

# „Qualitätszuschläge“ in der kassenärztlichen Honorierung von Hausärzten

Bedrohung der Qualität der Versorgung

## *Pay-for-Performance as “Quality Surplus Cash” in Family Practice*

*A Challenge to the Quality of Health Care*

Stefan Sachtleben<sup>1</sup>

**Hintergrund:** Das medizinische Arbeitsfeld der Hausarztpraxis ist dadurch charakterisiert, dass klar abgrenzbare und benennbare Diagnosen selten vorkommen; eine Niedrigprävalenzumgebung vorliegt. Dies hat Einfluss auf die Vorhersagekraft von Untersuchungen, insbesondere auf die Zahl der falsch-positiven Befunde. Diese nehmen in ihrem Anteil an allen positiven Befunden zu, und sie nehmen absolut gesehen nochmals zu, wenn viele, wenig gezielte Untersuchungen gemacht werden. Eine Honorierung, die in diesem Arbeitsfeld zur Durchführung vieler Untersuchungen Anreize setzt, lässt viele falsch-positive Befunde entstehen. Die Honorierung mittels sog. „Qualitätszuschläge“ ist eine solche Honorierung.

**Methode:** Für 4 technische Untersuchungen, für die es Qualitätszuschläge gibt, wird für die bundesdeutsche Durchschnitts-Praxis und unter Nutzung bekannter Testcharakteristika (Sensitivität und Spezifität) und anzunehmender Prävalenzen, die Zahl falsch-positiver Befunde bei der Erbringung dieser Leistungen errechnet.

**Ergebnisse:** Technische Untersuchungen in der hausärztlichen Niedrigprävalenzumgebung tragen das Risiko, dass pro einem richtig-positiven Patienten fast zwei falsch-positive Patienten „erzeugt“ werden. Die tatsächliche Anzahl der falsch-positiven Patienten dürfte sogar höher liegen.

**Schlussfolgerungen:** Eine Vergütung, die die massenweise Anwendung von Untersuchungen in der Hausarztpraxis fördert, ist nicht versorgungsgerecht und gefährdet über das Entstehen vieler falsch-positiver Befunde die Qualität in diesem Versorgungsbereich.

**Background:** FP-environment is different from hospital-environment. Clear-cut and well defined diagnoses are rare, i. e. the prevalence of severe diseases is low. These epidemiologic features influence predictive values of diagnostic tests, i. e. the rate of false-positive findings will increase. Any payment system which gives incentives for an increase in using an investigation will therefore increase the number of false-positive findings. The recent introduction of the so called “Quality Surplus Cash” in the German payment system endangers the quality of care in family medicine.

**Methods:** We applied published data on sensitivity and specificity of four technical investigations to the prevalences of the diseases which should be identified by these tests, to calculate the rate of false-positive findings.

**Results:** Technical investigation in FPs' low-prevalence-environment carries the risk of producing nearly two false-positive findings for each correct-positive finding. The true number of false-positive findings might be considerably higher than those found by calculations considering best case.

**Conclusion:** A payment system employing incentives for a high number of technical investigations in family practice might endanger quality of care.

*Keywords: fee-for-service, reimbursement, family practice, quality of care*

*Schlüsselwörter: Honorierungssystem, Einzelleistung, Hausarztpraxis, Qualität der Versorgung*

<sup>1</sup> niedergelassen in Einzelpraxis in Pirmasens

## Hintergrund

In seiner 7. Sitzung am 27./28. August 2008 hat der erweiterte Bewertungsausschuss unter Punkt 5 im Sitzungsprotokoll „... ein weiteres Honorarvolumen ...“ beschlossen, dass dem Hausarzt „... zur Verfügung steht, wenn er über die entsprechenden Qualifikationsvoraussetzungen verfügt“. Dieses Honorarvolumen wird heute „Qualitätszuschlag“ genannt (siehe Tabelle 1).

Auf Basis dieser sog. Qualitätszuschläge wird ein Budget für die jeweilige, dann selbst als Einzelleistung abzurechnende Leistung berechnet. Dazu werden die in Tabelle 1 genannten Beträge mit der Zahl der im Quartal versorgten Patienten multipliziert. Kommt man mit der Erbringung der Einzelleistung für einen Leistungsbereich über die Budgetgrenze, dann werden diese Leistungen nicht mehr honoriert. Bleibt man unterhalb der Budgetgrenzen, dann hat man „Geld verschenkt“. Es besteht damit ein erheblicher ökonomischer Druck, das Budget auszuschöpfen, weil es sich hier um ein Geldvolumen von bis zu 10 % des Umsatzes einer entsprechenden Praxis handelt. Diesem ökonomischen Druck nachzugeben, hat Auswirkungen, die dieser Aufsatz aufzeigen möchte.

Die Frage stellt sich, ob unter diesem „Druck“ etwas Positives in Bezug auf die Qualität der Versorgung resultiert, bzw. resultieren kann. Dazu ist der Arbeitshintergrund der Hausarztpraxis zu betrachten.

### Hausarztumgebung ist Niedrigprävalenzumgebung

Es gibt zahlreiche Literatur, die sich mit der Frage beschäftigt, inwiefern sich der hausärztlich, allgemeinmedizinische Bereich von dem spezialmedizinischen, vor allem stationären Bereich unterscheidet [1]. Die Vielzahl der genannten Unterschiede lassen sich auf ein Kernkriterium zurückführen: die unterschiedliche Zusammensetzung der Patientengruppen. Hausärzte sehen mehr und andere Patienten als Spezialärzte. Was ist der Unterschied?

Green [6] hat diesen Unterschied auf der Basis von Befragungsdaten in der Bevölkerung deutlich dargestellt (siehe Abbildung 1). Von 1000 Personen haben ca. 80 % Beschwerden, 23 % werden

Nr.	Leistungsbereich	GOP	Betrag in Euro
1.	Sonografie	33000 bis 33002, 33010 bis 33012, 33040 bis 33044, 33050 bis 33052, 33060 bis 33062, 33076, 33080, 33081, 33090 bis 33092	3,50
2.	Psychosomatik	35100 und 35110	3,00
3.	Prokto-/Rektoskopie	03331 bzw. 04331	1,00
4.	Kleinchirurgie	02300 bis 02302	1,50
5.	Langzeit-EKG	03322 bzw. 04322	1,00
6.	Langzeit-Blutdruckmessung	03324 bzw. 04324	1,00
7.	Spirometrie	03330 bzw. 04330	1,00
8.	Ergometrie	03321 bzw. 04321	1,50
9.	Chirotherapie	Chirotherapie GOPen des Abschnitts 30.2 des EBM	1,00

**Tabelle 1** Fallwertzuschlag zum Regelleistungsvolumen für Ärzte des hausärztlichen Bereiches für qualitätsgebundene Leistungen [3].

beim Allgemeinmediziner vorstellig, lediglich knapp 1 % bedürfen der Krankenhausversorgung. Diese Zahlen bestätigen sich auch in anderen Studien [13]. Die hausärztliche Patientenpopulation besteht ganz offensichtlich – im Verhältnis von etwa 25 : 1 – aus anderen Patienten als die Population der Krankenhauspatienten. In der hausärztlichen Population finden sich viele Patienten mit Beschwerden – aber nur wenige mit ernsten und eindeutigen, d. h. gut definierten Diagnosen. Denn meist bleiben die Beschwerden im vordiagnostischen Bereich und lassen sich durch symptomatische Therapie behandeln. In der spezialärztlichen Population sind dagegen ernste oder eindeutige Diagnosen sehr viel häufiger, denn diese Population ist von der Hausarztmedizin gerade deswegen „ausgesiebt“ worden.

Wegen der Unterschiede in der Prävalenz von Krankheit in den beiden Populationen, wird der spezialmedizinische Bereich als Hochprävalenzbereich und der allgemeinmedizinische als Niedrigprävalenzbereich bezeichnet. Die Prävalenzen eines bestimmten medizinischen Arbeitsgebietes haben weitreichende Konsequenzen auf Verwertbarkeit technischer Untersuchungen in

diesem Bereich. Diese Konsequenzen werden weithin unterschätzt.

### Sensitivität und Spezifität und die Vorhersagewerte

Die Leistungsfähigkeit eines Tests wird üblicherweise mit Sensitivität und Spezifität definiert. Sensitivität ist die Fähigkeit eines Tests, Kranke als krank zu erkennen. So bedeutet eine Sensitivität von 80 %, dass aus einem Kollektiv von 100 Kranken (!) 80 korrekt als „krank“ erkannt werden (Richtig-Positive), jedoch 20 fälschlicherweise als „gesund“ bezeichnet werden (Falsch-Negative). Spezifität von 80 % bedeutet, dass aus einem Kollektiv von 100 Gesunden (!) 80 korrekt als „gesund“ erkannt werden (Richtig-Negative), jedoch 20 fälschlicherweise als „nicht gesund“ (Falsch-Positive). Erweitert kann man sagen: Von der Sensitivität eines Tests hängt ab, wie viele Kranke übersehen werden. Von der Spezifität eines Tests hängt ab, wie viele Gesunde fälschlicherweise als krank bezeichnet werden. Sensitivität und Spezifität sind jedoch auf Basis von Kollektiven berechnet, die zu 100 % aus Kranken oder eben Gesunden bestehen. In der ärztlichen Arbeit kommen solche Kollektive nicht vor. Normale

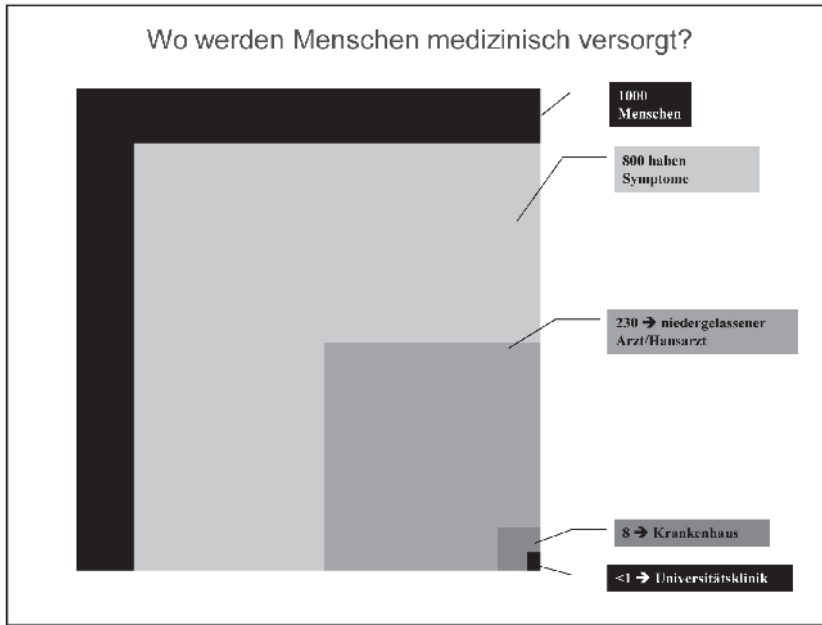


Abbildung 1 Wo werden Menschen mit Beschwerden medizinisch versorgt? [nach 6]

Leistung:	Sono	LZ-RR	LUFU	B-EKG
<b>Maximale Erbringung:</b>	180	104	135	60

Tabelle 2 Maximale Erbringungsanzahl der technischen Tests innerhalb des Budgets „Qualitätszuschläge“ in einer 800-Scheine-Hausarztpraxis.

Leistung	Max. Erbringung pro Praxis	Sensitivität und Spezifität	Erbracht insgesamt von 60.000 Praxen	Richtig-Positive	Falsch-Positive
Sono	180	65 %/90 %	10.800.000	70.200	1.069.200
LZ-RR	104	90 %/90 %	6.240.000	1.010.880	511.680
LUFU	135	90 %/90 %	8.100.000	729.000	729.000
B-EKG	60	65 %/80 %	3.600.000	23.400	712.800
<b>Alle zusammengefasst</b>			<b>28.740.000</b>	<b>1.833.480</b>	<b>3.022.680</b>

**Der Tabelle liegen folgende Prävalenzen zugrunde:**

Sonographie: 1 % – geschätzt, für sonografisch zu detektierende Erkrankungen

LZ-RR: 18 % für Hypertonie aus [7]

LUFU: 10 % für COPD aus [12]

B-EKG: 1 % für KHK aus [8]

Quellen der Sensitivitäten und Spezifitäten:

Sonografie: geschätzt

LZ-RR: geschätzt

LUFU: aus [9]

B-EKG: aus [11]

Spezifitäten wurden mit 90 %, bzw. 80 % beim B-EKG, sehr hoch angenommen; in der Realität dürften sie eher niedriger liegen. Dies hätte auf die dargestellten Zahlen den Effekt, dass die Falsch-Positiven weiter zunehmen.

Patientenkollektive sind gemischt – und es gibt viele Grenzbefunde zwischen „krank“ und „gesund“.

Bayes (Englischer Mathematiker und presbyterianischer Pfarrer, 1702–1761) hat gezeigt, dass die Leistungsfähigkeit eines Tests bei Anwendung wesentlich und zusätzlich zu Sensitivität und Spezifität davon abhängt, mit welcher Prävalenz die gesuchte Erkrankung in dem Untersuchungskollektiv vorkommt. Deswegen ist auch die Leistungsfähigkeit einer medizinischen technischen Untersuchung davon abhängig, ob sie im spezialmedizinischen Hochprävalenz- oder im allgemeinmedizinischen Niedrigprävalenzbereich durchgeführt wird. Das Ergebnis eines stationär durchgeführten EKGs ist zum Beispiel anders zu bewerten als das eines hausärztlich durchgeführten EKGs. Die Leistungsfähigkeit von Tests unter unterschiedlichen Prävalenzen wird nach dem Bayes'schen Theorem mit den positiven/negativen Vorhersagewerten beschrieben.

**Methode**

Aus Tabelle 1 geht hervor, dass es sich bei den genannten Qualitätszuschlägen unter anderem um 6 technische Tests

Tabelle 3 Ergebnisse von 60.000 Hausärzten mit durchschnittlich 800 Scheinen, pro Quartal.

**Dr. med. Stefan Sachtleben ...**

... Jahrgang 1957, niedergelassen in Einzelpraxis in Pirmasens seit 1994. Schwerpunkte: Geriatrie, hausärztliche Psychiatrie, Suchtmedizin mit Methadonabgabe. Lehrbeauftragter Allgemeinmedizin an der med. Fakultät Homburg/Saar. Mitglied im Beirat des Vorstandes des Hausärzterverbandes Rheinland-Pfalz. Organisator des Tages der Allgemeinmedizin in Homburg/Saar.

handelt. Prinzipiell gilt das im Text Dargestellte für alle Untersuchungsmethoden – und somit auch für alle hier im Zusammenhang mit dem „Qualitätszuschlag“ aufgeführten Untersuchungen. Jedoch liegen für Langzeit-EKG und Rektoskopie die entsprechenden Zahlen zur Testcharakteristik (s. o.) nicht vor, sodass sich die Argumentation auf die 4 folgenden Untersuchungen (Tests) bezieht: Sonografie, Langzeit-Blutdruckmessung (LZ-RR), Spirografie (LUFU) und Ergometrie (B-EKG). Wie oft diese in der oben dargestellten Budgetlogik abgerechnet werden dürfen, stellt Tabelle 2 beispielhaft für eine 800-Scheine-Praxis dar. Die 800-Scheine-Praxis ist die bundesdeutsche Durchschnitts-Hausarztpraxis.

Unter Kombination von Testcharakteristika zu diesen 4 Untersuchungen, der anzunehmenden Prävalenz der mit den Untersuchungen gesuchten Erkrankungen sowie unter Ansatz der 800-Schein-Praxis wird die Zahl der richtig- und der falsch-positiven Befunde errechnet.

**Ergebnisse**

Bei Annahme von 60.000 Hausärzten mit im Schnitt 800 Scheinen, die ihre Budgets ausschöpfen, zeigen sich die in Tabelle 3 wiedergegebenen Zahlen pro Quartal. Es ist zu erwarten, dass innerhalb eines einzigen Quartales fast 29 Millionen Personen im Niedrigprävalenzbereich einer technischen Untersuchung unterzogen würden. Von diesen werden über 3 Millionen fälschlicherweise als „nicht gesund“ bezeichnet. Als Schaden der Massenuntersuchungen sind also über 3 Millionen weitere Abklärungen von Gesunden notwendig – dies jeweils pro Quartal.

Das Verhältnis von „gesund“ zu „vermeintlich nicht gesund“ beträgt über alle Untersuchungen 9 : 1. Von 9 Gesunden, die unnötig untersucht werden, wird fälschlicherweise 1 als Kranker bezeichnet werden (29 Millionen Untersuchte, davon fast 1,8 Millionen Kranke = 27 Millionen Gesunde; dies gegenüber 3 Millionen Falsch-Positiven).

Bezogen auf die tatsächlich Kranken, die Richtig-Positiven, siehe Tabelle 3. Es werden 1,8 Millionen Kranke gefunden. Diesen stehen gut 3 Millionen Falsch-Positiven gegenüber: Unter den als krank erscheinenden (4,8 Millionen) sind 1,8 Millionen wirklich krank. D. h. auf 1 wirklich Kranken kommen über 1,5 vermeintlich Kranke.

**Diskussion**

Dies bedeutet, dass auf 1 wirklich Kranken fast 2 vermeintlich Kranke „erzeugt“ werden. Tatsächlich jedoch dürfte die Zahl höher liegen, da bei diesen Beispielrechnungen vergleichsweise hohe Prävalenzen und sehr hohe Spezifitäten angenommen wurden. Zudem wurde hier fälschlich unterstellt, dass jeder gefundene „wirklich positive Fall“ auch einer ist, der a) therapeutische Konsequenzen hat und b) diese dann noch mit Erfolg für den Patienten zu ziehen sind. Nimmt man einmal realistisch an, dass dies aber nur in 10 % oder 20 % der Fall ist, so ist die Relation von vermeintlich Kranken zu real Kranken, die mit Nutzen behandelt werden können, schon bei 1 : 15 bis 1 : 10.

Diese Zahlen zeigen, dass die hochfrequente Anwendung technischer Untersuchungen dem hausärztlichen Niedrigprävalenzbereich nicht nur nicht gerecht wird, sondern gesundheitssystematisch kontraproduktiv ist. Es ist versor-

gungsgerechter dem Primärversorger zu ermöglichen, eine Untersuchung nur dann einzusetzen, wenn er durch sorgfältiges klinisches Eingrenzen die Wahrscheinlichkeit für das Vorliegen einer testbaren Erkrankung erhöht hat. Hierzu gibt es üppige Literatur [4, 10, 11, 5]. Die Folge eines solchen Vorgehens wird sein, dass technische Untersuchungen in der Hausarztpraxis sehr viel seltener, aber mit geringerem Schaden erbracht werden. Technische Tests als Einzelleistung zu vergüten, blockiert eine solche kritische hausärztliche Arbeitsweise, da Geräte so nicht kostendeckend betrieben werden können. Es wäre versorgungsgerechter und dem Niedrigprävalenzbereich angemessener, technische Tests über z. B. Vorhaltepauschalen zu vergüten.

Aktuell hat der Barmer-GEK Arztreport 2010 [2] erneut die weltweit einmalig hohe Arzt-in-Anspruchnahmerate in Deutschland dargestellt. Die hier vorgestellten Zahlen zeigen auch, dass die unkritische Anwendung technischer Tests ein Motor dieser ungewöhnlichen Entwicklung sein könnte.

Eine andere Frage ist die generelle Bedeutung technischer Untersuchungen in der Hausarztpraxis, da pathologische Befunde in ihrer Bedeutung für den einzelnen Patienten schwer einzuschätzen sind. Hier soll aber nicht der technischen „Barfußmedizin“ das Wort geredet werden, denn die hohen Zahlen der „Richtig-Negativen“ unter Niedrigprävalenzbedingungen erlauben dem Hausarzt besonders sichere Aussagen, wenn die technischen Tests keine pathologischen Veränderungen zeigen – d. h., wenn die Patienten gesund sind. Bei der Interpretation pathologischer Veränderungen jedoch, müssen Selbstkritik und strenges klinisches Denken im Vordergrund stehen.

**Interessenkonflikte:** keine angegeben.

**Korrespondenzadresse:**

Dr. med. Stefan Sachtleben  
Kaiserstr. 2a  
66955 Pirmasens  
E-Mail: stefan.sachtleben@t-online.de

**Literatur**

1. Abholz HH, Donner-Banzhoff N. Epidemiologische und biostatistische Aspekte der Allgemeinmedizin. In: Kochen M. Allgemeinmedizin und Familienmedizin. 3. Aufl. Stuttgart: Georg Thieme Verlag, 2006; 507–521
2. Barmer-GEK. [www.Barmer.de](http://www.Barmer.de) (Zugriff am 23.01.2010)
3. Bewertungsausschuss-Erweiterter. [www.kbv.de](http://www.kbv.de) (Zugriff am 23.01.2010 auf <http://www.kbv.de/8157.html>)
4. Donner-Banzhoff N. Wie stellt der Allgemeinarzt eine Diagnose? *Z Allg Med* 1999; 75: 744–749
5. Doust J. Using probabilistic reasoning. *BMJ* 2009; 339: 1080–1082
6. Green LA. The ecology of medical care – revisited. *N Eng J Med* 2001; 344: 2021–25
7. Hydra-Studie. Multimorbidität wird unterschätzt. *Dtsch. Ärztebl* 2002; 99: A-1933
8. Kolenda KD. Sekundärprävention der koronaren Herzkrankheit: Effizienz nachweisbar. *Dtsch. Ärztebl* 2005; 26: A-1889
9. Schneider A. Diagnostic accuracy of spirometrie in primary care. *BMC pulmonary medicine* 2009; 9: 31
10. Schneider A. Stepwise diagnostic work-up in general practice as a consequence of the Bayesian reasoning. *Zae FQ* 2006; 100: 121–7
11. Sönnichsen A, Donner-Banzhoff N. Thoraxschmerzen – hausärztliches Dilemma zwischen Fehl- und Überdiagnostik. *Z Allg Med* 2006; 82: 223–230
12. Nationale Versorgungsleitlinie „COPD“. [www.versorgungsleitlinien.de](http://www.versorgungsleitlinien.de) (Zugriff am 23.01.2010)
13. White I. In: Rosser SM. *Evidence-based Family Medicine*. Hamilton-London: B. C. Decker Inc., 1998; 20