichtlich ihrer Messbarkeit und Eignung als P4P-Grundlage. Der langfristige Erfolg von P4P hängt, neben der Auswahl geeigneter Indikatoren zur Messung des Leistungsniveaus von der Akzeptanz der Programme durch die LeistungserbringerInnen ab. Die Wahl des Anreizmodells und der Bemessungsgrundlagen hängt deshalb davon ab, welche Ziele mit dem Programm verfolgt werden. Erfahrungen haben gezeigt, dass sich eine Kombination von monetären und nicht-monetären Anreizen positiv auf die Qualität auswirken kann.

Internationale Erfahrungen zeigen, dass P4P-Programme dazu beitragen können die Behandlungsqualität zu optimieren, die Transparenz im Gesundheitswesen zu steigern und Kosten zu reduzieren. Als herausfordernd wird erachtet, dass P4P-Programme mit Selektionseffekten behaftet sein können, welche eventuell die medizinische Praxis beeinflussen, dass Messergebnisse eventuell nicht aussagekräftig sind und dass die Akzeptanz der ÄrztInnen nicht gegeben ist.

Im Folgenden wird versucht die eingangs definierte Forschungsfrage

- *Welche Aussagen zur Evidenz im Hinblick auf die Wirksamkeit von P4P-Projekten können mithilfe von aktuellen wissenschaftlichen Daten getroffen werden?*

zu beantworten.

Zusammenfassend lassen sich folgende Aussagen zur wissenschaftlichen Evidenz der Wirksamkeit von P4P treffen:

- Die Heterogenität der bis dato durchgeführten P4P-Programme erschwert Aussagen zur Evidenz.

- Die Evidenz ist oftmals nur auf Ebene einzelner Qualitätsindikatoren nachweisbar.
- Bisherige Evaluationsergebnisse zeigen widersprüchliche Ergebnisse im Hinblick auf die Wirksamkeit. Diese konnte somit in den Studien noch nicht zweifelsfrei nachgewiesen werden.
- Die Wirksamkeit ist je nach chronischer Erkrankung unterschiedlich.
- Aufgrund der teilweise in Kombination verwendeten monetären und nicht-monetären Anreize kann der Anteil der finanziellen Anreize an der Wirksamkeit bislang noch nicht eindeutig nachgewiesen werden.
- P4P kann ein wirksames Instrument zur Optimierung der Versorgungsqualität sein, wenn Versorgungsdefizite eruiert werden konnten.
- Bei der Interpretation der Evaluationsergebnisse muss der Einfluss der Komplexität medizinischer Versorgung sowie der systematischen Optimierungs- und Umgehungsreaktionen berücksichtigt werden¹³.
- P4P stellt nur eine der Möglichkeiten der Optimierung der Qualität im Gesundheitswesen dar.

Wird angedacht bei der Implementierung von P4P-Programmen auf die Inhalte von bereits umgesetzten Programmen zurückzugreifen, so muss beachtet werden, dass eine Anpassung an die nationalen Erfordernisse und Umgebungsfaktoren unumgänglich ist. In diesem Zusammenhang müssen – basierend auf den Erfahrungswerten - Faktoren berücksichtigt werden, welche für eine erfolgreiche Implementierung vielversprechend sind. Explizit sei hier die mangelnde Transparenz der Versorgungsleistung im österreichischen Gesundheitssystem erwähnt, welche bei der Konzipierung von Qualitätsmessungsstrategien als erschwerender Faktor erachtet wird.

Zusammenfassend kann nun festgehalten werden, dass ein eindeutiger Beleg zur Wirksamkeit von P4P durch die Analyse der Studienergebnisse nicht gegeben ist. Dennoch kann P4P zu einer optimaleren Ressourcenallokation und Versorgungsqualität beitragen, wenn die Limitationen des Instruments bekannt sind und bei der Konzipierung der Programme Beachtung finden.

¹³ Vgl. Kapitel 3.4

Literaturverzeichnis

- Eijkenaar F., Emmert M., Scheppach M. et al.: Effects of pay for performance in health care, a systematic review of systematic reviews. *Health Policy*, Vol. 110 (2), 2013, p. 115-130
- Emmert M.: Pay for Performance im Gesundheitswesen, Ein Ansatz für das österreichische Gesundheitswesen. Präsentationsunterlagen, Linzer Forum, 23.10.2013
- Emmert M.: Pay for Performance (P4P) im Gesundheitswesen, Ein Ansatz zur Verbesserung der Gesundheitsversorgung. *Schriften zur Gesundheitsökonomie* 14, HERZ, Burgdorf, 2008
- Ogundeji Y.K. et al.: The effectiveness of payment for performance in health care, a meta-analysis and exploration of variation in outcomes.
- Scheppach M., Emmert M., Schöffski O.: Pay for Performance (P4P) im Gesundheitswesen, ein Leitfaden für eine erfolgreiche Einführung. *Schriften zur Gesundheitsökonomie* 19, HERZ, Burgdorf, 2011
- Schrapppe M.: P4P, aktuelle Einschätzung und Integration in umfassende Vergütungsansätze. Präsentationsunterlagen, 12. Deutscher Kongress für Versorgungsforschung, Berlin, 23.10.2013
- Veit CH., Hertle D., Bungard S. et al.: 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, Institut für Qualität & Patientensicherheit, 2012

Anhang

Tabelle 4: Checkliste für die Ausgestaltung und Implementierung eines P4P-Programms

Phase 1: Analyse	
	1. Ist die Organisation dafür bereit?
	2. Sollte eine Zusammenarbeit mit anderen Zählern erfolgen?
	3. Wann und wie sollten die LeistungserbringerInnen in die Diskussion einbezogen werden?
Phase 2: Ausgestaltung	
	4. Auf welche LeistungserbringerInnen sollte das Programm zunächst ausgerichtet werden?
	5. Was sind die Vor- und Nachteile einer individuellen als auch Gruppenbelohnung?
	6. Sollten nur Boni gezahlt werden oder sollten auch Strafzahlungen verhängt werden?
	7. Wie sollte der Bonus ausgestaltet sein?
	8. Sollten absolute Ziele oder die relative Verbesserung zu Vorjahreswerten als Basis für den Anreiz herangezogen werden?
	9. Wie sollte die stufenweise Einführung von P4P in unserem Fall erfolgen?
	10. Wie erfolgt die Finanzierung?
	11. Wie viel Geld sollte für die Belohnung der Leistung eingesetzt werden?
	12. Welche Messgrößen sind angebracht?
	13. Wie erfolgt die Dokumentation der ärztlichen Leistung?
Phase 3: Implementierung	
	14. Wie sollte die Risikoanpassung der Indikatoren erfolgen, um eine Selektion zu verhindern?
	15. Sollten die P4P-Leistungsdaten der Öffentlichkeit zur Verfügung gestellt werden?
	16. Sollte das Programm an die einzelnen Arztgruppen und Abteilungen angepasst werden?
Phase 4: Bewertung	
	17. Wie kann die Effektivität des Programms bewertet werden?
	18. Welche unerwarteten Konsequenzen gilt es zu berücksichtigen?

Quelle: Eigene Darstellung in Anlehnung an Scheppach et al. 2011