

P4P – Pay for Performance, qualitätsorientierte Vergütung – ein weiterer Aufruf zur Gelassenheit

„ist die Verknüpfung von Qualität und Vergütung der richtige Weg, „denn so wird dem Faktor Qualität noch mehr Geltung verliehen und damit eine noch bessere Behandlungsqualität ... garantiert.“⁵

Die Fragen sind stets dieselben, wenn man überlegt neue Dinge ins Gesundheitswesen einzuführen¹: Seien es neue Arzneimittel, Medizintechnik, operative Verfahren und anderes mehr, aber auch bei anderen Interventionen in das Gesundheitswesen, ob Organisationsformen, Finanzierungsformen, Versicherungsmodelle wie Selbstbehälte etc. etc.

erreichte sie 2004 Großbritannien, 2007 das Gutachten des deutschen Sachverständigenrates zur Begutachtung der Entwicklung im Gesundheitswesen⁴ und in den letzten beiden Jahren auch die österreichische Diskussion.

Rezent war Mitte September P4P Thema einer vom Gesundheitsfonds eines Bundeslandes organisierten Veranstaltung. Der Autor war zugegen und hat einerseits darauf hingewiesen, dass die Belege für P4P äußerst dürfzig sind und andererseits darauf, dass man hochgebildete Professionals besser nicht wie Ratten motivieren und behandeln sollte. In der Zusammenfassung zur Veranstaltung wurde nun gemeldet⁵:

Was weiß man über die einzuführende Technologie?

- wirkt sie in kontrollierten Situationen (z.B. Studienseitungen) – die Frage der Wirksamkeit (Efficacy)
- wirkt sie im „Alltagsleben“ – die Frage der Effektivität (Effectiveness)
- wie steht es um Neben- und Wechselwirkungen – die Frage der Sicherheit (Safety)
- wie ist das Verhältnis von Kosten und Nutzen, ist die Technologie ihr Geld wert – die Frage der Kosten- effektivität (Cost-Effectiveness).

„Qualitätsbezogene Vergütung“

Die Verknüpfung von Qualitätsanforderungen mit monetären Leistungen ist nichts Neues. International wird dieses Instrument der qualitätsbezogenen Vergütung schon länger diskutiert und zahlreiche Studien haben diesbezüglich positive Effekte gezeigt. Im Rahmen der am vergangenen Montag stattgefundenen Vernetzungskonferenz der PlattformQ hat Prof. Matthias Schrappe von der Universität Köln in seinem Vortrag die unterschiedlichen Aspekte der qualitätsbezogenen Vergütung beleuchtet. In der anschließenden Podiumsdiskussion mit VertreterInnen des österreichischen Gesundheitswesens wurden die Einsatzmöglichkeiten solcher Modelle im Gesundheitssystem diskutiert.

Für Gesundheitslandesrätin und Vorsitzende der Gesundheitsplattform Kristina Edlinger-Plöder ist die Verknüpfung von Qualität und Vergütung der richtige Weg, „denn so wird dem Faktor Qualität noch mehr Geltung verliehen und damit eine noch bessere Behandlungsqualität für die steirischen Patientinnen und Patienten garantiert.“

Auch für den Obmann der Steiermärkischen Gebietskrankenkasse, Josef Pesserl, liegt die Bedeutung dieser Qualitätsoffensive auf der Hand: „Die Steigerung der Qualität kommt letztlich immer den Patientinnen und Patienten zugute. Daher müssen alle sinnvollen Initiativen in diese Richtung begrüßt und forciert werden.“

The image shows the front cover of a Health Affairs Policy Brief. The title 'Health Policy Brief' is at the top, followed by the subtitle 'Pay-for-Performance'. The cover includes the Health Affairs logo and a small photograph of Robert Wood Johnson Foundation. The text on the cover discusses the new payment systems reward doctors and hospitals for improving the quality of care, but studies to date show mixed results. It also includes sections on 'WHAT'S THE ISSUE?' and 'WHAT'S THE BACKGROUND?' with detailed text and small images.

„Dienstleister dafür zu bezahlen bessere Outcomes für die Patienten zu erzielen sollte diese Outcomes verbessern. In der Theorie – aber tatsächlich haben die Studien über diese Programme nur gemischte Ergebnisse hervorgebracht.“⁶

„Unter Pay-for-Performance versteht man eine spezielle strategische Form der Vergütung. Pay-for-Performance-Verfahren koppeln die Vergütung von Versorgern im Gesundheitswesen an das von diesen erbrachte, durch Kennzahlen dargestellte Leistungsniveau. Durch differenzierende finanzielle Anreize sollen Versorger motiviert werden, ihre Versorgungspraxis im Hinblick auf Qualität und Effizienz des Gesundheitssystems nachhaltig und umfassend zu optimieren und weiterzuentwickeln. (BQS 2012)

Wenn man dazu kaum etwas sagen kann – und es ist kein sonst unbekommbarer Notfall, der das Risiko der Anwendung rechtfertigte – dann empfiehlt sich wohl „abwartendes Beobachten“.

Das Gesundheitsmanagement ist ein Ort ständig neuer und wechselnder Moden². Ihren Ursprung haben sie meist in den Vereinigten Staaten und es ist nicht ohne Ironie, wie die guten Ideen der angewandten Gesundheitsökonomie gerade aus jenem Land importiert werden, das ein einzigartiger Ausreißer bei der Höhe der Gesundheitsausgaben ist. Natürlich gibt es ein Übergewicht an Publikationen von dort, natürlich sind die Dinge spannend und anregend. Aber abgesehen davon, dass auch in den USA die Moden aufsteigen und wieder verblassen, ist doch interessant, wie groß die Neigung ist, aus einem ganz anders organisierten – stark geldgetriebenem³ – System Anregungen aufzunehmen.

Ich entsinne mich gut eines Gespräches mit dem werten Kollegen Herrn Dr. U. vor etlichen Jahren, da stellten wir uns die Frage, wie lange es dauern würde, bis die P4P Welle auch Europa und Österreich erreichen werde: Tatsächlich

„Für die Wirksamkeit von P4P-Projekten fehlt trotz langjähriger Praxis nach wie vor eine verlässliche Evidenzbasis. Es stellt sich die Frage, ob das Thema damit als erledigt gelten muss oder ob es dennoch Sinn macht, P4P als Steuerungsinstrument in Betracht zu ziehen.“⁷



Ol. Velt, D. Hertle, S. Bungard, A. Trömler, V. Daniske, B. Meyer-Höfmann
Pay-for-Performance im Gesundheitswesen:
Sachstandsbericht zu Evidenz und Realisierung
sowie Darlegung der Grundlagen
für eine künftige Weiterentwicklung

Ein Gutachten im Auftrag des Bundesministeriums für Gesundheit, 2012

„So betonen die Wissenschaftler, dass das komplexe Pay-for-Performance-Instrument erst dann gewählt werden soll, wenn die Möglichkeiten der vorhandenen anderen Steuerungsinstrumente ausgeschöpft sind und zweifelsfrei die Wirksamkeit von Pay-for-Performance-Projekten nachgewiesen wurde.“⁸

Tatsächlich aber mehren sich die warnenden Hinweise zur P4P Mode in der Literatur. Sie zeigen auch das aus klinischen Studien bekannte Phänomen: Je besser und größer die Studien, desto geringer bzw. nicht vorhanden sind die Effekte, die sich in den frühen euphorischen kleinen Studien zeigten.

Einige dieser warnenden Hinweise werden im Folgenden angeführt, systematische Reviews, neue Einzelstudien und auch Kommentare. Systematische Reviews sind nach einer bestimmten Vorgangsweise des Suchens, des Analysierens und des Bewertens hergestellte Zusammenfassungen der Literatur.

Deutschland 2012 BQS Gutachten

Mitte August 2012 wurde das vom deutschen Bundesministerium für Gesundheit bei BQS (BQS Institut für Qualität und Patientensicherheit) in Auftrag gegebene Gutachten „Pay-for-Performance im Gesundheitswesen: Sachstandsbericht zu Evidenz und Realisierung sowie Darlegung der Grundlagen für eine künftige Weiterentwicklung“ präsentiert⁷. Es wird hier als erstes angeführt, weil doch die Meisten gerne auf Deutsch lesen.

„Staatssekretär Thomas Ilka sagte zur Veröffentlichung des Gutachtens: „Das Gutachten bietet einen guten Überblick über die nationalen und internationalen Entwicklungen im Bereich Pay-for-Performance und stellt eine wichtige Diskussionsgrundlage für Weiterentwicklungen in diesem Bereich dar. Es macht aber auch deutlich, dass vor einer breiteren Einführung solcher Verfahren noch Forschungsbedarf besteht.“

„So betonen die Wissenschaftler, dass das komplexe Pay-for-Performance-Instrument erst dann gewählt werden soll, wenn die Möglichkeiten der vorhandenen anderen Steuerungsinstrumente ausgeschöpft sind und zweifelsfrei die Wirksamkeit von Pay-for-Performance-Projekten nachgewiesen wurde.“⁸

Das Gutachten enthält ein systematisches Review der Literatur bis 2011. Um an die Eingangsfragen anzuschließen, das Review findet kaum Hinweise, dass P4P wirkt, dass es effektiv ist und dass es kosteneffektiv wäre.

Die Zusammenfassung zu Beginn des Gutachtens ist noch relativ versöhnlich:

„Eine kritische Sichtung der Literatur mit den Schwerpunkten hausärztlicher Bereich und Krankenhausbereich sowie Kosteneffektivität von P4P kommt - auch unter Berücksichtigung der neueren Literatur von 2010 und 2011 - zum gleichen Ergebnis wie andere Reviews zuvor: Bislang ist es in Studien noch nicht gelungen, zweifelsfrei die Wirksamkeit von P4P-Projekten nachzuweisen. Zwar gibt es Belege für die Effektivität kombinierter Verfahren unter Einschluss nicht-finanzieller Anreize. Der Anteil der finanziellen Anreize am Erfolg ließ sich bislang jedoch nicht eindeutig nachweisen. Dies widerspricht der allgemeinen Erfahrung, dass Geld durchaus ein sehr starker Anreiz ist und Änderungen der Vergütungssysteme stets intensive Reaktionen nach sich ziehen.“

Teil B des Gutachtens „Teil B Wissenschaftliche Evidenz“ beschreibt auf etwas über 40 Seiten das systematische Review (Evidenz von P4P im hausärztlichen Bereich, Evidenz von P4P im Krankenhausbereich, P4P in der stationären Rehabilitation, Langzeiterfahrungen mit P4P). Es darf grundsätzlich darauf verwiesen werden, nur einige Auszüge seien angeführt:

Aus den Schlussfolgerungen:

„Für die Wirksamkeit von P4P-Projekten fehlt trotz langjähriger Praxis nach wie vor eine verlässliche Evidenzbasis. Es stellt sich die Frage, ob das Thema damit als erledigt gelten muss oder ob es dennoch Sinn macht, P4P als Steuerungsinstrument in Betracht zu ziehen.“

Der Mangel an Evidenz überrascht umso mehr, als die alltägliche Erfahrung dafür spricht, dass finanzielle Anreize sehr mächtige Steuerungswirkungen entfalten können und dass diese im Rahmen der bisherigen Vergütungssysteme regelmäßig nachgewiesen werden konnten (z. B. Byrnes und Fifer 2010; Gosden et al. 2000). Dies geht so weit, dass bereits in der Öffentlichkeit vor einer Übermacht finanzieller Anreize gewarnt wird (Baumann 2012). Daher besteht auch bei den interviewten Experten mit großer Mehrheit die Ansicht, dass künftig P4P-Vergütungsformen eine zunehmende Rolle spielen werden. Verschiedene Autoren entwerfen konkret die Möglichkeiten künftiger Projekte (z. B. SVR 2007; Malzahn et al. 2011), und diejenigen, die im Rahmen dieses Gutachtens über ihre Projekte berichteten, waren keineswegs durch den Mangel an Evidenz irritiert. Auch in den USA und anderen Ländern werden sehr intensiv neue P4P-Projekte auf den Weg gebracht, gleichzeitig begleitet von klarer Erfolgsgewissheit der Protagonisten und ebenso deutlichem Zweifel der Gegner.

Der Disput wird nur beigelegt werden können, indem durch eine verbesserte Gestaltung der P4P-Projekte sowie durch eine bessere Evaluation derselben die ausstehende wissenschaftliche Evidenz geliefert wird und entweder die Wirksamkeit nachgewiesen werden kann oder gezeigt wird, dass auch optimale Projekte in der Routine letztlich unwirksam bleiben.“

„diejenigen, die im Rahmen dieses Gutachtens über ihre Projekte berichteten, waren keineswegs durch den Mangel an Evidenz irritiert.“ – das klingt doch vertraut, vom klassischen Aderlass bis zur Alternativmedizin.

Das systematische Review ist nur ein Teil des umfangreichen Gutachtens. Die restlichen rund 300 Seiten umfassen die theoretischen Aspekte und in deutscher Gründlichkeit auch die Vorgangsweisen, wie man P4P einführt incl. einer ausführlichen Analyse der rechtlichen Grundlagen und Notwendigkeiten.

„Die Evidenzlage ist nicht ausreichend um den Einsatz von finanziellen Anreizen zur Verbesserung der Qualität der allgemeinmedizinischen Versorgung zu unterstützen oder nicht zu unterstützen. Umsetzungen sollten vorsichtig erfolgen und die Anreize sollten sorgfältig gestaltet werden und evaluiert.“⁹

P4P in der hausärztlichen Betreuung

2011 wurde ein Review der Cochrane Collaboration veröffentlicht „Der Effekt von finanziellen Anreizen auf die Qualität der Versorgung durch Hausärzte“⁹.

Von den vielen Studien, die sie in der Literatursuche fanden, konnten die Autoren schließlich nur sieben für die Auswertung verwenden und auch diese hatten methodische Probleme.

Für die allgemeine Leserschaft fassen sie zusammen:

„The use of different ways of paying primary care physicians in an attempt to increase quality of care, including the use of financial incentives to directly reward 'performance' and 'quality', is increasing in a number of countries. There are many examples in the US, and the Quality and Outcomes Framework (QoF) for general practitioners (GPs) in the UK is an example of a major system-wide reform. Despite the popularity of these schemes, there is currently little rigorous evidence of their success in improving the quality of primary health care, or of whether such an approach is cost-effective relative to other ways to improve the quality of care. The aim of this review is to examine the effect of changes in the method and level of payment on the quality of care provided by primary care physicians (PCPs). There is insufficient evidence to support or not support the use of financial incentives to improve the quality of primary health care. Implementation should proceed with caution and incentive schemes should be carefully designed and evaluated.“

Um an die Eingangsfragen anzuschließen, das Review findet geringe Antworten, dass P4P wirkt, dass es effektiv ist und dass es kosteneffektiv wäre.

Im Septemberheft der *Annals of Family Medicine* wurde erneut ein systematisches Review zum P4P für die englischen Hausärzte publiziert¹⁰. Die Autoren fanden fast 600 Artikel bei ihrer Literatursuche und haben schließlich 49 als relevant für das Review identifiziert.

CONCLUSIONS: Observed improvements in quality of care for chronic diseases in the framework were modest, and the impact on costs, professional behavior, and patient experience remains uncertain. Further research is needed into how to improve quality across different domains, while minimizing costs and any unintended adverse effects of payment for performance schemes. Health care organizations should remain cautious about the benefits of similar schemes¹².

„In view of our findings that observed improvements in care from a very large payment-for-performance scheme in the United Kingdom were modest, with uncertainty about possible adverse effects, we recommend that policy makers continue to exercise caution about implementing similar schemes. Consideration should be given to improving different dimensions of quality, including user experience and equity. Costs should be monitored and balanced against benefits. Wherever possible, schemes should be designed in collaboration with health service researchers to evaluate the benefits of minor differences in system design. Payment for performance is still an imperfect approach to improving primary care, and should be considered as only one option alongside alternative quality improvement methods.“

Methodisch ist es nicht einfach, eine komplexe Intervention wie P4P zu evaluieren. Das BQS Gutachten beschreibt hier ausführlich die Ansätze und Probleme. Ob die konkrete Intervention Ursache einer Veränderung ist oder andere Ursachen mitspielen oder eigentlich verantwortlich sind, ist oft schwer zu trennen: Das Umfeld kann sich geändert haben, allgemeine Trends können bestehen, die ohnehin wirksam geworden wären.

Serumaga et. al haben die Zeittrends in einer Studie zum englischen Ansatz berücksichtigt¹¹. Sie untersuchten an Hand von über 470.000 Patienten, bei denen zwischen 2000 und 2007 Bluthochdruck diagnostiziert wurde, ob das 2004 eingeführte - wahrlich nicht billige - P4P zu zusätzlichen positiven Veränderungen der Ergebnisse führte. Sie meinen, das wäre nicht der Fall gewesen:

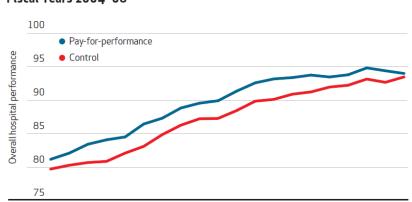
"After accounting for secular trends, no changes in blood pressure monitoring (...), control (...), or treatment intensity (...) were attributable to pay for performance. Pay for performance had no effect on the cumulative incidence of stroke, myocardial infarction, renal failure, heart failure, or all cause mortality in both treatment experienced and newly treated subgroups.

Conclusions: Good quality of care for hypertension was stable or improving before pay for performance was introduced. Pay for performance had no discernible effects on processes of care or on hypertension related clinical outcomes. Generous financial incentives, as designed in the UK pay for performance policy, may not be sufficient to improve quality of care and outcomes for hypertension and other common chronic conditions."

Die Frage des Zeitverlaufs von P4P Programmen und des allgemeinen Trends ist früher zu wenig berücksichtigt worden. Je größer die Studie und je länger der Beobachtungszeitraum desto geringer sind die Auswirkungen der Programme. Die Abbildung unten ist dem auf Seite 1 zitierten Policy Brief von Health Affairs entnommen und betrifft Krakenhäuser.

RESUME

Average Overall Performance in Pay-for-Performance and Control Hospitals,



SOURCE Rachel M. Werner, Jonathan T. Kolstad, Elizabeth A. Stuart, and Daniel Polksy, "The Effect of Pay-for-Performance in Hospitals: Lessons for Quality Improvement," *Health Affairs* 30, no. 4 (2011): 690-8. Data from Hospital Compare: survey of patients' hospital experiences [Internet]. Department of Health and Human Services. **NOTES** Performance is averaged across the three conditions: acute

„Die beobachteten Verbesserungen in der Qualität in Betreuung chronischer Krankheiten im Programm waren bescheiden und die Auswirkungen auf die Kosten, das professionelle Verhalten und die Patientenerfahrungen bleiben ungewiss.“¹⁰

„Wir fanden keine Belege, dass das größte Pay for Performance Programm für Krankenhäuser zu einer Verringerung der 30 Tage Mortalität führte. Erwartungen an verbesserte Outcomes bei Programmen wie dem Premier HQID sollten daher bescheiden sein.“¹²

THE NEW ENGLAND JOURNAL OF MEDICINE

SPECIAL ARTICLE

The Long-Term Effect of Premier Pay for Performance on Patient Outcomes

Ashish K. Jha, M.D., M.P.H., Karen E. Joynt, M.D., M.P.H., E. John Orav, Ph.D., and Arnold M. Epstein, M.D.

ABSTRACT

BACKGROUND
Pay for performance has become a central strategy in the drive to improve health care. We assessed the long-term effect of the Medicare Premier Hospital Quality Incentive Demonstration (HQID) on patient outcomes.

METHODS
We used Medicare data to compare outcomes between the 252 hospitals participating in the Premier HQID and 3363 control hospitals participating in public reporting alone. We examined 30-day mortality among more than 6 million patients who had acute myocardial infarction, congestive heart failure, or pneumonia or who underwent coronary-artery bypass grafting (CABG) between 2003 and 2009.

THE NEW ENGLAND JOURNAL OF MEDICINE

SPECIAL ARTICLE

Effect of Nonpayment for Preventable Infections in U.S. Hospitals

Grace M. Lee, M.D., M.P.H., Ken Kleinman, Sc.D., Stephen B. Soumerai, Sc.D., Alison Tse, Sc.D., David Cole, B.M., Scott K. Fruland, M.D., Teresa Horan, M.P.H., Richard Platt, M.D., Charlene Gay, B.A., William Kassler, M.D., M.P.H., Donald A. Goldmann, M.D., John Jernigan, M.D., and Ashish K. Jha, M.D., M.P.H.

ABSTRACT

BACKGROUND
In October 2008, the Centers for Medicare and Medicaid Services (CMS) discontinued additional payments for certain hospital-acquired conditions that were deemed preventable. The effect of this policy on rates of health care-associated infections is unknown.

„Obwohl es substantielle finanzielle Anreize bot, hat das MassHealth P4P Programm die Qualität während der ersten Jahre der Umsetzung nicht verbessert.“¹³

„Wir fanden keine Evidenz, dass das Programm der CMS von 2008, die Bezahlung für Infektionen bei zentralvenösen Kathetern und katheterassoziierten Harnwegsinfekten zu reduzieren, irgendeine messbare Auswirkung auf die Infektionsraten der US-Spitäler hatte.“¹⁴

Rezentes zu P4P in Spitälern

Das Premier Hospital Quality Incentive Demonstration (PHQID) wird auch im BQS Gutachten beschrieben. 252 Krankenhäuser nahmen an diesem freiwilligen Programm von CMS teil. Die Qualitätsmessungen bezogen sich auf fünf „Tracerdiagnosen“: akuter Myokardinfarkt, Herzinsuffizienz, ambulant erworbene Pneumonie, koronare Bypassoperation sowie Hüft- und Kniegelenkersatz. Erfasst wurden ganz überwiegend Prozessindikatoren sowie einige wenige Ergebnisindikatoren (Mortalität, Wiederaufnahme, schwere Komplikationen). Das PHQID, so BQS, sei ist insofern ein besonders bemerkenswertes Projekt, weil es nach Mehrotra et al. (2009) das einzige evaluierte Projekt im Krankenhausbereich wäre (bis 2009), dessen Daten mit einer Kontrollgruppe verglichen werden konnten.

Im April wurde im New England Journal of Medicine der langfristige Effekt dieses Programmes untersucht¹². Auch hier sind die Ergebnisse nicht überzeugend.

„Results

At baseline, the composite 30-day mortality was similar for Premier and non-Premier hospitals (...). The rates of decline in mortality per quarter at the two types of hospitals were also similar (...), and mortality remained similar after 6 years under the pay-for-performance system (...). We found that the effects of pay for performance on mortality did not differ significantly among conditions for which outcomes were explicitly linked to incentives (acute myocardial infarction and CABG) and among conditions not linked to incentives (congestive heart failure and pneumonia) ($P = 0.36$ for interaction). Among hospitals that were poor performers at baseline, mortality was similar in the two groups of hospitals at the start of the study (...), with similar rates of improvement per quarter (...) and similar mortality rates at the end of the study (...).

Conclusions:

We found no evidence that the largest hospital-based pay-for-performance program led to a decrease in 30-day mortality. Expectations of improved outcomes for programs modeled after Premier HQID should therefore remain modest.

Im Juni Heft von Health Services Research publizierten Ryan und Blustein Ergebnisse aus dem P4P Programm des Staates Massachusetts¹³. Wieder sind Krankenhäuser die Zielgruppe, das Programm bezieht sich auf die Medicaid Versicherten. Verwendet werden Daten von 2004 bis 2009, das P4P wurde 2008 eingeführt. Untersucht wurden die Auswirkungen auf Qualität der Betreuung von Lungenentzündungen und der post-operativen Infektionen.

“Principal Findings. Estimates from our preferred specification, including hospital fixed effects, trends, and the control for measure completeness, indicate small and nonsignificant program effects for pneumonia (- 0.67 percentage points, $p > .10$) and SIP (- 0.12 percentage points, $p > .10$). Sensitivity checks indicate a similar pattern of findings across specifications.

Conclusions. Despite offering substantial financial incentives, the MassHealth P4P program did not improve quality in the first years of implementation.”

Gesundheitsökonomisch spannend ist der Ansatz, dann nicht zu zahlen, wenn die Qualität scheinbar unzureichend ist, statt Prämien für Qualität zu zahlen. Es gehört zu den Paradoxien der Verhaltensökonomie, dass 1000 Dollar weniger mehr schmerzen als 1000 Dollar mehr freuen.

Das „Centers for Medicare & Medicaid Services“ (CMS), das die Versicherung für vor allem ältere Bürger der USA (Medicare) verwaltet, hat 2008 auf Grund einer gesetzlichen Vorgabe eine alternative Strategie eingeführt, die in finanziellen Strafen bestand: Spitäler sollten kein entsprechendes zusätzliches Entgelt erhalten, wenn Diagnosen anfallen, die im Zuge des Aufenthaltes erworben wurden. Das gilt für 10 vermeidbare Komplikationen, z.B. auch für Infektionen bei zentralvenösen Kathetern und katheterassoziierten Harnwegsinfekten.

Im Oktober berichtete eine Autorengruppe der Universität Harvard im New England Journal of Medicine, was sie denn zu dieser Anreizmethode herausgefunden haben¹⁴: Keine sichtbaren Ergebnisse.

„Results

A total of 398 hospitals or health systems contributed 14,817 to 28,339 hospital unit-months, depending on the type of infection. We observed decreasing secular trends for both targeted and nontargeted infections long before the policy was implemented.

There were no significant changes in quarterly rates of central catheter-associated bloodstream infections (incidence-rate ratio in the postimplementation vs. preimplementation period, 1.00; $P = 0.97$), catheter-associated urinary tract infections (incidence-rate ratio, 1.03; $P = 0.08$), or ventilator-associated pneumonia (incidence-rate ratio, 0.99; $P = 0.52$) after the policy implementation. Our findings did not differ for hospitals in states without mandatory reporting, nor did it differ according to the quartile of percentage of Medicare admissions or hospital size, type of ownership, or teaching status.

Conclusions

We found no evidence that the 2008 CMS policy to reduce payments for central catheter-associated bloodstream infections and catheter-associated urinary tract infections had any measurable effect on infection rates in U.S. Hospitals.

Wieder sei auf das BQS Gutachten verwiesen, in dem die früheren Studien zu P4P für Krankenhäuser gut zusammengefasst sind.

Wie lauteten die Eingangsfragen?

Was weiß man über die einzuführende Technologie?:

- wirkt sie in kontrollierten Situationen (z.B. Studiensituationen) – die Frage der Wirksamkeit (Efficacy)
- wirkt sie im „Alltagsleben“ – die Frage der Effektivität (Effectiveness)
- wie steht es um Neben- und Wechselwirkungen – die Frage der Sicherheit (Safety)
- wie ist das Verhältnis von Kosten und Nutzen, ist die Technologie ihr Geld wert – die Frage der Kosteneffektivität (Cost-Effectiveness).

„Mehr Forschung, die sich mit der Effizienz von P4P beschäftigt ist notwendig“¹⁵

Eur J Health Econ (2012) 13:755–767
DOI 10.1007/s10198-011-0298-4

ORIGINAL PAPER

Economic evaluation of pay-for-performance in health care: a systematic review

Martin Emmert · Frank Eijkenaar · Heiko Kunter ·
Adeleida Susanne Esslinger · Oliver Schöffski

Received: 17 August 2010/Accepted: 25 May 2011/Published online: 10 June 2011
© Springer-Verlag 2011

„Wir folgern, obwohl es einen offensichtlichen Bedarf nach größerer Qualität und Effizienz im System gibt, dass es Gründe gibt, P4P Initiativen zu diesem Zeitpunkt nicht zu empfehlen. Zumaldest solange bis die vielen Komplexitäten der Honorierung der Dienstleister, der Zusammenarbeit öffentlich/privat und die Fragen des Patientenzugangs zum System geklärt sind“¹⁶

The Potential Role of Pay-for-Performance in Irish Health Care

Anne Nolan
Jacqueline O'Reilly
Samantha Smith
Aoife Brick

RENEWAL SERIES
PAPER 4

December 2011

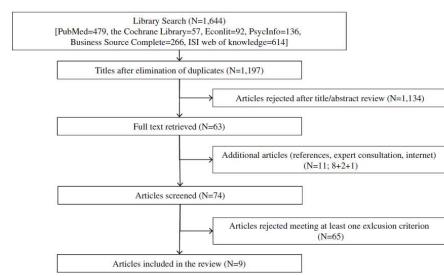
„Insgesamt liefern die Ergebnisse eine warnende Botschaft über die Effektivität von P4P Programmen zur Steigerung der Qualität der Versorgung.“¹⁸

Kosteneffektivität von P4P

Es scheint, man weiß noch nicht genug zu P4P, eher schlägt das Pendel in sehr geringe Effekte und völlig ungewisse Kosteneffektivität. Wenn man die Nutzen kaum messen kann, sind Fragen der Effizienz schwierig.

Eine Berliner Autorengruppe um Schöffski hat den Versuch unternommen letztes Jahr im European Journal of Health Economics und folgert¹⁵:

„Conclusion: The results show that evidence on the efficiency of P4P is scarce and inconclusive. P4P efficiency could not be demonstrated. The small number and variability of included studies limit the strength of our conclusions. More research addressing P4P efficiency is needed.“



Rezente Berichte und Arbeitspapiere zu P4P

Ehe auf die Frage der möglichen Neben- und Wechselwirkungen von P4P eingegangen wird, sollen noch Hinweise aus der sogenannten grauen Literatur geboten werden.

In Irland hat das Economic Research Institute (ESRI) letztes Jahr einen Bericht zur möglichen Rolle von P4P im irischen Gesundheitswesen vorgelegt¹⁶. Auch sie beschreiben die verschiedenen P4P Projekte und Studien. Die Schlussfolgerungen klingen vertraut:

„P4P is becoming increasingly common in international health-care systems despite the lack of evidence on its effectiveness and a lack of consensus on how to design and implement such programmes. It is therefore important to understand the implications of existing P4P programmes before recommending their introduction in the Irish context.

We provide an overview of the literature relating to large scale P4P schemes, focusing on programmes that have been instituted by national public sector organisations. The literature highlights the poor quality of evidence on P4P schemes, stemming partly from technical challenges inherent in evaluating P4P schemes. Notwithstanding these technical difficulties, the available evidence does not provide a clear answer to the question of whether P4P should be implemented. Limitations include difficulties in obtaining valid performance indicators, unintended consequences, and the absence of evidence on cost effectiveness of P4P schemes. We assess how a P4P scheme would interact with the payment structures already in place in the Irish health care system. We conclude that while there is an obvious need for greater efficiency and quality in the system, there are reasons why P4P initiatives are not recommended at this stage at least until the many complexities in provider reimbursement, public/private interaction, and patient access to the system are resolved.“

Nicht viel anders liest sich ein australischer Report zu Anreizen für Spitalsärzte¹⁷.

Autoren von McMaster veröffentlichten letztes Jahr in der Working Paper Serie des National Bureau of Economic Research eine interessante Arbeit¹⁸ über das „natürliche Experiment“ in Ontario, mit dem Ärzten gewisse P4P Anreize geboten werden. Die Frage lautete unter anderem, ob es nicht Unterschiede in der Reaktion der Ärzte auf die Anreize gibt, je nachdem welche Tätigkeiten mit Anreizen gefördert werden. Das scheint der Fall zu sein. Die grundsätzliche Botschaft der Autoren ist skeptisch:

„Insgesamt liefern die Ergebnisse eine warnende Botschaft über die Effektivität von P4P Programmen zur Steigerung der Qualität der Versorgung.“

„Explicit financial incentives, especially pay-for-performance (P4P) incentives, have been extensively employed in recent years by health plans and governments in an attempt to improve the quality of health care services. This study exploits a natural experiment in the province of Ontario, Canada to identify empirically the impact of pay-for-performance (P4P) incentives on the provision of targeted primary care services, and whether physicians' responses differ by age, practice size and baseline compliance level. We use an administrative data source which covers the full population of the province of Ontario and nearly all the services provided by practicing primary care physicians in Ontario. With an individual-level data set of physicians, we employ a difference-in-differences approach that controls for both "selection on observables" and "selection on unobservables" that may cause estimation bias in the identification. We also implemented a set of robustness checks to control for confounding from the other contemporary interventions of the primary care reform in Ontario. The results indicate that, while all responses are of modest size, physicians responded to some of the financial incentives but not the others. The differential responses appear related to the cost of responding and the strength of the evidence linking a service with quality.

Overall, the results provide a cautionary message regarding the effectiveness of pay-for-performance schemes for increasing quality of care.“

„Wann haben finanzielle Anreize mehr Nutzen als Schaden: Eine Checkliste“¹⁹



Paul Glasziou

BMJ

BMJ 2012;344:e427 doi:10.1136/bmj.e427 (Published 14 August 2012)

Page 1 of 5
ANALYSIS

When financial incentives do more good than harm: a checklist
Financial incentives can sometimes improve the quality of clinical practice, but they may also be an expensive distraction. Paul Glasziou and colleagues have devised a checklist to help prevent their premature or inappropriate implementation

„Trotz des Mangels an robuster Evidenz, dass dieses System im Gesundheitswesen klinisch effektiv ist, fahren Finanziers damit fort, überall eine Intervention umzusetzen von der es keine Belege gibt, dass sie irgendwo funktioniert hat.“²⁰

Wirkungen, Neben und Wechselwirkungen

Im August haben Paul Glasziou (ehemaliger Direktor des „Centre for Evidence-Based Medicine“ in Oxford von 2003-2010 und nun wieder in Australien) und Kollegen in der englischen Ärztezeitung, dem BMJ, einen Aufsatz veröffentlicht mit dem Titel „Wann haben finanzielle Anreize mehr Nutzen als Schaden: Eine Checkliste“¹⁹. Im Wissen um all die Probleme und dem gleichzeitigen Drängen nach P4P schlagen die Autoren neun Fragen in einer Checkliste vor, gegliedert in Sinnhaftigkeit und Umsetzung.

Teil A: Ist ein finanzieller Anreiz hier angemessen

1. Verbessert die gewünschte klinische Handlung den Patienten Outcome?
2. Würde ein unerwünschtes klinisches Handeln ohne die Intervention weiter fortgeführt werden?
3. Stehen valide, verlässliche und praktikable Maße für das erwünschte klinische Verhalten zur Verfügung?
4. Wurden die Hürden und die „Befähiger“ zur Verbesserung des klinischen Verhaltens untersucht?
5. Werden die finanziellen Anreize funktionieren, und das besser als andere Interventionen zur Verhaltensänderung und warum?
6. Wird der Nutzen eindeutig eventuelle nicht beabsichtigte schädliche Auswirkungen überwiegen, und das zu akzeptablen Kosten?

Teil B: Einführung

7. Stehen Systeme und Strukturen zur Verfügung, die man für eine Veränderung braucht?
8. Wie viel soll gezahlt werden, an wen und für wie lange?
9. An wen werden die finanziellen Anreize ausbezahlt?

Jede dieser Fragen wird im Beitrag kommentiert und mit Hintergrundliteratur belegt. Im Anhang „Application of financial incentives checklist to some real examples“ gibt es einige praktische Beispiele. Ein Beitrag zum Nachdenken.

Prominente Kommentatoren waren vom BMJ eingeladen, den Beitrag von Glasziou et al. zu ergänzen: Das Ehepaar Steffie Woolhandler und David Himmelstein, beide Allgemeinmediziner, Forscher und unermüdliche Kämpfer für ein besseres Gesundheitswesen in den Staaten. Sie kommentieren gemeinsam mit Dan Ariely und das mit Verve „Warum Pay for Performance vielleicht unvereinbar mit Qualitätsverbesserung ist“²⁰. Auf weniger als zwei Seiten führen sie gut begründet die ganzen Fragezeichen zu P4P an:

- den Verlust intrinsischer Motivation durch externe Anreize
- das Problem des Case-Mix und die Möglichkeit, das System zu „spielen“
- das Problem des Messens und des Zeithorizonts zwischen Leistung und Outcome
- das Problem der Risikoadjustierung und der Möglichkeit, Diagnosen zahlenmäßig „anzuhaben“ und damit besser dazustehen
- die Möglichkeit „intensiveren“ Kodierens (als Beispiel: statt „Pneumonie und Herzinsuffizienz“- „Aspirationspneumonie mit akuter chronischer Herzinsuffizienz“, das brächte in den USA nicht nur eine dreifach höhere Honorierung, sondern auch einen höheren Risikoscore des Patienten)
- das Problem der Prozessindikatoren und deren oft schlechten Vorhersagekraft für den Outcome.
- das weitgehende Fehlen der Einbeziehung sozialer Charakteristika der betreuten Patienten, die aber höchst relevant sein könnten.

Die Autoren meinen schließlich: „Zukünftige Studien mögen einmal die klinische Effektivität von Pay for Performance belegen, aber angesichts der fragwürdigen Annahmen wird das vielleicht nie der Fall sein. Trotz des Mangels an robuster Evidenz, dass dieses System im Gesundheitswesen klinisch effektiv ist, fahren Finanziers damit fort, überall eine Intervention umzusetzen von der es keine Belege gibt, dass sie irgendwo funktioniert hat.“

Schlussfolgerungen

Zurückkommend auf die eingangs gestellten Fragen:

Die Belege für die Effektivität und die Kosteneffektivität von Pay for Performance (P4P) Programmen sind nicht gegeben. Eher gibt es Hinweise auf eine Unwirksamkeit.

Wären es ein neues Arzneimittel, eine Impfung, eine neue Operation oder ein neues medizintechnisches Gerät mit einer ähnlich schlechten eher negativen Studienlage, man würde von EBM'lern, HTA'lern, Public Healther'n oder Gesundheitsökonomen gleichsam an die Wand genagelt werden, wollte man die Einführung fordern.

Natürlich sind komplexe Interventionen in das Gesundheitswesen schwieriger zu evaluieren. Jedoch, angesichts der gegebenen Studienlage P4P als „richtigen Weg“ zu bezeichnen, ist schon interessant.

Mein lieber pensionierter Bürokollege und „Lektor“ Klaus Zulauf ist im Jänner überraschend verstorben. Er fehlt sehr. Den anderen Querlesern danke ich wie immer herzlich. Für Hinweise auf Fehler bin ich dankbar.

Aber Gesundheitsmanagement ist anders – überall.

P4P Projekte werden weiter entwickelt werden und werden weiter daherkommen. Allein was alles aus dem Affordable Care Act der USA an Ideen entsprungen wird, es wird uns einholen.

Die Gläubigen in den Verwaltungen und den Finanzierungsinstitutionen werden sich nicht von Evidenzen aufhalten lassen.

Und die Dienstleister werden sich schon adaptieren und sich durchwursteln. So wie durch den täglichen Stau auf der Osttangente. Schade um die Energie wird es halt sein.

Daher: „P4P – Pay for Performance, qualitätsorientierte Vergütung – ein weiterer Aufruf zur Gelassenheit“.

Markus Narath (office@narath.at)

- 1 Die Checkliste - Entscheidungen im Gesundheitswesen http://www.oeaw.ac.at/ita/hta/checkliste_170303.pdf und Narath M. Entscheidungen im Gesundheitswesen: Checkliste. Österreichische Krankenhauszeitung 2004 (1).
- 2 Marmor Theodor. Fads, Fallacies and Foolishness in Medical Care Management and Policy, World Scientific Publishing, 2007
- 3 Maggie Mahar. Money-Driven Medicine: The Real Reason Health Care Costs So Much. HarperCollins. 2006
- 4 Gutachten 2007 Kooperation und Verantwortung - Voraussetzungen einer zielorientierten Gesundheitsversorgung Kapitel 5.3.3.: Qualitätsbezogene Vergütung (pay for performance, P4P): finanzielle Anreize für Qualitätsverbesserung. <http://dipbt.bundestag.de/dip21/btd/16/063/1606339.pdf>
- 5 <http://www.gesundheitsportal-steiermark.at/Gesundheitsfonds/Aktuell/Seiten/SALUS-2012.aspx>
- 6 Health Policy Brief: Pay-for- Performance," Health Affairs, October 11, 2012.
- 7 Veit Ch., Hertle D., Bungard S., Trümmer A., Ganske V., Meyer-Hofmann B. Pay-for-Performance im Gesundheitswesen: Sachstandsbericht zu Evidenz und Realisierung.sowie Darlegung der Grundlagen für eine künftige Weiterentwicklung. Ein Gutachten im Auftrag des Bundesministeriums für Gesundheit, 2012 http://www.bundesgesundheitsministerium.de/fileadmin/dateien/Pressemitteilungen/2012/2012_03/120817_PM_58_Anlage_Gutachten_BQS_0_1.pdf
- 8 http://www.bmg.bund.de/fileadmin/dateien/Pressemitteilungen/2012/2012_03/120817_PM_58_Pay-for-performance-Gutachten.pdf
- 9 Scott A, Sivey P, Ait Ouakrim D, Willenberg L, Naccarella L, Furler J, Young D. The effect of financial incentives on the quality of health care provided by primary care physicians. *Cochrane Database Syst Rev*. 2011 Sep 7;(9):CD008451. PMID: 21901722
- 10 Gillam SJ, Siriwardena AN, Steel N. Pay-for-Performance in the United Kingdom: Impact of the Quality and Outcomes Framework-A Systematic Review. *Ann Fam Med*. 2012;10(5):461-8 PMID: 22966110
- 11 Serumaga B, Ross-Degnan D, Avery AJ, Elliott RA, Majumdar SR, Zhang F, Soumerai SB. Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study. *BMJ*. 2011 Jan 25;342:d108. PMID: 21266440
- 12 Jha AK, Joynt KE, Orav EJ, Epstein AM. The long-term effect of premier pay for performance on patient outcomes. *N Engl J Med*. 2012;366:1606-15 PMID: 22455751
- 13 Ryan AM, Blustein J. The effect of the MassHealth hospital pay-for-performance program on quality. *Health Serv Res*. 2011;46(3):712-28 PMID: 21210796
- 14 Lee GM, Kleinman K, Soumerai SB, Tse A, Cole D, Fridkin SK, Horan T, Platt R, Gay C, Kassler W, Goldmann DA, Jernigan J, Jha AK. Effect of nonpayment for preventable infections in U.S. hospitals *N Engl J Med*. 2012;367:1428-37 PMID: 23050526
- 15 Emmert M, Eijkenaar F, Kemter H, Esslinger AS, Schöffski O. Economic evaluation of pay-for-performance in health care: a systematic review. *Eur J Health Econ*. 2012 Dec;13(6):755-67. PMID: 21660562
- 16 Nolan Anne, O'Reilly Jacqueline, Smith Samantha, Brick Aoife. The Potential Role of Pay for Performance in Irish Health Care. The Economic and Social Research Institute. Dublin December 2011
- 17 Scott, A. Ait Ouakrim, D. Using Financial Incentives to Improve the Performance of Hospital Clinicians: An Evidence Check rapid review brokered by the Sax Institute 2011
- 18 Li Jinhui, Jeremiah Hurley, Philip DeCicca, and Gioia Buckley. Physician Response to Pay-for-Performance: Evidence from a Natural Experiment. NBER Working Paper No. 16909 March 2011
- 19 Glasziou PP, Buchan H, Del Mar C, Doust J, Harris M, Knight R, Scott A, Scott IA, Stockwell A. When financial incentives do more good than harm: a checklist. *BMJ*. 2012;345:e5047. PMID: 22893568
- 20 Woolhandler S, Ariely D, Himmelstein DU. Why pay for performance may be incompatible with quality improvement. *BMJ*. 2012 Aug 13;345:e5015. PMID: 22893567 auch unter: <http://www.pnhp.org/sites/default/files/docs/nc/bmj-incentives-edit.pdf>



Steffie Woolhandler und David Himmelstein