



Gerald Lux, Jürgen Wasem

## Die statistische Performance eines Zwei-Jahres-Modells im morbiditätsorientierten Risikostrukturausgleich

Das Bundesversicherungsamt verknüpft für die Kalkulation im morbiditätsorientierten Risikostrukturausgleich Morbiditätsdaten eines Basisjahres mit den versicherten-bezogenen Leistungsausgaben im Folgejahr. Der Beitrag erweitert diese Perspektive, indem er modellhaft Morbiditätsinformationen aus zwei Kalenderjahren verknüpft, um damit die Krankheitsgeschichte mit in die Berechnung der Zuweisungen und der relevanten Kostengewichte einbeziehen zu können. Mit diesem neuartigen Vorgehen soll beispielhaft empirisch getestet werden, ob sich daraus eine verbesserte Erklärungskraft des Gesamtmodells ergeben kann. Die Effekte der Modellanpassungen werden auf Basis unterschiedlicher Gütemaße bewertet.

### Hintergrund

Der im Jahr 2009 eingeführte morbiditätsorientierte Risikostrukturausgleich soll die Rahmenbedingungen für einen funktionierenden Wettbewerb zwischen den gesetzlichen Krankenversicherungen schaffen. Es erfolgt ein finanzieller Ausgleich zwischen den Krankenkassen, indem die unterschiedliche Morbiditätsstruktur und damit die unterschiedlichen Kostenrisiken berücksichtigt werden und die Krankenkassen risikoadjustierte Zuweisungen aus dem Gesundheitsfonds auf Basis des morbiditätsorientierten RSA erhalten.

Der morbiditätsorientierte RSA verfolgt dabei zwei Ziele. Zum einen soll ein systematischer Wettbewerbsvorteil beziehungsweise -nachteil allein aufgrund einer Konzentration kranker, insbesondere chronisch kranker Versicherter vermieden werden. Zum anderen soll ein funktionsfähiger Wettbewerb zwischen Krankenkassen ermöglicht werden, der zu einer verbesserten Qualität der Versorgung und einer verbesserten Wirtschaftlichkeit führt (Bundestagsdrucksache 14/6432).

Die Basis für die Zuweisungen aus dem Gesundheitsfonds stellen die für die einzelnen Morbiditätsgruppen errechneten Kostengewichte dar. Das Bundesversicherungsamt (BVA) verknüpft im morbiditätsorientierten Risikostrukturausgleich Morbiditätsdaten eines Basisjahres mit versichertenbezogenen Leistungsausgaben im Folgejahr. Die ermittelten Kostengewichte im linearen Regressionsverfahren werden lediglich auf Morbiditätsdaten eines einzelnen Kalenderjahres für ausgewählte Morbiditätsgruppen gebildet (BVA 2008).

Schon aus datenschutzrechtlicher Sicht ist es dem BVA nicht möglich, eine verbreiterte Informationsbasis – wie beispielsweise die Morbiditätsinformationen aus zwei Kalenderjahren – zu verwenden. Somit ist dem Bundesversicherungsamt eine Differenzierung in Versicherte mit und ohne mehrjährige Krankheitsgeschichte bisher nicht möglich. Momentan bestehen keine veröffentlichten Analysen, die eine Beurteilung der Krankheitsgeschichte mit Blick auf den morbiditätsorientierten RSA und die für die Zuweisungen relevanten Kostengewichte enthalten.

### Fragestellung

Die folgenden Analysen beschäftigen sich mit der Frage, ob sich die Kostenniveaus von Versicherten mit einer zweijährigen Krankheitsgeschichte auf HMG-Ebene (HMG = Hierarchisierte Morbiditätsgruppe) von denen der Versicherten, die erstmalig die entsprechende HMG erreichen, unterscheiden. In diesem Zusammenhang wäre zu prüfen, ob mögliche HMG-Differenzierungen nach bestehender Krankheitsgeschichte zu relevanten Verteilungseffekten auf Subpopulationsebenen und zu einer verbesserten Erklärungskraft des Gesamtmodells führen können.

Zu vermuten ist, dass krankheitsspezifische Unterschiede in der Kostenintensität der Diagnostik und Therapie mit Blick auf die Krankheitsgeschichte bestehen. Erstmalig diagnostizierte Krankheiten mit aufwändigen und kostenintensiven diagnostischen Verfahren sowie hohem Aufwand in der Initialtherapie könnten höhere Kostenniveaus aufweisen als Krankheiten mit bereits bestehender Geschichte. Es wäre ebenfalls mög-

lich, dass eine bestehende Krankheitsgeschichte durch das bereits fortgeschrittene Krankheitsstadium höhere Therapiekosten und höhere Komplikationsraten aufweist als bei einem noch leichteren Krankheitsstadium im Rahmen einer Erstdiagnose.

### Verwendete Datenbasis

Die für das Projekt genutzte Datenbasis bestand aus mehr als drei Millionen Versicherten einer bundesweit tätigen Krankenkasse. Für die Analysen wurden Morbiditätsdaten der Versicherten aus den Jahren 2006 und 2007 sowie Leistungsausgaben des Jahres 2008 verwendet. Um Modell-Verzerrungen zu vermeiden, konnten aus dem Gesamtversichertenbestand nur Versicherte mit lückenlosem Versicherungszeitraum und deshalb vollständiger Morbiditätserfassung im Jahr 2006 und 2007 (Ausnahme: Neugeborene) berücksichtigt werden (eine unterjährige Versicherungszeit im Jahr der Leistungsausgaben 2008 war jedoch möglich).

Die Population wurde an die Alters- und Geschlechtsstruktur der Gesetzlichen Krankenversicherung (GKV) in 2008 (auf Basis der KM6) adjustiert, um eine entsprechende GKV-Repräsentativität zu gewährleisten. Die Analysen im Status-quo-Modell wurden sowohl auf dem nichtadjustierten als auch auf dem adjustierten Datensatz durchgeführt. Die einzelnen Modellanpassungen wurden jedoch nur auf dem GKV-adjustierten Datensatz durchgeführt, um aussagekräftigere Ergebnisse mit Blick auf die GKV treffen zu können.

Das Durchschnittsalter des GKV-adjustierten Datensatzes ( $n = 2.775.973$  Versicherte) liegt bei 43,12 Jahren (im nichtadjustierten Datensatz bei 47,20 Jahren), wobei die weiblichen Versicherten mit 44,79 Jahren ein durchschnittlich höheres Alter aufweisen als die männlichen Versicherten. Frauen sind im Modelldatensatz – wie auch in der GKV – nach Adjustierung mit 53 Prozent stärker vertreten als Männer mit 47 Prozent.

Eine Analyse der in den Regressionen berücksichtigten Leistungsausgaben ergab die in Tabelle 1 ausgewiesenen Ergebnisse. Die durchschnittlichen

Leistungsausgaben liegen im adjustierten Datensatz bei 1.904 Euro je Versicherten, wobei Frauen mit 1.962 Euro ein höheres Ausgabenniveau aufweisen als männliche Versicherte mit 1.839 Euro. Die Streuung der durch männliche Versicherte verursachten Leistungsausgaben liegt jedoch mit einer Standardabweichung von 7.262 Euro deutlich höher als bei weiblichen Versicherten (5.596 Euro). Aufgrund der starken Varianz in den Leistungsausgaben und einem Anteil von mehr als sieben Prozent von Versicherten komplett ohne Leistungsausgaben liegt die Standardabweichung höher als der Mittelwert – anders als dies bei einer Normalverteilung der Fall wäre. Etwa 10,46 Prozent der männlichen Versicherten weisen keinerlei Leistungsausgaben auf, während diese Quote bei Frauen mit 3,96 Prozent deutlich niedriger liegt.

**Tabelle 1: Leistungsausgaben nach Geschlecht**

	Leistungsausgaben (nach GKV-Adjustierung)		
	weiblich	männlich	gesamt
RSA-relevante Gesamtleistungsausgaben: Mittelw. (Std.-Abw.)	1.962 Euro (5.596 Euro)	1.839 Euro (7.262 Euro)	1.904 Euro (6.431 Euro)
Versichertenanteil ohne Leistungsausgaben	3,96 Prozent	10,46 Prozent	7,01 Prozent

Quelle: eigene Berechnungen

Die durchschnittlichen Leistungsausgaben im nichtadjustierten Datensatz liegen mit 2.144 Euro deutlich höher als im adjustierten Datensatz – in erster Linie bedingt durch das mit 47,20 gegenüber 43,12 Jahren höhere Durchschnittsalter.

## Methodisches Vorgehen

### Modellansätze

Zunächst wurde das BVA-Verfahren für das Ausgleichsjahr 2010 als Referenzgröße in Modell 1 umgesetzt. Auf Basis dieser Ergebnisse wurden die Gruppierungen der Versicherten aus zwei Morbiditätsjahren (2006

und 2007) aggregiert und anschließend in Versicherte mit und ohne Krankheitsgeschichte differenziert. Dabei wurde die Krankheitsgeschichte in unterschiedlichen Varianten definiert. Im Modell 2.0 musste der Versicherte dieselbe HMG bereits im Vorjahr erreicht haben, um die Klassifizierung „mit Krankheitsgeschichte“ zu erhalten.

In einem weiteren Modell (Modell 2.1) wurde die Krankheitsgeschichte über dieselbe beziehungsweise eine niedrigere HMG desselben HMG-Hierarchiebaumes des BVA definiert. Somit könnte der Versicherte dieselbe HMG oder aber eine leichtere Form der Erkrankung bereits im Vorjahr aufweisen, um eine bestehende Krankheitsgeschichte zu begründen. In einem dritten Modell reichte es aus, wenn der Versicherte eine HMG desselben HMG-Hierarchiebaumes im Vorjahr erreicht hatte (Modell 2.2).

**Tabelle 2: Initialer Eingangsraum der Modelle**

---

Modelle	Modellbeschreibung	HMG-Anzahl	AGG-Anzahl	EMG-Anzahl	relevantes Signifikanzniveau der HMGn (in Prozent)
Modell 1 (Status quo)	BVA-Verfahren des Ausgleichsjahres 2010	112	40	6	5
Modell 2.0	Differenzierung der HMGn nach bestehender gleicher HMG im Vorjahr 2006 ja/nein	224	40	6	1
Modell 2.1	Differenzierung der HMGn nach bestehender leichter/gleicher HMG derselben Hierarchie im Vorjahr 2006 ja/nein	224	40	6	1
Modell 2.2	Differenzierung der HMGn nach bestehender HMG derselben Hierarchie im Vorjahr 2006 ja/nein	224	40	6	1

---

Quelle: eigene Darstellung

Alle neuen Regressionsmodelle (zur Kurzbeschreibung der Modelle siehe die Tabellenübersicht) mit den Leistungsausgaben im Jahr 2008 als Zielvariable wurden solange wiederholt, bis nur signifikante Morbiditätsgruppen verblieben.

In allen Modellen erfolgte eine Aufspaltung der 112 HMGn des BVA aus dem Jahr 2007 in 224 HMGn in 2007, indem die Eingruppierung eines Versicherten in die jeweilige HMG dahingehend unterschieden wurde, ob zusätzlich eine Krankheitsgeschichte im Jahr 2006 vorlag oder nicht. Die Definition, ob ein Versicherter das Kriterium der Krankheitsgeschichte erfüllte oder nicht, ist im vorangegangenen Abschnitt für die drei Modellanpassungen beschrieben worden. Die neu implementierten HMG-Differenzierungen (224 HMGn statt der originären 112 HMGn im ursprünglichen BVA-Modell) mussten ein Signifikanzniveau von einem Prozent unterschreiten. Bei Überschreiten des relevanten Signifikanzniveaus oder bei ausgewiesenem Abschlag (negatives Kostengewicht) einer HMG-Differenzierung wurden jeweils beide Differenzierungen derselben HMG aus dem Modell entfernt und durch die ursprüngliche HMG ohne Differenzierung im Regressionsmodell ersetzt. Für die bereits validierten 112 HMGn des BVA galt dann das Signifikanzniveau von fünf Prozent als ausreichend.

Für alle Modelle gilt weiterhin das Prinzip eines HMG-Zuschlagsmodells – ermittelte Abschläge in Form negativer Kostengewichte sorgen dafür, dass die entsprechenden HMG-Differenzierungen wieder durch die ursprüngliche HMG in der Regression ersetzt werden. Auch im BVA-Modell sind keine negativen Kostengewichte für Morbiditätsgruppen möglich – diese würden einem finanziellen Abschlag entsprechen, wenn ein Versicherter die entsprechende Morbiditätsgruppe besetzt (BVA 2009).

Für die Analysen wurde auf eine Teilung des Datensatzes in eine Lernmenge (zur Modellbildung) und eine Testmenge (zur Modellanwendung) mit Blick auf die geringen Fallzahlen in den HMGn verzichtet. Die Effekte der Modellanpassungen wurden im Anschluss auf Basis unterschiedlicher Gütemaße bewertet.

## Modellperformance – R<sup>2</sup>/CPM/MAPE

Das Bestimmtheitsmaß (R<sup>2</sup>) wird zur Beurteilung der Erklärung der Varianzen in den Leistungsausgaben verwendet und beschreibt den Anteil der Varianz in der Zielvariablen an der Gesamtvarianz, der allein auf die erklärenden Variablen des Modells zurückzuführen ist (siehe Formel 1).

### Formel 1: Bestimmtheitsmaß (R<sup>2</sup>)

$$R^2 = 1 - \left( \sum_i (a_i - \hat{a}_i)^2 \right) / \left( \sum_i (a_i - \bar{a})^2 \right), \text{ mit}$$

⟨ a<sub>i</sub> = tatsächliche Leistungsausgaben des Versicherten i

⟨  $\hat{a}_i$  = prognostizierte Leistungsausgaben des Versicherten i

⟨  $\bar{a}$  = durchschnittliche tatsächliche Leistungsausgaben über alle Versicherten

⟨ i = Index durchläuft die Werte von 1 bis n (n = Anzahl aller Versicherten)

Quelle: Schäfer (2011)

Je mehr sich das R<sup>2</sup> dem Wert 1,0 nähert, desto höher ist die Erklärungskraft des entsprechenden Modells. Dabei ist jedoch zu beachten, dass neben der bereits beschriebenen Anfälligkeit des R<sup>2</sup>-Wertes gegenüber Kostenausreißern ein übermäßig ausgeweitetes Set von unabhängigen Variablen zu einer künstlichen Erhöhung des Bestimmtheitsmaßes führt, da eine Aufnahme zusätzlicher Regressoren zu einer Erhöhung des Bestimmtheitsmaßes führt – unabhängig davon, ob die zusätzlichen unabhängigen Variablen tatsächlich einen Beitrag zur Erklärungskraft liefern. Aus diesem Grund wurde das R<sup>2</sup> unter Berücksichtigung der zusätzlichen Regressoren in Form des korrigierten Bestimmtheitsmaßes ermittelt.

Das Cummings Prediction Measure (CPM) wurde zur weniger starken Gewichtung von Kostenausreißern und Hochkostenfällen in der Modellbeurteilung genutzt. Es unterscheidet sich vom Bestimmtheitsmaß in der Nicht-Quadrierung der versichertenindividuellen Abweichungen



von Prognose und tatsächlichem Wert – stattdessen werden die betragsmäßigen versichertenindividuellen Abweichungen verwendet (siehe Formel 2). Der CPM liegt dabei in einem Wertebereich von 0 bis 1 und kann theoretisch in extremen Fällen sogar negative Werte annehmen. Auch hier würden Werte näher an 1 einer tendenziell höheren Erklärungskraft entsprechen (Schäfer 2011).

### Formel 2: Vorhersagemaß von Cumming

$$\text{CPM} = 1 - (\sum_i |a_i - \hat{a}_i|) / (\sum_i |a_i - \bar{a}|), \text{ mit}$$

⟨  $a_i$  = tatsächliche Leistungsausgaben des Versicherten  $i$

⟨  $\hat{a}_i$  = prognostizierte Leistungsausgaben des Versicherten  $i$

⟨  $\bar{a}$  = durchschnittliche tatsächliche Leistungsausgaben über alle Versicherten

⟨  $i$  = Index durchläuft die Werte von 1 bis  $n$  ( $n$  = Anzahl aller Versicherten)

Quelle: Schäfer (2011)

Der Mean Absolute Prediction Error (MAPE) stellt die betragsmäßig durchschnittliche Abweichung von versichertenindividuellen Prognosen und versichertenindividuellen Leistungsausgaben dar (siehe Formel 3) – je kleiner der MAPE, desto besser ist die Performance des Modells (Schäfer 2011).

### Formel 3: Mittlerer absoluter Vorhersagefehler

$$\text{MAPE} = (\sum_i |a_i - \hat{a}_i|) / n, \text{ mit}$$

⟨  $a_i$  = tatsächliche Leistungsausgaben des Versicherten  $i$

⟨  $\hat{a}_i$  = prognostizierte Leistungsausgaben des Versicherten  $i$

⟨  $i$  = Index durchläuft die Werte von 1 bis  $n$  ( $n$  = Anzahl aller Versicherten)

Quelle: Schäfer (2011)

## Verteilungswirkungen

Für die Ermittlung von Verteilungswirkungen der einzelnen Modellanpassungen wurden unterschiedliche Subpopulationen analysiert. Das relevante Messinstrument war hierbei die Predictive Ratio (PR) als Relation von prognostizierten und tatsächlichen Leistungsausgaben. Die PR ergeben sich gemäß folgender Formel als Quotienten aus summierten Prognosen und summierten tatsächlichen Leistungsausgaben der jeweils zu analysierenden Subpopulationen (siehe Formel 4).

### Formel 4: Predictive Ratio

$$PR = \sum \hat{a}_i / \sum a_i, \text{ mit}$$

⟨  $a_i$  = tatsächliche Leistungsausgaben des Versicherten  $i$

⟨  $\hat{a}_i$  = prognostizierte Leistungsausgaben des Versicherten  $i$

⟨  $i$  = Index durchläuft die Werte von 1 bis  $n$  ( $n$  = Anzahl aller Versicherten)

Quelle: Schäfer (2011)

Je näher der Wert bei 1,0 liegt, desto besser ist die Deckungsquote der Versichertenpopulation, wobei eine PR größer als 1,0 eine Überschätzung der Leistungsausgaben darstellt und umgekehrt. Die PR werden getrennt nach Geschlecht, nach Altersklassen und auf HMG-Ebenen bewertet, um mögliche Verteilungseffekte innerhalb des Modells identifizieren zu können. Für alle Testebenen wurde jeweils der GKV-adjustierte Modelldatensatz verwendet.

## Effekte der Modellanpassungen auf die Kostengewichte

Die neu ermittelten Kostengewichte auf Basis der HMG-Differenzierungen der Modellanpassungen wurden beispielhaft für einige ausgewählte HMGn den Kostengewichten im Status quo auf Basis des Modelldatensatzes gegenübergestellt und beurteilt.

## Analyse der Modelleffekte

### Modellperformance im Status quo

#### Zum Bestimmtheitsmaß $R^2$

Das Modell im Status quo erreicht auf dem nicht an die GKV adjustierten Analysedatensatz ein unkorrigiertes  $R^2$  von 25,56 Prozent – das Modell auf der GKV-Stichprobe erreichte im BVA ein  $R^2$  von 23,35 Prozent, sodass die BVA-Logik im Modelldatensatz einen deutlich größeren Anteil der prospektiven Leistungsausgaben erklären kann. Das  $R^2$  verbessert sich durch die GKV-Adjustierung noch einmal deutlich von 25,56 Prozent auf 27,17 Prozent – der Grund für diese deutliche Abweichung zum Vergleichswert der GKV-Stichprobe (23,35 Prozent) liegt in der spezifischen Versichertenstruktur des Modelldatensatzes.

Während die GKV-Stichprobe ein Versichertenkollektiv aus unterschiedlichen Kassenarten darstellt, wurden die vorliegenden Analysen auf dem homogenen Versichertenklientel einer einzelnen Krankenkasse erstellt. Der zusätzliche Anstieg des  $R^2$ -Wertes aufgrund der GKV-Adjustierung ist damit erklärbar, dass durch die Adjustierung eine stärkere Gewichtung des männlichen Versichertenklientels erfolgte, das  $R^2$ -Werte von etwa 30 Prozent aufweist, während weibliche Versicherte lediglich Werte von etwas mehr als 22 Prozent erreichen (siehe Tabelle 3). Diese anteiligen  $R^2$ -Werte wurden durch geschlechtsspezifische Filterungen ermittelt. Es wurden auch separate geschlechtsspezifische Modelle berechnet, die sich im Ergebnis nur geringfügig von den anteiligen geschlechtsspezifischen Werten unterscheiden und an dieser Stelle deshalb nicht zusätzlich aufgeführt werden.

**Tabelle 3: Modellperformance im Status quo**

Modell Status quo	R <sup>2</sup> -Werte (in Prozent)	korrigiertes R <sup>2</sup> (in Prozent)	CPM (in Prozent)	MAPE (in Euro)
BVA	23,35	n. a.	n. a.	n. a.
nicht adjustiert	25,56	25,56	21,55	1.903
adjustiert	27,17	27,16	26,03	1.738
Frauen (nicht adjustiert)	22,31	22,30	19,37	1.899
Frauen (adjustiert)	22,36	22,35	22,67	1.761
Männer (nicht adjustiert)	29,00	28,99	24,97	1.909
Männer (adjustiert)	30,41	30,40	29,59	1.712

Quelle: eigene Berechnung

### Korrigiertes R<sup>2</sup>

Die korrigierten R<sup>2</sup>-Werte sowohl des adjustierten (27,16 Prozent) als auch des nichtadjustierten Datensatzes (25,56 Prozent) liegen nah an den nichtkorrigierten R<sup>2</sup>-Werten (27,17 Prozent beziehungsweise 25,56 Prozent), sodass eine Überfrachtung an Regressoren mit nur sehr geringem zusätzlichem Erklärungskraftanteil nicht vorliegen konnte. Auch die geschlechtsspezifischen Werte lassen diese Schlussfolgerung zu.

### CPM-Wert

Der CPM sowohl des nichtadjustierten als auch des adjustierten Datensatzes liegt jeweils niedriger als der vergleichbare R<sup>2</sup>-Wert (21,55 Prozent gegenüber 25,56 Prozent beziehungsweise 26,03 Prozent gegenüber 27,17 Prozent). Dies deutet darauf hin, dass die Zuweisungsmodelle des BVA insbesondere bei Versicherten mit höheren Leistungsausgaben größere Anteile der tatsächlichen Leistungsausgaben mit den Leistungsausgabenprognosen erklären können, da die versichertenindividuellen Abweichungen beim R<sup>2</sup> durch die Quadrierung stärker ins Gewicht fallen als bei der Ermittlung des CPM.

Zu erkennen ist, dass die Differenz von R<sup>2</sup> und CPM im adjustierten Datensatz mit 1,14 Prozent (= 27,17 Prozent minus 26,03 Prozent) deutlich

geringer ausfällt als im nichtadjustierten Datensatz mit 4,01 Prozent (= 25,56 Prozent minus 21,55 Prozent). Auch die geschlechtsspezifischen Analysen weisen kleinere Differenzen von  $R^2$  und CPM-Wert im adjustierten gegenüber dem nichtadjustierten Modell auf, wobei bei weiblichen Versicherten der CPM im adjustierten Modell mit 22,67 Prozent gegenüber 22,35 Prozent sogar höher ausfällt als das vergleichbare  $R^2$ . Bei männlichen Versicherten sinkt die Abweichung von  $R^2$  zu CPM-Wert von 4,03 Prozent (= 29,00 Prozent minus 24,97 Prozent) auf 0,82 Prozent (30,41 Prozent minus 29,59 Prozent).

Die Tatsache, dass alle adjustierten Modelle im Status quo zu einer Annäherung des CPM-Wertes an das jeweilige  $R^2$  – und bei weiblichen Versicherten sogar zu einem insgesamt höheren CPM-Wert gegenüber dem  $R^2$ -Wert – führen, ist allein auf die GKV-Altersadjustierung zurückzuführen. Diese sorgte für eine Reduktion des Durchschnittsalters und damit zu einer Reduktion von Verstorbenen beziehungsweise Hochkostenfällen im Modell, deren Prävalenzen mit steigendem Alter korrelieren. Aus diesem Grund verringern sich die Unterschiede in den statistischen Gütemaßen des CPM und des  $R^2$  – beide Gütemaße erreichen jedoch im adjustierten Modell insgesamt und geschlechtsspezifisch höhere Werte.

### Mean Absolute Prediction Error – MAPE

Der MAPE des nichtadjustierten Modells liegt im Status quo bei 1.903 Euro und im adjustierten Modell bei 1.738 Euro – weibliche Versicherte erreichen einen MAPE von 1.899 Euro (nichtadjustiert) beziehungsweise 1.761 Euro (adjustiert) und Männer weisen einen MAPE von 1.909 Euro (nichtadjustiert) beziehungsweise 1.712 Euro (adjustiert) auf. Auch hier ist erkennbar, dass die Altersadjustierung jeweils zu einer Verringerung der Abweichungen von Prognose zu tatsächlichen Leistungsausgaben geführt hat.

### Modellperformance – Effekte der Modellanpassungen

Die HMG-Differenzierungen des Modells 2.0 sorgen für eine Verbesserung von  $R^2$  und CPM von 0,58 beziehungsweise 0,29 Prozentpunkten.

Die Modelle 2.1 und 2.2 erreichen R<sup>2</sup>-Werte von 27,54 Prozent beziehungsweise 27,51 Prozent und somit Verbesserungen von nur 0,37 beziehungsweise 0,34 Prozentpunkten. Die CPM-Werte liegen jedoch mit 26,31 Prozent und 26,29 Prozent nah am CPM des Modells 2.0 (26,32 Prozent).

**Tabelle 4: Modellperformances der Modellanpassungen im Vergleich**

Modelle	R <sup>2</sup> -Werte (in Prozent)	korrigiertes R <sup>2</sup> (in Prozent)	CPM (in Prozent)	Vergleich R <sup>2</sup> zu Modell 1 (in Prozent)
Modell 1 (Status quo)	27,17	27,16	26,03	–
Modell 2.0	27,75	27,74	26,32	0,58
Modell 2.1	27,54	27,53	26,31	0,37
Modell 2.2	27,51	27,50	26,29	0,34

Quelle: eigene Berechnung

Die relativ deutlichen Abweichungen im R<sup>2</sup> auf der einen und die sehr geringen Differenzen im CPM-Wert auf der anderen Seite sprechen dafür, dass insbesondere Modell 2.0 einen größeren Erklärungsanteil für die Leistungsausgaben von Versicherten mit höherem Kostenniveau leistet – bei einer weniger starken Gewichtung der versichertenindividuellen Abweichungen von Prognosewert und tatsächlichen Leistungsausgaben im CPM unterschieden sich die Modelle 2.0, 2.1 und 2.2 kaum. Wegen der besseren Ergebnisse bei R<sup>2</sup> und CPM stellt das Modell 2.0 die überlegene Modellanpassung gegenüber den Modellen 2.1 und 2.2 dar. Der Vorteil des Modells 2.0 in der Modellperformance wurde für die Gesamtbeurteilung der Modelle um den Aspekt der Umverteilungswirkungen auf Subpopulationsebenen erweitert.

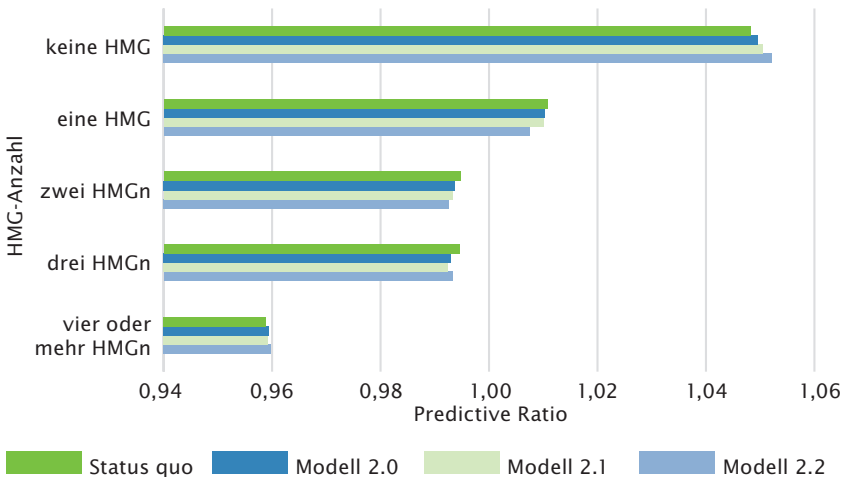
### Verteilungswirkungen

Eine detaillierte Analyse der Verteilungswirkungen auf Subpopulationsebenen ergab, dass in allen Modellvarianten praktisch keine Umver-

teilung zwischen den Geschlechtern erfolgt – auch altersklassenspezifisch waren keine relevanten Umverteilungseffekte zu beobachten.

Eine Analyse der PR nach RSA-relevanter Morbiditätslast in Form der HMG-Anzahl je Versicherten zeigte nur minimale Effekte. Alle Modellanpassungen sorgen im Vergleich zum Status quo für eine leichte Verstärkung der Überschätzung für Versicherte ohne HMG – für Versicherte mit HMGn lassen die Ergebnisse keine modellspezifischen Aussagen zu Umverteilungswirkungen zu (siehe Abbildung 1).

**Abbildung 1: Predictive Ratios nach HMG-Anzahl**



Quelle: eigene Berechnung

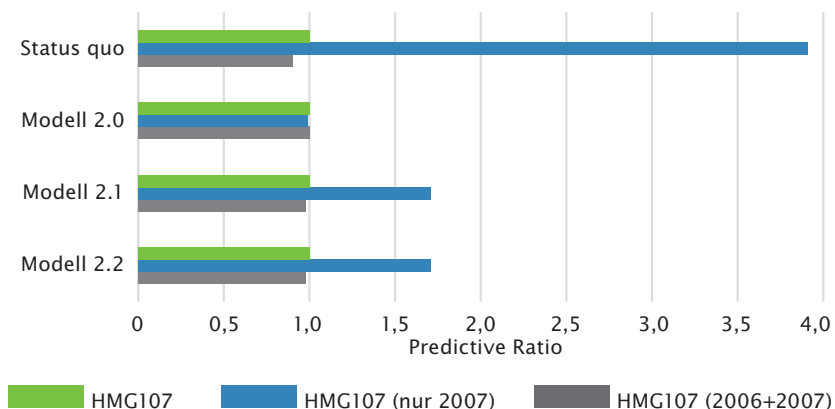
Für eine beispielhafte Darstellung unterschiedlicher Effekte auf einzelne HMGn wurden jeweils zwei HMGn mit hohen Fallzahlen beziehungsweise mit geringen Prävalenzen vergleichend gegenübergestellt. Dafür wurden die Hypertonie (HMG091), die Koronare Herzkrankheit (KHK) (HMG084) und die Mukoviszidose (HMG107) – eine genetisch bedingte Stoffwechselerkrankung – sowie der schwere angeborene Herzfehler

(HMG087) ausgewählt, die anschließend auch für die Darstellung der Effekte auf die Kostengewichte verwendet wurden.

Die PR der einzelnen HMGn ändern sich kaum. Die Umverteilungen finden in erster Linie innerhalb der jeweiligen HMG-Zuweisungen statt. Für die beiden HMGn mit geringen Fallzahlen (HMG087 und HMG107) zeigen sich deutliche Umverteilungseffekte für Versicherte in den zwei Gruppen der HMG-Differenzierungen. Für alle Versicherte mit der HMG107 (Mukoviszidose) liegt die PR im Status quo bei 100,47 Prozent – im Modell 2.0 kann dieser Wert auf 100,26 Prozent verbessert werden. Die PR der Versicherten in den beiden HMG-Differenzierungen des Modells 2.0 verbessern sich im Vergleich zum Status quo jedoch deutlich. Versicherte ohne Krankheitsgeschichte weisen im Status quo eine PR von 3,91 und damit eine deutliche Überschätzung auf, während für Versicherte mit Krankheitsgeschichte eine systematische Unterschätzung von etwa zehn Prozent vorliegt (PR: 0,90).

**Abbildung 2:** Predictive Ratios HMG107 – Mukoviszidose

---



Quelle: eigene Berechnung

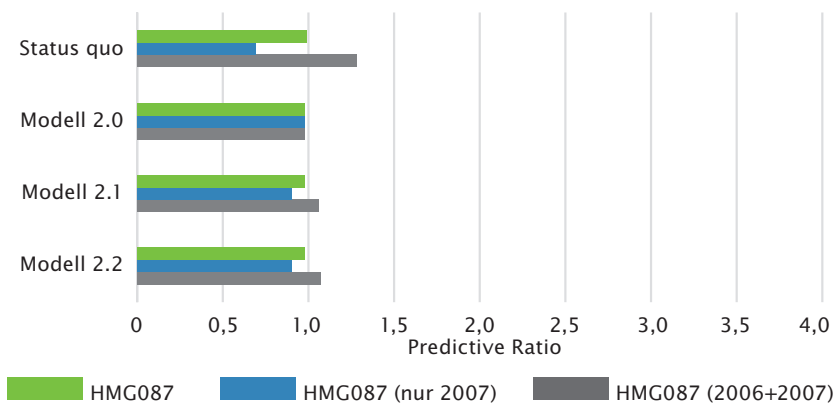
---



Im Modell 2.0 ändern sich diese beiden PR von 3,91 auf 0,99 beziehungsweise von 0,90 auf 1,03 – die Über- und Unterschätzungen reduzieren sich somit innerhalb der HMG107 deutlich. Eine Begründung für die deutliche Überschätzung für Versicherte ohne Krankheitsgeschichte könnte eine tendenziell geringe Therapieintensität in einem frühen Krankheitsstadium sein – die deutliche Unterdeckung für Versicherte mit Krankheitsgeschichte könnte durch einen steigenden Therapiebedarf und erhöhte Komplikationsraten im fortschreitenden Krankheitsverlauf erklärt werden.

Auch für den schweren angeborenen Herzfehler (HMG087) zeigen sich starke Umverteilungseffekte innerhalb der HMG. Die PR für alle Versicherten mit HMG087 reduziert sich zwar leicht im Modell 2.0 gegenüber dem Status quo von 0,987 auf 0,981, aber die PR für Versicherte ohne Krankheitsgeschichte erhöht sich von 0,694 auf 0,979 und die systematische Überschätzung von 28 Prozent bei Versicherten mit Krankheitsgeschichte (PR 1,285) wandelt sich in eine leichte Unterschätzung von knapp zwei Prozent (PR 0,984).

**Abbildung 3: Predictive Ratios HMG087 – Schwerer angeborener Herzfehler**

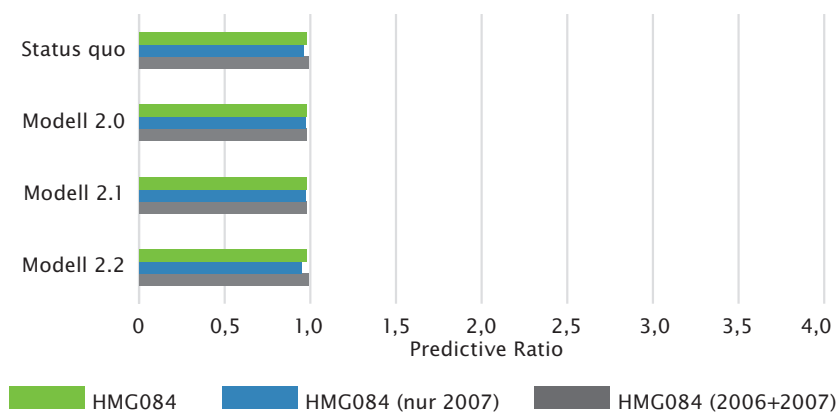


Quelle: eigene Berechnung

Die Effekte für die Beispiel-HMGn mit höheren Prävalenzen fallen eher gering aus. Während die PR von Versicherten mit HMG084 (KHK) im Status quo und Modell 2.0 bei jeweils 0,976 liegt, verbessert sich zwar die PR von Versicherten ohne Krankheitsgeschichte von 0,955 auf 0,972, aber die PR für Versicherte mit Krankheitsgeschichte reduziert sich von 0,986 auf 0,978, sodass für die HMG084 kein eindeutiger Vorteil des Modells 2.0 gegenüber dem Status quo abgeleitet werden kann.

**Abbildung 4:** Predictive Ratios HMG084 – Koronare Herzkrankheit

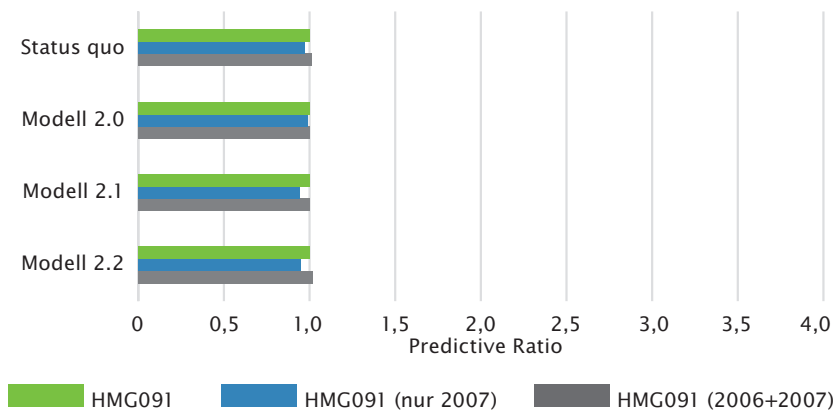
---



Quelle: eigene Berechnung

Die PR der HMG091 (Hypertonie) verändert sich im Modell 2.0 gegenüber dem Status quo kaum (99,823 Prozent gegenüber 99,816 Prozent) – getrennt nach Krankheitsgeschichte verbessern sich jedoch die PR. Für Versicherte ohne Krankheitsgeschichte verbessert sich die PR von 0,966 auf 0,990 – bei Versicherten mit Historie kann die PR von 1,009 auf 1,001 gesenkt werden.

**Abbildung 5:** Predictive Ratios HMG091 – Hypertonie



Quelle: eigene Berechnung

### Effekte auf die Kostengewichte

Die HMG-Differenzierungen in Abhängigkeit der Krankheitsgeschichte weisen in allen Modellen für mehr als 90 Prozent der HMGn signifikante Ergebnisse auf. Die Differenzen zwischen den jeweiligen HMG-Differenzierungen fallen jedoch sehr unterschiedlich aus.

In der Tabelle 5 sind die einzelnen HMG-Differenzierungen aufgeführt, die in den Modellen entweder keine ausreichende Signifikanz aufweisen oder aber zu negativen Kostengewichten geführt haben.

Durch die Modellanpassungen finden nur relevante Umverteilungen innerhalb einer HMG statt – relevante Umverteilungen zwischen HMGn sind hingegen nicht zu erkennen. Teilweise zeigen Gruppen mit und teilweise ohne vorliegende Krankheitsgeschichte höhere Kostengewichte auf. Da sich die Kostengewichtseffekte der einzelnen Modellanpassungen sehr ähneln, das Modell 2.0 allerdings die beiden anderen Modelle dominiert, wurden an dieser Stelle nur die jeweiligen Ergebnisse des Modells 2.0 dargestellt.

**Tabelle 5:** Nicht berücksichtigte HMG-Differenzierungen in den einzelnen Modellen

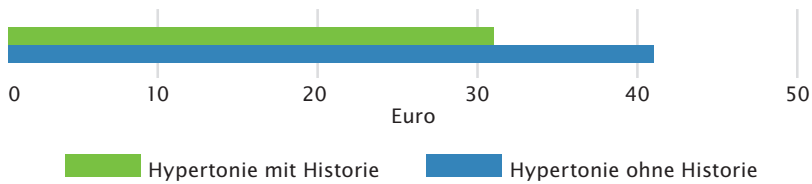
HMG-Differenzierung	Modell 2.0	Modell 2.1	Modell 2.2
HMG143 (Hämophilie ohne Dauermedikation Frauen)	nicht signif.	nicht signif.	nicht signif.
HMG159 (Pathologische Fraktur des Humerus, der Schulter, der Tibia und/oder Fibula)	nicht signif.	nicht signif.	nicht signif.
HMG136 (Nierenfunktionsstörung)	nicht signif.	nicht signif.	nicht signif.
HMG106 (Artherosklerose, arterielles Aneurysma und sonstige, nicht näher bezeichnete Krankheiten der Arterien und Arteriolen)	nicht signif.	nicht signif.	nicht signif.
HMG165 (Andere iatrogene Komplikationen)	nicht signif.	nicht signif.	nicht signif.
HMG089 (Hypertensive Herz- und Nierenerkrankung oder Enzephalopathie)	nicht signif.	nicht signif.	nicht signif.
HMG110 (Chronische obstruktive Bronchitis/ Emphysem (Alter < 18 Jahre))	nicht signif.	nicht signif.	nicht signif.
HMG042 (Postmenopausale Osteoporose)	nicht signif.	nicht signif.	nicht signif.
HMG161 (Traumatische Amputation)	nicht signif.	nicht signif.	nicht signif.
HMG043 (Hämophilie ohne Dauermedikation Männer)	–	–	nicht signif.
HMG176 (Transplantation der Niere)	negativer Koeffizient	negativer Koeffizient	negativer Koeffizient
HMG029 (Sonstige virale Hepatitis (ohne HAV), akut oder nicht näher bezeichnet)	–	–	negativer Koeffizient

Quelle: eigene Berechnung

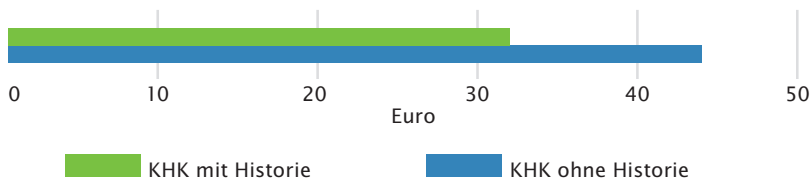
Für die Mukoviszidose und den schweren angeborenen Herzfehler fallen die Kostengewichtsunterschiede auf Basis der HMG-Differenzierungen deutlich stärker aus, was vermutlich auch auf die geringen Fallzahlen zurückzuführen ist, wobei die Kostengewichte der HMG-Differenzierungen

jeweils entsprechende Signifikanzen aufweisen. Für Versicherte mit Mukoviszidose und bestehender Krankheitsgeschichte liegt das monatliche Kostengewicht mit 1.838 Euro deutlich höher als bei Versicherten ohne Krankheitsgeschichte (217 Euro). Für den schweren angeborenen Herzfehler kann ein gegenläufiger Effekt mit einem höheren Kostengewicht für Versicherte ohne Krankheitsgeschichte gegenüber Versicherten mit Historie (577 Euro gegenüber 172 Euro) festgestellt werden. Für die Hypertonie und die KHK fallen die Unterschiede deutlich geringer aus – bei bestehender Krankheitsgeschichte liegen die Kostengewichte etwa jeweils 30 Prozent unter denen ohne Historie (siehe Abbildung 6 bis 9).

**Abbildung 6:** HMG091 Hypertonie – Kostengewichte



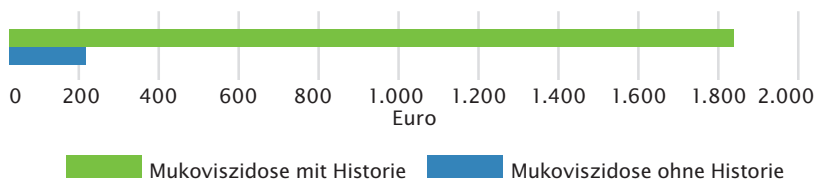
**Abbildung 7:** HMG084 Koronare Herzkrankheit – Kostengewichte



Quelle: eigene Berechnung

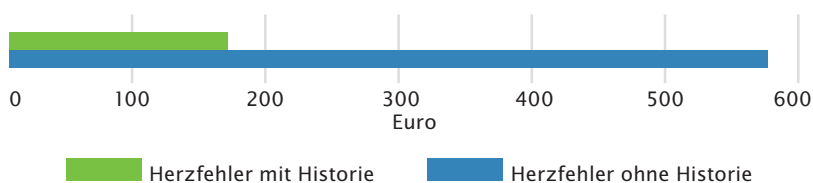
**Abbildung 8:** HMG107 Mukoviszidose – Kostengewichte

---



**Abbildung 9:** HMG087 Schwerer angeborener Herzfehler – Kostengewichte

---



Quelle: eigene Berechnung

### Fazit und Ausblick

Die Ergebnisse zeigen, dass eine Verbreiterung der Morbiditäts-Informationsbasis und eine HMG-Differenzierung in Krankheitsgeschichte „ja/nein“ zum Teil deutliche Verteilungseffekte innerhalb der einzelnen HMGn erzeugen und dort zu zielgerichteteren Zuweisungen führen würde. Die Effekte der HMG-Differenzierungen fallen jedoch krankheitsspezifisch unterschiedlich aus – so führt eine bereits bestehende Krankheitsgeschichte bei einigen HMGn zu einem höheren Kostengewicht als bei nicht identifizierter Historie und umgekehrt.

Die Modellqualität verbessert sich in allen getesteten Modellen im Vergleich zum Status quo relativ deutlich, wobei das Modell 2.0 sowohl mit Blick auf die Modellperformance als auch mit Blick auf die Verteilungseffekte auf Subpopulationsebenen am vorteilhaftesten und der dahinterstehende methodische Ansatz im Vergleich zu den Modellen 2.1 und

2.2 am pragmatischsten wäre, wenngleich die Krankheitsgeschichte vieler Versicherter aufgrund des Hierarchisierungsprinzips im BVA-Modell im Modell 2.0 nicht erfasst wird und somit nur teilweise Krankheitsgeschichten erfasst werden können. Eine differenzierte krankheitsspezifische Analyse zur Identifikation von weiteren Optimierungsmöglichkeiten des Modellansatzes 2.0 wäre insbesondere aus medizinisch fachlicher Sicht notwendig, um weitere Modellverbesserungen zu ermöglichen.

Alle Modellanpassungen erzeugen keine relevanten geschlechts- oder altersklassenspezifischen Umverteilungen; auch zwischen einzelnen HMGn sind keine relevanten Umverteilungen zu beobachten. Die Analysen von  $R^2$ - und CPM-Werten – und der deutlich stärkere positive Effekt auf das  $R^2$  im Vergleich zum CPM-Wert – zeigen, dass die Modellanpassungen insbesondere Leistungsausgaben für Versicherte mit höheren Kostenniveaus erklären können. Ob und inwieweit die Modellanpassungen Verteilungseffekte zwischen einzelnen Krankenkassen erzeugen können, kann auf Basis des vorliegenden Datensatzes einer Kassenart nicht beurteilt werden und würde vergleichende Analysen von Daten unterschiedlicher Kassen erfordern.

Aufgrund bestehender Datenschutzvorgaben wäre die Umsetzung eines Modells mit zweijährigen Morbiditätsinformationen nicht realisierbar. Hierzu wäre eine Anpassung der gesetzlichen Rahmenbedingungen notwendig, um eine Verbreiterung der Informationsbasis zuzulassen und damit Modellverbesserungen zu ermöglichen.

## Literatur

Bundesversicherungsamt (2008): So funktioniert der neue Risikostrukturausgleich im Gesundheitsfonds. S. 2-10.

Bundesversicherungsamt (2009): Dokument 01\_Festlegung\_Klassifikation\_JA2010. S. 8. [www.bundesversicherungsamt.de/cln\\_115/nn\\_1440668/DE/Risikostrukturausgleich/Festlegungen/Festlegung\\_\\_Klassifikationsmodell\\_\\_30092009.html](http://www.bundesversicherungsamt.de/cln_115/nn_1440668/DE/Risikostrukturausgleich/Festlegungen/Festlegung__Klassifikationsmodell__30092009.html) (Download 5. Juli 2011).

Schäfer, T. (2011): Statistische Bewertung unterschiedlicher Hierarchievarianten im Klassifikationssystem für den Risikostrukturausgleich. [www.bundesversicherungsamt.de/cln\\_115/nn\\_1046668/DE/Risikostrukturausgleich/Weiterentwicklung\\_20RSA/Weiterentwicklung\\_\\_node.html?\\_\\_nnn=true](http://www.bundesversicherungsamt.de/cln_115/nn_1046668/DE/Risikostrukturausgleich/Weiterentwicklung_20RSA/Weiterentwicklung__node.html?__nnn=true) (Download 5. Juli 2011).