

Franz-Josef Fischer/Dr. Jaan Sidorov/Benjamin Lüßem

# Pay for Performance

## Erfahrungen mit der Bewertung medizinischer Leistungsqualität aus den USA

Mit der Einführung des Krankenhausstrukturgesetzes wurde eine neuerliche Diskussion hinsichtlich einer adäquaten und justiziablen Bewertung der medizinischen Leistungsqualität entfacht und gesetzlich verankert. Der folgende Beitrag stellt Erfahrungen mit leistungsorientierten Vergütungsmodellen am Beispiel des Pay-for-Performance-Ansatzes aus den USA vor, diskutiert ihre Übertragbarkeit auf die Kliniklandschaft in Deutschland und deren Limitierung.

Das deutsche Gesundheitssystem wandelt sich stetig. Die demographische Entwicklung, die Zunahme älterer und multimorbider Patienten und sich verändernde gesetzliche Rahmenbedingungen sind die großen Herausforderungen im Spannungsfeld zwischen Wirtschaftlichkeit, zunehmend hohen Qualitätsanforderungen, etwa in Bezug auf Transparenz, Performance-Qualität und Patientensicherheit.

Neben der Frage der Finanzierbarkeit der Gesundheitsversorgung, hierbei vor allem der stationären Versorgung als größtem Ausgabenblock, wird durch den Gesetzgeber mehr und mehr die Thematik der Patienten-Versorgungsqualität aufgegriffen und in den Fokus gestellt. So wird aktuell mit der Einführung des Krankenhausstrukturgesetzes (KHSG) die Weiterentwicklung einer qualitätsorientierten Krankenhausplanung, -finanzierung sowie -vergütung vorangetrieben. Diese bezieht sich sowohl auf die Struktur- und Prozessqualitätsanforderungen (u.a. Sicherstellung einer qualitativ hochwertigen Versorgung unter Berücksichtigung der Erreichbarkeit, Mindestmengenregelung), aber auch in verstärktem Maße auf die Quantifizierung von Ergebnisqualität. Demzufolge sollen nach dem seit 1. Januar 2016 in Kraft getretenem KHSG einzelne Leistungen mit nachgewiesener hoher Qualität mit Zuschlägen honoriert und entsprechend nachgewiesene unterdurchschnittliche Qualität mit Abschlägen belegt werden.

Das Modell, dass sich die Vergütung einer medizinischen Leistung an der Qualität eben dieser orientieren sollte, ist dabei nicht neu. Bereits seit einigen Jahren wird im deutschen Gesundheitssystem von verschiedenen Experten und Fachgesellschaften die Diskussion über die Einführung einer Vergütungsregelung nach dem Prinzip der separaten Honorierung einer quantifizierbaren Ergebnisqualität nach US-amerikanischem Vorbild („Pay-for-Performance“) geführt.

Denn gerade im US-amerikanischen Gesundheitswesen ist das Prinzip des „Pay-for-Performance“ (P4P) bereits seit Mitte der 90er-Jahre des vorigen Jahrhunderts im Rahmen

von sogenannten Managed-Care-Programmen etabliert. Ziel dabei ist sowohl die Verbesserung der Finanzsituation der Kostenträger wie auch die Stärkung der Patienten- und Versichertenrechte hinsichtlich der Überprüfung von Notwendigkeit, Umfang und Qualität der medizinischen Leistungserstellung.

### Grundkonzeption von P4P-Modellen in den USA

Der Leitgedanke der P4P-Modelle in den USA ist eine direkte Einflussnahme auf eine effiziente Mengensteuerung medizinischer Interventionen nach den Prinzipien der externen Evidenz und der patientenindividuellen Bedarfssituation. Dabei ist die Ausgestaltung der P4P-Modelle jedoch nicht einheitlich. Sowohl in Vergütungsform, Umfang und Bewertung der medizinischen Intervention und der Anzahl der eingebundenen Leistungserbringer (Ärzte, Ärztgruppen, Kliniken etc.) unterscheiden sich die verschiedenen Modelle zum Teil erheblich. Die Gründe dafür liegen unter anderem in der Struktur und dem polypolistischen Aufbau des US-amerikanischen Versicherungsmarktes, der mit mehr als 2 000 sogenannten Managed-Care-Organisationen (auch oft als „Healthplans“ bezeichnet) und einer weitgehenden Freiheit in der Vertragsführung (etwa Ein- und Ausschluss ausgewählter Risiken und Leistungsbereiche) viele Möglichkeiten für unterschiedliche P4P-Modelle bietet. Jenseits der bundesstaatlichen Grundversorgung via Medicare (Bürger über 65 Jahren), der neu eingeführten bundesstaatlichen Grundversorgung („Obama Care“) und den einzelstaatlichen Programmen für Einkommensschwache sowie Kinder und Jugendliche bis 18 Jahre (Medicaid) besaß und besitzt das US-amerikanische Versicherungssystem keine einheitliche Struktur. Dadurch entstanden unzählige Kontrakte und Detailvereinbarungen zwischen den in den USA überwiegend gut organisierten sogenannten Provider-Groups (Ärzte, Apotheken-Ketten, aber auch Kliniken und Klinik-Gruppen etc.). Aufgrund der großen Wahlfreiheit hin-

sichtlich des Einschlusses von individuellen Risiken, Versicherungsleistungen und der Dominanz arbeitgeberfinanzierten Versicherungspolices (sogenannte „fringe benefits“) entwickelte sich ein sehr komplexes „Contracting“ zwischen den einzelnen Healthplans und den jeweiligen Leistungsbringern. Während die rund 7 500 US-amerikanischen Kliniken eher in einem tendenziell – zumindest in Ballungsräumen – überbesetzten Markt agieren, konnten Mediziner ihre im Gesamtsystem eher überschaubare Zahl (nur ca. 500 000 Ärzte für eine Gesamtpopulation von rund 300 Millionen Einwohnern) durch ihren hohen Organisationsgrad einkommenssteigernd in neue Versichertenmodelle einbringen. Die Vergütung einer besonderen Qualität von medizinischen Leistungen entwickelte sich nur im Bereich der stationären Versorgung eher durch die Nachfragemacht von Kostenträgern. Im Bereich der niedergelassenen Ärzte hingegen bedeutete das Angebot von P4P zunächst eine Option für zusätzliche Vergütungen durch sinnvoll vereinbarte Mehrleistungen und Qualitätsziele bei tendenzieller Knappheit des „ökonomischen Gutes“ ärztliche Versorgung.

Die Erfahrungen aus mehr als 20 Jahren mit US-amerikanischen P4P-Modellen können nach Ansicht der Autoren mit dem Ziel einer Anpassung an europäische Versorgungsstrukturen und der Umsetzungsmöglichkeiten einer qualitätsorientierten Vergütung nur limitiert einbezogen werden. Gerade die weitreichende Liberalität des US-Versicherungsmarktes im Hinblick auf individuelle Prämien-gestaltungen (Risikoselektion), den Umfang des Versicherungsschutzes und den Einbezug von einzelnen, ausgewählten Leistungsanbietern in die Versorgungskette bedarf hinsichtlich einer Umsetzung sinnvoll erscheinender Konzepte einer kritischen Würdigung. So dürfte es bei der Implementation von P4P-Modellen in den deutschen Versicherungsmarkt als langfristige Zielperspektive eher um die Beseitigung von Überkapazitäten im stationären Sektor gehen. Die dauerhafte Schaffung budgeterhöhender zusätzlicher Einkommensmöglichkeiten nach US-amerikanischem Vorbild für die ärztliche Seite wird für deutsche P4P-Modelle wegen allgemeiner Budgetrestriktionen weniger relevant werden. Wegen der Unverzichtbarkeit der bestehenden Kliniken der Schwerpunkt- und Maximalversorgung im deutschen Gesamtsystem und in der regionalen (oft gesundheitspolitisch abgesicherten) Aufgabenteilung dieser Kliniken werden sich P4P-Modelle in erster Linie als Gefahr für die Systemrelevanz von Kliniken der Grund- und Regelversorgung einstufen lassen. Dies dürfte im Ergebnis zur Unterstützung von Einkaufsmodellen der Krankenkassen in eher elektiven Leistungsbereichen führen. Dafür steht etwa das niederländische Beispiel der DPC (Fallpauschalen), die bereits in eine Gruppe von DPC für Akut-Fälle und eine Gruppe für elektive Versorgungen getrennt sind. Während die erste Gruppe hinsichtlich der Entgelte – zugunsten der Kliniken – durch feste Pauschalen vergütet wird und nicht verhandelbar ist, existiert für die Pauschal-Entgelte der elektiven Versorgungen eine Verhandlungsfreiheit zwischen Klinik und Kostenträger. Hier kommen dann P4P-Überlegungen zum Zuge, die – bei deutlich unterlebensgener Qualität – für die einzelne Klinik in Vergütungsnachtei-

len, aber auch in einen Totalverlust des gesamten im Rahmen des P4P-Modells aufgerufenen elektiven Leistungsbereiches münden können.

### Die Eckpunkte von P4P-Modellen im US-amerikanischen Versicherungssystem

Wenn die Ergebnis-Qualität medizinischer Interventionen (und nur diese Betrachtung macht hinsichtlich tatsächlich eintretender Steuerungseffekte Sinn) zum Gegenstand von Vergütungsmodellen gemacht werden soll, ist die zentrale Frage nach der Justiziabilität und der damit verbundenen Messung von anerkannten Qualitätskriterien, die auch tatsächlich von den Kliniken beeinflusst werden können. Dies ist umso bedeutender, da diverse Risiken für negative Outcomes bereits durch den Patienten und seine spezifische Krankheitssituation (etwa Komorbiditäten, Alter) in der Aufnahmesituation unabhängig von der Versorgungsstufe bestehen können („present-on-admission“). Werden diese Faktoren nicht klar von den klinikseitig zu verantwortenden Behandlungsergebnissen unterschieden, besteht die Gefahr, dass Kliniken für eine Abwehr komplexer Behandlungsrisiken auch noch monetär belohnt werden. Allerdings konnten diese Herausforderungen in den USA auf die bereits bestehenden Verfahren beispielsweise zur Bestimmung von „Physician Malpractice“ im Rahmen von weit verbreiteten und zum Teil umfangreichen Schadensersatzklagen anlässlich medizinischer Behandlungen aufsetzen, die für die Leistungserbringer erhebliche, zum Teil existenzbedrohende Dimensionen erreichen konnten (und können). Auch bei diesen Verfahren gilt es, ähnlich wie im Bereich des P4P, folgende justiziablen Bewertungen vorzunehmen:

- In welcher Prognose wurde der Patient in die Klinik aufgenommen? Welche Komorbiditäten und Komplikationen lagen bereits zu Beginn der Behandlung vor?
- Gab es einen statistisch signifikanten Zusammenhang zwischen Intervention und Outcome? (Kausalität)
- Wie hoch war der statistische Erwartungswert für ein bestimmtes medizinisches Outcome vor dem Hintergrund einer unterschiedlichen Morbiditätslast der Patienten?
- Haben finanzielle und/oder personelle Restriktionen einen Einfluss auf das Behandlungsergebnis ausgeübt (zum Beispiel Mortalität bei Aufnahme am Wochenende)?

Während im Bereich der Schadensersatzforderungen auch in den USA immer auf eine individuelle Situation abgehoben wird, stellen P4P-Modelle häufig populationsbezogene Outcomes in den Mittelpunkt von Vergütungsmodellen. Beispiele sind etwa vermiedene stationäre Aufnahmen, reduzierter Einsatz von redundanten Arzneimitteln und Blutprodukten, Eingrenzung diagnostischer Leistungen und eventuell vermeidbare Überdiagnostik in einer vorher mit den Leistungserbringern genau bestimmten Population.

Die in das P4P-Modell eingeschlossene Population wird dann zu Beginn des jeweiligen Behandlungs-Programms entsprechend ihrer Morbiditätslast (Komorbiditäten) risikoadjus-

tiert bewertet. Anhand von umfangreichen Datenbanken lassen sich dann risikoadjustierte prädiktive Aussagen über die Wahrscheinlichkeit der jeweiligen Behandlungsergebnisse machen. Kommt es zu Abweichungen, kann ermittelt werden, ob ein nicht erreichtes gewünschtes Behandlungsergebnis tatsächlich durch die medizinische Intervention (den Arzt oder die Klinik und deren jeweilige Prozessperformance) oder aber eher durch die ungünstige Prognose des Patienten verursacht worden ist. Dem Ausschluss des Zufalls innerhalb der Bewertungsalgorithmen kommt dabei eine zentrale Bedeutung zu. Nicht risikoadjustierte Bewertungen von Ergebnisqualitäten, die bereits in einem nominalen Wertevergleich (Chi-Quadrat-Test etc.) keine Signifikanz zeigen, haben innerhalb dieser P4P-Modelle keinerlei Bedeutung, da diese Outcomes dann absolut nicht justizibel sind, sondern allenfalls tendenzielle Informationen enthalten.

Eine weitere wichtige Determinante der Belastbarkeit von Messverfahren zur medizinischen Prozess- und Ergebnisqualität sind die gewählten Ergebniskriterien selbst sowie die Art der Erhebung dieser Kriterien. Diese Debatte könnte in Deutschland in der Tat durch Erfahrungen aus anderen Gesundheitssystemen befruchtet werden, ohne dass es dazu allzu großer Grundsatzdiskussionen über die Qualität von Gesundheitssystemen selbst bedarf.

Da ist zum einen die immer wieder gemachte Erfahrung, dass die Datenbereitstellung im Rahmen von medizinischen Qualitätsbewertungen definitiv ohne subjektive und sequentielle Filter und möglichst ohne Manipulationsmöglichkeiten erfolgen muss.

Vor dem Hintergrund dieser unbedingt einzuhaltenden Grundsätze haben sich Routinedaten (idealerweise Abrechnungsdaten) als die weitaus verlässlichste Option etabliert. Diese lassen nach Risikoadjustierung eine weitgehend unverzerrte Messung von Prozess- und Ergebnisqualität in großen Kollektiven zu. Der Gegenentwurf dazu – die Beobachtung (inklusive Befragung) zu ausgewählten Einzelkriterien für eine begrenzte Fragestellung – hat sich international in einer justiziblen Perspektive wegen des teilweise umfangreichen und oft fehleranfälligen Erhebungsaufwandes und der immer wieder gegebenen Manipulationsmöglichkeiten nicht durchsetzen können, obwohl sich insbesondere europäische Gesundheitssysteme damit sehr auseinandergesetzt haben.

In den US-amerikanischen P4P-Modellen haben sich – unter der Prämisse einer möglichst weitgehenden Erhebung von Komorbiditäten und Komplikationen bereits bei Patientenaufnahme – hinsichtlich der Bestimmung der Qualität medizinischer Interventionen im stationären Bereich vor allem folgende Kriterien bewährt:

- Risikoadjustierte Inhouse-Mortalität (nicht 30-Tage-Mortalität)
- Risikoadjustierte Verweildauer (Abweichung vom risikoadjustierten Erwartungswert)
- Risikoadjustierte stationäre Wiederaufnahme nach Entlassung („Readmissions“)

Diese Kennzahlen zeichnen sich – bei statistischer Signifikanz der Ergebnisse und entsprechender „present-on-admission“ bezogener Datenerhebung – hinsichtlich der Verantwortungszuordnung durch eine hohe Justiziabilität und unproblematische Plausibilisierung aus.

Die gewählte Risikoadjustierung muss daher in der Lage sein, bezüglich der Kausalität eine klare Abgrenzung zwischen einem Risiko, das vom Patienten ausgeht, einem Behandlungsrisiko und einer dritten Risikodimension (klinikbezogenes Risiko) zu unterscheiden. Wenn dies nicht möglich ist, können valide Messungen einer Ergebnisqualität nicht durchgeführt werden. Damit steht und fällt die gesamte Debatte um qualitätsbezogene Vergütungsformen.

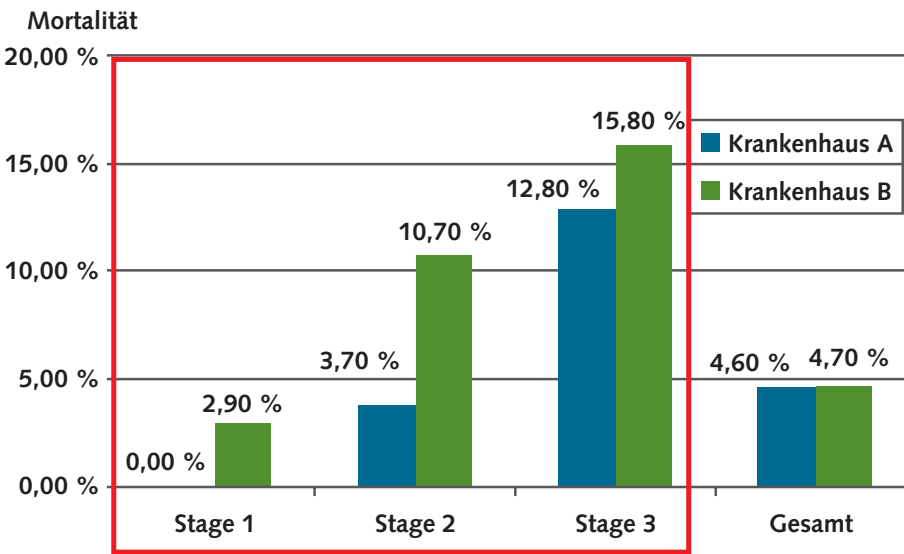
### Beispiel: Inhouse-Mortalität als kritisches Qualitätskriterium bei unzureichender Risikoadjustierung

Die Inhouse-Mortalität nimmt einen zentralen Stellenwert in P4P-Modellen zur Ergebnisbewertung medizinischer Interventionen ein. Gerade dieses Kriterium ist aber auf eine möglichst umfassende Risikoadjustierung angewiesen, um überhaupt eine Aussagefähigkeit zu erlangen. Zudem ist die Verteilung derart linksschief, dass bei kleineren Kollektiven (weniger als 200 Patienten) selbst eine Zunahme um mehr als 1 Prozent noch nicht als statistisch signifikant zu bewerten ist.

Anhand des nachfolgenden Beispiels werden zwei Kliniken gleicher Versorgungsstufe und mit vergleichbarem Leistungsspektrum bewertet. Diese weisen bei risikoadjustierter Betrachtung deutliche und signifikante Qualitätsunterschiede hinsichtlich der Inhouse-Mortalität in allen Stufen der medizinischen Fallschwere auf. In der undifferenzierten Gesamtbetrachtung liegen die Mortalitätskennziffern in einer vergleichbaren Größenordnung; Unterschiede zwischen den Kliniken sind nicht zu erkennen. Nach Risikoadjustierung durch eine möglichst umfangreiche Erfassung von Komorbiditäten bei Aufnahme der Patienten lässt sich nun aber ein wesentlich präziseres Bild der „Performance“ der Klinik abbilden. Der eingesetzte Grouper ermittelt unter Beachtung aller relevanten Komorbiditäten der Patienten drei Hauptgruppen mit unterschiedlicher Prognose hinsichtlich einer Mortalität:

- **Stage 1:** Patienten haben weder in der Hauptdiagnose noch in den Nebendiagnosen eine Mortalitätserwartung über 1 Prozent. Meist ist nur ein Organsystem betroffen (planbare Routinebehandlungen, bei denen eigentlich kein Mortalitätsrisiko gegeben ist). Diese Gruppe umfasst in vielen europäischen Kliniken bis zu 60 Prozent des gesamten Patientenkollektivs.
- **Stage 2:** Patienten haben in der Hauptdiagnose oder in einer der Nebendiagnosen eine Mortalitätserwartung bis zu 3 Prozent, im Einzelfall bis zu 5 Prozent. Diese Gruppe stellt in europäischen Kliniken ca. 25 Prozent der Gesamtkollektive dar (komplexere aber immer noch planbare Eingriffe, bei denen bereits ein höheres Komplika-

**Abbildung 1:** Risikoadjustierter Vergleich der Inhouse-Mortalität über alle Fallschwere-Stadien



tionsrisiko besteht, wie zum Beispiel gelenkchirurgische Eingriffe, therapeutische Herz-Katheter-Untersuchungen u.a.).

- **Stage 3:** Patienten haben in der Hauptdiagnose eine sehr hohe Mortalitätserwartung (meist über 8 Prozent; im Einzelfall auch wesentlich höher). Diese Patienten stellen echte medizinische Notfälle dar, bei denen schnell gehandelt werden muss (ca. 15 Prozent der Kollektive).

Im Ergebnis hat also die Klinik B in allen Fallschwere-Stadien eine signifikant höhere Inhouse-Mortalität zu verzeichnen. Bei unzureichender Risikoadjustierung wird dieses Ergebnis nicht evident und kann nicht zur weiteren Steigerung der Patientensicherheit genutzt werden (siehe ► **Abbildung 1**).

den USA sowie auch in einigen Ländern der EU entwickelten Steuerungs- und Messverfahren zur Bestimmung der medizinischen Prozess- und Ergebnisqualität könnten auch für die Umsetzung der Vorgaben des KHSG („Qualität wird vergütungsrelevant“) genutzt werden. Im Ergebnis könnten im deutschen Klinikmarkt P4P-Modelle den Weg für Einkaufsmodelle öffnen. Daher ist eine umfassende Risikoadjustierung der Behandlungsfälle den Kliniken unbedingt anzuraten.

**Fazit**

Eine quantifizierbare Prozess- und Ergebnisqualität kann bei adäquater Risikoadjustierung Bestandteil von Vergütungsmodellen im Bereich der stationären Versorgung sein. Unverzichtbare Voraussetzung für die Etablierung justiziabler Messverfahren ist die Erhebung aller Komorbiditäten und Risikofaktoren unmittelbar bei Patientenaufnahme („present-on-admission“). Dies könnte für die stationäre Versorgung in Deutschland den Abschied von einer vorwiegend erlösorientierten Kodierung bedeuten.

Die im Laufe der Jahre im Rahmen von P4P-Modellen in

**Anschrift der Verfasser**

Franz-Josef Fischer/Benjamin Lüßem, Medstat S.A. Luxembourg, Niederlassung Deutschland, Kaiser-Wilhelm-Ring 41, 40545 Düsseldorf/  
Dr. Jaan Sidorov, Harrisburg, Pennsylvania, USA

„Ich schaue zweimal hin, wenn es um das Thema Krankenhausbau geht. Wer ohne Beeinträchtigung des laufenden Krankenhausbetriebs bauen will, sollte über Modulbau nachdenken.“

**AUF DEN ZWEITEN BLICK DIE ERSTE WAHL.**

Genaueres Hinsehen lohnt sich – Modulbau ist mehr als eine Alternative zu konventionellen Bauweisen. Denn wir bauen anders: zuverlässig, modern und innovativ.

- **70 % kürzere Bauzeit** und **Terminsicherheit** durch witterungsunabhängige, industrielle Produktion
- **Saubere und leise Baustelle** durch hohen Vorfertigungsgrad
- Planungs- und Investitionssicherheit durch **Festpreisgarantie**
- **Geringe Life Cycle Costs** durch integrale Planung und nachhaltige Bauweise
- **Kontrollierte Qualität** durch zertifizierte Werksfertigung

Fixe Kosten. Fixe Termine. Fix fertig. [www.alho.com](http://www.alho.com)

