



Forschungsnahe
Beratungsgesellschaft
im Gesundheitswesen GmbH

**Beurteilung des Screening-Verfahrens
der Krankenversicherer in der Schweiz
zur Identifikation von Überarztung**

**ForBiG GmbH
in Kooperation mit**

**Prof. Dr. Jürgen Wasem
Lehrstuhl für Medizinmanagement**

**Gutachten beauftragt von:
Verein Ethik und Medizin Schweiz**

Version 2

Essen, den 26.07.2010

Inhaltsverzeichnis

1.	Projekthintergrund	5
2.	Status Quo des Vorgehens bei Überarztung in der Schweiz	6
3.	Diskussion des Wirtschaftlichkeitsprüfungsablaufs	10
4.	Risikoadjustierung	12
4.1.	Risikostrukturausgleich: Schweiz vs. Deutschland	13
4.2.	Datenbasis und Methodik der Modellsimulationen.....	15
4.3.	Ergebnisse der Modellanalysen	18
4.4.	Trade-off des 130 %-Kriteriums im Screeningverfahren	24
5.	Gefahr der versteckten Rationierung durch die ANOVA-Methode	26
6.	Limitationen	30
7.	Diskussion	32
8.	Fazit	35
9.	Bearbeiter des Projektes	36
10.	Quellenverzeichnis	37

I. Darstellungsverzeichnis

Abbildung 1: Ablaufschema Wirtschaftlichkeitsprüfung.....	7
Abbildung 2: Reporting Statistikjahre 2002 - 2006	8
Abbildung 3: R ² -Werte der untersuchten Modelle - prospektiv -	15
Abbildung 4: Predictive Ratios in Abhängigkeit der HMG-Anzahl im Bereich Kardiologie.....	19
Abbildung 5: Predictive Ratios von Patientenkohorten im Bereich Kardiologie (homogene Morbiditätslasten)	20
Abbildung 6: Predictive Ratios von Patientenkohorten im Bereich Kardiologie (heterogene Morbiditätslasten)	22
Abbildung 7: Predictive Ratios in Abhängigkeit der Kohortengrößen.....	24
Tabelle 1: Identifikation der Fachrichtung Kardiologie über HMGn.....	17
Tabelle 2: Risikofaktoren der 10 Patientenkohorten.....	21
Tabelle 3: Korrelationsstärke: Kohortengröße – Residuum der Kostenschätzungen	23

II. Abkürzungsverzeichnis

ANOVA = Analysis of Variance

BVA = Bundesversicherungsamt

EMG = Erwerbsminderungsgruppe

HMG = Hierarchische Morbiditätsgruppen

Morbi-RSA = Morbiditätsorientierter Risikostrukturausgleich

PCG = Pharmaceutical Cost Groups

PR = Predictive Ratio

PZN = Pharmazentralnummer

RSA = Risikostrukturausgleich

1. Projekthintergrund

Die Krankenversicherer in der Schweiz nutzen ein gesetzlich verankertes Kontrollverfahren zur Identifikation von Fällen der Überarztung. Der Tatbestand gilt als offensichtlich, wenn behandelnde Ärzte in ihren Kosten mehr als 30 % über den durchschnittlichen Kosten der Vergleichsärzte ihrer Kontrollgruppe liegen. Die Kontrollgruppe wird über die Variablen Facharztgruppe und Kanton definiert und die Kosten der Ärzte werden über das Verfahren der ANOVA um die Effekte des Alters und des Geschlechts ihrer Patienten bereinigt, um einen Vergleich zu ermöglichen.

Liegt ein Arzt trotz Bereinigung der Alters- und Geschlechtseffekte seiner Patienten mehr als 30 % über den Durchschnittskosten vergleichbarer Ärzte (derselben Fachrichtung und desselben Kantons), gilt dies als Beweismittel für Überarztung und führt zu einer genaueren Ursachenforschung, die im weiteren Verlauf zu entsprechenden Konsequenzen für den behandelnden Arzt führen kann.

Die Kritik an diesem Screening-Verfahren der Krankenversicherer liegt in der unzureichenden Risikoadjustierung der Kosten beim Vergleich der Ärzte untereinander, sodass tatsächliche Vergleichbarkeit nicht gegeben sein könnte. Ein weiterer Kritikpunkt ist die versteckte Rationierung in der Grundversorgung, die dieses Prüfverfahren ex ante mit sich bringen könnte.

2. Status Quo des Vorgehens bei Überarztung in der Schweiz

Die Krankenversicherer in der Schweiz nutzen ein gesetzlich verankertes Kontrollverfahren zur Identifikation von Fällen der Überarztung, das sich aus dem 6. Abschnitt „Kontrolle der Wirtschaftlichkeit und Qualität der Leistungen“ des Bundesgesetzes über die Krankenversicherung KVG vom 18. März 1994 ergibt. Die Leistungserbringer unterliegen der gesetzlichen Qualitätssicherung¹ und werden auf die Wirtschaftlichkeit der Leistungen beschränkt:

„Der Leistungserbringer muss sich in seinen Leistungen auf das Maß beschränken, das im Interesse der Versicherten liegt und für den Behandlungszweck erforderlich ist.“²

Dadurch erhalten die Krankenversicherungen den gesetzlichen Auftrag zur Prüfung der Wirtschaftlichkeit der Leistungserbringer. Für die Wirtschaftlichkeitsprüfung im Bereich der ambulanten ärztlichen Versorgung wurde 1970 die Rechnungstellerstatistik angewendet, deren Grundlage ein Vergleich der Durchschnittskosten pro behandeltem Versicherten zwischen Ärzten der gleichen Spezialisierung und derselben Region ist.^{3,4}

Seit 2004 wurde die Rechnungstellerstatistik durch eine univariate Varianzanalyse (ANOVA-Methode, nach der englischen Bezeichnung *analysis of variance*) ergänzt, die eine zusätzliche Adjustierung der Kosten nach Alters- und Geschlechtsstruktur der Patientenkollektive der einzelnen Ärzte berücksichtigt.⁵

Santésuisse, der Branchenverband der schweizerischen Krankenversicherer im Bereich der sozialen Krankenversicherung⁶, kann Ärzte kontaktieren, die auf Basis der genannten Berechnungsgrundlage „statistisch auffällig“ werden. Dies ist in der Regel der Fall, wenn die Durchschnittskosten der Patienten des

¹ Vgl. Artikel 58 KVG vom 18.03.1994 (Stand am 1. Januar 2010)

² Vgl. Artikel 56 Absatz 1 KVG vom 18.03.1994 (Stand am 1. Januar 2010)

³ Vgl. Amstutz R (2005), S. 3.

⁴ Vgl. Kraft P (2005), S. 4-5.

⁵ Vgl. Roth H-R, Stahel W (2005), S. 1-17.

⁶ Vgl. santésuisse (2010)

individuellen Arztes mindestens 130 % des Referenzwertes (Durchschnittskosten der Patienten aller Ärzte der gleichen Facharztgruppe) übersteigen.⁷

Der Ablauf der Wirtschaftlichkeitsprüfung erfolgt in einem mehrstufigen Prozess, der (laut *santésuisse*) eine Klärung bzw. eine Kostensenkung zum Ziel hat. Dieses Wirtschaftlichkeitsverfahren konfrontiert betroffene Ärzte im Extremfall mit Regressforderungen (siehe Abbildung 1).

Abbildung 1: Ablaufschema Wirtschaftlichkeitsprüfung



Quelle: *santésuisse* (2009), S. 2.

In Abbildung 2 sind die Statistikjahre 2002 bis 2006 des Reporting SWV Tools der Maßnahmen der Wirtschaftlichkeitsprüfung abgebildet. Mit Einführung der ANOVA-Methode im Jahr 2004 erfolgte im Vergleich zum Vorjahr fast eine Verdopplung der „statistisch auffälligen“ Ärzte.

⁷ Vgl. *santésuisse* (2009), S. 1.

Abbildung 2: Reporting Statistikjahre 2002 - 2006

AUSGANGSLAGE	2002	2003	2004	2005	2006
TOTAL ÄRZTE	15 252	16 583	17 228	17 599	17 931
ANZAHL STATISTISCH AUFFÄLLIGE ÄRZTE	1367 8,96%	1297 7,82%	2335 13,55	2655 15,09%	2599 14,49%
MASSNAHMEN					
AUFKLÄRUNG/ INFORMATIONSBRIEF	231 1,51%	213 1,28%	884 5,13%	717 4,07%	596 3,32%
GESPRÄCH	6 0,04%	19 0,11%	63 0,37%	130 0,74%	113 0,63%
EINGABE (VERGLEICH/ PVK/GERICHT)	28 0,18%	35 0,21%	52 0,30%	163 0,93%	130 0,73%

Quelle: Wirtschaftlichkeitsprüfungen – Reporting SWV Tool – Statistikjahr 2002, 2003, 2004, 2005 und 2006

Quelle: Vgl. Brändli C (2009), S. 9.

Trotz der Kritik der Ärzte, dass die Methodik für die Identifikation der auffälligen Ärzte irreführend⁸⁸ sei und daher nicht für eine initiale Wirtschaftlichkeitsbeurteilung geeignet ist, ist die Wirtschaftlichkeitsprüfung bei den kantonalen Schiedsgerichten und dem Eidgenössischen Versicherungsgericht ein anerkanntes Verfahren.

In einem wissenschaftlichen Gutachten zum Vergleich verschiedener Instrumente zur Beurteilung der von Schweizer Ärzten verursachten Behandlungskosten erscheinen die Zweifel der Ärzte an der Methodik aus drei Gründen potentiell berechtigt⁸⁹:

1. Die wesentlichen Determinanten der ärztlich verursachten Kosten werden in der Analyse nicht berücksichtigt. Insbesondere die (in anderen Ländern bereits etablierte) Zusammensetzung und die Morbiditätslast des Patientenlientels (z. B. Anteil der Chroniker). Des Weiteren könnten in der Schweiz folgende Kriterien von Relevanz sein: der Praxisstandort (Stadt, Agglomeration oder Land), Art der Medikamentenabgabe (keine Abgabe, Notfallmedikation, Praxisapotheke), Alter des Praxisinhabers sowie die Notfalldienstleistung (dichotome Ausprägung: ja/nein).

⁸⁸ Vgl. Romanens M et al. (2009), S. 51-52.

⁸⁹ Vgl. Schwenkglenks, M (2010), S. 12-13.

2. Die Auswirkungen des Abdeckungsgrads fachlicher Leistungen (z. B. Röntgen) und der Zeitpunkt der Überweisung des Patienten an den Facharzt oder Spital werden nicht angemessen berücksichtigt. Die induzierten Medikamenten-, Labor und Physiotherapiekosten werden sowohl in der Rechnungstellerstatistik als auch in der ANOVA-Methode berücksichtigt, hingegen die induzierten Kosten bildgebender Verfahren, Facharztekosten, ambulanten und stationären Spitalkosten und Spitexleistungen keine Berücksichtigung finden.
3. Im Schweizer Krankenversicherungssystem unterscheiden sich die Patientenkollektive in Bezug auf die gewählten Franchisen und die dadurch verbundene Abrechnungsproblematik (Anteil der verursachten Kosten, die dem Krankenversicherer lediglich bekannt sind). Im Rahmen der Wirtschaftsprüfung werden zwar alle Kosten berücksichtigt, aber in den Bereichen, in denen der Arzt keinen Einfluss auf die Franchisenverteilung und Selbstzahleranteile hat, darf es nicht zu einem negativen Ergebnis bei der Wirtschaftlichkeitsprüfung der ärztlichen Arbeitsweise führen.

3. Diskussion des Wirtschaftlichkeitsprüfungsablaufs

„Statistisch auffällige“ Ärzte werden im Einzelfall unter Einbezug der von Erkenntnissen früherer Abklärungen oder weiterer zugänglicher Informationen (z. B. Zahlstellenregister, FMH-Ärzteindex) durch Experten von *santésuisse* beurteilt. Berücksichtigt werden verfügbare Daten zur Ausbildung des Arztes und seinen Schwerpunkten sowie Informationen zum Patientenkontext (Alters- und Geschlechtsverteilung).¹⁰

Im Jahr 2004 wurde die gesamtschweizerische Tarifstruktur TARMED eingeführt. Seitdem liefern Krankenversicherer Ihre Tarifdaten in den *santésuisse*-Tarifpool, der eine effiziente Unterstützung bei einer gezielten Analyse des Abrechnungsverhaltens eines Arztes ermöglicht und für die Beurteilung der Wirtschaftlichkeit Verwendung findet. Praxisbesonderheiten und Tarifmissbrauch können so auffallen und ermöglichen die Ableitung von Maßnahmen zur Verbesserung der wirtschaftlichen Behandlungsweise.¹¹

Sind erhöhte Durchschnittskosten je Patient im Beurteilungsverfahren nicht erklärbar, wird der Arzt aufgefordert, seine Kosten im Sinne anerkannter Praxisbesonderheiten in seiner Stellungnahme nachvollziehbar darzulegen.¹²

Welche Determinanten in den Stellungnahmen anerkannt sind und in welchem individuellen Auslegungsspielraum diese Anwendung finden, stellt eine der Unsicherheiten der Ärzte in der Wirtschaftlichkeitsprüfung dar.¹³

Bei einer möglichen Rückerstattungspflicht entspricht die Rückerstattungsforderung grundsätzlich dem festgestellten unwirtschaftlichen Mehraufwand (Differenz zwischen dem Fallwert des geprüften Arztes und dem Fallwert, welcher der Sicherheits- und Toleranzzone zuzüglich weiterer Praxisbesonderheiten entspricht). Auftragsrechtlich hat sich der Arzt an die allgemeinen ökonomischen Prinzipien der sparsamen Mittelverwendung zu halten, doch das Wirt-

¹⁰ Vgl. Brändli C (2009), S. 8.

¹¹ Vgl. Brändli C, (2009), S. 7.

¹² Vgl. Brändli C (2009), S. 8.

¹³ Vgl. Romanens M et al. (2009), S. 51.

schaftlichkeitsgebot des KVG ist strenger ausgelegt. Mit dem Art. 56 Abs. 1. KVG im Widerspruch stehende Leistungen müssen nicht alle direkt medizinisch unzweckmäßig sein. Ein Arzt der in solchen Fällen eine Leistung dem Patienten erfüllt, die dem Patienten nutzt, muss dennoch das Honorar an die KV zurückzahlen.¹⁴

Bei der Beurteilung ist die Beweislast dem Arzt übertragen. Bei Fallwerten von 120 bzw. 130 Indexpunkten wäre eine Rückforderung aufgrund des statistischen Vergleiches allein ein großes Willkürisiko. Die Tatsachenvermutungsgrenze liegt höher und differiert je medizinischem Fachgebiet. Ärzte mit breitem Leistungsspektrum (z. B. allgemeine bzw. allmeinen inneren Medizin) sollten Indexwerten zwischen 150 und 160 erklären und belegen können. Spezialisierte Ärzte sollten die Erklärung bereits bei Indexwerten um 140 erbringen können. Im Übergangsbereich zwischen der Toleranzgrenze und der Schwelle der Tatsachenvermutung kann es dennoch Unwirtschaftlichkeit geben. Eine Diskussion eines möglichen Trade-Offs wurde im Kapitel 4.4 vorgenommen.

Laut santésuisse zeigt die Erfahrung aus den letzten Jahrzehnten, dass viele Ärzte nach der Beobachtungszeit die Kosten pro Patient senken bzw. eine Verbesserung der Wirtschaftlichkeit erzielen. Die Vermutung der versteckten Rationierung wird detailliert in Kapitel 5 diskutiert.¹⁵

Es bleibt zu hinterfragen, ob die Kriterien, die zur Beurteilung der individuellen Stellungnahmen der Ärzte herangezogen werden, einen objektiven Charakter haben oder nicht lediglich auf rein wirtschaftlichen Kriterien beruhen. Werden Kriterien für alle Ärzte an gleichen Fakten gemessen und mit demselben Verfahren bewertet, oder gibt es eine regionale Willkür¹⁶ („schwere“ Fälle, besondere Leistungsangebote, hoher Anteil an Chronikern¹⁷)? Wie groß ist der Anteil an weichen Faktoren? Ist eine objektive und faire Beurteilung möglich?¹⁸

¹⁴ Vgl. Eugster D (2009), Punkt 102-103.

¹⁵ Vgl. Eugster D (2009), Punkt 72-77.

¹⁶ Vgl. Eugster D (2009), Punkt 93-97.

¹⁷ Vgl. Eugster D (2009), Punkt 89.

¹⁸ Vgl. Romanens M et al. (2009), S. 51.

Ferner könnten die Einigungen im Vergleich, lediglich die Folge von möglicherweise überzogenen Rückforderungen im gerichtlichen Verfahren sein, mit denen ein Vergleich nahezu erzwungen wird. Dennoch weist santésuisse im Jahr 2006 lediglich 40 Fälle auf, bei denen die Rückforderungen gerichtlich gelten gemacht worden sind und verweist somit auf eine nicht missbräuchliche Nutzung des Rechtsweges.¹⁹

4. Risikoadjustierung

Das Screeningverfahren zur Identifikation von Überarztung basiert im ersten Schritt auf der Beurteilung der Abweichungen von standardisierten Kosten in Form von Mittelwerten und tatsächlichen Kosten der behandelnden Ärzte und berücksichtigt dabei nur eine teilweise Risikoadjustierung. Eine Risikoadjustierung sorgt für eine Vergleichbarkeit von Kostenniveaus, die unterschiedlichen Risiken unterliegen. Die Güte einer Risikoadjustierung hat somit direkten Einfluss auf die Möglichkeit der Vergleichbarkeit und in vorliegendem Fall auf die Zielgenauigkeit des Screeningverfahrens. Eine mögliche unzureichende Risikoadjustierung durch alleinige Berücksichtigung von Alter, Geschlecht und Facharzttrichtung je Kanton in der ANOVA kann auf der einen Seite zu einem großen Teil als ungerechtfertigt kostenauffällig identifizierter Ärzte und somit zu ungerechtfertigten finanziellen Konsequenzen nach dem daraus evtl. folgenden Prüfungsverfahren der betreffenden Ärzte sorgen sowie zu einer Nicht-Identifikation von tatsächlichen Fällen von Überarztung. Auf der anderen Seite könnte die Sorge der behandelnden Ärzte vor der Einstufung „kostenauffällig“ im Screeningverfahren zu einer versteckten Rationierung führen (siehe Kapitel 5). Beide Aspekte sollen im Rahmen der Analyse des bisherigen Verfahrens beurteilt werden.

¹⁹ Vgl. Brändli C (2009), S. 9.

4.1. Risikostrukturausgleich: Schweiz vs. Deutschland

Der Risikostrukturausgleich (RSA) zwischen Krankenversicherern hat in der Schweiz eine noch stärkere Bedeutung als in Deutschland aufgrund der Möglichkeit der gesetzlichen Krankenversicherungen, private Zusatzversicherungen direkt zu vertreiben und der Möglichkeit der Konglomeratbildungen. Trotzdem ist die Ausgestaltung des RSA in der Schweiz eher pragmatisch mit einem finanziellen regionalen Ausgleich (nach Kantonen) über die Merkmale Geschlecht und Alter – aufgeteilt in insgesamt 30 Gruppen, sodass deutliche Anreize zur Risikoselektion bestehen.²⁰ Auch im RSA existiert die Diskussion, weitere Morbiditätsvariablen mit einzubeziehen. Alternative Modelle wurden durch die zusätzliche Berücksichtigung von Pharmaceutical Cost Groups (PCG) und/oder die Berücksichtigung eines Hospitalisierungskennzeichens im Vorjahr analysiert. PCGs stellen dabei Arzneimittelgruppen dar, die auf entsprechend vorhandene Morbiditäten schließen lassen. Sie könnten zu einer deutlichen höheren Zielgenauigkeit im RSA führen, da eine Manipulierbarkeit seitens der Krankenversicherer praktisch nicht gegeben ist.

Die Problematik der Vergleichbarkeit und damit der Risikoadjustierung stellt sich auch in Deutschland bereits mehrfach in unterschiedlichen Sparten des Gesundheitswesens. So werden aktuell im Bereich der ambulanten ärztlichen Vergütung mögliche Formen der morbiditäts- oder wenigstens regionsbezogenen Risikoadjustierung diskutiert. Die mangelnde Risikoadjustierung sorgt hier für systematische Unter- und Überdeckungen bei den ärztlichen Vergütungen.

Die Ausgestaltung des morbiditätsorientierten Risikostrukturausgleiches (Morbi-RSA) zwischen Krankenkassen zur Schaffung von Wettbewerb erfolgte ebenfalls unter dem Aspekt der Morbiditätsorientierung und löste am 01.01.2009 den bisherigen RSA mit Berücksichtigung von Alters-, Geschlechts-, Erwerbsminderungs- und Disease Management Programmvariablen ab, der in seiner Ausgestaltung zu deutlichen Wettbewerbsverzerrungen

²⁰ Vgl. Leu R E, Beck K (2006), S. 20 ff.

führte und systematische Unter- oder Überdeckungen nicht hinreichend verhindern konnte. Der RSA unter Verwendung direkter Morbiditätsinformationen erreicht in einem linearen Regressionsmodell eine deutlich höhere Erklärungskraft in Form eines höheren R^2 -Wertes²¹, während der RSA vor 2009 auf Basis eines Zellansatzes mit standardisierten Kosten je Zelle errechnet wurde. Der „alte“ RSA beinhaltete somit – genauso wie die ANOVA-Methodik im Screeningverfahren – keine weiteren Morbiditätsinformationen.

Die Ergebnisse einer ANOVA können mit den Ergebnissen einer linearen Regression verglichen werden. Der Vorteil des Regressionsansatzes besteht in der beliebigen Kombinierbarkeit von Variablen unterschiedlicher Skalierung und der anteiligen Ermittlung der Kosteneffekte je Merkmal, sodass jedem einzelnen Merkmal ein Kostengewicht zugeschrieben werden kann. Das ANOVA-Modell weist jeder Zelle als Merkmalskombination ein Kostengewicht zu. Eine zusätzliche Berücksichtigung von Diagnosen und/oder Arzneimittel/-gruppen wäre möglich – sollten Variablen anderer Skalierung (z. B. stetige Variablen) in der Analyse einbezogen werden, müsste eine Umstellung auf ein lineares Regressionsmodell erfolgen.

Bereits bestehende Analysen in prospektiven Modellen des linearen Regressionsansatzes im Morbi-RSA belegen eine deutliche Verbesserung des Modells durch die zusätzliche Berücksichtigung von Morbiditätsvariablen. In Abbildung 3: R^2 -Werte der untersuchten Modelle - prospektiv - erreicht das Alters- und Geschlechtsmodell (inklusive Erwerbsminderungsvariablen) lediglich ein R^2 von 6 %. Das Modell mit der höchsten Erklärungskraft (R^2 von 24 %) wird durch die Hinzunahme von Arzneimitteln in Form von Rx-Gruppen und stationären Diagnosen (Modell RxGroups+IPHCC) erreicht. In die Zielvariable der bereits bestehenden Analysen waren neben den Kosten im ambulanten Bereich auch Kosten im stationären Bereich und den Arzneimitteln eingeflossen und die bisherigen Modelle wurden prospektiv (Morbiditätsdaten des Vorjahres, Leistungsausgaben des Folgejahres) ausgerichtet.

²¹ Vgl. IGES Lauterbach Wasem (2004), S. 12, 13.

Abbildung 3: R²-Werte der untersuchten Modelle - prospektiv -

**R²-Werte der untersuchten Modelle – prospektiv –
(RSA-berücksichtigungsfähige Sachleistungen)**

Zeithorizont	Modell	R ² (in Prozent)
prospektiv	RSA S.Q.	6 %
	ACG	9 %
	ACG-PM	12 %
	CDPS	12 %
	HCC	15 %
	PCG+DCG	15 %
	RxGroups+IPHCC	24 %

Anm.: RSA S.-Q.: ohne Trennung der Rechtskreise, ohne Berücksichtigung von RSA-wirksamer DMP-Einschreibung, ohne Risikopool, Ausgaben für Sachleistungen ohne Zahnmedizin.

Quelle: IGES Lauterbach Wasem (2004), S. 12.

Vergleichbare Resultate auf Basis von schweizerischen Krankenkassen-Statistiken wurden bereits für die Schweiz generiert.²²

Das vorliegende Gutachten enthält neben einer qualitativen Analyse auch eine datenbasierte quantitative Analyse zur Beurteilung der Screeningverfahrens. Somit erfolgte die Analyse einer alters-, geschlechts-, regions- und fachbereichsbezogenen ANOVA und eine ANOVA mit zusätzlicher direkter Morbiditätsorientierung. Durch Gegenüberstellung von prädizierten Kosten und tatsächlichen Kosten in Form von Predictive Ratios (PR) in beiden Modellen für simulierte Patientenkohorten und deren Abweichungsanalysen können Rückschlüsse auf das bisher gültige Screeningverfahren zur Identifikation von Überarztlung in der Schweiz ermöglicht werden, wobei eine PR von 1,0 einer Punktlandung von prädizierten und tatsächlichen Kosten entsprechen würde.

4.2. Datenbasis und Methodik der Modellsimulationen

Die vorliegende Datenbasis deutscher gesetzlicher Krankenkassen besteht aus versichertenbezogenen soziodemografischen, Morbiditäts- und Leistungskostendaten, die im Rahmen des RSA zur Ermittlung der kassenspezifischen Zuweisungen genutzt werden. Sie sind ebenfalls die Basis für die Ermittlung

²² Vgl. Beck K, Trottmann M (2007), S. 336 ff.

der standardisierten Kosten durch das Bundesversicherungsamt (BVA), indem über ein lineares Regressionsmodell die risikoadjustierten Kostengewichte für 40 Alters- und Geschlechtsgruppen (AGG) sowie für 112 hierarchisierte Morbiditätsgruppen (basierend auf 80 Krankheiten) und 6 Erwerbsminderungsgruppen (Versicherte mit Erwerbsminderungsstatus) ermittelt werden.²³ Die Daten enthalten auch regionale Informationen in Form von Bundeslandkennzeichen, die somit auch eine regionale Risikoadjustierung ermöglichen würden. Insgesamt konnte für die vorliegenden Simulationen auf eine Datenbasis von etwa 300.000 Versicherten zurückgegriffen werden.

Ziel der Simulationen war die Ermittlung und Analyse von PR für simulierte Patientenkohorten in unterschiedlichen Fachrichtungen (identifiziert über die Morbiditäten der Versicherten) in einer ANOVA-Analyse nach Alter, Geschlecht und eine ANOVA-Analyse unter zusätzlicher direkter Morbiditätsorientierung durch Berücksichtigung der 112 Morbiditätsgruppen des BVA. Diese 112 Morbiditätsgruppen berücksichtigen neben ICD10-Diagnosen auch Arzneimittelinformationen in Form von verordneten Pharmazentralnummern (PZN) zur Validierung der Diagnosen bzw. zur Identifikation klinisch relevanter Fälle, sodass durch die zusätzliche Berücksichtigung der Morbiditätsgruppen sowohl der Bereich der Arzneimittel als auch der Bereich der Morbiditäten erfasst wird. Wie im Schweizer Screeningverfahren würden anschließend die tatsächlichen Kosten der Patientenkohorten mit den ermittelten PR als primäres Kriterium der Überarztung geprüft werden.

Für die Analysen wurden Modelle mit und ohne Konstante sowie unter Berücksichtigung von Alter, Geschlecht (eingeteilt in 40 AGGn) und Fachrichtung (Filterung nach Krankheitsbildern der entsprechenden Fachrichtungen) sowie in einem weiteren Schritt mit zusätzlichen Morbiditätsvariablen berechnet. Bei den Leistungsausgaben wurden nur die Kosten im ambulanten Bereich und im Bereich der Arzneimittel in der Zielvariable berücksichtigt, da auch das Screeningverfahren auf Kostenauffälligkeiten in diesen Bereichen abzielt. Da

²³ Vgl. BVA (2010)

lediglich ein Gesamteffekt der direkten Morbiditätsorientierung von Interesse ist und nicht die jeweilige Stärke der einzelnen Kostengewichte und deren Signifikanz und Bedeutung bereits im BVA-Verfahren belegt wurde, blieben für diese Modellsimulation die Signifikanzniveaus der einzelnen Gruppen unberücksichtigt. Es wurden nur Versicherte berücksichtigt, die volljährige Versichertenzeiten und somit vollständige Morbiditätsinformationen aufwiesen. Als Regionsvariable hätte alternativ zu den Schweizer Kantonen das Bundeslandkennzeichen als Filtervariable genutzt werden, jedoch wären die Datenbasen der einzelnen Bundesländer zu klein gewesen, um valide Aussagen treffen zu können. Im Vorfeld durchgeführte Analysen zeigten, dass die Regionsvariable in Form des Bundeslandkennzeichens auch keinen signifikanten Einfluss auf die Kostenprognose der Versicherten hatte. Zunächst waren Analysen für drei unterschiedliche Fachrichtungen (Diabetologie, Kardiologie und Orthopädie) vorgesehen – letztlich konnten nur Analysen im Bereich Kardiologie erstellt werden, da nur in diesem Bereich ausreichende Fallzahlen vorhanden waren. Für die Abbildung der Fachrichtung Kardiologie wurde der Versichertenbestand nach entsprechenden Krankheitsbildern in Form von HMGn gefiltert, so dass nur Versicherte mit kardiologisch relevanten Erkrankungen aus Sicht des Morbi-RSA in die Analyse einfließen konnten. Folgende HMGn waren für die Identifikation eines Versichertenbestandes im Fachbereich Kardiologie relevant:

Kardiologie (Identifikation auf Basis der HMGn):

Tabelle 1: Identifikation der Fachrichtung Kardiologie über HMGn

Gruppe/ Position	Bezeichnung
HMG080	Herzinsuffizienz
HMG081	Akuter Myokardinfarkt
HMG082	Instabile Angina pectoris und andere akute ischämische Herzerkrankungen
HMG083	Angina pectoris / Z. n. altem Myokardinfarkt
HMG084	Koronare Herzkrankheit / andere chronisch-ischämische Erkrankungen des Herzens
HMG086	Erworbene Erkrankungen der Herzklappen und rheumatische Herzerkrankungen
HMG087	Schwere angeborene Herzfehler
HMG088	Andere angeborene Herzfehler
HMG089	Hypertensive Herz- und Nierenerkrankung oder Enzephalopathie

HMG090	Hypertensive Herzerkrankung
HMG091	Hypertonie
HMG092	Näher bezeichnete Arrhythmien

Quelle: BVA (2010)

Auf Basis der auf diese Weise identifizierten Versicherten wurden anschließend die bereits beschriebenen ANOVA-Simulationen im Status Quo und eine um Morbiditätsvariablen erweiterte ANOVA-Simulation durchgeführt. In einem weiteren Schritt wurden per Zufallsstichprobe 10 Patientenkohorten mit je 1.000 Versicherten erstellt. Für diese Patientenkohorten wurden anschließend die PR auf Basis der ANOVA-Ergebnisse beider Modelle ermittelt. Die zu erwartenden Kosten der Patientenkohorten wurden dabei den tatsächlichen Kosten gegenübergestellt. Ein Vergleich der PR beider Modelle (ANOVA im Status Quo und erweitertes ANOVA-Modell) ermöglichte anschließend eine Einschätzung der Screeningverfahrens hinsichtlich Zielgenauigkeit und Fehlerwahrscheinlichkeit und die Beurteilung einer möglichen Modelloptimierung.

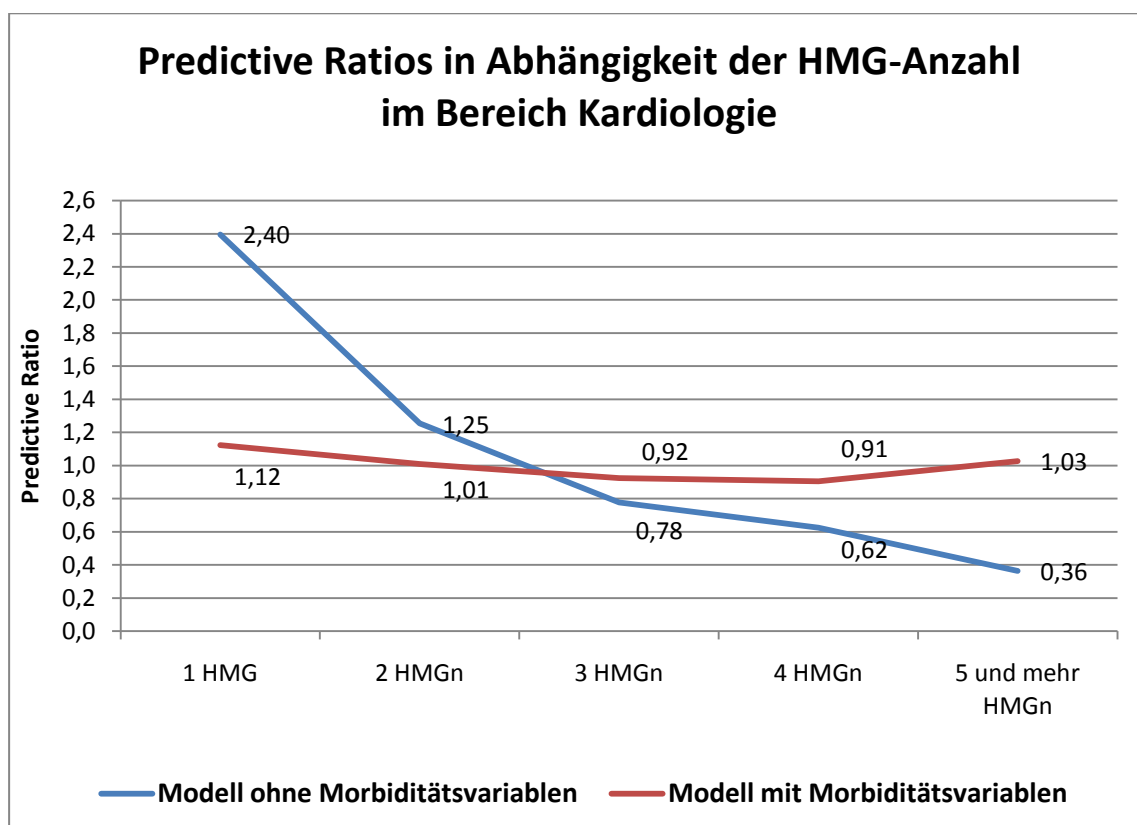
In einer weiteren Analyse wurde der Zusammenhang von Patientenkohortengröße und PR ermittelt. Damit konnte die variierende Prognosegüte für Ärzte mit unterschiedlichen Patientenbeständen simuliert werden. Für die Simulationen wurden Kohorten unterschiedlicher Größen – beginnend bei 500 Patienten bis hin zu Kohorten von 100.000 Patienten – über Stichproben mit Zurücklegen gebildet. Im Anschluss wurde der Korrelationskoeffizient als Zusammenhangsmaß zwischen Kohortengrößen und den absoluten Residuen von Kostenschätzung zu tatsächlichen Kosten ermittelt.

4.3. Ergebnisse der Modellanalysen

Die beispielhafte Analyse für den Fachbereich Kardiologie ergab einen direkten Zusammenhang von höherer Morbiditätslast in Form von HMG-Zahlen und sinkenden PR-Werten in den entsprechenden Patientenkohorten im ANOVA-Modell des Status Quo. Dafür wurden die Patienten in HMG-Clustern zusammengefasst und für jede HMG-Klasse wurde die durchschnittliche PR ermittelt. Die Gegenüberstellung von prädizierten Kosten und tatsächlichen Kosten in

Form von PR beider Modelle und deren Abweichungsanalysen können Rückschlüsse auf das bisher gültige Screeningverfahren zur Identifikation von Überarztung in der Schweiz ermöglichen (siehe Kapitel 4.1), wobei eine PR von 1,0 einer Punktlandung von prädizierten und tatsächlichen Kosten entsprechen würde. In der Patientengruppe mit nur 1 relevanten HMG liegt die PR im Modell des Status Quo bei 2,4 und sorgt somit für eine deutliche Überschätzung der Kosten – das um direkte Morbiditätsinformationen erweiterte Modell erreicht eine PR von 1,12. In den Patientengruppen mit höherer Morbiditätslast pendelt die PR im erweiterten Modell um den Wert von 1, während das Status Quo-Modell in der Tendenz mit zunehmender Morbiditätslast der Patienten die Kosten immer stärker unterschätzt.

Abbildung 4: Predictive Ratios in Abhängigkeit der HMG-Anzahl im Bereich Kardiologie



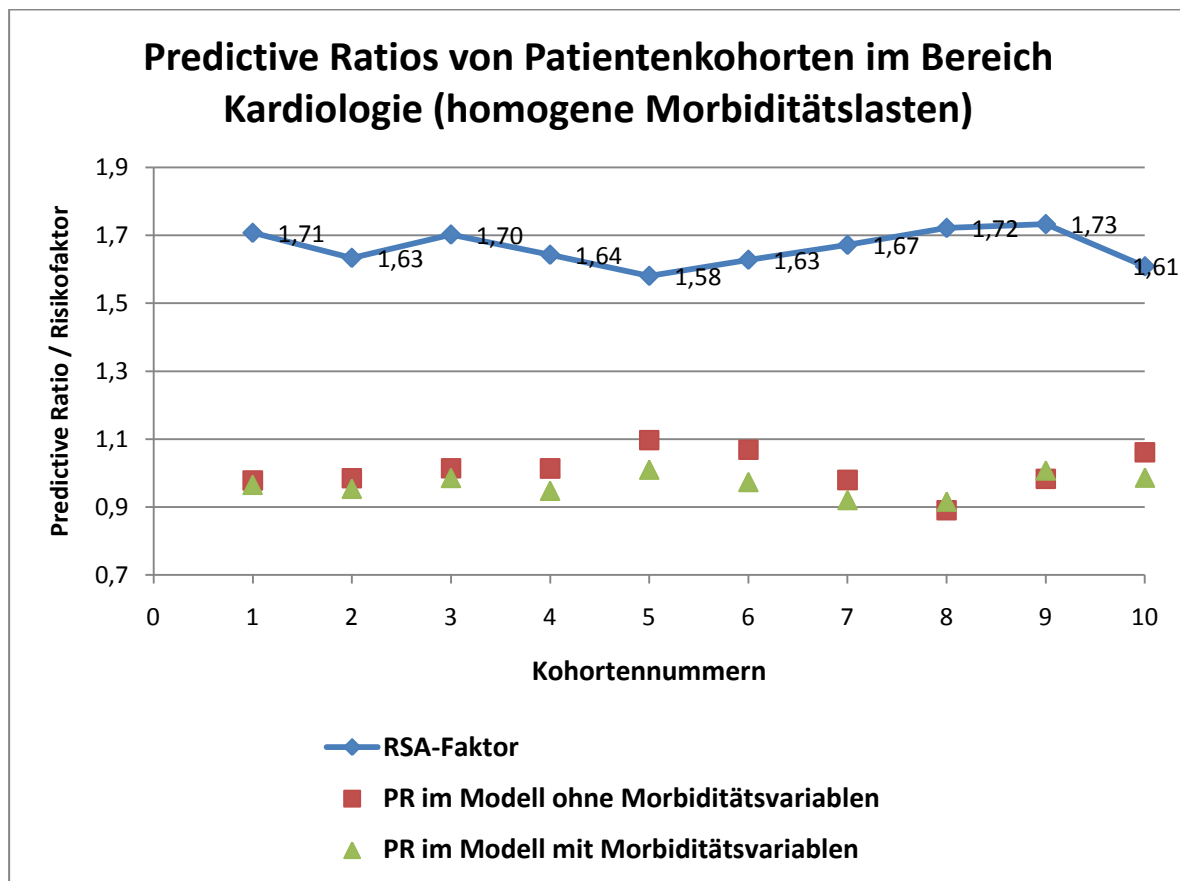
Quelle: Eigene Berechnungen

Bei Patientengruppen mit 4 und mehr HMGn sinkt die PR unter den Wert von 0,7 – in diesen Fällen würde das im Screeningverfahren verwendete Kriterium der 130 %-Grenze als primäres Indiz für eine Überarztung überschritten werden. Ärzte, deren Patienten durchschnittlich mehr als 4 HMGn aufweisen,

würden in der vorliegenden Simulation als „kostenauffällig“ identifiziert werden. Die PR verbessern sich im erweiterten Modell deutlich durch Hinzunahme von direkten Morbiditätsvariablen in Form von Diagnosen und Arzneimittelinformationen.

Die Analyse der PR der 10 per Zufallsstichprobe generierten Patientenkohorten ist in Abbildung 5: Predictive Ratios von Patientenkohorten im Bereich Kardiologie dargestellt. Für jede der 10 Kohorten ist die PR des erweiterten Modells mit dem des Status Quo-Modells verglichen worden.

Abbildung 5: Predictive Ratios von Patientenkohorten im Bereich Kardiologie (homogene Morbiditätslasten)



Quelle: Eigene Berechnungen

Es zeigt sich, dass bei simulierten Patientenkohorten mit ähnlicher Morbiditätsstruktur und Kostenstruktur keine signifikanten Unterschiede in der Prädiktionsgüte im direkten Vergleich der beiden Modelle sichtbar sind. Erkennbar ist jedoch, dass mit höherem RSA-Faktor der Kohorte die PR des Status Quo-Modells tendenziell unter 1 liegt und vice versa. Die Kohorten sind in Tabelle

2: Risikofaktoren der 10 Patientenkohorten hinsichtlich ihrer Morbiditätslast aus Sicht des deutschen Morbi-RSA in Form des Risikofaktors gegenübergestellt worden, wobei ein Risikofaktor von 1 den durchschnittlichen Kosten eines Versicherten in der deutschen Gesetzlichen Krankenversicherung entspricht. Würden alle behandelnden Ärzte relativ durchschnittlich in der Morbiditätsstruktur der Patienten sein, wäre das augenblickliche Modell ausreichend, um eine Überarztung zu identifizieren.

Tabelle 2: Risikofaktoren der 10 Patientenkohorten

Kohortenummer	Risikofaktor
Kohorte 1	1,71
Kohorte 2	1,63
Kohorte 3	1,70
Kohorte 4	1,64
Kohorte 5	1,58
Kohorte 6	1,63
Kohorte 7	1,67
Kohorte 8	1,72
Kohorte 9	1,73
Kohorte 10	1,61

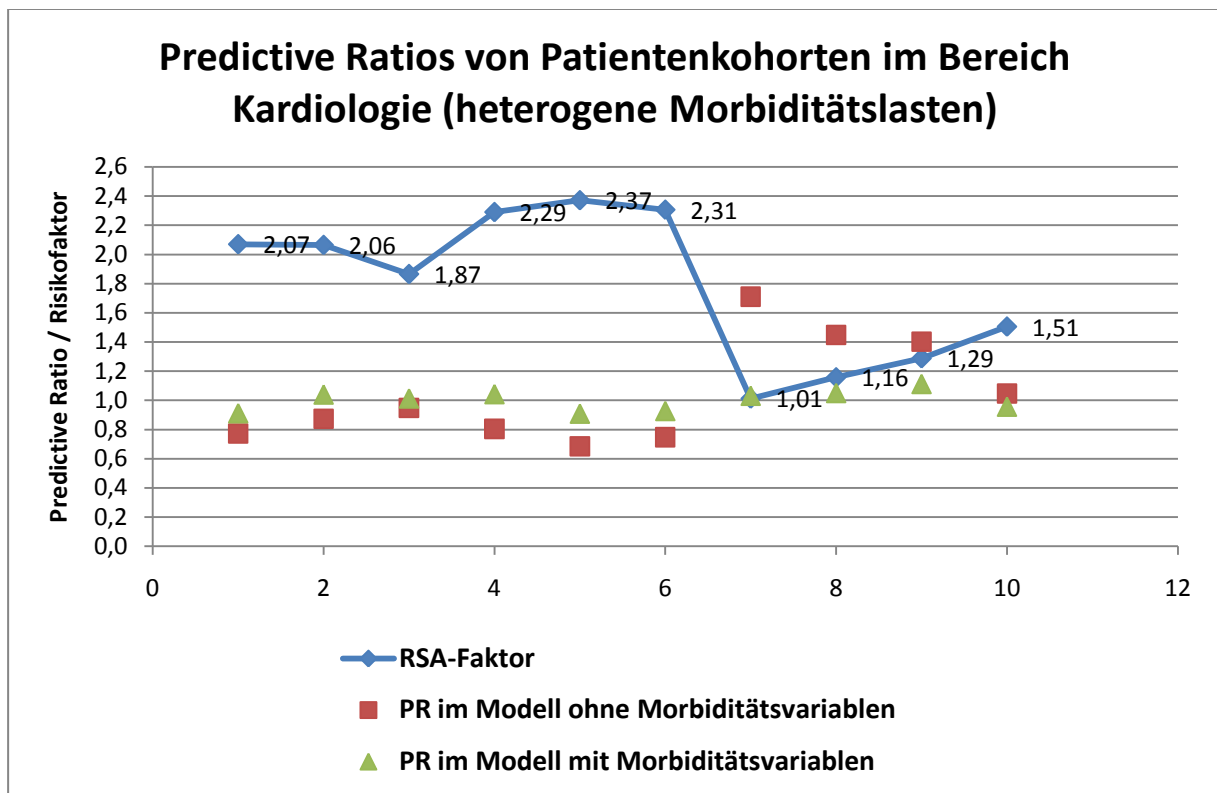
Quelle: Eigene Berechnungen

Predictive Ratios von Kohorten mit variierenden Morbiditätslasten

In einer weiteren Analyse wurden die Patientenstichproben dahingehend geändert, dass im Vorfeld signifikant unterschiedliche Risikofaktoren der Patientenkohorten entstanden.

Die Kritik des augenblicklichen Modells begründet sich in der mangelnden Prädiktionsgüte bei Ärzten mit vom Durchschnitt abweichender Morbiditätsstruktur – deshalb wurden Patientenklientele mit unter- und überdurchschnittlicher Morbiditätslast simuliert. In Abbildung 6: Predictive Ratios von Patientenkohorten im Bereich Kardiologie (heterogene Morbiditätslasten) sind die PR der erneut simulierten Patientenkohorten mit heterogener Morbiditätslast dargestellt.

Abbildung 6: Predictive Ratios von Patientenkohorten im Bereich Kardiologie (heterogene Morbiditätslasten)



Quelle: Eigene Berechnungen

Die Schwankungen der PR im Status Quo-Modell fallen in diesen Simulationen deutlich stärker aus und bekräftigen die Aussage, dass tendenziell für Patientenkohorten mit unterdurchschnittlicher Morbiditätslast eine Überschätzung der Kosten erfolgt und vice versa. Daraus folgt im Umkehrschluss, dass Ärzte mit tendenziell weniger morbidem Klientel, aber auffällig hohen Kosten weniger leicht identifiziert werden können. Auf der anderen Seite würden Ärzte mit tendenziell morbideren Patienten relativ schnell das 130 %-Kriterium überschreiten können.

Zusammenhang von Kohortengrößen und Prognosegüte

Neben dem Einfluss der Morbiditätslast der Patientenkohorte auf die Prädiktionsgüte des Modells hat auch die Größe der Patientenkohorte eine Bedeutung für die Kostenschätzungen.

Die Ergebnisse in Tabelle 3: Korrelationsstärke: Kohortengröße – Residuum der Kostenschätzungen zeigen mit einem Korrelationskoeffizienten von etwa -0,5 auf Basis beider Modelle einen deutlichen Zusammenhang von Größe der Patientenkohorten und Abweichung der Kostenschätzung von den tatsächlichen Kosten in Form der Residuen auf. Mit zunehmender Kohortengröße sinken die Residuen und es erfolgen genauere Schätzungen. Damit würden Ärzte im Screeningverfahren neben ihrer praxisindividuellen Morbiditätsstruktur auch in Abhängigkeit der Größe ihres Patientenbestandes unterschiedlich gut erfasst und beurteilt werden. Die Ergebnisse sind mit einem Signifikanzwert von 0,02 bzw. 0,03 für beide Modelle signifikant. Die Analysen wurden auf Basis von 36 Patientenkohorten mit steigenden Patientenzahlen von 500 bis 100.000 Patienten für beide Modelle (ANOVA im Status Quo und ANOVA mit Morbiditätserweiterung) durchgeführt.

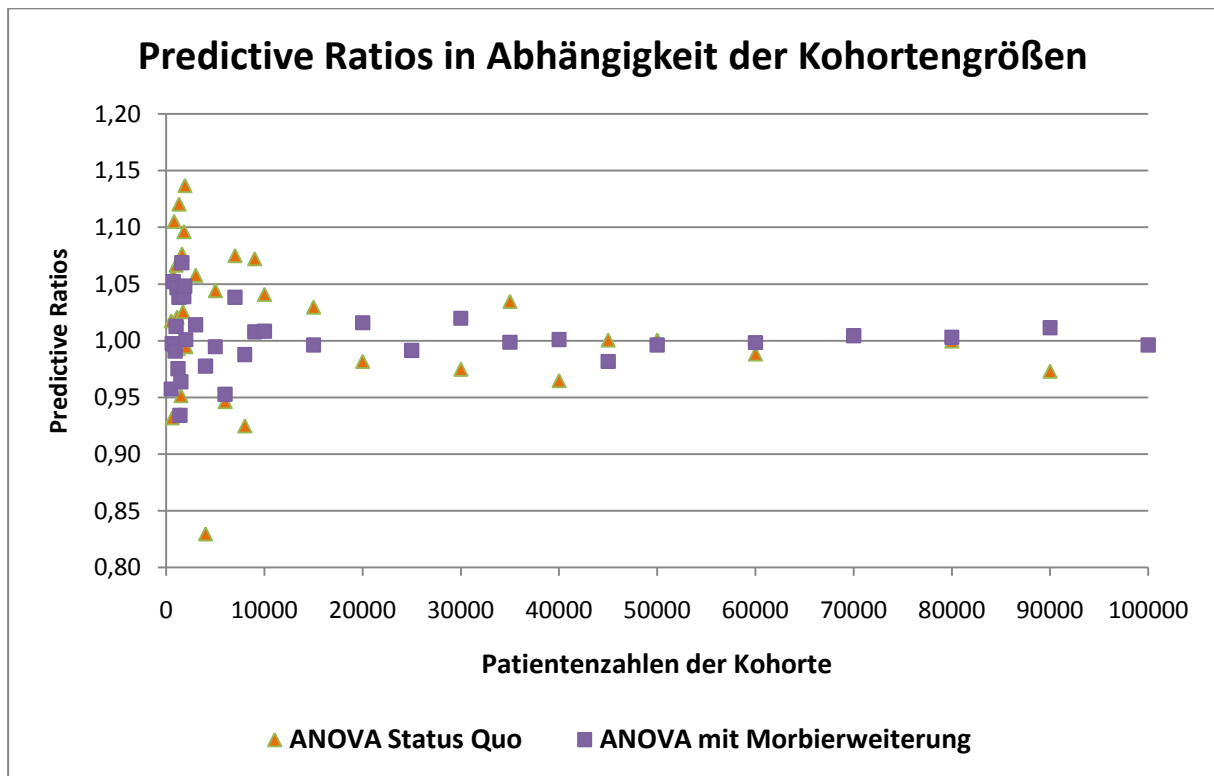
Tabelle 3: Korrelationsstärke: Kohortengröße – Residuum der Kostenschätzungen

	Korrelationsstärke: Kohortengröße - Prädiktionsresiduum	Signifikanz
ANOVA Status Quo	-0,50	,002
ANOVA mit Morbibezug	-0,47	,003

Quelle: Eigene Berechnungen

Die Streuung der einzelnen Abweichungen von Kostenschätzungen und tatsächlichen Kosten je Kohorte ist in Abbildung 7: Predictive Ratios in Abhängigkeit der Kohortengrößen dargestellt.

Abbildung 7: Predictive Ratios in Abhängigkeit der Kohortengrößen



Quelle: Eigene Berechnungen

In beiden Modellen erreichen die Kostenschätzungen bei etwa 50.000 Patienten eine PR von 1,0 und unterliegen bei noch weiter steigenden Kohortengrößen nur noch minimalen Schwankungen.

4.4. Trade-off des 130 %-Kriteriums im Screeningverfahren

Die willkürliche Festlegung des Schwellenwertes von 130 % der Durchschnittskosten als primäres Kriterium für eine vorliegende Überarztung lässt sich final nur mit einer ausreichend großen realen Datenbasis beurteilen. Die vorliegenden Analysen auf Daten deutscher gesetzlicher Krankenversicherungen können lediglich die grundsätzlichen Verbesserungspotentiale durch eine Verbreiterung des Variablen-Eingangsraumes aufzeigen und die Notwendigkeit einer direkten Morbiditätsorientierung belegen. Die Empfehlung einer optimalen Schwellenwertgrenze erscheint aus folgenden weiteren Gründen nicht sinnvoll:

1. Zunächst werden sich facharztgruppenspezifisch unterschiedliche Varianzen in den Leistungsausgaben ergeben – somit können je nach Fachrichtung unterschiedliche Schwellenwerte sinnvoll sein. Fachrichtungen mit typischerweise eher größeren Schwankungen in den Behandlungskosten (wegen evtl. höherer Komplikationsraten o. ä.) würden eine höhere Obergrenze erforderlich machen und vice versa.
2. Der Schwellenwert wird in jeder Ausprägung dafür sorgen, dass Ärzte fälschlich als „kostenauffällig“ identifiziert werden bzw. Ärzte mit tatsächlichen Kostenauffälligkeiten unentdeckt bleiben. Eine Anhebung des bisherigen Schwellenwertes von 130 % würde somit zu einer Reduktion fälschlich identifizierter Ärzte und zu einer Reduktion der identifizierten, tatsächlich kostenauffälligen Ärzten führen und vice versa. Dieser Trade-off ist zu bewerten und der Schwellenwert entsprechend der eigenen Präferenz zu setzen. Erst eine Modellerweiterung könnte dafür sorgen, dass beide Fehlerquellen reduziert werden und somit der Trade-off von falsch auffällig und falsch unauffällig identifizierten Ärzten an Bedeutung verliert.

5. Gefahr der versteckten Rationierung durch die ANOVA-Methode

Wie die meisten westlichen Industriegesellschaften steht auch das Schweizer Gesundheitssystem vor der Herausforderung, rasch wachsende Nachfrage nach Gesundheitsleistungen und innovative Entwicklungen beim medizinischen Angebot mit einer nur begrenzten Zahlungsbereitschaft im Rahmen des durch Zwangsbeiträge finanzierten Obligatoriums des Krankenkassensystems in Übereinklang zu bringen.

Mit den Begrifflichkeiten der Rationalisierung, der expliziten Rationierung und der impliziten bzw. versteckten Rationierung steht ein geeigneter Begriffsapparat zur Verfügung, die möglichen gesellschaftlichen Strategien zum Umgang mit diesen Herausforderungen zu charakterisieren:²⁴

- Unter Rationalisierung soll in diesem Kontext die Verringerung des Einsatzes von Ressourcen (also: geringere Kosten) ohne Einbußen bei der Zielerreichung (also: den Outcomes) verstanden werden. In diesem Begriffsverständnis liegt Rationalisierung etwa vor, wenn auf Doppeluntersuchungen, die keine zusätzlichen Informationen generiert haben, verzichtet wird. Auch eine Verlagerung einer bislang im Spital erbrachten Maßnahme im ambulanten Bereich bei gleicher Qualität und gleichen Nebenwirkungen aber geringeren Kosten wäre Rationalisierung. Schließlich würde auch das Vorenthalten eines teureren aber gleich nützlichen neuen Testes oder Arzneimittels im Vergleich zur Standardtherapie als Rationalisierung angesehen werden können.
- Demgegenüber wollen wir von Rationierung sprechen, wenn Leistungen vorenthalten und insoweit Ressourcen eingespart werden, obwohl diese Leistungen nützlich sind. Hierbei hat sich eine Unterscheidung zwischen expliziter und impliziter bzw. versteckter Rationierung als sinnvoll erwiesen:

²⁴ Vgl. dazu und im Folgenden: Strech D et al. (2008) S. 94-109.

- Von expliziter Rationierung soll gesprochen werden, wenn Zuteilungsentscheidungen nach von einer durch die Rechtsordnung dazu legitimierten Instanz (z. B. können dies je nach gesetzlichen Zuständigkeiten Parlament, Gesundheitsministerium, Arzneimittelkommission, Krankenkassenverband sein) ausdrücklich festgelegten, allgemein verbindlichen Kriterien „oberhalb“ der individuellen Beziehung zwischen Leistungserbringer (Arzt, Krankenhaus) und Patient erfolgen und infolge dessen bestimmte Leistungen entweder vom Leistungskatalog des Pflichtversicherungssystems entweder komplett ausgeschlossen werden oder etwa Indikationen für ihren Einsatz eingeschränkt werden oder nur eine teilweise Kostenübernahme vorgesehen wird. Wird etwa eine neue Operationstechnik nicht in das Verzeichnis der von den Ärzten abrechenbaren Leistungen aufgenommen und kann im Rahmen der Pflichtversicherung deswegen nicht zulasten der Krankenversicherung abgerechnet werden, obwohl entsprechende Studien unumstritten einen Zusatznutzen gegenüber dem bisherigen Standardverfahren gezeigt haben, liegt eine Leistungsbeschränkung im Sinne der expliziten Rationierung vor. Kriterien für solche expliziten Rationierungen können etwa geringe Dringlichkeit, Eigenverantwortlichkeit, geringer Nutzen oder geringe Kosteneffektivität der ganz oder teilweise ausgegrenzten Leistungen sein. Von den nützlichen Maßnahmen wird also ein Teil nicht finanziert, um das beschriebene Spannungsverhältnis zwischen Nachfrage, Angebot und Finanzierbarkeit aufzulösen.
- Von impliziter bzw. versteckter Rationierung wird demgegenüber der Leistungsanspruch der Versicherten (also etwa: der Leistungskatalog, das Verzeichnis erstattungsfähiger Medikamente oder Behandlungen) nicht angetastet. Vielmehr werden finanzielle Rahmenbedingungen und Anreizsysteme bei den Leistungserbringern (Krankenhaus, Arzt) installiert, die dazu führen, dass diese

einem Ungleichgewicht zwischen dem, was nützlich ist, und den zur Verfügung stehenden finanziellen Ressourcen erfahren. Das „klassische Beispiel“ ist ein fixes Budget, das einem Krankenhaus oder Arzt zur Verfügung steht, welches so bemessen ist, dass nicht alle nützlichen Leistungen davon erbracht werden können. Der Arzt bzw. das Krankenhaus muss in dieser Situation auf der individuellen Arzt-Patienten-Ebene durch Vorenthalten von Leistungen das Leistungsvolumen auf das Niveau des verfügbaren Finanzvolumens führen. Diese Rationierung ist implizit und versteckt, weil es keine öffentlichen Kriterien und Entscheidungen gibt, sondern im Gegenteil die öffentliche Erwartung im Gegenteil darauf abzielt, dass gerade nicht rationiert wird, sondern alle Leistungen, die im jeweiligen Fall einen Zusatznutzen stiften können, auch erbracht werden.

Die implizite bzw. versteckte Rationierung gilt aus mehreren Gründen als ethisch problematisch: Sie ist intransparent, sie führt zu Ungleichheit, Willkür der sie umsetzenden Ärzte kann nicht ausgeschlossen werden.

Wirtschaftlichkeitsprüfungen, die vorsehen, dass das Behandlungsgeschehen eines Arzt, der eine einen statistischen Durchschnitt seiner Fachgruppe um einen bestimmten Anteil übersteigende Leistungsmenge erbracht hat, überprüft wird mit dem möglichen Ergebnis einer Rückzahlung von Honoraren an die Krankenversicherung stellen eine Form der impliziten bzw. versteckten Rationierung dar. Denn sie führen dazu, dass die Ärzte sich in einer Abwägungssituation befinden, ob sie Leistungen, die sie ärztlich für sinnvoll halten, erbringen und riskieren, deswegen möglicherweise eine Honorarrückzahlung leisten zu müssen, oder ob sie die Leistung, obwohl für sinnvoll gehalten, mit Blick auf die möglichen finanziellen Konsequenzen unterlassen.

Dieser Effekt ist grundsätzlich der Wirtschaftlichkeitsprüfung der beschriebenen Art immanent und der Regulator, der die Regeln für die Wirtschaftlichkeitsprüfung festlegt, kann ihr grundsätzlich nicht entkommen. Umso wichtiger

ist es aber, dass die Regeln für die Wirtschaftlichkeitsprüfungen so festgesetzt werden, dass ein sachgerechter Umgang mit dem Rationierungsproblem ermöglicht wird. Dazu gehört insbesondere, dass der Benchmark, an dem gemessen wird, ob „unwirtschaftlich“ gehandelt wird, sachgerecht ermittelt wird. Wie der empirische Teil dieser Untersuchung gezeigt hat, ist genau dieses nicht der Fall: Die Adjustierung des Durchschnittes der Menge ärztlicher Leistungen, an denen der einzelne Arzt gemessen wird, nur mit den Kriterien Alter und Geschlecht im Rahmen der ANOVA-Methode ist genau nicht sachgerecht. Da die Patientenkollektive sehr unterschiedlich sind, wird es Ärzte geben, die über ein Patientenkontingent verfügen, das auch nach Adjustierung bezüglich Alter und Geschlecht erheblich morbiditer ist, daher also einen erheblich höheren Behandlungsbedarf hat, als der Durchschnitt der Patienten. Diese Ärzte kommen also bereits sehr rasch in die Gefahr, bei Erbringung von ihnen als notwendig angesehener Leistungen das Auffälligkeitskriterium zu überschreiten. Diese Ärzte sind also sehr rasch einer potenziellen Gefahr einer Forderung nach Honorarrückzahlung ausgesetzt. Folglich werden diese Ärzte in besonderer Weise die Abwägung zu treffen haben, sich dem Risiko einer solchen Rückzahlungsforderung auszusetzen oder aber ärztlich von ihnen als notwendig angesehene Leistungen nicht zu erbringen.

Das System eines nur über Alter und Geschlecht für die unterschiedlichen Risikostrukturen der Patientenkollektive adjustierten Benchmarks im Rahmen der Wirtschaftlichkeitsprüfung ärztlicher Leistungserbringung führt daher zu einer in diesem Umfang vermeidbaren impliziten bzw. versteckten Rationierung. Zu fordern ist daher, dass die Rationierungseffekte möglichst begrenzt werden – es geht ja darum, Unwirtschaftliches, nicht aber Notwendiges zu begrenzen. Aus diesem Grunde muss die Adjustierung für die unterschiedlichen Patientenkollektive die Morbidität der Patienten berücksichtigen. Wie die empirischen Ausführungen gezeigt haben, stehend entsprechende, geeignete Methoden zur Verfügung.

6. Limitationen

Die Analysen dieses Gutachtens unterliegen einigen Limitationen, die die Aussagekraft der erzielten Ergebnisse relativieren können. Den Gutachtern lagen keine Echtdateien aus der Schweiz zur Verfügung. Aus diesem Grund mussten für die Simulationen Daten deutscher gesetzlicher Krankenkassen verwendet werden, die sich in ihrer Struktur von denen in der Schweiz unterscheiden.

Die Daten waren keine facharztbezogenen Daten niedergelassener Ärzte – die Daten enthielten auch keine Facharztvariable, sodass die Facharztkohorten anhand von Krankheitsbildern der Versicherten gebildet werden mussten. Diese Krankheitsbilder basieren auf den Festlegungen des BVA in Deutschland und können sich im Einzelnen von denen in der Schweiz unterscheiden.

Die Kostenstrukturen und Kostenniveaus im deutschen Gesundheitswesen bei den ambulanten Ärzten können sich von denen in der Schweiz unterscheiden, sodass die Simulationsergebnisse lediglich Tendenzen abbilden können. Die im ambulanten und Arzneimittelbereich auftretenden Kosten basieren auf den Definitionen der sogenannten Satzarten, die vom BVA für die deutschen Krankenkassen vorgegeben wurden. Welche einzelnen Konten in die Leistungsausgaben in welcher Form eingehen und welche Einzelkosten nicht oder nur teilweise berücksichtigt werden, kann sich von den im Schweizer Screeningverfahren genutzten Kostendaten unterscheiden.

Die Leistungsausgaben der Versicherten sind summierte Jahreswerte – sowohl über alle behandelnden Ärzte und Facharztgruppen hinweg - und keine arztbezogenen Kosten. Somit ist sowohl eine Aufspaltung in einzelne Facharztgruppen als auch eine Verteilung auf einzelne Ärzte nicht möglich.

In den Analysen wurden nur Versicherte berücksichtigt, die bei vorliegender facharztrelevanter Krankheit Leistungsausgaben verursacht haben – inwieweit diese Leistungsausgaben ursächlich durch die jeweilige Erkrankung verursacht wurde oder aber andere Morbiditäten der Auslöser waren, lässt sich anhand der vorliegenden Daten nicht beurteilen.

Es wurden nur Versicherte berücksichtigt, die vollständige Versichertenzeiten aufwiesen. Demnach wurden keine Verstorbenen berücksichtigt - dies war für eine Tendenzanalyse aber auch nicht notwendig.

Die Fallzahlen der Simulationen im Fachbereich Kardiologie liegen bei nur 22.000 Versicherten. Aus diesem Bestand wurden Stichproben zur Simulation einzelner Patientenkohorten gezogen. Einige Morbiditätsgruppen in der ANOVA wiesen keine ausreichende Signifikanzniveaus auf – aufgrund der bereits im BVA-Verfahren belegte Signifikanz der jeweiligen Gruppen und der Reduktion der Ergebnisinterpretation auf Tendenzanalysen, blieben die Signifikanzniveaus in den Analysen unberücksichtigt.

Die Simulationen heterogener Morbiditätslasten in den Patientenkohorten erfolgte anhand der RSA-Faktoren der Versicherten. Der RSA-Faktor ergibt sich wiederum aus den vom BVA als RSA-relevant definierten Krankheiten für Deutschland. Für die Schweiz kann die Definition von Krankheitsbildern und deren Relevanz für den RSA abweichend sein.

Der Ablauf der Verfahrensanhörung und der Grad der Objektivität bei der Beurteilung der Kostenauffälligkeit nach dem Screeningverfahren ist den Gutachtern nicht im Einzelnen bekannt, sodass über die Qualität des Anhörungsverfahrens nur anhand der vorliegenden Quellen geurteilt werden kann.

7. Diskussion

Die Diskussion um mögliche Risikoadjustierung ist keine neu aufkommende Thematik, sondern wurde sowohl in Deutschland als auch in der Schweiz bereits in mehreren Sparten des Gesundheitswesens behandelt. In der Schweiz existieren im Rahmen des RSA bereits Diskussionen um eine präzisere Methodik, um Fehlanreize und Prädiktionsungenauigkeiten zu vermeiden.

Das augenblickliche Verfahren einer demografischen ANOVA zur primären Identifikation von Überarztung ist neben bereits existierenden Publikationen auch in vorliegendem Gutachten als unzureichend einzuschätzen. Beck K, Trottmann M (2007) zeigten bereits auf Basis von schweizerischen Krankenkassen-Statistiken, dass eine Verbesserung der Erklärungskraft (R^2) durch die Berücksichtigung von Morbiditätsvariablen erzielt werden kann. Das Gutachten Schwenkglens M. (2010) weist ebenfalls auf unzureichend berücksichtigte Variablen hin, die aufgrund der Literaturlage und den Ergebnissen des statischen Berichts sowie weiterführenden Überlegungen zielführend wären. Das Ergebnis führt analog zu dem vorliegenden Gutachten zu dem Schluss, dass bei der Wirtschaftlichkeitsbeurteilung der schweizerischen Ärzte die Morbiditätsstruktur der Patientenkollektive sowie alle veranlassten Kosten berücksichtigt werden sollten. Vorhandene Instrumente der santésuisse können dieses nicht ausreichend leisten.

Eine Risikoadjustierung bei ambulanten Ärzten lediglich auf Basis von Alter und Geschlecht sorgt dafür, dass der behandelnde Arzt in einem nicht vertretbaren Umfang das Morbiditätsrisiko trägt. Selbst innerhalb einer Facharztgruppe unterscheidet sich die Morbiditätslast zwischen Arztpraxen erheblich und somit würde das augenblickliche Verfahren zu sachlich nicht zu rechtfertigenden Umverteilungen führen.²⁵ Die Ärzte haben nur geringen Einfluss auf die Morbiditätslast ihrer Patienten durch Auswahl ihrer Behandlungsfälle. Somit ist es zwangsläufig so, dass einzelne Leistungserbringer entweder zu kos-

²⁵ Vgl. Wasem J (2006), S. 291.

tenintensive oder aber zu kostengünstige Behandlungsfälle aufweisen²⁶ und somit in den Zwang der versteckten Rationierung geraten.

Weiterführende Informationen zur wissenschaftlichen Analyse der Rationierung in der Schweiz bieten bereits die Studien von Hurst S A et al. (2006) und Hurst S A et al. (2007).

Die Relativierung der Bedeutung des Screeningverfahrens mit der Betonung auf „primäres Kriterium“ eines Prüfprozesses erscheint mit Blick auf die Gefahr der versteckten Rationierung nicht sinnvoll. Zwar kann mit dem nachgelagerten Anhörungsverfahren und der fallweisen Prüfung die Quote der primär fälschlich als kostenauffällig identifizierten Ärzte deutlich reduziert werden – die Problematik der versteckten Rationierung ist nachgelagert jedoch nicht lösbar. Zusätzlich existieren auf der einen Seite trotz einer nachgelagerten Prüfung weiterhin nicht als kostenauffällig identifizierte Ärzte aufgrund ihrer unterdurchschnittlichen Morbiditätslast ihrer Patienten. Auf der anderen Seite ist eine vollständige Entlastung der im primären Schritt fälschlicherweise als kostenauffällig eingestuften Ärzte durch eine nachgelagerte Prüfung evtl. nicht möglich. Für eine abschließende Beurteilung müssten den Gutachtern Details der objektiven Prüfungskriterien vorliegen.

Die Festlegung eines Schwellenwertes im Screeningverfahren müsste aus Sicht der Gutachter getrennt nach „Fachrichtung“ erfolgen, da die unterschiedlichen fachbereichsspezifischen Kostenvarianzen dafür sorgen können, dass unterschiedliche Schwellenwerte notwendig wären. Deshalb ist eine Beurteilung des 130 %-Schwellenwertes auf Validität nicht pauschal möglich. Grundsätzlich erscheint diese Festlegung – insbesondere mit Blick auf die mangelnde Risikoadjustierung – willkürlich gesetzt zu sein. Eine Empfehlung für eine Schwellenwertanpassung kann nicht erfolgen, sondern müsste mit Blick auf den Trade-off von fälschlich als kostenauffällig identifizierten und fälschlich als nicht kostenauffällig identifizierten Ärzten abgewogen werden. Eine Reduktion beider unerwünschter Effekte könnte lediglich durch ein verfeinertes

²⁶ Vgl. Schröder W und Ryll A (2003), S. 149.

Screeningverfahren durch Einbezug direkter Morbiditätsinformationen erreicht werden.

Das augenblickliche Verfahren sorgt dafür, dass Ärzte mit für ihren Fachbereich überdurchschnittlicher Patientenmorbidity (und damit erhöhtem Behandlungsbedarf) tendenziell eher als kostenauffällig identifiziert werden und vice versa. Zusätzlich würden bei Ärzten mit kleineren Patientenkohorten kostenintensivere Behandlungsfälle tendenziell schneller zu einer positiven Kostenauffälligkeitsprüfung führen als bei größeren Kohorten, in denen einzelne Hochkostenfälle abgepuffert werden können. Somit führt das augenblickliche Verfahren insbesondere bei Ärzten mit überdurchschnittlicher Patientenmorbidity und/oder kleinen Patientenkohorten zu falsch positiven Kostenauffälligkeitsprüfungen – Ärzte mit tendenziell unterdurchschnittlicher Patientenmorbidity und größeren Patientenbeständen werden vom Screeningverfahren trotz tatsächlich überdurchschnittlich hohem Kostenniveau, das erst nach Risikoadjustierung ermittelt werden kann, nicht identifiziert.

8. Fazit

Die Analysen dieses Gutachtens zeigen – ebenso wie bereits bestehende Publikationen – die Notwendigkeit einer morbiditätsorientierten Risikoadjustierung bereits im Screeningverfahren auf.

Variierende Kohortengrößen und heterogene Morbiditätslasten behandelnder Ärzte sorgen gleichermaßen für schwankende Prädiktionsgüten und somit zu falsch kostenauffällig und falsch kostenunauffällig identifizierten Ärzten. Insbesondere Ärzte mit kleineren Patientenkohorten und für ihren Fachbereich überdurchschnittlicher Patientenmorbidity werden durch das augenblickliche Screeningverfahren benachteiligt und vice versa.

Die Kostenprädiktionen im erweiterten Modell weisen signifikant bessere PR auf als das Modell im Status Quo. Die Gutachter konnten kein Gegenargument für eine direkte Morbiditätsorientierung identifizieren. Auch in anderen Sparten des Gesundheitswesens im europäischen Ausland – wie z. B. im Bereich des RSA der Krankenversicherung – werden morbiditätsorientierte Modelle mit verbesserter Zielgenauigkeit erfolgreich eingesetzt. Auch in der Schweiz ist die Erweiterung des RSA durch Berücksichtigung von PCGs oder aber einer Hospitalisierungsvariable aus dem Vorjahr geplant und folgt damit indirekt ebenfalls einer Morbiditätsorientierung.

Die Erklärungen, dass das Screeningverfahren lediglich ein „primäres“ Kriterium und somit einen ersten Schritt in einem Prüfprozess darstellt, ist mit Blick auf die versteckte Rationierung nicht haltbar, da auch ein nachgelagertes Prüfverfahren die Gefahr der versteckten Rationierung nicht vermeidet. Aus Sicht der Gutachter ist bereits im Screeningverfahren eine direkte Morbiditätsorientierung notwendig, um sowohl die versteckte Rationierung als auch Wettbewerbsverzerrungen und Fehlanreize bei den behandelnden Ärzten zu vermeiden.

9. Bearbeiter des Projektes

Das Projekt wurde als Kooperationsprojekt von Professor Dr. Jürgen Wasem und der ForBiG GmbH durchgeführt. Die ForBiG GmbH (www.forbig.de) ist eine „Ausgründung“ des Lehrstuhls für Medizinmanagement der Universität Duisburg-Essen und hat ihren Arbeitsschwerpunkt im Bereich der Budgetplanung, des Risikomanagements und Versorgungsmanagements der gesetzlichen Krankenversicherungen.



Helmut Dahl
(ForBiG GmbH)



Gerald Lux
(ForBiG GmbH)



Prof. Dr. Jürgen Wasem
(Universität Duisburg-Essen)

10. Quellenverzeichnis

Amstutz R (2005): Wirtschaftlichkeitsverfahren: Das Ziel heisst Verhaltensänderung. infosantésuisse, S. 3.

Brändli C (2009): DIE WIRTSCHAFTLICHKEITSPRÜFUNGEN VON SANTÉSUISSE, S. 6-9.

Beck K, Trottmann M (2007): Der Risikostrukturausgleich und die langfristigen Profite der Risikoselektion: Wie erfolgreich sind verschiedene Formeln? In: Göppfarth D et al. (Hrsg.) Jahrbuch Risikostrukturausgleich 2007. Sankt Augustin: Asgard-Verlag Dr. Werner Hippe GmbH. S. 329-355.

BVA (2010):

http://www.bundesversicherungsamt.de/cln_115/nn_1440668/DE/Risikostrukturausgleich/Festlegungen/Festlegung__Klassifikationsmodell__30092009.html [Zugriff am 06.07.2010]

Eugster D (2009): Überarztung aus juristischer Sicht. Ein Konzert mit Misstönen, Punkte 72-77, 89, 93-97, 102-103.

Hurst S A et al. (2006): Prevalence and determinants of physician bedside rationing: data from Europe. J Gen Intern Med 2006 Nov;21(11):1138-43.

Hurst S A et al. (2007): Physicians' views on resource availability and equity in four European health care systems. BMC Health Serv Res 2007 Aug 31;7(1):137.

IGES Lauterbach Wasem (2004): Klassifikationsmodelle für Versicherte im Risikostrukturausgleich – Endbericht, Untersuchung im Auftrag des Bundesministeriums für Gesundheit, Bonn, S. 12, 13.

Kraft P (2005): Wirtschaftlichkeitsverfahren bauen auf verlässliche Grundlagen. infosantésuisse, S. 4-5.

- Leu R E, Beck K (2006): Risikoselektion und Risikostrukturausgleich in der Schweiz, Gutachten zuhanden der Techniker Krankenkasse Hamburg, S. 20 ff.
- Romanens M et al. (2009): Krankenkassen, santésuisse und Wirtschaftlichkeitsverfahren, die keine sind... Schweizerische Ärztezeitung 2009;90:1/2, S. 51-52.
- Roth H-R, Stahel W (2005): Die ANOVA-Methode zur Prüfung der Wirtschaftlichkeit von Leistungserbringern nach Artikel 56 KVG. Gutachten zu Handen von santésuisse, S. 1-17.
- santésuisse (2009): Wirtschaftsprüfungen bei frei praktizierenden Ärzten (Fassung vom 29.01.2009). Positionspapier santésuisse, S. 1-3.
- santésuisse (2010):
http://www.santesuisse.ch/de/dyn_output.html?content.void=3270&navid=120 [Zugriff am 06.07.2010]
- Schräder W, Ryll A (2003): Pauschalierende Vergütungssysteme in der integrierten Versorgung. In: Integrierte Versorgung. Entwicklungsperspektiven für Praxisnetze, C. Tophoven and L. Lieschke, Editors. 2003, Deutscher Ärzteverlag: Köln. S. 131-169.
- Schwenkglens M (2010): Vergleich verschiedener Instrumente (Rechnungsstellerstatistik der santésuisse und Praxisspiegel der Trustcenter) zur Beurteilung der von Schweizer Ärzten in der Grundversorgung verursachten Behandlungskosten. Vorläufiger Statistischer Bericht (Version 0.8. vom 20. Mai 2010), S. 12-13.
- Strech D et al. (2008): Ärztliches Handeln bei Mittelknappheit. In: Ethik in der Medizin, 20, Heft 2, 2008, 94-109. Online: DOI 10.1007/s00481-008-0550-x.
- Wasem J (2006): Zur Weiterentwicklung der vertragsärztlichen Gesamtvergütung. In: Die BKK, 94, 06/2006, S. 289-291.