

**Bericht des Instituts des Bewertungsausschusses
zu Vollständigkeits- und Repräsentativitätsprüfungen
gemäß Beschluss des 184. BA vom 20. Mai 2009 und
den Ermittlungen der Strukturveränderungen der Morbidität
zwischen den Jahren 2006 und 2008**

Autorenliste

Frau Peggy David
Frau Kerstin Leipnitz
Frau Hanna Tillmanns
Herr Dr. Andreas Hinz
Herr Dirk Horenkamp-Sonntag
Herr Sebastian Krüger
Herr Dr. Jörn Quedenau
Herr Peter Reschke

unter Mitarbeit von

Frau Dr. Kati Dege
Frau Bella Naftaliewa
Herr Dr. Jörg Betzin
Herr Stefan Biélka
Herr Martin Hanek

Das Institut dankt Frau Dr. Klein, Frau Merins, Frau Wolff, Herrn Bomke, Herrn Staffeldt, Herrn Dr. Ryll und den Teilnehmern der Arbeitsgruppe Grouperanpassung für die vielen wertvollen Hinweise und Anregungen.

Adresse und Kontakt

Institut des Bewertungsausschusses
Wilhelmstraße 138
10963 Berlin
info@institut-ba.de

Inhaltsverzeichnis

1	Einleitung.....	10
1.1	Auftrag des Instituts	10
1.2	Gegenstand der Analysen und Aufbau des Berichts.....	10
1.3	Zeitlicher Ablauf und Danksagung	11
2	Vollständigkeits- und Repräsentativitätsprüfung	12
2.1	Datengrundlagen des Bewertungsausschusses und seines Instituts.....	12
2.1.1	Kurze Darstellung der Versichertenstichprobe.....	12
2.1.2	Externe Datenquellen	18
2.1.3	Technische Aspekte der Datengrundlage	19
2.1.4	Mögliche Probleme mit der Datengrundlage	19
2.2	Technische Prüfung der Datenlieferungen an den Bewertungsausschuss.....	22
2.2.1	Bilaterale Datenfreigabe durch die Datenlieferanten	22
2.2.2	Datenfreigabe durch die AG IT	22
2.2.3	Ausschluss von Versicherten aus der Grouper-Schnittstelle	23
2.2.4	Ausschluss von Diagnosen aus der Grouper-Schnittstelle.....	24
2.3	Vollständigkeitsprüfungen.....	24
2.3.1	Anzahl der Versicherten	24
2.3.2	Leistungsbedarf	29
2.3.3	Unterjährig Versicherte	33
2.3.4	Ärztliche Inanspruchnahme.....	36
2.3.5	Kassenfusionen	37
2.4	Repräsentativitätsprüfungen.....	38
2.4.1	Stadt-/Land-Verteilung	39
2.4.2	Zuzugs-/Fortzugsregionen	40
2.4.3	Demographische Struktur	41
2.4.4	Externe Plausibilisierung der Relativgewichte.....	45
2.4.5	Externe Plausibilisierung der Veränderungsdaten	46
2.4.6	Externer Vergleich zum Niveau des Punktzahlvolumens.....	48
2.5	Fremdkassenzahlungsausgleich.....	48
2.5.1	Darstellung des Fremdkassenzahlungsausgleichs	49
2.5.2	Theoretische Einflüsse des fehlenden Fremdkassenzahlungsausgleichs	49
2.5.3	Auswertungen zum Einfluss des Fremdkassenzahlungsausgleiches	50
2.6	Entwicklung der Diagnosen je Versicherten	52
2.7	Prüfung der epidemiologischen Repräsentativität der regionalen Datengrundlage	53
2.8	Zusammenfassung der Vollständigkeits- und Repräsentativitätsprüfungen.....	54
2.9	Analysen zur Schätzgenauigkeit bei der Bestimmung von morbiditätsbedingten Veränderungsdaten.....	58
3	Berechnung von Veränderungsdaten.....	59
3.1	Hintergrund zur Modellierung.....	59
3.1.1	Darstellung der Varianten A, F, G und H.....	59

3.1.2	Erläuterung eines „kompletten“, „reduzierten“ oder „komprimierten“ Modells	60
3.1.3	Erläuterung „hierarchisiert“ vs. „nicht hierarchisiert“	62
3.1.4	Erläuterung „zeitgleich“ vs. „prospektiv“	65
3.1.5	Datengrundlage des Bewertungsausschusses	66
3.2	Kennzahlen und Gütemaße der Modellierung	69
3.2.1	Kennzahlen der Modellierung	70
3.2.2	Gütemaße der Modellierung	73
3.3	Berechnung der Veränderungsrate	75
3.3.1	Allgemeines Vorgehen	75
3.3.2	Informationsverluste bei prospektiven Modellen	77
3.3.3	Änderungen in der Klassifikation der CC 130	79
3.3.4	Ergebnisse der berechneten Varianten	79
3.4	Sonderauswertungen	80
3.4.1	Interregionale Stabilität von Relativgewichten	81
3.4.2	Intertemporale Stabilität der Relativgewichte	85
4	Das beschlossene Modell	89
4.1	Abgrenzung des Leistungsbedarfs zur Berechnung der Relativgewichte im beschlossenen Modell	89
4.2	Darstellung der Veränderungsrate nach Alter und Geschlecht	90
4.3	Darstellung des beschlossenen Modells zur Berechnung des diagnosenbezogenen Anteils	92
4.3.1	Klassifikation und Hierarchisierung im beschlossenen Modell	94
4.3.2	Kennzahlen	97
4.3.3	Gütemaße	100
4.3.4	Predictive Ratios	105
4.3.5	Prävalenzen der Risikoklassen	121
4.3.6	Relativgewichte	125
4.3.7	Vergleich des beschlossenen Modells einerseits mit einem kompletten Modell und andererseits mit einem komprimierten Modell mit einer anderen Leistungsabgrenzung	128
5	Ausblick auf die zukünftige Datenlieferungen	131
Anhang 1	Technische Aspekte der Datengrundlage	133
Anhang 2	Datensatzbeschreibungen der Sonderauswertung zum Fremdkassenzahlungsausgleich (Lieferung der Kassenseite)	140
Anhang 3	Datensatzbeschreibungen der Sonderauswertung zum Fremdkassenzahlungsausgleich (Lieferung der Kassenärztlichen Bundesvereinigung)	143
Anhang 4	Umfang der bilateralen Prüfungen zwischen Datengeber und Datenstelle	150
Anhang 5	Umfang der Plausibilitätsprüfungen von Datenstelle und AG IT	155
Anhang 6	Simulation zu Unterjährigkeit von Versicherten	158
A 6.1	Hintergrund und mögliche Auswirkungen	158
A 6.2	Unterjährigkeit in der Datengrundlage (3-Mio-Stichprobe)	159
A 6.3	Simulation der Auswirkungen von Unterjährigkeit auf die Regressionsergebnisse (Fehler 1)	159

A 6.3.1	Zeitgleiches Modell.....	159
A 6.3.2	Prospektives Modell	162
A 6.3.3	Zusammenfassung.....	165
A 6.4	Simulation der Auswirkungen von Unterjährigkeit auf die Berechnung des geschätzten LB / Morbiditätsindex (Fehler 2)	166
Anhang 7	Hierarchietabelle für reduzierte und komprimierte Modelle.....	168
Anhang 8	Nicht zur Diagnosekodierung verpflichtete Arztgruppen	169
Anhang 9	Ausschlussgründe und ihre Häufigkeiten	172
Anhang 10	Unterjährig Versicherte (Versichertenzeit < 356 Tage)	175
Anhang 11	Illustration des Problems fehlender Versichertendaten	176
Anhang 12	Verarbeitung von unterjährigen Versicherten.....	178
Anhang 13	Aufteilung der Altersgruppen in der KM6	180
Anhang 14	Vergleich der Altersstruktur in Versichertenstichprobe und KM6 (2006 und 2007).....	181
Anhang 15	Externer Vergleich der Relativgewichte bei einer Klassifikation nach Alter, Geschlecht, EM-Status.....	184
Anhang 16	Weitere Auswertungen zum Kapitel „Fremdkassenzahlungsausgleich“	185
Anhang 17	Überblick über die Erarbeitung der Varianten F und G	187
Anhang 18	Kalibrierungsstrategie für hierarchische komplette Modelle	188
Anhang 19	Kalibrierungsstrategie für nicht hierarchische komplette Modelle.....	192
Anhang 20	Kalibrierungsstrategie für reduzierte und komprimierte Modelle	194
Anhang 21	Berechnung der regionalen Relativgewichte mit dem durchschnittlichen Leistungsbedarf der 4 KVen.....	195
Anhang 22	Übersicht der Zuordnung von Risikokategorien (CC) zu ACC.....	197
Anhang 23	Schaubild zur Berechnung von Veränderungsraten	205
Anhang 24	Formeln zur Berechnung der Veränderungsrate	206
Anhang 25	Triggerung der CC 130 durch die folgenden GOPs:	208
Anhang 26	Korrelationskoeffizienten der RCC 24 und RCC 28 im beschlossenen Modell.....	209
Anhang 27	Extra Anhang MPE	210

Abbildungsverzeichnis

Abbildung 1:	Datenmodell der Versichertenstamm- und Abrechnungsdaten der Datenstelle des Bewertungsausschusses	15
Abbildung 2:	Datenlieferung an den Bewertungsausschuss	17
Abbildung 3:	Vergleich der Anteile der Altersklassen an der Gesamtpopulation zwischen der Stichprobe und der KM6 - Gesamt 2008	42
Abbildung 4:	Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Frauen – 2008	43
Abbildung 5:	Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Männer – 2008	44
Abbildung 6:	Vergleich der Relativgewichte der Versichertenstichprobe und des BVA	46
Abbildung 7:	Schaubild der Methodik der externen Plausibilisierung der Veränderungsraten	47

Abbildung 8: Beispiel eines Hierarchiebaums	63
Abbildung 9: Erstes Beispiel für Hierarchisierung im reduzierten und komprimierten Modell	64
Abbildung 10: Zweites Beispiel für Hierarchisierung im reduzierten und komprimierten Modell	65
Abbildung 11: Darstellung des Zeithorizontes bei der Schätzung des Leistungsbedarfs	66
Abbildung 12: Darstellung des Zeithorizonts bei der Berechnung der Kosten-/Relativgewichte	66
Abbildung 13: Formel zur Berechnung des Prognosefehlers	73
Abbildung 14: Formel zur Berechnung des adjustierten R^2	74
Abbildung 15: Versichertenkollektiv im Jahr 1 und Jahr 3	78
Abbildung 16: Versicherte zur Berechnung von Relativgewichten	78
Abbildung 17: Berechnung der Risikowerte der Versicherten	79
Abbildung 18: Hierarchisierung innerhalb der ACC 2 und ACC 3 im beschlossenen Modell	94
Abbildung 19: Hierarchisierung innerhalb der ACC 5 im beschlossenen Modell	95
Abbildung 20: Hierarchisierung innerhalb der ACC 8 im beschlossenen Modell	95
Abbildung 21: Hierarchisierung innerhalb der ACC 12 im beschlossenen Modell	96
Abbildung 22: Hierarchisierung innerhalb der ACC 19 im beschlossenen Modell	97
Abbildung 23: Hierarchisierung innerhalb der ACC 27 im beschlossenen Modell	97
Abbildung 24: Formel zur Berechnung des mittleren absoluten Prognosefehlers	101
Abbildung 25: Formel zur Berechnung des adjustierten R^2	101
Abbildung 26: Formel zur Berechnung des Prognosefehlers	103
Abbildung 27: Histogramm der Prognosefehler je Versicherten in der Validierungsstichprobe .	104
Abbildung 28: Histogramm der Prognosefehler je Versicherten in der Validierungsstichprobe .	105
Abbildung 29: Vergleich des tatsächlichen und geschätzten durchschnittlichen Leistungsbedarfs in Abhängigkeit von der Höhe der geschätzten Punkte	107
Abbildung 30: Vergleich des tatsächlichen und geschätzten durchschnittlichen Leistungsbedarfs in Abhängigkeit von der Höhe der geschätzten Punkte mit Zoom auf die 5 unteren Kostenklassen	108
Abbildung 31: Vergleich des tatsächlichen und geschätzten durchschnittlichen Leistungsbedarfs in Abhängigkeit von der Höhe der tatsächlichen Punkte	110
Abbildung 32: grafische Darstellung der Predictive Ratios je TCC	114
Abbildung 33: grafische Darstellung der Predictive Ratios je RCC	119
Abbildung 34: Pseudonymisierungsvorgang	136
Abbildung 35: Matching und Bestimmung aktueller Krankenkassenangaben	139
Abbildung 36: Differenzen der Kostengewichte zwischen den zeitgleichen Modellen in Punkten	161
Abbildung 37: Differenzen der Kostengewichte zwischen den prospektiven Modellen in Punkten	164
Abbildung 38: Differenzen der Kostengewichte zwischen den zeitgleichen und den prospektiven Modellen	165
Abbildung 39: Anzahl der quartalsweise übermittelten Versichertenstammdatensätzen am Beispiel einer Kasse	176
Abbildung 40: Zahl der Abrechnungsfälle je Quartal am Beispiel einer Kasse	177

Abbildung 41: Unterjährige Versicherte.....	178
Abbildung 42: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Gesamte Bevölkerung – 2006.....	181
Abbildung 43: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Frauen – 2006	181
Abbildung 44: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Männer – 2006	182
Abbildung 45: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Gesamte Bevölkerung – 2007.....	182
Abbildung 46: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Frauen – 2007	183
Abbildung 47: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Männer – 2007	183
Abbildung 48: Vergleich der Relativgewichte der Versichertenstichprobe und des BVA nach EM-Status	184
Abbildung 49: Vergleich des Anteils der Diagnosen aus den drei nicht-WOP-KV der Datengrundlage zur Summe der WOP-KV für vdek (gemäß Datengrundlage BA).....	185
Abbildung 50: Vergleich des Anteils der Diagnosen aus den drei nicht-WOP-KV der Datengrundlage zur Summe der WOP-KV für BKK (gemäß Datengrundlage BA) .	186
Abbildung 51: Schaubild zur Berechnung von Veränderungsraten am Beispiel einer zeitgleichen Modellierung.....	205

Tabellenverzeichnis

Tabelle 1: Anzahl Versicherte KM6 und Versichertenstichprobe	26
Tabelle 2: Versichertenzahl: Verhältnis Versichertenstichprobe zu KM6 ($VS/KM6 \geq 20\%$)	27
Tabelle 3: Versichertenzahl: Verhältnis Versichertenstichprobe zu KM6 ($VS/KM6 < 20\%$).....	29
Tabelle 4: Leistungsbedarf: Verhältnis Versichertenstichprobe zu C4	31
Tabelle 5: Leistungsbedarf: Verhältnis Versichertenstichprobe zu C4, MGV und EGV.....	32
Tabelle 6: Verteilung der unterjährig Versicherten (Versicherungszeit < 360 Tage)	34
Tabelle 7: Beispiele zu Implausibilitäten in der ärztlichen Inanspruchnahme.....	36
Tabelle 8: Verteilung Stadt/Land	40
Tabelle 9: Relative Veränderung der Versichertenzahl	41
Tabelle 10: Verteilung Frauen und Männer.....	42
Tabelle 11: Vergleich Morbiditätsindex zwischen KM6 und Versichertenstichprobe	44
Tabelle 12: Vergleich der Veränderungsraten zwischen KM6 und Versichertenstichprobe	45
Tabelle 13: Vergleich der Veränderungsraten: Stichprobe und externe Datenquellen	47
Tabelle 14: Vergleich Punktzahlniveau Bund mit dem Punktzahlvolumen der Stichprobe.....	48
Tabelle 15: Anteil des Leistungsbedarfs außerhalb der Datengrundlage am Leistungsbedarf je WOP-KV	51
Tabelle 16: Anteil der Diagnosen außerhalb der Datengrundlage an den Diagnosen je WOP-KV für Ersatzkassen und die Betriebskrankenkassen	51

Tabelle 17: Zusammenfassung der Ergebnisse zu Auswirkungen des Fremdkassenzahlungsausgleichs (Quelle: Sonderabfrage durch die Ersatzkassen (Diagnosen) und KBV (Leistungsbedarf)).....	52
Tabelle 18: Durchschnittliche Anzahl Diagnosen, CC und HCC je Versicherten (Bericht aus dem Jahr 2008).....	53
Tabelle 19: Durchschnittliche Anzahl der Diagnosen je Versichertenjahr.....	53
Tabelle 20: Zusammenfassung Problemfelder	55
Tabelle 21: Übersicht zu den Variante A, F, G und H.....	60
Tabelle 22: Ausgestaltung der beauftragten Varianten.....	70
Tabelle 23: Kennzahlen der berechneten reduzierten Varianten.....	71
Tabelle 24: Kennzahlen der berechneten kompletten Varianten.....	72
Tabelle 25: Überblick über negative RCC in reduzierten Modellen	73
Tabelle 26: Gütemaße der beauftragten Varianten.....	75
Tabelle 27: Übersicht über die Ergebnisse der berechneten Varianten	80
Tabelle 28: Kennzahlen zur Berechnung der regionalen Relativgewichte (komplettes Modell)	82
Tabelle 29: Abweichungen der regionalen Relativgewichte von den Relativgewichten der vier KVen (komplettes Modell)	83
Tabelle 30: Kennzahlen zur Berechnung der regionalen Relativgewichte (reduziertes Modell)	84
Tabelle 31: Abweichungen der regionalen Relativgewichte von den Relativgewichten der vier KVen (reduziertes Modell)	84
Tabelle 32: Regionale Veränderungsraten	85
Tabelle 33: Kennzahlen zur Berechnung der intertemporalen Stabilität der Relativgewichte.....	86
Tabelle 34: Intertemporale Stabilität von Relativgewichten (Berechnung der Relativgewichte 2008 mit dem durchschnittlichen Leistungsbedarf 2008).....	87
Tabelle 35: Intertemporale Stabilität von Relativgewichten (Berechnung der Relativgewichte 2008 mit dem durchschnittlichen Leistungsbedarf 2007).....	88
Tabelle 36: Relativgewichte für die Veränderungsrate nach Alter und Geschlecht	91
Tabelle 37: Kennzahlen zum beschlossenen Modell	98
Tabelle 38: Kennzahlen der beschlossenen Ausgestaltung im Vergleich mit weiteren Ausgestaltungen.....	100
Tabelle 39: Gütemaße der beschlossenen Ausgestaltung.....	102
Tabelle 40: Gütemaße der beschlossenen Ausgestaltung im Vergleich mit weiteren beauftragten Ausgestaltungen.....	103
Tabelle 41: Vergleich des durchschnittlichen tatsächlichen und geschätzten Leistungsbedarfs in Abhängigkeit von der Höhe der geschätzten Punkte.	106
Tabelle 42: Vergleich des durchschnittlichen tatsächlichen und geschätzten Leistungsbedarfs in Abhängigkeit von der Höhe des tatsächlichen Leistungsbedarfs.	109
Tabelle 43: Predictive Ratios je TCC.....	114
Tabelle 44: Predictive Ratios je RCC.....	119
Tabelle 45: Predictive Ratios für Alters- und Geschlechtsgruppen.....	121
Tabelle 46: Prävalenzen je Risikoklasse in der Lernstichprobe und den Validierungsstichproben	125
Tabelle 47: Kosten- und Relativgewichte je Risikoklasse in der beschlossenen Ausgestaltung ..	128
Tabelle 48: Kennzahlen des beschlossenen Modells im Vergleich mit einem kompletten Modell der gleichen Ausgestaltung.....	129

Tabelle 49: Gütemaße des beschlossenen Modells im Vergleich mit einem kompletten Modell der gleichen Ausgestaltung.....	129
Tabelle 50: Kennzahlen des beschlossenen Modells im Vergleich mit einem komprimierten Modell anderer Abgrenzung des Leistungsbedarfs.....	130
Tabelle 51: Gütemaße des beschlossenen Modells im Vergleich mit einem komprimierten Modell anderer Abgrenzung des Leistungsbedarfs.....	130
Tabelle 52: Datensatzbeschreibung zur „Abfrage zur Zahl der Diagnosen“	142
Tabelle 53: Datensatzbeschreibung zur „Abfrage der C4-Daten“	145
Tabelle 54: Datensatzbeschreibung zur „Abfrage der Daten der Ersatzkassen“	149
Tabelle 55: Durchschnittliche Anzahl Versichertentage	159
Tabelle 56: Perzentile der Anzahl Versichertentage.....	159
Tabelle 57: Vergleich eines zeitgleichen Modells der Stichprobe „3-Mio-2007“ mit einem zeitgleichen Modell der Stichprobe „3-Mio-2007 ganzjährig“	160
Tabelle 58: Ergebnisse des t-Tests zum Vergleich der Kostengewichte	161
Tabelle 59: Risikokategorien mit auffälligen Abweichungen im Kostengewicht zwischen den beiden zeitgleichen Modellen	162
Tabelle 60: Vergleich eines prospektiven Modells der Stichprobe „3-Mio-2006/2008 prosp“ mit einem prospektiven Modell der Stichprobe „3-Mio-2006/2008 prosp ganzjährig“	163
Tabelle 61: Ergebnisse des t-Tests zum Vergleich der Kostengewichte	163
Tabelle 62: Risikokategorien mit auffälligen Abweichungen im Kostengewicht zwischen den beiden prospektiven Modellen	165
Tabelle 63: Vergleich des geschätzten und tatsächlichen LB in den Stichproben mit Kostengewichten der Stichprobe „3-Mio-2007“	167
Tabelle 64: Vergleich des geschätzten und tatsächlichen LB in den Stichproben mit Kostengewichten der Stichprobe „3-Mio-2007 ganzjährig“.....	167
Tabelle 65: Verteilung der Ausschlussgründe.....	174
Tabelle 66: Verteilung der unterjährig Versicherten (Versichertenzeit < 356 Tage)	175
Tabelle 67: Anteil des Leistungsbedarfs außerhalb der Datengrundlage am WOP-Leistungsbedarf je KV (Quelle: Sonderabfrage durch die KBV).....	185
Tabelle 68: Abweichungen der regionalen Relativgewichte von den Relativgewichten der vier KVen berechnet mit dem durchschnittlichen Leistungsbedarf der 4 KVen (komplettes Modell).....	195
Tabelle 69: Abweichungen der regionalen Relativgewichte von den Relativgewichten der vier KVen berechnet mit dem durchschnittlichen Leistungsbedarf der 4 KVen (reduziertes Modell).....	196
Tabelle 70: GOPs, die CC130 triggern	208
Tabelle 71: Korrelationskoeffizienten der RCC 24 und RCC 28 im beschlossenen Modell	209
Tabelle 72: Prognosefehler der beschlossenen Ausgestaltung im Vergleich mit weiteren beauftragten Ausgestaltungen.....	210
Tabelle 73: Predictive Ratios je TCC.....	212
Tabelle 74: Predictive Ratios je RCC.....	213
Tabelle 75: Predictive Ratios für Alters- und Geschlechtsgruppen.....	214

1 Einleitung

1.1 Auftrag des Instituts

Der Bewertungsausschuss hat in seiner 184. Sitzung am 20. Mai 2009 das Institut des Bewertungsausschusses beauftragt, Berechnungen und Überprüfungen zu den folgenden Themenkomplexen vorzunehmen:

- Berechnung von Veränderungsraten der morbiditätsbedingten Leistungsmenge (im Folgenden nur mit „Veränderungsrate“ bezeichnet) mithilfe des Klassifikationssystems
- Bewertung der interregionalen und intertemporalen Stabilität der Relativgewichte
- Berechnung von kassen- und KV-spezifischen Veränderungsraten nach Alter und Geschlecht
- Überprüfung der Datengrundlage auf Repräsentativität, Vollständigkeit und epidemiologische Aspekte.

1.2 Gegenstand der Analysen und Aufbau des Berichts

Die Darstellung der Datengrundlage des Bewertungsausschusses, einschließlich der Überprüfung auf Vollständigkeit und Repräsentativität, erfolgt im Abschnitt 1. Die epidemiologischen Aspekte werden in einem separaten Bericht dargestellt, hier wird nur kurz auf die Hauptschlussfolgerungen eingegangen (s. Abschnitt 2.6).

Eine Erläuterung zum Vorgehen bei der Berechnung von morbiditätsbedingten Veränderungsraten erfolgt am Anfang des Abschnitts 1. Die Darstellung von Veränderungsraten verschiedener Modelle sowie Kennzahlen zu den Modellen folgen in den weiteren Abschnitten. Im Abschnitt 3.4 erfolgt eine Sonderauswertung zur interregionalen und zeitlichen Stabilität der Relativgewichte. Die Darstellung der Berechnung von Kassen- und KV-spezifischen Veränderungsraten nach Alter und Geschlecht wurden in einer separaten Präsentation zur Simulation NVV 2010 berichtet

Im Abschnitt 1 wird das vom 15. Erweiterten Bewertungsausschuss beschlossene Modell zur Berechnung der morbiditätsbedingten Veränderungsrate dargestellt.

1.3 Zeitlicher Ablauf und Danksagung

Kurz nach der Beauftragung durch den BA stand der erste Teil der Datengrundlage zur Verfügung (ergänzte und zum Teil korrigierte Daten zu Versicherten und vertragsärztlichen Abrechnungsdaten auf Stichprobenbasis). Dieses Datenmaterial umfasst die Jahre 2006 und 2007.

Die Fortschreibung dieses Datenbestandes bis ins vierte Quartal 2008 wurde am Ende des Juli 2009 technisch abgenommen und war in auswertbarer Form (sog. „Grouper-Schnittstelle“) am 10. August 2009 fertiggestellt.

Einige der in Abschnitt 2.1.2 genannten weiteren Datenquellen wurden erst während der Berichtserstellung erschlossen.

Aus diesem Grund konnten nicht alle geplanten oder denkbaren Analysen für diesen Bericht abschließend durchgeführt werden, an den entsprechenden Stellen werden Abschätzungen gegeben. Mit dieser Berichtslegung kann somit noch nicht gewährleistet werden, dass die Darstellung der technischen Probleme abschließend wäre.

Dafür, dass dennoch derart umfangreiche Analysen in derart kurzer Zeit fertiggestellt werden konnten, sei den Beteiligten in der Datenstelle des BA (Colt Telecom, Trivadis), in der AG IT (GKV-Spitzenverband, Kassenärztliche Bundesvereinigung, Kassenverbände und Einzelkassen) sowie allen internen und externen Mitarbeitern des Instituts aufs Herzlichste gedankt.

2 Vollständigkeits- und Repräsentativitätsprüfung

Im ersten Teil dieses Abschnitts werden die Datengrundlage des Bewertungsausschusses sowie verwendete externe Datenquellen dargestellt. Nach der Beschreibung der technischen Prüfungen in 2.2, werden in 2.3 und 2.4 die Ergebnisse der Vollständigkeits- und Repräsentativitätsprüfung berichtet. Die Entwicklung der Fremdkassenzahlungsausgleiche, der Diagnosen je Versicherten sowie die epidemiologische Repräsentativität der regionalen Datengrundlage werden in den Abschnitten 2.5 bis 2.7 dargestellt. In 2.8 erfolgt eine Zusammenfassung der Datenprüfungen.

2.1 Datengrundlagen des Bewertungsausschusses und seines Instituts

Ziel dieses Abschnittes ist es, einen Überblick über die Datengrundlage des Bewertungsausschusses zu geben. Dabei werden die Bearbeitungsschritte zur Herstellung eines bearbeitbaren Datensatzes und mögliche Probleme bei diesen Schritten erläutert.

2.1.1 Kurze Darstellung der Versichertenstichprobe

An die Datenstelle des Bewertungsausschusses werden von den Trägerorganisationen des Bewertungsausschusses pseudonymisierte Versichertenstammdaten, Kostenträgerstammdaten sowie ebenfalls pseudonymisierte vertragsärztliche Abrechnungsdaten übermittelt (siehe Beschluss des Bewertungsausschusses, 174. Sitzung, 2009).

Die Übermittlung bezieht sich auf eine Vollerhebung von arzt- und patientenbezogenen Abrechnungsdaten einerseits und versichertenbezogenen Kassendaten andererseits gemäß der in dem Beschluss des Bewertungsausschusses aus der 89. Sitzung Abschnitt C 2. vom 13. Mai 2004 festgelegten vier KV-Bereiche und umfasst aktuell den Zeitraum 1/2006 bis 4/2008.

Übermittelt werden die Versichertenstammdaten für die Versicherten, die in den Bereichen der Kassenärztlichen Vereinigungen Bremen, Niedersachsen, Nordrhein und Thüringen in dem jeweiligen Erhebungsjahr wohnen. Die Datenlieferungen erfolgen quartalsbezogen.

Die Datenlieferungen an den Bewertungsausschuss bestehen aus den folgenden Datensätzen:

Satzart 090 – Versicherten-Pseudonyme:

Für jeden Versicherten wird eine PersonID vergeben. Zusätzlich wird jeder Krankenkasse eines Versicherten dauerhaft eine Betriebsnummer zugeordnet. Maßgeblich ist diejenige Krankenkasse, bei der der Versicherte zum Zeitpunkt des erstmaligen Einbezugs versichert war. Die PersonID und die Betriebsnummer einer Krankenkasse identifizieren einen Versicherten eindeutig.

Satzart 091 – Versicherten-Stammdaten:

Für jeden Versicherten wird je Quartal, in dem er versichert war, ein Datensatz geliefert. Die Stammdaten des Versicherten (Geburtsjahr, Geschlecht) werden dort wiederholt.

Satzart 092 – Ambulante Abrechnungen:

(KV-Fall): Berücksichtigt werden alle Behandlungsfälle von Versicherten gemäß SGB V des jeweiligen Leistungsquartals der Praxen des jeweiligen KV-Bereichs.

Satzart 093 – Diagnosen der ambulanten Behandlung:

(KV-Fall-Diagnosen): Je Diagnose der Behandlungsfälle aus Satzart 092 (KV-Fall) wird ein Datensatz geliefert.

Satzart 100 – Gebührennummern der ambulanten Behandlung:

Für jede verschiedene Gebührennummer der Behandlungsfälle aus Satzart 092 (KV-Fall) wird ein Datensatz geliefert.

Satzart 101 – Praxis-/Betriebsstättenverzeichnis:

Berücksichtigt und geliefert werden alle Praxen (für Quartal 1/2006 bis 2/2008) bzw. Betriebsstätten und Nebenbetriebsstätten (ab 3. Quartal 2008) des jeweiligen Leistungsquartals des jeweiligen KV-Bereichs. Je Leistungsquartal wird je Praxis bzw. (Neben-) Betriebsstätte des jeweiligen KV-Bereichs ein Datensatz geliefert.

Satzart 102 – Arztverzeichnis:

Berücksichtigt und geliefert werden die Ärzte des jeweiligen Leistungsquartals der Praxen (für Quartal 1/2006 bis 2/2008) bzw. Betriebsstätten (ab 3. Quartal 2008) des jeweiligen KV-Bereichs. Je Leistungsquartal wird je Arzt-/Praxis-(Betriebsstätten-) Kombination des jeweiligen KV-Bereichs ein Datensatz geliefert.

Satzart 103 – ICD-10-Stammdatei:

Berücksichtigt werden ICD-10-Diagnosen, die im Rahmen der Abrechnung im jeweiligen Kalenderjahr verwendet werden können und nach SGB V zulässig sind.

Satzart 105 – Gebührennummern-Stammdatei:

Für jede verschiedene Gebührennummer, die in einem Quartal des jeweiligen KV-Bereichs gültig ist, wird ein Datensatz geliefert.

Satzart 108 – Kostenträgerumschlüsselung:

Da aus technischen Gründen zu einer natürlichen Person nicht in allen Fällen in der Satzart 090 das auf der Krankenversichertenkarte gespeicherte Institutionskennzeichen geliefert werden konnte bzw. kann, muss das in der Satzart 092 gelieferte Krankenversichertenkarten-Institutionskennzeichen entsprechend dieser Satzart 108 umgeschlüsselt werden, um eine erfolgreiche Zusammenführung von Stammdaten und Leistungsdaten zu gewährleisten.

Satzart 109 – Kostenträgerverzeichnis (KT-Stamm):

Die Datenlieferanten teilen in dieser Satzart in Form einer Austauschlieferung für ihren Bereich jeweils für den gesamten Zeitraum ab 2004 zu jedem in Deutschland auf einer Krankenversichertenkarte vermerkten KVK-IK das zugehörige Kassensitz-IK zusammen mit einem Gültigkeitszeitraum mit.

Satzart 110 – Kostenträgerhistorie:

Zu jedem in der Satzart 109 gelieferten Kassensitz-IK wird in dieser Satzart 110 durch die Datenlieferanten in Form einer Austauschlieferung jeweils für den gesamten Zeitraum ab 2004 zusammen mit einem Gültigkeitszeitraum der Name und die Kassenart des Kostenträgers sowie ein eventueller Rechtsnachfolger geliefert.

Satzart 120 – Zuordnung der Abrechnungsnummerpseudonyme zu Betriebsstättenpseudonymen für das 3. Quartal 2008:

Zu jedem in der Satzart 101 für das 2. Quartal 2008 gelieferten Praxispseudonym werden diejenigen Betriebsstättenpseudonyme geliefert, welche derselben Praxis im Rahmen der Initialvergabe im 3. Quartal 2008 zugeordnet wurden.

In Abbildung 1 sind die Beziehungen zwischen den einzelnen Satzarten dargestellt.

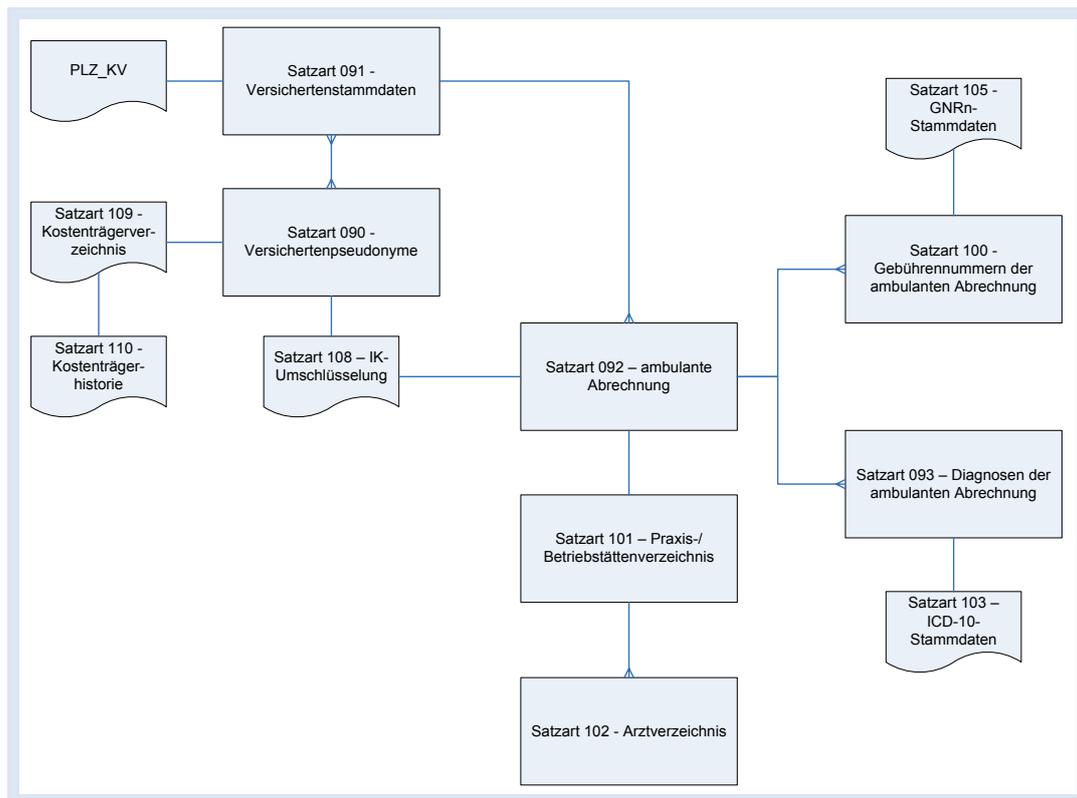


Abbildung 1: Datenmodell der Versichertenstamm- und Abrechnungsdaten der Datenstelle des Bewertungsausschusses

Um den Anforderungen des Datenschutzes gerecht zu werden, müssen die Versichertenstammdaten und die ambulanten ärztlichen Abrechnungsdaten pseudonymisiert zur Verfügung gestellt werden. Insbesondere sind die Personenmerkmale Krankenversicherungsnummer (KVK-KVNR) bzw. lebenslange Arztnummer (LANR) und das Organisationsmerkmal Abrechnungsnummer, Betriebsstättennummer bzw. Nebenbetriebsstättennummer (ANR, BSNR, NBSNR) nach einem vom Bundesamt für die Sicherheit in der Informationstechnik vorgegebenen sicheren Verfahren zu pseudonymisieren.

Die von den Krankenkassen erhobenen Versichertenstammdaten werden von den ehemaligen Spitzenverbänden der Krankenkasse gesammelt, über den GKV-Spitzenverband erneut pseudonymisiert und an die Datenstelle des Bewertungsausschusses übermittelt.

Für jeden Versicherten wird eine über den gesamten Zeitraum aller Datenlieferungen hinweg konstante PersonID geliefert. Dieser PersonID können mehrere pseudonymisierte Versichertennummern zugeordnet sein. Personen, welche

durch Umzug eine neue KVK-IK bekommen oder durch Statuswechsel eine neue Versichertennummer erhalten, sollen durch diese PersonID in Verbindung mit der eingefrorenen Betriebsnummer der Krankenkasse zusammengefasst werden. Die Daten werden so erhoben, dass mehrfache Zuordnungen von jeder gültigen Kombination aus KVK-IK und pseudonymisierter Versichertennummer zu einer Kombination aus einer PersonID und Betriebsnummer der Krankenkasse führen kann. Damit sind mehrere Datensätze je Versicherten möglich.

Von der Datenstelle des Bewertungsausschusses werden im Rahmen der Datenverarbeitung die Attribute KVK-IK und pseudonymisierte Versichertennummer durch die Kombination PersonID und Betriebsnummer ergänzt und anschließend um einen neuen kassenartenübergreifenden Surrogatschlüssel für den Versicherten durchnummeriert.

Die Kassenärztlichen Vereinigungen übermitteln ihre pseudonymisierten Abrechnungsdaten an die KBV. Die Pseudonymisierung erfolgt für die Krankenversicherungsnummer (KVK-KVNR), das Organisationsmerkmal Abrechnungsnummer, Betriebsstättennummer bzw. Nebenbetriebsstättennummer (ANR, BSNR, NBSNR) sowie (ab 3/2008) für die lebenslange Arztnummer (LANR).

Die KBV pseudonymisiert die Daten ein zweites Mal und gibt die Daten ebenfalls an die Datenstelle des Bewertungsausschusses weiter.

Die Datenlieferung ist zwischen den Datenlieferanten so abgestimmt, dass im Zuge der Datenerhebung dieses zweistufige Pseudonymisierungsverfahren umgesetzt ist. Sowohl bei der Datenerhebung (Krankenkasse, Kassenärztliche Vereinigung), als auch bei der Stelle der Datensammlung und Zusammenführung (ehem.) Spitzenverband, Kassenärztliche Bundesvereinigung) erfolgen Pseudonymisierungen der personen- oder organisationsspezifischen Merkmale. Die Kennworte, mittels derer die Pseudonymisierung durchgeführt wird, sind nur den jeweils auf der Stufe der Datenerhebung beteiligten Parteien bekannt.

Abbildung 2 zeigt den schematischen Verlauf der Datenerhebung:

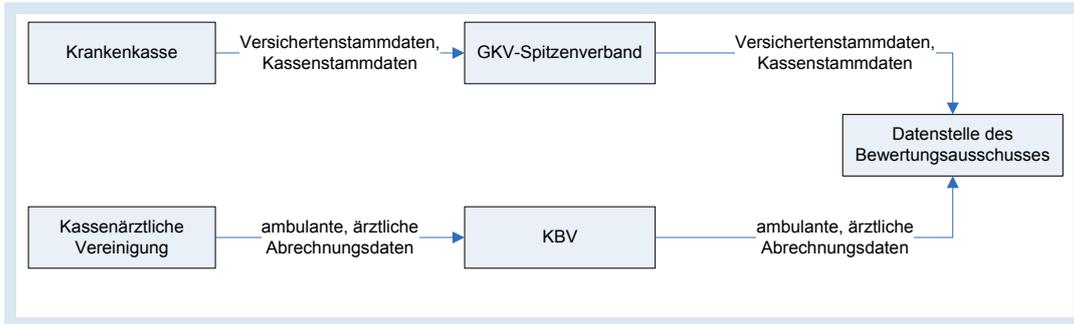


Abbildung 2: Datenlieferung an den Bewertungsausschuss

Der Datenstrom ist so organisiert, dass die Zusammenführung von Versichertenstammdaten der Krankenkassen mit den ambulanten ärztlichen Abrechnungsdaten in der Datenstelle des Bewertungsausschusses erfolgt. In Anhang 1 werden technische Aspekte der Datengrundlage sowie die Zuordnung von Versichertenstammdaten zu ärztlichen Leistungsdaten detaillierter dargestellt. Näheres hierzu ist in den technischen Dokumenten der Beschlüsse nachzulesen (siehe Beschluss des Bewertungsausschusses zu Datenlieferungen für das Jahr 2008 gemäß §§ 87, 87a-c SGB V an die Datenstelle des Bewertungsausschusses (174. Sitzung); Beschluss des Bewertungsausschusses zu Datenlieferungen für das Jahr 2007 an die Datenstelle des Bewertungsausschusses (148. Sitzung) Beschluss des Bewertungsausschusses zur Schaffung einer kurzfristigen Datengrundlage gemäß §§ 87 und 87a bis 87c SGB V (Teil C mit Wirkung zum 1. Juli 2007) (131. Sitzung); Arbeitsaufträge an die Datenstelle des BA).

In der Datenstelle werden die pseudonymisierten Daten vorgehalten, nach Vorgabe und in Zusammenarbeit mit der AG IT des Bewertungsausschusses plausibilisiert, freigegeben und technisch für die Verarbeitung durch das Institut des Bewertungsausschusses aufbereitet.

Aus den zusammengeführten und plausibilisierten Daten wird dem Institut des Bewertungsausschusses eine Stichprobenmenge von 33% aller Versichertenstamm- und den diesen Versicherten zugeordneten ambulanten ärztlichen Abrechnungsdaten in einer eigenen Datenbank zur Verfügung gestellt (4 KVen Stichprobe). Diese Daten stellen die Grundlage für die Datenverarbeitung des Instituts des Bewertungsausschusses dar. Die Daten enthalten auch Versicherte, bei denen keine ärztlichen Leistungen angefallen sind.

2.1.2 Externe Datenquellen

An verschiedenen Stellen der Bearbeitungen wurden für externe Plausibilisierungen oder Berechnungen von Morbiditätsveränderungen externe Daten herangezogen, die im Folgenden kurz erläutert werden sollen.

Daten der KM 6

- Statistik über die Zahl der Versicherten in der GKV jeweils zum Stichtag 1. Juli des Jahres.
- Daten sind gegliedert nach Einzelkasse, KV (nach Wohnortprinzip); Geschlecht und 17 Altersgruppen.
- Die Daten wurden für die Jahre 2006 bis 2008 verwendet.
- Durch das BMG bereitgestellt.

Bericht an das BMG

- Jährlicher Bericht des Bewertungsausschusses an das BMG über die Entwicklung der vertragsärztlichen Versorgung.
- Verwendet wurde die Version des Jahres 2007

C4-Daten

- Datengrundlage des Bewertungsausschusses zur Berechnung des erstmaligen Behandlungsbedarfs und zur Korrektur des erstmaligen Behandlungsbedarfs gemäß § 87c Abs. 4 SGB V
- Je Gebührenordnungsposition liegen Werte zur Frequenz, zum Leistungsbedarf vor Abstufung und nach Abstufung für die Jahre 2007 und 2008 vor.
- Die Datengrundlage ist getrennt nach Kasse und KV und wurde deutschlandweit erhoben.

Sonderauswertung zum Fremdkassenzahlungsausgleich

- Sonderauswertung auf den Daten der Ersatzkassen, der Betriebskrankenkassen und der Kassenärztlichen Bundesvereinigung
- Auswertung umfasst die Daten 2006 bis 2008. Da bei den meisten Ersatzkassen und einer Reihe von Betriebskrankenkassen schon in allen drei Jahren das Wohnortprinzip umgesetzt war, kann bei ihnen eine klare Trennung zwischen Wohnort-KV des Versicherten und Leistungserbringungs-KV erfolgen. Die ge-

lieferten Daten umfassen jeweils nur die Kassen, die das Wohnortprinzip umgesetzt hatten.

- Datensatzbeschreibungen zu den Datenlieferungen von Kassenseite finden sich im Anhang 2.
- Datensatzbeschreibungen zu den Datenlieferungen von der Kassenärztlichen Bundesvereinigung finden sich im Anhang 3.

Daten des BVA

- Versichertentage nach SA 40 je RSA-Gruppe sowie errechnete Kosten pro Versichertentag (Euro-pro-Tag Werte) je RSA-Gruppe
- alle am RSA teilnehmenden Krankenkassen
- Verwendet wurden die Daten der Jahre 2006 und 2007

Daten des Statistischen Bundesamtes

- Kreisfreie Städte und Landkreise nach Fläche und Bevölkerung auf Basis des Gemeindeverzeichnisses
- Informationen wurden zum Gebietsstand 31.12. des jeweiligen Jahres erhoben
- Verwendet wurden die Daten der Jahre 2006 und 2007

2.1.3 Technische Aspekte der Datengrundlage

(Die Ausführungen zu den technischen Aspekten der Datengrundlage finden sich in Anhang 1.)

2.1.4 Mögliche Probleme mit der Datengrundlage

Der folgende Abschnitt soll einen Überblick über mögliche Probleme der Datengrundlage geben.

Pseudonyme der Versicherten:

Da in den vergangenen Jahren mehr als 200 Krankenkassen und vier Kassenärztliche Vereinigungen unabhängig voneinander Pseudonyme für die Versicherten vergeben haben, ist die Zusammenführbarkeit der Angaben zum Versicherten mit seinen ärztlichen Angaben stets eine mögliche Schwachstelle im Verfahren.

Ordnungsnummern der Krankenkassen:

Die Krankenkassen geben in den übermittelten Versichertendaten als Ordnungsnummern u. a. Institutionskennzeichen (IKs) an, zum Teil regionsspezifische, zum Teil aus organisatorischen Gründen besondere IKs, insbesondere bei Fusionen im Zeitverlauf wechselnde. Die Kassenärztlichen Vereinigungen geben unabhängig davon ebenfalls IKs an.

Ein systematischer Aufbau von Überführungslisten in der Datenstelle des BA (sog. Satzarten 108 ff.) hat im Jahr 2008 begonnen, der Stand der Listen ist noch nicht befriedigend. Das Institut hat diese Listen mit Hilfe von Angaben des Deutschen Ärzteverlags, des BMG, der KBV und des GKV-SV überarbeitet, korrigiert und komplettiert. Eine Vollständigkeit ist damit noch nicht gewährleistet.

Unvollständigkeiten in der Dokumentation der ärztlicher Leistungen und der Diagnosen:

Da vertragsärztliche Abrechnungsdaten nur aus vier KV-Bereichen zur Verfügung gestellt werden, sind hinsichtlich der Versicherten aus diesen vier Regionen diejenigen Leistungen und Diagnosen nicht verfügbar, die außerhalb der vier KV-Bereiche erbracht bzw. gestellt wurden.

In Fällen des sog. „Ersatzverfahrens“ (die Chipkarte liegt bei Behandlung nicht vor) sind die Versichertennummer (und damit das Pseudonym) nicht verfügbar, diese Fälle können dem Versicherten in der Datengrundlage nicht mehr zugeordnet werden.

Beide Konstellationen führen zur Unvollständigkeit der Dokumentation zum Versicherten.

Fusionen zum Jahreswechsel:

Fusionieren Kassen im Erhebungszeitraum, so kann das Problem auftreten, dass die anfusionierte Kasse nicht nur durch ein neues KVK-IK gekennzeichnet ist, sondern eventuell auch für die Versicherten dieser Kasse völlig neue PersonIDs vergeben werden müssen. Der Versicherte wird dadurch im Datenkörper als ein anderer, neuer Versicherter geführt. Dies hat zur Folge, dass solche Versicherte nicht mehr für Auswertungen mit prospektiven Modellen herangezogen werden können, da bei diesen Modellen Daten aus zwei unterschiedlichen Jahren herangezogen werden müssen. Es entsteht ein Schwund in der statistischen Masse.

Unterjährige Fusionen:

In diesem Fall kann es ebenfalls vorkommen, dass die PersonIDs neu vergeben werden, d. h., der Versicherte wird während eines Jahres mit zwei PersonIDs im Datensatz geführt, die jeweils für den gesamten Zeitraum eines Jahres jeweils unvollständig sind. In der Dokumentation der Krankheiten des Versicherten entsteht ein Bruch.

Verzögerte Chipkartenversendung:

Falls eine Krankenkasse im Zuge einer Fusion die neuen Chipkarten nicht absolut zeitgleich mit der Fusion ausgibt (und der Versicherte die neue Karte auch sofort einsetzt) hat dies im Datensatz zur Folge, dass die Abrechnungsdaten (mit alter Chipkarte und alter PersonID vor Fusion) nicht mit den Versichertenstammdaten gematcht werden können. Einerseits laufen die Abrechnungsdaten „ins Leere“ und andererseits sinkt die Inanspruchnahme des Versicherten mit seinen neuen Versichertenstammdaten. Im weiteren Verlauf ist zu erwarten (nach Kartenausgabe und Karteneinsatz durch den Versicherten), dass der Leistungsbedarf wieder ansteigt.

Kassenwechsler:

Wechselt ein Versicherter die Kasse, ändert sich die KVK-IK, die pseudonymisierte Versichertennummer, die PersonID und die Betriebsnummer der Kasse. Der Versicherte wird so im Datensatz zu einer anderen Person (bzw. ggf. mit jeweils entsprechenden Versichertenzeiten während eines Jahres zu zwei unterjährig Versicherten).

Umzug, Wechsel in die PKV:

Versicherte, die durch Umzug ein neues KVK-IK erhalten, können über die Person ID zusammengefasst werden, allerdings nicht, wenn der Versicherte eine der vier Erhebungs-KVen verlässt (Umzug in andere KV oder ins Ausland oder Wechsel in die PKV). Dies hat zur Folge, dass der Versicherte im Datenbestand zu einem unterjährig Versicherten wird.

Statuswechsel:

Bei einigen Kassen besteht die Möglichkeit, dass es bei Statuswechsel des Versicherten zur Vergabe einer neuen Versichertennummer und einer neuen PersonID kommt. Auch hier werden aus einer physischen Person zwei unterjährig Versicherte erzeugt.

Alle eben genannten Faktoren, die die Aufspaltung einer physischen Person in zwei oder mehrere Versicherte im Datensatz mit sich bringen, haben zur Folge, dass im Datensatz mehr „Versicherte“ vorhanden sind, als anhand vergleichbarer Statistiken zu erwarten sind. Ein Problem entsteht hierdurch zunächst nicht, es wird routinemäßig mit dem Konzept der „Versichertenjahre“ gearbeitet.

Auf problematische Aspekte dieser Aufspaltungen geht Abschnitt 2.3.3 ein.

2.2 Technische Prüfung der Datenlieferungen an den Bewertungsausschuss

Im folgenden Abschnitt wird die technische Prüfung der Datengrundlage des Bewertungsausschusses dargestellt.

2.2.1 Bilaterale Datenfreigabe durch die Datenlieferanten

In einem ersten Schritt werden die Datenlieferungen der Kassen bzw. ihrer Verbände sowie der KBV bilateral zwischen dem jeweiligen Datenlieferanten und der Datenstelle überprüft (vgl. Qualitätssicherung der Satzarten 90 und 91 sowie der Satzarten 108 bis 110; Qualitätssicherung Satzart 92 bis 105). So soll sichergestellt werden, dass bei der Datenübergabe und beim Ladevorgang in der Datenstelle keine Differenzen aufgrund von technischen Gegebenheiten auftreten. Der Umfang der Prüfungen ist in Anhang 4 beschrieben.

2.2.2 Datenfreigabe durch die AG IT

In einem zweiten Schritt im Rahmen der technischen Prüfung der Datenlieferungen an die Datenstelle des Bewertungsausschusses hat die AG IT einen Katalog mit Prüfaufträgen für die Datenstelle des Bewertungsausschusses erstellt. Anhand der Prüfergebnisse wird der Datenkörper freigegeben oder bei Unstimmigkeiten der Prüfergebnisse weitere Zusatzauswertungen oder ggf. Neulieferungen der Daten beauftragt. Es werden Auswertungen zum zusammengeführten („gematchten“) und auch zum nicht-gematchten Datenbestand sowohl für die Vollerhebung als auch für die Stichprobenmenge durchgeführt (vgl. Anhang 5 auch Arbeitsaufträge an die Datenstelle des BA). Für die Stichprobenmenge werden die gleichen Auswertungen erstellt wie für den gematchten Datenbestand der Vollerhebung.

2.2.3 Ausschluss von Versicherten aus der Grouper-Schnittstelle

Die Datenlieferungen der Versichertenstammdaten an die Datenstelle des Bewertungsausschusses sind nicht immer fehlerfrei. So müssen in den Versichertenstammdaten (Satzart 091) aus verschiedenen Gründen Versicherte ausgeschlossen werden, weil die zu ihnen vorliegenden Versichertenstammdatensätze definierte Mindestanforderungen an die Datenqualität nicht erfüllen. Um Versichertenstammdatensätze auszuschließen, werden diese in der Datenbank nicht physikalisch gelöscht, sondern durch sogenannte Delete-Flags „markiert“.

Bei der Nutzung der Versichertenstichprobe für die Berechnung von morbiditätsbedingten Veränderungsdaten führen diese Löschemarkierungen dazu, dass die entsprechenden Datensätze nicht in die sog. Grouper-Schnittstelle übernommen werden.

Eine der wichtigsten Gründe für die Nichtverwendung liegt dann vor, wenn die Versicherten einer Kasse über eine längere Periode hinweg überhaupt keine zurechnungsfähigen ärztlichen Abrechnungen aufweisen. Dieser Sachverhalt kann insbesondere dadurch entstehen, dass

- die Kasse in einer Datenlieferung ihre Versicherten fehlerhaft pseudonymisiert hat oder
- die Kasse (ggfs. der Verband) der Datenstelle das aktuelle IK nicht gemeldet hat.

In diesem Fall wird die Kasse mit sämtlichen Versicherten für den gesamten Auswertungszeitraum (2006–08) von den weiteren Berechnungen ausgeschlossen. Damit wird verhindert, dass die Kasse im Zeitverlauf ihre Morbiditätsstruktur „wechselt“ (Versicherte ohne Inanspruchnahme weisen keine Diagnosen auf und erhalten für diesen Zeitraum einen Morbi-Index, der im Extremfall Null beträgt).

Ein periodenspezifischer Ausschluss von Versicherten wäre ebenfalls nachteilig für die morbiditätsbedingten Berechnungen, da die Grouper-Schnittstelle dann eine im Zeitverlauf wechselnde Zusammensetzung von Krankenkassen aufweisen würde. Auch dann wären verzerrende Effekte auf die Morbiditätsstruktur und Veränderungsrate möglich, ohne dass sie genau bezifferbar wären.

In der Mehrzahl der anderen „Löschemgründe“ wird ganz analog verfahren.

2.2.4 Ausschluss von Diagnosen aus der Grouper-Schnittstelle

Diagnosen von nicht zur Diagnosekodierung verpflichteten Ärzten werden in der Grouper-Schnittstelle ausgeschlossen und daher bei der Berechnung der Veränderungsdaten nicht berücksichtigt. (Eine Auflistung der Arztgruppen findet sich im Anhang 6). Darüber hinaus werden Diagnosen, die keine Arztgruppenkennung tragen und Diagnosen, die nicht der standardisierten Form entsprechen (ein Buchstabe gefolgt von zwei bis vier Zahlen), bei der Erstellung der Grouper-Schnittstelle nicht berücksichtigt.

2.3 Vollständigkeitsprüfungen

Wie bereits in Kapitel 2.1.1 beschrieben, wurde die Datengrundlage auf Basis einer Stichprobe der 4 KVen erstellt. Um die Vollständigkeit dieser Stichprobe zu untersuchen, werden im Folgenden Vergleiche mit den KM6- und C4-Daten durchgeführt (Erläuterungen der KM6- und C4-Daten in Kapitel 2.1.2). Die Kriterien, anhand derer die Vollständigkeit der Stichprobe überprüft wird, sind der Versichertenbestand, der Leistungsbedarf, die mögliche Unterjährigkeit der Versicherung, Kassenfusionen sowie ärztliche Inanspruchnahme.

Sowohl für die Vollständigkeits- als auch für die Repräsentativitätsprüfung ist zu bedenken, dass bei der Versichertenstichprobe mit insgesamt etwa 6 Millionen Versicherten die Nichterfassung von 60.000 Versicherten bereits eine Veränderung des Morbiditätsindex um 1% auslösen kann. Daraus resultiert die Notwendigkeit der Analysegenauigkeit, d. h., es ist eine Fehlerquote von weniger als 1% zu erreichen. Für die Vollständigkeits- und Repräsentativitätsuntersuchung bedeutet das, dass die Abweichungen zwischen der Versichertenstichprobe und den KM6-Daten dementsprechend gering sein müssen.

2.3.1 Anzahl der Versicherten

Zur Überprüfung der Vollständigkeit wird zuerst der Versichertenbestand in der auswertungsfähigen Versichertenstichprobe mit dem Versichertenbestand der KM6-Statistik verglichen. Ausgehend von der KM6-Statistik werden zunächst die Versichertenzenzahlen subtrahiert, die zu Kassen gehören, die für die Datenstelle des Bewertungsausschusses keine Daten zur Verfügung stellen. Dieser Sachverhalt trifft für alle Jahre auf einzelne AOK Kassen und für die Jahre 2006 und 2007 auch auf die Seekasse zu. Nach der Fusion der Seekasse mit der Bundesknappschaft zu Beginn des Jahres 2008 werden auch die Versichertenzenzahlen der See-

kasse als Untermenge der Knappschaft Bahn/See berücksichtigt. Aufgrund dieser Voraussetzung und der Festlegung, dass in der Versichertenstichprobe ein Drittel der Versicherten abgebildet sein müssen, wird eine leicht steigende Zahl an Versicherten in der Stichprobe von 5.802.345 (2006), 5.803.931 (2007) und 5.805.641 (2008) erwartet.

Demgegenüber werden durch die Datenstelle des Bewertungsausschusses mehr Versichertenstammdatensätze, als auf Basis der KM6-Statistik gefordert, zur Verfügung gestellt. Es ergibt sich hierbei die folgende Reihe 6.136.774 (2006), 6.066.150 (2007) und 5.986.334 (2008). Von diesen übergebenen Versichertenstammdatensätzen werden 163.334 (2006), 176.145 (2007) und 163.305 (2008) Versichertenstammdatensätze aufgrund von inhaltlichen Fehlern ausgeschlossen. Gründe für den inhaltlichen Ausschluss von Versichertenstammdatensätzen sind in Anhang 9 aufgeführt und sollen an dieser Stelle nicht wiederholt werden. Somit verbleiben 5.968.440 (2006), 5.890.005 (2007) und 5.823.029 (2008) Versichertenstammdatensätze, die in der Grouper-Schnittstelle Berücksichtigung finden. Das Verhältnis dieser Versichertenstammdatensätze zur Anzahl der nach KM6-Statistik erwarteten Anzahl an Versichertenstammdatensätzen der Versichertenstichprobe ist 102,9 % (2006), 101,5 % (2007) und 100,3 % (2008).

Die Tabelle 1 fasst die geschilderten Befunde zusammen.

	Anzahl der Versicherten		
	2006	2007	2008
KM6-Statistik	17.559.917	17.563.627	17.540.160
Unberücksichtigt aufgrund nicht vereinbarter Lieferung	152.882	151.832	123.237
Berücksichtigt	17.407.035	17.411.795	17.416.923
Erwartete Stichprobenmenge	5.802.345	5.803.931	5.805.641
Versichertenstichprobe	6.136.774	6.066.150	5.986.334
Unberücksichtigt aufgrund von Formalfehlern	168.334	176.145	163.305
Berücksichtigt (in VJ ¹)	5.968.440 (5.602.963,7)	5.890.005 (5.559.555,6)	5.823.029 (5.529.294,7)
Anteil der Versichertenstichprobe an erwarteter Stichprobenmenge in % (in VJ)	102,9 (96,6)	101,5 (95,8)	100,3 (95,2)

Tabelle 1: Anzahl Versicherte KM6 und Versichertenstichprobe

(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

Version: 4.04

Per Konstruktion sollte die Versichertenstichprobe ein Drittel der versicherten Personen der KM6 enthalten.² Um zu untersuchen, ob dies tatsächlich der Fall ist, wird in Tabelle 2 dieses Verhältnis für verschiedene Kassengrößen und KV-Bereiche betrachtet.³ Es wird deutlich, dass die Anzahl der Versicherten in etwa 33 % der KM6-Statistik entspricht. Dies gilt für alle Größenklassen und KV-Bereiche. Auch über die betrachteten Jahre bleibt das Verhältnis zwischen Versichertenstichprobe und KM6-Daten relativ konstant.

¹ VJ steht als Abkürzung für den Begriff „Versichertenjahre“. Dieses bedeutet, dass die Anzahl der Versicherten mit dem Anteil an Tagen im Jahr gewichtet werden, die diese jeweils versichert waren.

² Bei der Versichertenstichprobe handelt es sich um eine quartalsweise, bei der KM6-Statistik um eine jährliche Erhebung (Stichtag 30.06.). Um eine bessere Vergleichbarkeit zu gewährleisten, wird die Versichertenanzahl in der Stichprobe in Versichertenjahren angegeben.

³ Versicherte, denen keine KV-Nummer zugeordnet werden konnte, bleiben bei dieser Betrachtung unberücksichtigt.

Größenklasse/ KV-Bereich	Versichertenzahl: Versichertenstichprobe (VJ) zu KM6 in % (Anzahl der Kassen)		
	2006	2007	2008
1–300 Versicherte			
• <i>Gesamt</i>	33,60	33,52	33,67
• KV Bremen	33,28 (95)	33,90 (97)	33,76 (96)
• KV Niedersachsen	34,02 (55)	33,96 (53)	24,07 (53)
• KV Nordrhein	35,01 (59)	34,64 (58)	34,77 (57)
• KV Thüringen	32,53 (93)	32,29 (92)	32,71 (96)
301–1000 Versicherte			
• <i>Gesamt</i>	32,72	33,14	33,00
• KV Bremen	32,26 (13)	32,28 (15)	32,33 (16)
• KV Niedersachsen	31,78 (20)	32,62 (22)	33,06 (21)
• KV Nordrhein	34,15 (19)	34,25 (20)	33,15 (19)
• KV Thüringen	32,61 (18)	33,41 (15)	33,58 (13)
1001–3000 Versicherte			
• <i>Gesamt</i>	32,35	32,50	32,41
• KV Bremen	31,29 (13)	33,11 (14)	32,45 (12)
• KV Niedersachsen	32,53 (32)	32,44 (29)	32,57 (30)
• KV Nordrhein	32,66 (33)	32,53 (32)	32,23 (32)
• KV Thüringen	32,27 (15)	31,94 (14)	32,45 (17)
3001–10000 Versicherte			
• <i>Gesamt</i>	32,48	32,26	32,50
• KV Bremen	32,29 (7)	32,35 (6)	32,51 (7)
• KV Niedersachsen	32,55 (21)	32,23 (21)	32,13 (19)
• KV Nordrhein	32,85 (22)	32,60 (23)	33,03 (24)
• KV Thüringen	31,93 (16)	31,73 (17)	31,89 (14)
10001 und mehr Versicherte			
• <i>Gesamt</i>	32,70	32,51	32,31
• KV Bremen	32,58 (10)	32,45 (10)	32,42 (10)
• KV Niedersachsen	32,27 (44)	32,15 (45)	31,96 (46)
• KV Nordrhein	33,12 (41)	32,86 (42)	32,56 (42)
• KV Thüringen	32,50 (14)	32,37 (14)	32,43 (15)

Tabelle 2: Versichertenzahl: Verhältnis Versichertenstichprobe zu KM6 ($VS/KM6 \geq 20\%$)
(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

Die Kassen, die keine Versichertendaten geliefert haben und für die lediglich KM6-Daten zur Verfügung stehen, sind in der obigen Darstellung nicht berücksichtigt worden. Der umgekehrte Fall, d. h. dass Kassenangaben aus der Versichertenstichprobe nicht in der KM6-Statistik existieren, tritt bei den vorliegenden Datengrundlagen nicht auf. Ebenso unberücksichtigt blieben die Kassen, für die der Anteil der Versichertenstichprobe an den KM6-Daten weniger als 20 % beträgt. Für diese Kassen ist in Tabelle 3 das Verhältnis zwischen Versichertenstichprobe und KM6 für verschiedene Kassengrößen und KV-Bereiche dargestellt.

Größenklasse/ KV-Bereich	Versichertenzahl: Versichertenstichprobe (VJ) zu KM6 in % (Anzahl der Kassen)		
	2006	2007	2008
1–300 Versicherte			
• <i>Gesamt</i>	10,45	8,40	8,78
• KV Bremen	10,72 (14)	9,38 (11)	11,79 (9)
• KV Niedersachsen	18,70 (4)	13,85 (5)	14,84 (4)
• KV Nordrhein	2,67 (5)	2,03 (3)	2,48 (3)
• KV Thüringen	10,94 (8)	9,64 (11)	14,13 (8)
301–1000 Versicherte			
• <i>Gesamt</i>	17,18	15,89	3,76
• KV Bremen	16,18 (1)	–	12,29 (1)
• KV Niedersachsen	–	–	0,23 (1)
• KV Nordrhein	17,72 (2)	13,93 (1)	4,07 (3)
• KV Thüringen	–	19,60 (1)	–
1001–3000 Versicherte			
• <i>Gesamt</i>	–	0,37	–
• KV Niedersachsen	–	0,30 (1)	–
• KV Nordrhein	–	0,42 (1)	–
3001–10000 Versicherte			
• <i>Gesamt</i>	5,20	10,26	6,84
• KV Niedersachsen	4,38 (2)	5,52 (2)	5,80 (2)
• KV Nordrhein	8,26 (1)	14,54 (2)	10,06 (1)
10001 und mehr Versicherte			
• <i>Gesamt</i>	12,47	12,59	–
• KV Niedersachsen	3,73 (2)	4,43 (2)	–
• KV Nordrhein	17,86 (1)	19,45 (1)	–

Tabelle 3: Versichertenzahl: Verhältnis Versichertenstichprobe zu KM6 ($VS/KM6 < 20\%$)
(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

2.3.2 Leistungsbedarf

Weiterer Gegenstand zur Untersuchung der Vollständigkeit der Versichertenstichprobe ist der Leistungsbedarf pro Versicherten. Zur Berechnung des Leistungsbedarfs je Versicherten wird als Zähler in beiden verwendeten Datensätzen

(C4-Daten und Daten des Bewertungsausschusses) die Jahressumme der Punktzahlen nach Überführung in die Beschlusslage des Bewertungsausschusses für das Jahr 2009 zuzüglich der in Punkte umgerechneten Euro-Leistungen verwendet. Als Nenner dienen in der Versichertenstichprobe die Versichertenjahre, im C4-Datensatz die Versichertenzahl gemäß KM6-Statistik. Tabelle 4 zeigt den Vergleich zwischen Versichertenstichprobe und C4-Daten für verschiedene Kassengrößen und KV-Bereiche. Die Kassen, für die der Anteil der Versichertenstichprobe an den KM6-Daten weniger als 20 % beträgt, und Kassen, für die lediglich C4-Daten, jedoch keine Daten in der Versichertenstichprobe zur Verfügung stehen, sind in dieser Darstellung nicht berücksichtigt worden.

Größenklasse/ KV-Bereich	Leistungsbedarf: Versichertenstichprobe (VJ) zu C4-Daten in %	
	2007	2008
1–300 Versicherte		
• <i>Gesamt</i>	91,31	88,07
• KV Bremen	98,13	96,25
• KV Niedersachsen	90,12	81,86
• KV Nordrhein	88,00	89,61
• KV Thüringen	87,17	84,07
301–1000 Versicherte		
• <i>Gesamt</i>	88,32	90,65
• KV Bremen	90,84	93,14
• KV Niedersachsen	87,17	87,69
• KV Nordrhein	90,40	93,27
• KV Thüringen	85,15	88,88
1001–3000 Versicherte		
• <i>Gesamt</i>	92,73	93,90
• KV Bremen	92,33	97,36
• KV Niedersachsen	89,97	88,70
• KV Nordrhein	95,56	97,42
• KV Thüringen	91,91	93,16
3001–10000 Versicherte		
• <i>Gesamt</i>	94,16	92,31
• KV Bremen	93,36	86,76
• KV Niedersachsen	91,29	90,29
• KV Nordrhein	96,28	94,50
• KV Thüringen	93,52	92,99
10001 und mehr Versicherte		
• <i>Gesamt</i>	95,45	94,82
• KV Bremen	96,25	96,81
• KV Niedersachsen	94,64	94,54
• KV Nordrhein	95,18	94,97
• KV Thüringen	99,08	94,80

Tabelle 4: Leistungsbedarf: Verhältnis Versichertenstichprobe zu C4
(Quelle: C4 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

(Hinweis: Für die Berechnungen wird der an die Beschlusslage des Jahres 2009 angepasste Leistungsbedarf verwendet.)

Für die Jahre 2007 und 2008 zeigt Tabelle 5 den Vergleich zwischen dem Leistungsbedarf in der Versichertenstichprobe und dem Leistungsbedarf in den C4-Daten aggregiert auf KV-Bereiche, wobei nun zwischen MGV und EGV unterschieden wird. Ähnlich wie in der obigen Darstellung bleiben die Kassen, für die der Anteil der Versichertenstichprobe an den KM6-Daten weniger als 20 % beträgt und Kassen, für die lediglich C4-Daten, jedoch keine Daten in der Versichertenstichprobe zur Verfügung stehen, unberücksichtigt.

Größenklasse/ KV-Bereich	Leistungsbedarf: Versichertenstichprobe (VJ) zu C4 in %		
	MGV	EGV	Gesamt
2007			
• Gesamt	95,39	94,59	95,26
• KV Bremen	94,99	94,77	94,95
• Niedersachsen	93,69	92,78	93,54
• KV Nordrhein	94,41	93,34	94,22
• KV Thüringen	98,88	97,90	98,72
2008			
• Gesamt	93,86	95,13	94,09
• KV Bremen	94,43	98,02	95,14
• Niedersachsen	93,35	93,73	93,42
• KV Nordrhein	94,13	93,41	94,00
• KV Thüringen	93,49	94,76	93,71

Tabelle 5: Leistungsbedarf: Verhältnis Versichertenstichprobe zu C4, MGV und EGV
(Quelle: C4 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)
(Hinweis: Für die Berechnungen wird der an die Beschlusslage des Jahres 2009 angepasste Leistungsbedarf verwendet.)

Auffällig ist, dass der MGV-Anteil des KV-Bereiches Thüringen von 98,88 % (97,90 %) im Jahr 2007 auf 93,49 % (94,76 %) im Jahr 2008 sinkt⁴. Für alle übrigen KV-Bereiche bleibt die Verteilung über die Jahre hinweg relativ konstant.

⁴ Im Nachgang zur Berichtslegung hat sich ergeben, dass auch Unvollständigkeiten in den C4-Daten der KV Thüringen zu einer hochausfallenden Abdeckungsquote in 2007 geführt haben.

In der Gesamtbetrachtung sinkt der Anteil der Versichertenstichprobe von 95,26 % auf etwa 94,09 %, sodass sich ein Schwund in der Ausschöpfung der C4-Daten von 1,17 % ergibt. Dieser Schwund an Leistungsbedarf zieht auch einen Schwund an Diagnosen mit sich, was zu einer Verzerrung der morbiditätsbedingten Veränderungsrate 2007/08 führen kann. Das Ausmaß der Verzerrung wird durch die unbekannt Struktur der nicht verfügbaren Diagnosen bestimmt, wobei eine Verzerrung von -1,17 % nur im schlimmsten Fall eintritt und damit die unwahrscheinliche Obergrenze bildet. Außerdem ist zu bedenken, dass die C4-Daten für das Jahr 2007 eine geringere Qualität als im Jahr 2008 aufwiesen und somit der Ausgangswert für die Abschätzung der maximalen Verzerrung ungenau sein kann. Für die Veränderungsrate 2007/08 ergibt sich demnach eine geringfügige negative Verzerrung, deren Höhe nicht abschätzbar ist⁵. Für die morbiditätsbedingte Veränderungsrate 2006/07 können keine Aussagen getroffen werden, da die C4-Daten für das Jahr 2006 nicht vorliegen.

Die Versichertenstichprobe bildet die C4-Daten nicht zu 100 % ab, da in der Stichprobe die Inanspruchnahme von Leistungen außerhalb der vier KVen nicht berücksichtigt werden kann. Für Thüringen, zum Beispiel, machen diese sog. Fremdkassenzahlungsausgleiche für das Jahr 2008 5,14 % des gesamten Leistungsbedarfs aus (vgl. Kapitel 2.5). Unter Berücksichtigung dieser „Lücke“ von etwa 5 % entspricht der Anteil des Leistungsbedarfs in der Versichertenstichprobe zu 100 % dem Leistungsbedarf in der C4-Statistik. Zu beachten ist, dass es sich hier um Überschlagsrechnungen handelt und eine Addition der beiden Werte mit Vorsicht zu genießen ist.

2.3.3 Unterjährig Versicherte

Ein Sachverhalt, der zu Problemen in der Datenvollständigkeit führen kann, sind unterjährig versicherte Personen. Es existiert eine Vielzahl von Ursachen für Unterjährigkeit in der Versicherungsdauer. Dieser Bericht geht den folgenden Gründen nach: 1. Geburt, 2. Sterbefall, 3. Kassenwechsler, 4. unterjährige Fusionen, 5. Wechsel in die PKV, 6. Umzug, 7. Statuswechsel, 8. Quartalsdatensätze zur Versichertenzeit nicht (oder nicht verarbeitbar) übermittelt (vgl. Kapitel 2.1.4), 9. Quartalsdatensätze mit fälschlicherweise verkürzten Versichertenzeiten. Während die ersten beiden Gründe im vorliegenden Datensatz identifiziert

⁵ Die erst nach Berichtslegung offenbar gewordenen Unvollständigkeiten (vgl. Fußnote 4) stützen die Hypothese einer negativen Verzerrung nicht mehr.

werden können, sind die Punkte 3.-7. nicht zu unterscheiden. Zu den Punkten 8 und 9 wurden Sonderuntersuchungen in Angriff genommen.

Tabelle 6 gibt einen Überblick, wie die Versicherten in den Jahren 2006, 2007 und 2008 auf ganzjährig und unterjährig Versicherte sowie die zu identifizierenden Untergruppen verteilt sind.⁶ Hier wird deutlich, dass der Anteil der unterjährig Versicherten von 2006 auf 2007 um etwa einen Prozentpunkt sinkt und von 2007 auf 2008 um 4,5 Prozentpunkte steigt. Der Prozentsatz der Geburten, Sterbefälle sowie Sterbefälle im 1. Lebensjahr bleibt in allen drei Jahren relativ konstant, sodass der Rückgang vor allem durch die Gruppe „Sonstige“ begründet werden kann. Während nämlich 2006 etwa 11,9 % und 2007 etwa 10,9 % der Versicherten aus sonstigen Gründen unterjährig versichert sind, sind es 2008 mehr als 15,4 %.

Anzahl Versicherte	2006	2007	2008
Absolut			
Gesamt	5.968.440	5.890.005	5.823.029
(in VJ)	(5.602.963,7)	(5.559.555,6)	(5.529.294,7)
• Ganzjährig Versicherte	5.156.443	5.150.797	4.822.034
• Unterjährig Versicherte	811.997	739.208	1.000.995
(in VJ)	(446.721,4)	(408.939,5)	(709.120,7)
Prozentual			
• Ganzjährig Versicherte	86,395 %	87,450 %	82,810 %
• Unterjährig Versicherte	13,605 %	12,550 %	17,190 %
(in VJ)	(7,485 %)	(6,943 %)	(12,178 %)
1. Geburten	0,708 %	0,700 %	0,719 %
2. Sterbefälle	0,967 %	0,955 %	1,025 %
3. Sterbefälle im 1. Lebensj.	0,001 %	0,002 %	0,001 %
4. Sonstige	11,928 %	10,894 %	15,445 %

Tabelle 6: Verteilung der unterjährig Versicherten (Versichertenzeit < 360 Tage)
(Quelle: Daten des BA; eigene Berechnungen und Darstellung)

⁶ Unterjährigkeit liegt vor, wenn der Versicherte weniger als 360 Tage im Jahr versichert ist.

Das Ergebnis der **Sonderuntersuchung zu Quartalsdatensätzen mit fälschlicherweise verkürzten Versichertenzeiten** (Grund Nr. 9) ergab, dass der starke Anstieg der Unterjährigen von 2007 auf 2008 auf Ersatzkasse 1, die für das Jahr 2008 mehr als 80 % unterjährig Versicherte ausweist, zurückzuführen ist. Eine genauere Betrachtung dieser Kasse zeigt, dass sie für das erste Quartal 2008 für fast 80 % ihrer Versicherten nur 81 Versichertentage ausweist, was auf einen Dokumentationsfehler schließen lässt. Demnach werden 2008 für einen Großteil der Versicherten der Ersatzkasse 1 356 anstatt 366 Versichertentage ausgewiesen. Tabelle 66 in Anhang 10 zeigt die Verteilung der unterjährig Versicherten, wenn man eine Versichertenzeit von weniger als 356 Tagen als Definition für Unterjährigkeit (anstatt 360 Tagen) voraussetzt.

Nennenswerte Verzerrungen auf die Veränderungsrate sind hierdurch nicht zu erwarten, die Analyse der unterjährig Versicherten wurde dadurch allerdings behindert. Eine detaillierte Plausibilitätsprüfung der Versicherungszeiten war in der bisherigen Arbeit der Datenstelle noch nicht vorgesehen, sodass dieser Fehler erst sehr spät erkannt wurde.

Die **Sonderuntersuchungen zu fehlenden Versichertenzeiten** (Grund Nr. 8) wurden einerseits als exemplarische Prüfungen bei fusionierten Kassen durchgeführt, andererseits über eine Auswertung, zu welchen Anteilen bei den Versicherten Abrechnungsfälle vorliegen, die nicht durch Versicherungszeit im Quartal gedeckt sind (vgl. Anhang 11).

Beide Prüfstrategien ergaben, dass nicht alle Kassen die Versichertendatensätze stets vollständig übermittelt hatten. Im Extremfall wurde von einer Kasse mit mehreren 10.000 Versicherten in einem Quartal ein einziger Versicherter gemeldet. Die Ausfälle scheinen bei der Kassenart 3 und Kassenart 1 am stärksten zu sein, das Jahr 2007 ist etwas stärker betroffen als Vorjahr und Folgejahr.

Der genaue Anteil der aus diesem Grund „unterjährigen“ Versicherten ist noch nicht ermittelt, er muss allerdings im einstelligen Prozentbereich – gemessen an den Versicherten aller Kassen zusammen – liegen⁷ und könnte bei den beiden genannten Kassenarten zeitweise über zehn Prozent ihrer Versicherten ausmachen.

⁷ Der Anteil der nicht durch Versicherungszeiten abgedeckten Abrechnungsfälle der Stichprobenversicherten liegt in den drei untersuchten Jahren zwischen 1 und 2 Prozent.

In den **Fällen 3, 4, 6 und 7 (Kassenwechsel usw.)** kommt es zur Aufspaltung einer physischen Person in mehrere Versicherte, was bei der Berechnung von Veränderungsdaten problematisch werden kann. Dies wird in Anhang 12 anhand von Beispielen näher erläutert.

Da die Sonderuntersuchungen zu fehlenden Versichertenzeiten (Grund Nr. 8) zu keiner Quantifizierung geführt haben, sind die Anteile, die zu Aufspaltungen von Personen führen, leider nicht bezifferbar. Gesichert ist, dass die Gründe 3 ff. von 2006 auf 2007 abnehmen, möglicherweise auch von 2007 und 2008.

Ein Rückgang in der Zahl der unterjährig Versicherten kann eine geringfügige positive Verzerrung auf die Veränderung des Morbiditätsindex haben, was bisher aber noch nicht genau abgeschätzt werden konnte.

Im Nachgang zur ursprünglichen Berichtslegung wurde durch das Institut eine Simulation durchgeführt, welche den Einfluss von Unterjährigkeit darstellen soll. Diese Simulation findet sich im Anhang 6.

2.3.4 Ärztliche Inanspruchnahme

Wenn man die dokumentierte ärztliche Inanspruchnahme der einzelnen Kassen über für die Jahre 2006, 2007 und 2008 untersucht, wird deutlich, dass sich für den Großteil der Kassen die Quote der Nicht-Inanspruchnahme über die Jahre hinweg relativ konstant verhält. Lediglich für vier der insgesamt 207 Kassen zeigen sich temporäre Unregelmäßigkeiten. Um diesen Sachverhalt detaillierter darzustellen, zeigt Tabelle 7 die Anzahl der Versicherten ohne ärztliche Inanspruchnahme sowie die daraus resultierenden möglichen Verzerrungen der morbiditätsbedingten Veränderungsdaten.

Kasse	Versicherte ohne Inanspruchnahme (im Verhältnis zu Versicherten insgesamt)			Mögl. Verzerrungen der Veränderungsdaten	
	2006	2007	2008	2006/07	2007/08
Kassenart 1	8.058 (14,45 %)	23.086 (32,80 %)	8.530 (11,00 %)	-0,215 %	0,282 %
Kassenart 1	4.114 (15,14 %)	10.552 (31,93 %)	4.819 (12,68 %)	-0,092 %	0,122 %
Kassenart 2	113.068 (12,76 %)	102.050 (11,84 %)	111.170 (13,27 %)	0,133 %	-0,200 %
Kassenart 2	83.368 (10,96 %)	76.460 (10,27 %)	86.231 (11,84 %)	0,086 %	-0,191 %

Tabelle 7: Beispiele zu Implausibilitäten in der ärztlichen Inanspruchnahme
(Quelle: Daten des BA; eigene Berechnungen und Darstellung)

Bei den ersten beiden Kassen (Kassenart 1) entspricht die Anzahl der in 2007 neuen Versicherten der jeweiligen Kasse in etwa den Versicherten ohne dokumentierte ärztliche Inanspruchnahme, was sich vermutlich auf ein Zusammenführungsproblem in 2007 zurückführen lässt. Im Folgejahr 2008 liegt der Anteil der ärztlichen Inanspruchnahme wieder auf dem Niveau 2006. Bei den folgenden beiden Kassen (Kassenart 2) sind beim Vergleich der drei ausgewerteten Jahre zwar die Quoten recht ähnlich, jedoch sind Schwankungen in der absoluten Anzahl der Versicherten ohne Inanspruchnahme zu erkennen. Bei einer der beiden war durch Analysen der AG IT schon aufgefallen, dass nicht alle ärztlichen Abrechnungen und Diagnosen mit den Versicherten zusammengeführt werden konnten – mit leicht zunehmender Tendenz dieser Fehlermarge in 2008.

Legt man die Stichprobe von 6 Millionen Versicherten zugrunde, könnte das zuerst genannte Problem bei den beiden Kassen der Kassenart 1 mögliche Verzerrungen in der Veränderungsrate bewirken, die sich auf bis zu -0,31 Prozentpunkte bezüglich der Rate 2006/07 und +0,40 Prozentpunkte bezüglich der Rate 2007/08 belaufen. Das Problem, welches für die beiden Kassen der Kassenart 2 besteht, führt hingegen zu gegensätzlichen Verzerrungen, die für die Veränderungsrate 2006/07 nicht nennenswert sind und für die Rate 2007/08 etwa -0,39 Prozentpunkte betragen. Beides zusammen genommen führt zu der Einschätzung, dass die Rate 2007/08 hierdurch nicht beeinflusst oder nur geringfügig verzerrt wird.

2.3.5 Kassenfusionen

Während des Betrachtungszeitraumes treten Fusionen zwischen Kassen auf, die unterschiedliche Probleme hinsichtlich der Vollständigkeit der Datenlieferungen hervorrufen können. Beispielsweise können übervollständige Lieferungen auftreten. Dies bedeutet, dass die aufnehmende Kasse die Daten für die anfusionierte Kasse schon für einen Zeitraum liefert, der vor dem Fusionstermin liegt. Somit liegen die Daten doppelt vor: richtig für die anfusionierte Kasse vor der Fusion und falsch für die aufnehmende Kasse vor der Fusion (Überschusslieferung). Dies hat zur Folge, dass zu den falsch gelieferten Versicherten-Datensätzen keine Leistungsdaten gefunden werden können und ein Matching (richtigerweise) nicht möglich ist. Durch die Qualitätssicherung der AG IT und des InBA wird diese Übervollständigkeit in aller Regel erkannt und behoben.

Aufgrund von technischen Problemen bei Fusionen, z. B. Zusammenführung verschiedener EDV-Systeme, können außerdem technisch nicht hinreichende Datenlieferungen, z. B. formal fehlerhafte Pseudonymisierung, entstehen. Dies hat

ebenfalls zur Folge, dass Abrechnungsfälle und Diagnosen nicht mit den Versichertendaten zusammengeführt werden können und die Versicherten u. U. als Non-User erscheinen. Auch solche Fehler werden jedoch durch die Qualitätssicherung der AG IT und des InBA in aller Regel erkannt und durch Neulieferungen bzw. ggf. Ausschluss der Versicherten behoben.

Bei Kassenfusionen kann zudem das Problem der Zerteilung eines Versicherten in zwei technisch getrennte Versicherte auftreten. Dies hat zur Folge, dass für einen Versicherten auf das Jahr betrachtet zwei unvollständige Datensätze im Datenkörper vorhanden sind (vgl. Kapitel 2.1.4 sowie Anhang 11).

Des Weiteren kann eine verzögerte Chipkartenausgabe bzw. ein verspäteter Einsatz der neuen Karte durch den Versicherten dazu führen, dass der Leistungsbedarf des Versicherten kurzfristig im Datenkörper nicht adäquat abgebildet wird, da ein Matching der Daten nicht möglich ist.

Für die Jahre 2006 und 2007 wurden für die fusionierten Kassen Prüfungen hinsichtlich der Versichertenzahl und des Leistungsbedarfs pro Versicherten pro Quartal durchgeführt. Es wurden verschiedene „kleine“ Fehllieferungen festgestellt. Zudem konnte eine „größere“ Fehllieferung bei einer Kassenfusion im Quartal IV/2006 ausgemacht werden, die ein Absinken der Versichertenzahl um 20.000 auf 30.000 Versicherte ausmacht (Ursprungskasse 40.000 Versicherte, anfusionierte Kasse ca. 10.000 Versicherte).

Die Prüfung hinsichtlich des durchschnittlichen Leistungsbedarfs pro Versicherten und Quartal für die Jahre 2006 und 2007 ergab bei allen fusionierten Kassen den typischen Jahresverlauf mit erhöhten Bedarfen im ersten und vierten Quartal des entsprechenden Jahres. Ein nennenswerter Verlust des Leistungsbedarfs der anfusionierten Versicherten konnte nicht festgestellt werden.

2.4 Repräsentativitätsprüfungen

Um von einer repräsentativen Untersuchung sprechen zu können, muss die Versichertenstichprobe in der Verteilung ausgewählter Kriterien einer klar definierten Grundgesamtheit entsprechen. Dazu ist zum einen festzulegen, welche Kriterien ausgewählt werden und zum anderen zu untersuchen, inwieweit die Versichertenstichprobe die vier gewählten Regionen (4 KV-Bereiche) widerspiegelt. Darüber hinaus ist zu überprüfen, wie sich die vier KVen im Vergleich zur Gesamtheit der 17 KVen (Bund) verhalten. Zum Vergleich dienen externe Daten des Bundesversicherungsamtes sowie die der KM6-Statistiken (Erläuterungen der

BVA-Daten und der KM6 in Kapitel 2.1.2). Die Kriterien, die den Repräsentativitätsprüfungen zugrunde liegen, sind die Stadt-/Land-Verteilung, die Klassifizierung in Zuzugs- und Fortzugsregionen, die Entwicklung der demographischen Struktur, sowie die berechneten Relativgewichte und Veränderungsraten.

2.4.1 Stadt-/Land-Verteilung

Wenn die Verteilung der Bevölkerung auf städtische und ländliche Regionen in der Versichertenstichprobe andere Entwicklungen aufzeigt als in der gesamten GKV, könnte dies die Repräsentativität der Daten und damit die zuverlässige Berechnung der Veränderungsraten beeinflussen. Um dieses mögliche Problem auszuschließen, wird die Stadt-/Land-Verteilung in der Versichertenstichprobe mit der Verteilung in den vier KVen sowie im gesamten Bundesgebiet verglichen.

Da in den vorliegenden KM6-Daten eine Zuordnung der Versicherten zu Städten bzw. Kreisen fehlt, ist eine Klassifizierung in städtische und ländliche Regionen nicht möglich. Der Vergleich der Versichertenstichprobe basiert daher auf externen Daten des Statistischen Bundesamtes, welche Informationen zur Einwohnerzahl und Fläche auf Kreisebene enthalten. Diese Angaben erlauben eine Berechnung der Bevölkerungsdichte und somit eine Einteilung der Kreise und kreisfreien Städte in ländliche (weniger als 200 Einwohner pro km²) und städtische Regionen (mehr als 200 Einwohner pro km²).⁸ Da die Daten vom Statistischen Bundesamt zum Gebietsstand 31.12. eines jeden Jahres erhoben werden, werden zum besseren Vergleich die Informationen der Versichertenstichprobe aus dem vierten Quartal verwendet.⁹

Tabelle 8 zeigt die prozentualen Anteile der Stadt- bzw. Landbevölkerung basierend auf der Versichertenstichprobe sowie den externen Informationen zum gesamten Bundesgebiet und den 4 KVen. Es ist deutlich zu erkennen, dass unabhängig von der Datenbasis etwa zwei Drittel der Bevölkerung in städtischen Regionen wohnen. Sowohl für die Versichertenstichprobe als auch für das gesamte Bundesgebiet zeigt diese Verteilung auch über die betrachteten Jahre keine grö-

⁸ Zu den ländlichen Regionen zählen dünn besiedelte Räume (<100 Einwohner/km²) und gering verdichtete Räume (<200 Einwohner/km²) gemäß BBR (Bundesamt für Bauwesen und Raumordnung). Bezieht man sich bei der Definition ländlicher Regionen lediglich auf dünn besiedelte Räume erhält man für alle drei verwendeten Datensätze ebenfalls eine sehr ähnliche Stadt-/Land-Verteilung.

⁹ Versicherte, denen keine KV-Nummer zugeordnet werden konnte, bleiben bei dieser Betrachtung unberücksichtigt..

ßeren Veränderungen. Sowohl die Abweichungen zwischen den betrachteten Jahren als auch die Abweichungen zwischen Versichertenstichprobe und KM6 betragen deutlich weniger als einen Prozentpunkt, sodass hier eine angemessene Analysegenauigkeit erreicht wird. Aus diesem Grund kann dieser mögliche Problembereich keine nennenswerte verzerrende Wirkung auf die Berechnung der Veränderungsrate haben und muss daher nicht weiter berücksichtigt werden.

	Anteil der Stadt-/Landbevölkerung in %					
	2006		2007		2008	
	Stadt	Land	Stadt	Land	Stadt	Land
Statistisches Bundesamt (Bund)	65,11	34,89	64,66	35,34	–	–
Statistisches Bundesamt (4 KV)	66,22	33,78	66,30	33,70	–	–
Versichertenstichprobe in VJ	65,37	34,63	65,29	34,71	65,41	34,59

Tabelle 8: Verteilung Stadt/Land

(Quelle: Daten des Statistischen Bundesamtes und des BA; eigene Berechnungen und Darstellung)

Zu bedenken ist, dass die Bevölkerung in der Versichertenstichprobe zwar repräsentativ auf Stadt und Land verteilt ist, dies jedoch nicht zwangsläufig die Repräsentativität der demographischen Verteilung bedingt. Dieser Aspekt wird in Kapitel 2.4.3 überprüft.

2.4.2 Zuzugs-/Fortzugsregionen

Die Repräsentativität der Versichertendaten könnte auch durch Unterschiede in der Mobilität der Bevölkerung beeinträchtigt werden. Wenn nämlich bestimmte Regionen durch einen starken Zuzug bzw. Fortzug gekennzeichnet sind und sich die mobile Bevölkerung in ihrer Morbiditätsstruktur systematisch von der weniger mobilen unterscheidet, könnte das direkte Relevanz für die Berechnung der Veränderungsrate haben.

Um diesen Aspekt zu untersuchen, wird die Bevölkerungsentwicklung der Versichertenstichprobe mit der Bevölkerungsentwicklung in den KM6-Daten verglichen. Die KM6-Statistik in Tabelle 9 zeigt, dass in den Jahren 2006 bis 2008 die Teilmenge der vier KVen praktisch identische Veränderungen in der Versichertenanzahl aufweist wie die Gesamtheit der 17 KVen. Die Versichertenstichprobe deutet hingegen auf leicht abweichende Veränderungsrate hin. Während die KM6-Statistiken einen Anstieg der Bevölkerung um 0,02 % von 2006 auf 2007 anzeigen, sinkt die Anzahl der Personen in der Versichertenstichprobe um 0,77.

Auch für das Jahr 2008 sind abweichende Veränderungsdaten zu beobachten. Die Abweichung zwischen Versichertenstichprobe und KM6-Statistiken beträgt weniger als einen Prozentpunkt, sodass auch hier eine akzeptable Analysegenauigkeit erzielt werden kann und keine nennenswerten Verzerrungen auf die Schätzung der Veränderungsdaten zu erwarten sind.

	Veränderung der Versichertenzahl in %		
	2006	2007	2008
KM6 (Bund)	-0,25	0,02	-0,10
KM6 (4 KV)	-0,24	0,01	-0,10
Versichertenstichprobe in VJ		-0,77	-0,54

Tabelle 9: Relative Veränderung der Versichertenzahl

(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

Auch wenn Zu- und Fortzüge die Entwicklung der Bevölkerungszahl nur wenig beeinflussen, könnten sie dennoch auf die demographische Zusammensetzung der Bevölkerung einwirken und damit die Repräsentativität der Daten beeinträchtigen. Dieser Aspekt wird im nachfolgenden Kapitel analysiert.

2.4.3 Demographische Struktur

Um zu überprüfen, ob die verwendete Versichertenstichprobe auch in Bezug auf die demographische Zusammensetzung und deren Entwicklung repräsentativ für die 4 KVen und das gesamte Bundesgebiet ist, werden die Geschlechts- und Altersstruktur miteinander verglichen. Aufgrund der nicht vollständigen Übereinstimmung in der Stadt-/Land-Verteilung zwischen den vier verfügbaren KV-Bereichen und Deutschland insgesamt, vor allem aber aufgrund der notwendigen Ausschlüsse von Versicherten für die Berechnungen von Veränderungsdaten muss dieser abschließende Abgleich durchgeführt werden.

Aus Tabelle 10 wird ersichtlich, dass in allen drei Statistiken eine sehr ähnliche Verteilung von Frauen und Männern zu beobachten ist. Frauen machen etwa 53 % der Gesamtbevölkerung aus, wobei dieser Anteil auch über die betrachteten Jahre hinweg relativ konstant bleibt.

	Anteil Frauen und Männer in %					
	2006		2007		2008	
	Frauen	Männer	Frauen	Männer	Frauen	Männer
KM6 (Bund)	53,08	46,92	53,06	46,94	53,01	46,99
KM6 (4 KV)	52,97	47,03	52,94	47,06	52,90	47,10
Versichertenstichprobe in VJ	53,05	46,95	53,04	46,96	53,08	46,92

Tabelle 10: Verteilung Frauen und Männer

(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

Zur Überprüfung der Repräsentativität der Altersstruktur werden die Anteile bestimmter Altersklassen an der Gesamtpopulation auf Basis der Versichertenstichprobe und der KM6 (4 KVen und Bund) für die Jahre 2006 bis 2008 berechnet und gegenübergestellt.¹⁰ Die Ergebnisse für das Jahr 2008 sind in Abbildung 3 dargestellt. (Die Darstellungen für die Jahre 2006 und 2007 befinden sich in Anhang 14.) Es ist zu erkennen, dass die Altersstruktur in der Versichertenstichprobe im Wesentlichen sowohl der Altersstruktur in den 4 KVen als auch der im Bundesgebiet entspricht.

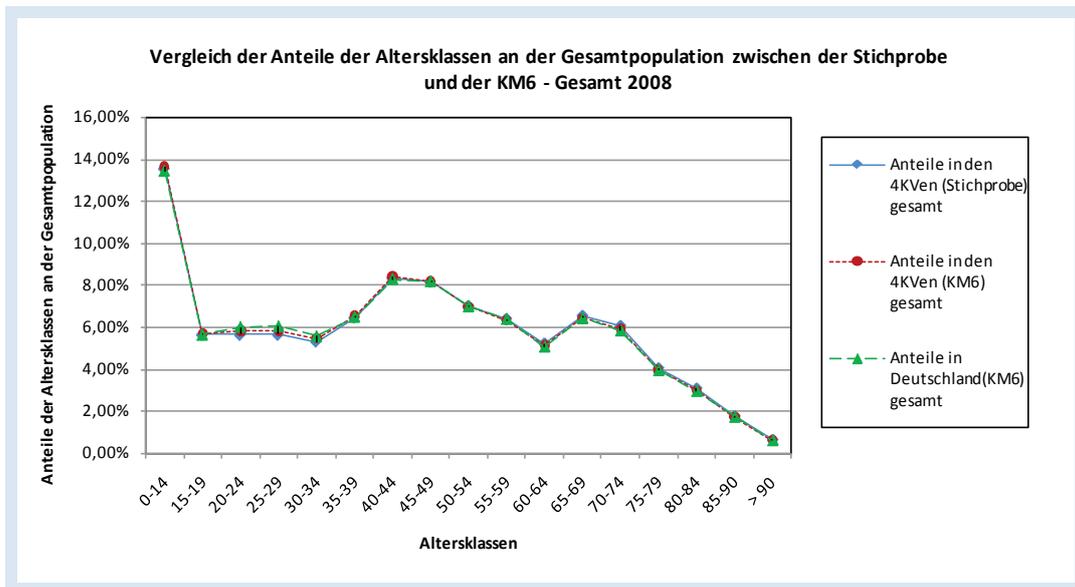


Abbildung 3: Vergleich der Anteile der Altersklassen an der Gesamtpopulation zwischen der Stichprobe und der KM6 - Gesamt 2008

(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

¹⁰ Die Altersklassen richten sich nach den Gruppierungen in der KM6-Statistik (siehe Anhang 13).

In Abbildung 4 und Abbildung 5 (und entsprechend in Anhang 15) kann gezeigt werden, dass auch bei getrennter Betrachtung von Frauen und Männern die Altersklassen in allen drei Datengrundlagen sehr ähnlich verteilt sind. Demnach kann die Versichertenstichprobe bezüglich Geschlechts- und Altersstruktur als repräsentativ angesehen werden, sowohl für die 4 KVen als auch für die Gesamtheit der 17 KVen.

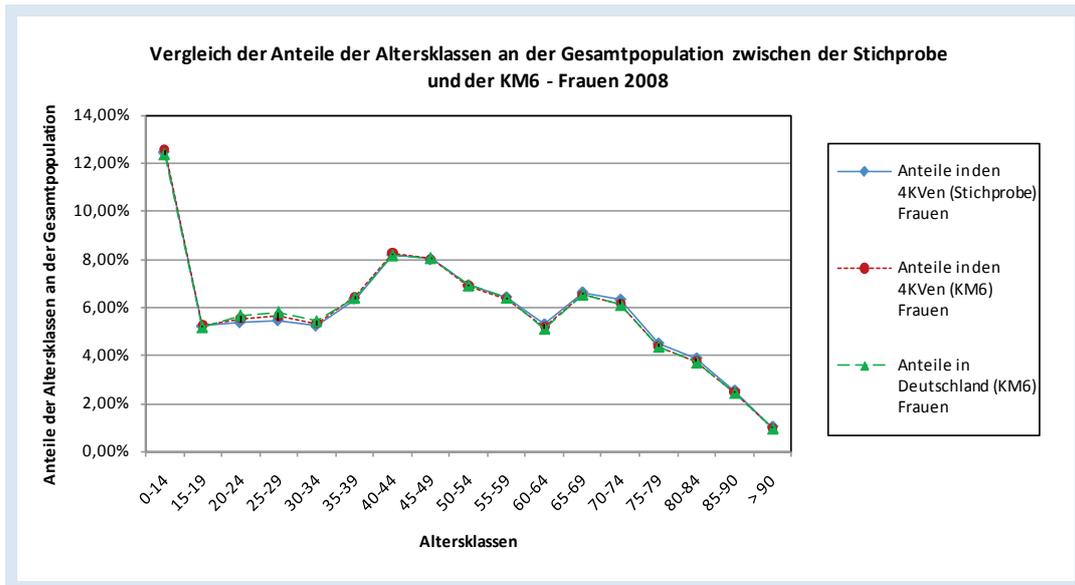


Abbildung 4: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Frauen – 2008
(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

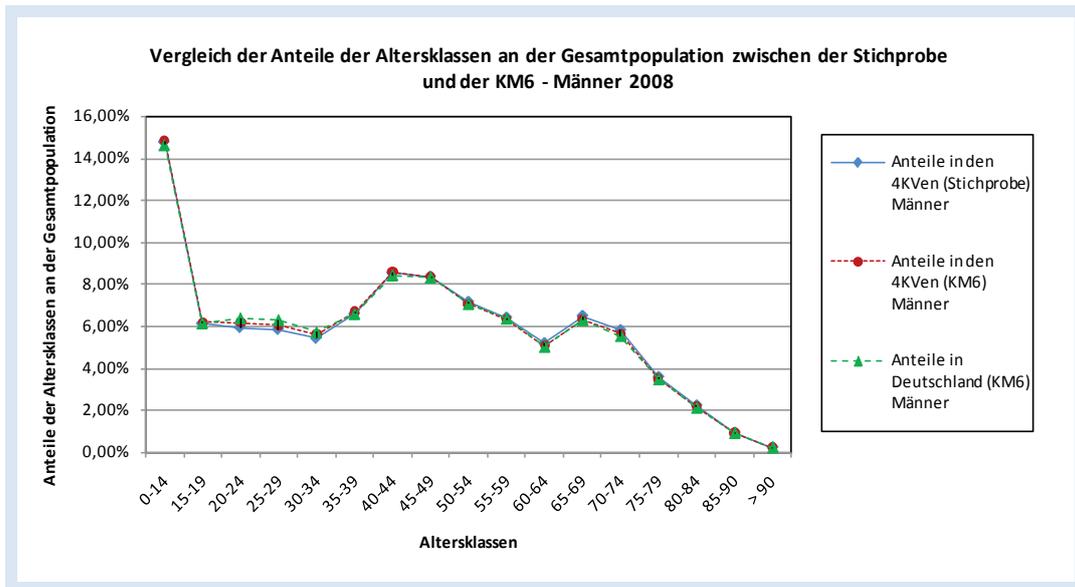


Abbildung 5: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Männer – 2008
(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

Um diesen Sachverhalt expliziter darzustellen, werden anhand von Relativgewichten, die auf Basis der Versichertenstichprobe für die verschiedenen Altersklassen gemäß KM6-Statistik berechnet werden, Morbiditätsindizes für alle drei Datengrundlagen ermittelt¹¹. Tabelle 11 zeigt die Indizes auf Basis der KM6-Statistiken 4 KV und Bund in Relation zum Morbiditätsindex, der sich für die Versichertenstichprobe ergibt. Je näher die Werte zusammen liegen, desto besser wird die Morbidität in den 4 KV bzw. Deutschland durch die Stichprobe abgebildet. Der Vergleich im Querschnitt zeigt sehr gute Übereinstimmungen.

	Vergleich des demographisch bedingten Morbiditätsindex berechnet anhand von Relativgewichten (2007 fix)		
	KM6 (Bund)	KM6 (4 KV)	Versichertenstichprobe
2006	0,988	0,990	0,994
2007	0,993	0,994	1,000
2008	0,997	0,998	1,005

Tabelle 11: Vergleich Morbiditätsindex zwischen KM6 und Versichertenstichprobe
(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

¹¹ Für eine genauere Darstellung der Berechnung von Relativgewichten, siehe Kapitel 3.3 .

Auf Basis dieser berechneten Morbiditätsindizes werden die Veränderungsdaten 2006/07 sowie 2007/08 berechnet.

	Vergleich der demographischen Veränderungsdaten berechnet anhand von Relativgewichten (2007 fix) in %		
	KM6 (Bund)	KM6 (4 KV)	Versichertenstichprobe
2006/07	0,421	0,391	0,561
2007/08	0,418	0,394	0,559

Tabelle 12: Vergleich der Veränderungsdaten zwischen KM6 und Versichertenstichprobe (Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

Beide demographisch bedingten Veränderungsdaten der GKV insgesamt liegen bei vier Promille, die regional beschränkten Daten (vier KV) knapp darunter. Die entsprechenden Daten aus der Versichertenstichprobe (vier KV) fallen in beiden Jahren gleich aus und liegen um gut einen Promillepunkt höher als im Bundesdurchschnitt.

2.4.4 Externe Plausibilisierung der Relativgewichte

Für den Vergleich von alters- und geschlechtsspezifischen Relativgewichten werden zunächst auf Basis der Versichertenstichprobe die durchschnittlichen Leistungspunkte je Versichertentag und Altersklasse (gegliedert nach SA 40) berechnet. Außerdem werden anhand der Daten des BVA die Euro-pro-Tag-Werte der Hauptleistungsbereiche „Ärzte“ und Dialyse vor Herausnahme der „Risikopoolfälle“ über alle DMP-Gruppen je Erwerbsminderungsstatus und Altersklasse zusammengefasst. Da die auf diese Weise berechneten Kostengewichte unterschiedliche Einheiten haben (Punkte und Euro), werden für beide Datenquellen Relativgewichte berechnet. Dazu wird jedes Kostengewicht ins Verhältnis zum durchschnittlichen Kostengewicht gesetzt. Außerdem werden die Relativgewichte jeweils auf Basis der Versichertenstruktur des BVA ermittelt, sodass Verzerrungen aufgrund unterschiedlicher Populationen vermieden werden können. Abbildung 6 zeigt den Vergleich der berechneten Relativgewichte für die Altersgruppen 0–90 sowie für Frauen und Männer. Es zeigt sich, dass die alters- und geschlechtsspezifischen Gewichte im Wesentlichen mit denen des BVA übereinstimmen.

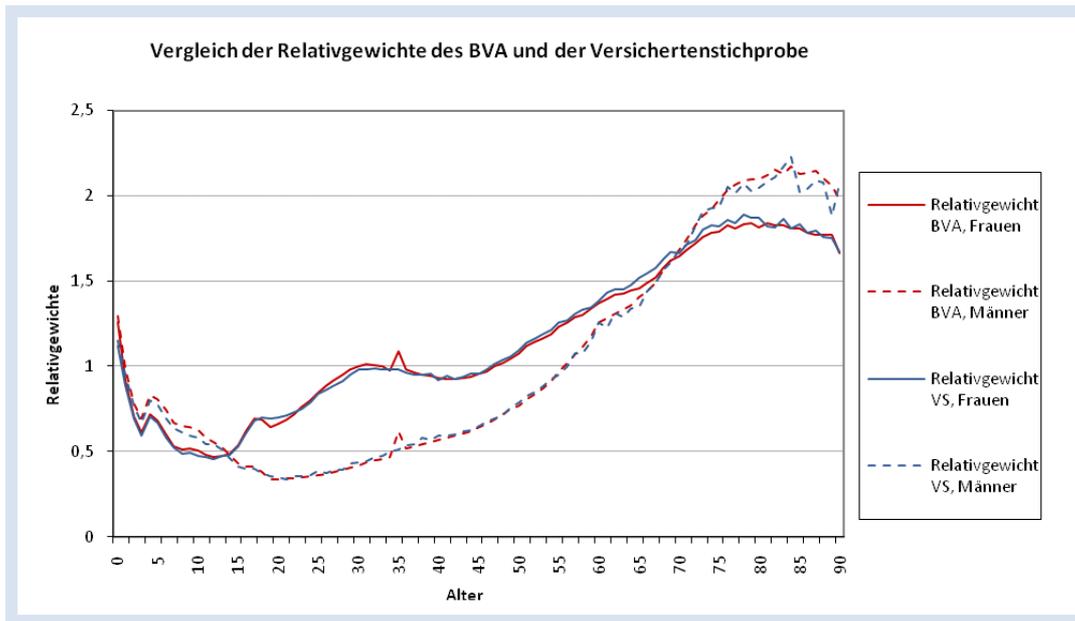


Abbildung 6: Vergleich der Relativgewichte der Versichertenstichprobe und des BVA
(Quelle: Daten des BVA, Daten des BA; eigene Berechnungen und Darstellung)

Die gleiche Vorgehensweise wird auch für die Berechnung der Relativgewichte angewendet, wenn man zusätzlich zum Alter und Geschlecht den Erwerbsminderungsstatus als Klassifizierungskriterium wählt. Abbildung 48 in Anhang 15 zeigt die Ergebnisse auf Basis der Versichertenstichprobe und des BVA. Für Versicherte ohne Erwerbsminderungsrente lässt sich für beide Datengrundlagen ein sehr ähnlicher Verlauf der berechneten Relativgewichte erkennen. Da die Berechnung für Versicherte mit Erwerbsminderungsrentnerstatus auf Basis der Versichertenstichprobe ohne Glättung erfolgt, ist der für diese Relativgewichte erkennbare gezackte Verlauf zu erklären.

2.4.5 Externe Plausibilisierung der Veränderungsdaten

Um auch die mithilfe der Versichertenstichprobe berechneten Veränderungsdaten auf Repräsentativität zu überprüfen, werden sie mit den Veränderungsdaten verglichen, die vollständig oder zumindest teilweise auf Basis externer Daten berechnet wurden. Als externe Datenquellen dienen hier die Daten des Bundesversicherungsamtes (BVA) sowie die auch schon zuvor verwendete KM6-Statistik.

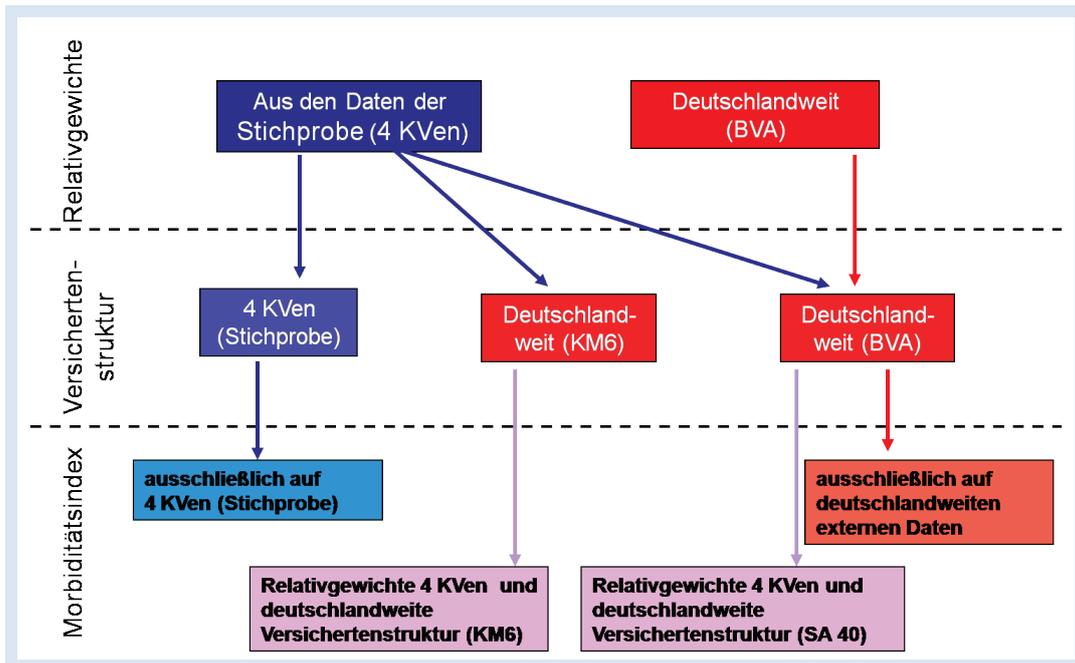


Abbildung 7: Schaubild der Methodik der externen Plausibilisierung der Veränderungsraten
(Quelle: eigene Darstellung)

Wie in Abbildung 7 dargestellt, werden zunächst Veränderungsraten anhand der Relativgewichte und der Versichertenstruktur der Stichprobe (beides in blau dargestellt) und damit nur für die 4 KVen berechnet. Anschließend erfolgt die Berechnung der Veränderungsraten für das gesamte Bundesgebiet. Dazu werden zwar wieder die Relativgewichte der Stichprobe verwendet, die Versichertenstruktur wird aber nun aus den zwei externen Datenquellen BVA und KM6 bezogen (in lila dargestellt). Im letzten Schritt werden die deutschlandweiten Veränderungsraten ausschließlich auf Basis externer Daten berechnet, d. h. sowohl die Relativgewichte als auch die Versichertenstruktur werden aus Daten des BVA bezogen (in rot dargestellt).

Vergleich der Veränderungsraten 2006/07			
Ausschließlich Stichprobe (4 KVen)	Relativgewichte: Stichprobe, Versichertenstruktur: KM6 (Bund)	Relativgewichte: Stichprobe, Versichertenstruktur: BVA(Bund)	Ausschließlich BVA (Bund)
0,56 %	0,38 %	0,53 %	0,47 %

Tabelle 13: Vergleich der Veränderungsraten: Stichprobe und externe Datenquellen
(Quelle: Daten des BA, KM6, BVA; eigene Berechnungen und Darstellung)

Aus Tabelle 13 wird ersichtlich, dass bei der Verwendung der Relativgewichte der Versichertenstichprobe und einer deutschlandweiten Versichertenstruktur (KM6 oder BVA) die Veränderungen der Morbiditätsrate von 2006 auf 2007 ähnlich sind.¹² Die Abweichungen liegen deutlich unter einem Prozentpunkt, sodass auch hier eine akzeptable Messgenauigkeit erreicht wird.

2.4.6 Externer Vergleich zum Niveau des Punktzahlvolumens

Um das Punktzahlvolumen der Versichertenstichprobe mit dem bundesweiten Punkteniveau vergleichen zu können, werden die Versichertenzahlen der KM6 mit den jeweiligen Kostengewichten der Alters- und Geschlechtsgruppe multipliziert und über alle Gruppen addiert. Die resultierende Summe wird dann mit der Summe des Leistungsbedarfs 2007 Bund aus dem Bericht des Bewertungsausschusses verglichen, wobei der Euro-Leistungsbedarf mit einem Punktwert von 3,5001 Cent umgerechnet wurde. Tabelle 14 zeigt das Ergebnis dieses Vergleichs.

Summe Punkte		Stichprobe/BMG
Versichertenstichprobe	Bericht an das BMG	
735.517.974.533	791.983.258.138	0,9287

Tabelle 14: Vergleich Punktzahlniveau Bund mit dem Punktzahlvolumen der Stichprobe (Quelle: Daten des BA und Bericht an das BMG; eigene Berechnungen und Darstellung)

2.5 Fremdkassenzahlungsausgleich

Im Beschluss des 184. Bewertungsausschusses wird unter Nr. 6., dritter Spiegelstrichpunkt, das Institut beauftragt, Auswirkungen der nicht berücksichtigten Inanspruchnahme von Leistungen außerhalb der vier KVen auf die Validität der Datengrundlage und die Belastbarkeit der ermittelten Veränderungsdaten zu prüfen.

Im Folgenden soll im ersten Teil dargestellt werden, warum die Inanspruchnahme nicht vollständig in der Datengrundlage abgebildet ist, welche theoretischen Einflüsse dieses auf die Berechnungen haben kann und welche empirischen Hinweise sich finden lassen.

¹² Eventuelle Abweichungen zu den morbiditätsbedingten Veränderungsdaten in Tabelle 12 sind dadurch zu erklären, dass in der vorliegenden Tabelle die auf der 3 Mio. Stichprobe basierenden Relativgewichte verwendet wurden.

2.5.1 Darstellung des Fremdkassenzahlungsausgleichs

Geht ein Versicherter mit Wohnort in Niedersachsen¹³ in Bremen zum Arzt, so löst dieses einen sogenannten Fremdkassenzahlungsausgleich aus. Demzufolge rechnet die KV Bremen den Fall des niedersächsischen Versicherten mit der KV Niedersachsen ab.

Die Datengrundlage des Bewertungsausschusses (vgl. Kapitel 2.1) umfasst alle Versicherten mit Wohnort in Niedersachsen, Bremen, Nordrhein und Thüringen und alle Leistungen, die in den vier KVen erbracht wurden. Daher ist das oben genannte Beispiel in der Datengrundlage abgebildet. Geht der niedersächsische Versicherte allerdings in Hamburg zum Arzt, so sind die dort gestellten Diagnosen und der abgerechnete Leistungsbedarf in der Datengrundlage des Bewertungsausschusses nicht vorhanden und der Versicherte hat ein unvollständiges Diagnosen- und Leistungsspektrum.

2.5.2 Theoretische Einflüsse des fehlenden Fremdkassenzahlungsausgleichs

Die unvollständige Darstellung von Diagnosen und Leistungsbedarf bei einem Teil der Versicherten kann sowohl bei der Berechnung der Relativgewichte als auch bei den Berechnungen zu morbiditätsbedingten Veränderungsdaten zu Verzerrungen führen. Im Folgenden sollen beide Aspekte getrennt voneinander erläutert werden.

Berechnung von Relativgewichten

Bei der Berechnung von Relativgewichten hat eine unvollständige Darstellung von Diagnosen und Leistungsbedarf Auswirkungen auf unterschiedliche Teilbereiche.

Die unvollständige Darstellung von Diagnosen kann dazu führen, dass den betroffenen Versicherten weniger Risikokategorien zugewiesen werden, als es bei vollständiger Diagnosedarstellung der Fall gewesen wäre.

Bei der eigentlichen Berechnung der Relativgewichte gibt es mehrere Effekte, die hier getrennt voneinander dargestellt werden sollen.

¹³ Die Darstellung orientiert sich an der Situation ab 1. Januar 2009, d. h. mit vollständiger Umsetzung des Wohnortprinzips.

Wenn die unvollständige Darstellung des Leistungsbedarfs einerseits und der Risikokategorien andererseits nicht „parallel“ zueinander läuft, führt eine vollständiger Darstellung der Risikokategorien in Zusammenhang mit einer unvollständigeren Darstellung des Leistungsbedarfs tendenziell zu einer Unterschätzung der Kostengewichte. Eine vollständiger Darstellung des Leistungsbedarfs im Zusammenspiel mit einer unvollständigeren Darstellung der Risikokategorien führt dazu im Gegensatz tendenziell zu einer Überschätzung der Kostengewichte.

Berechnung der Veränderungsraten

Die unvollständige Darstellung von Diagnosen der Versicherten führt zu einer Unterschätzung des Leistungsbedarfs dieser Versicherten. Da bei der Berechnung von Veränderungsraten die Relativgewichte auf einem Jahr berechnet und dann fixiert werden, hat diese Unterschätzung keine Auswirkungen, solange der Anteil der nicht verfügbaren Diagnosen in den betrachteten Jahren konstant bleibt.

Zu untersuchen ist daher vordringlich, ob es Hinweise darauf gibt, dass sich der Anteil der nicht verfügbaren Diagnosen über die drei betrachteten Jahre hinweg ändert.

2.5.3 Auswertungen zum Einfluss des Fremdkassenzahlungsausgleiches

Um mögliche Einflüsse des fehlenden Fremdkassenzahlungsausgleichs in der Datengrundlage des Bewertungsausschusses abschätzen zu können, hat das Institut des Bewertungsausschusses die Kassenärztlichen Bundesvereinigung, die Ersatzkassen und die Betriebskrankenkassen um zusätzliche Datenabfragen gebeten. Eine Darstellung der Datenlieferung erfolgt im Abschnitt 2.1.2.

Um eine Einschätzung zu erhalten, inwieweit die unvollständige Darstellung des Leistungsbedarfs zu Verzerrungen bei der Berechnung der Relativgewichte führt, wird der Anteil des Leistungsbedarfs außerhalb der Datengrundlage des Bewertungsausschusses am Gesamtleistungsbedarf der vier KVen in der Datengrundlage in Tabelle 15 dargestellt.

	2007	2008
Anteil des Leistungsbedarfs außerhalb der Datengrundlage (alle Kassen)	3,97 %	4,08 %

Tabelle 15: Anteil des Leistungsbedarfs außerhalb der Datengrundlage am Leistungsbedarf je WOP¹⁴-KV

(Quelle: Sonderabfrage durch die KBV)

In Anhang 16 wird diese Auswertung je KV (Tabelle 67) dargestellt.

Zur Abschätzung, ob die untermittelt dargestellte Darstellung von Diagnosen durch FKZ-Fälle außerhalb der Datengrundlage zu einer Verzerrung der Veränderungsrate führt (im oberen Abschnitt beschrieben), werden Daten der Ersatzkassen und der Betriebskrankenkassen herangezogen.

Tabelle 16 zeigt den Anteil der Diagnosen außerhalb der Datengrundlage an den Diagnosen gesamt je WOP-KV und Jahr für die Ersatzkassen und für die Betriebskrankenkassen. (Es wurde jeweils die Zahl der Diagnosendatensätze je Versicherten ausgewertet, nicht die Zahl der distinkten Diagnosen)

WOP-KV	Anteil der FKZ-Diagnosen an den Diagnosen gesamt je WOP-KV					
	Ersatzkassen			BKK		
	2006	2007	2008	2006	2007	2008
Bremen	2,84 %	2,51 %	2,45 %	1,52 %	1,52 %	1,52 %
Niedersachsen	4,30 %	4,02 %	3,82 %	3,20 %	3,20 %	3,20 %
Nordrhein	2,70 %	2,58 %	2,47 %	1,95 %	1,95 %	1,95 %
Thüringen	3,58 %	3,25 %	3,01 %	3,22 %	3,22 %	3,22 %

Tabelle 16: Anteil der Diagnosen außerhalb der Datengrundlage an den Diagnosen je WOP-KV für Ersatzkassen und die Betriebskrankenkassen

(Quelle: Sonderabfrage durch die Ersatzkassen und die Betriebskrankenkassen)

Um einen Vergleich der Sonderabfrage durch die Ersatzkassen und der Sonderabfrage der Betriebskrankenkassen mit der Datengrundlage des Bewertungsausschusses zu ermöglichen, ist im Anhang 16 (Abbildung 49 und Abbildung 50) ein Vergleich des Anteils der Diagnosen von außerhalb der WOP-KVen, aber innerhalb der KVen der Datengrundlage des Bewertungsausschusses, an den Diagnosen insgesamt dargestellt.

¹⁴ WOP-KV bezeichnet die jeweilige KV, in der ein Versicherter wohnt. Die Vereinbarung der Gesamtvergütung erfolgt jeweils für die WOP-KV.

Tabelle 17 fasst die Ergebnisse zu den Auswirkungen des Fremdkassenzahlungsausgleiches zusammen. Zwischen den Jahren 2007 und 2008 zeigen die Daten des vdek einen leichten Rückgang der Dokumentationslücke, während der Daten der BKK diese Verringerung nicht bestätigen. Wenn die Daten der nicht betrachteten Kassen einen ähnlichen Verlauf wie die Daten des Vdek haben, könnte dieses eine leicht positive Verzerrung der morbiditätsbedingten Veränderungsrate nach sich ziehen.

Die Verringerung der Dokumentationslücke zwischen 2006 und 2007 verläuft bei den Daten des vdek und den Daten der BKK gleichartig. Daher liegt die Vermutung nahe, dass eine mögliche Verzerrung zwischen 2006 und 2007 kaum höher als +0,2 Prozentpunkte ausfallen könnte.

	2006	2007	2008
Anteil des Leistungsbedarfs außerhalb der Datengrundlage (alle Kassen)	(k.A.)	3,97 %	4,08 %
Anteil des Leistungsbedarfs außerhalb der Datengrundlage (Ersatzkassen)	4,39 %	4,32 %	4,32 %
Anteil der Diagnosen außerhalb der Datengrundlage (Ersatzkassen)	3,34 %	3,14 %	2,99 %
Anteil der Diagnosen außerhalb der Datengrundlage (BKK)	2,54 %	2,34 %	2,39 %

Tabelle 17: Zusammenfassung der Ergebnisse zu Auswirkungen des Fremdkassenzahlungsausgleichs (Quelle: Sonderabfrage durch die Ersatzkassen (Diagnosen) und KBV (Leistungsbedarf))

2.6 Entwicklung der Diagnosen je Versicherten

Im „Bericht zur Schätzung der Morbiditätsveränderung 2008/2009 und zur Repräsentativität und Plausibilität der Datengrundlage des Bewertungsausschusses“ des Instituts des Bewertungsausschusses wurde im Abschnitt V unter 9.3 die folgende Auswertung in Tabelle 18 dargestellt.

Jahr	2004	2005	2006	2007
Anzahl der Diagnosen pro Fall	8,13	8,64	9,13	9,60
Veränderung in Prozent zum Vorjahr		6,33 %	5,64 %	5,14 %

Jahr	2004	2005	2006	2007
Anzahl der CC je Versicherten bei Variante A vor Hierarchisierung mit 5 - Steller ohne Interaction - Terms	6,44	6,77	6,93	7,03
Veränderung in Prozent zum Vorjahr		5,12 %	2,36 %	1,44 %
Anzahl der HCC je Versicherten bei Variante A nach Hierarchisierung mit 5 - Steller ohne Interaction - Terms	5,60	5,88	6,01	6,09
Veränderung in Prozent zum Vorjahr		5,00 %	2,21 %	1,33 %

Tabelle 18: Durchschnittliche Anzahl Diagnosen, CC und HCC je Versicherten (Bericht aus dem Jahr 2008)

Um einen ersten Eindruck der Entwicklung dieser Kennzahlen zu geben, wird in Tabelle 19 die Entwicklung der Diagnosen zwischen den Jahren 2006 und 2007 und 2007 und 2008 dargestellt.

Jahr	2006	2007	2008
Anzahl der Diagnosendatensätze je Versichertenjahr	27,2	28,9	30,9
Veränderung in Prozent zum Vorjahr		6,33 %	6,74 %

Tabelle 19: Durchschnittliche Anzahl der Diagnosen je Versichertenjahr

2.7 Prüfung der epidemiologischen Repräsentativität der regionalen Datengrundlage

Die Prüfung der epidemiologischen Repräsentativität erfolgt in einem separaten Bericht, der noch nicht vollständig ausgearbeitet ist. Die Untersuchungen weisen zwei Schwerpunkte auf:

1. einen Vergleich der vertragsärztlichen Diagnosen in der Datengrundlage des BA mit externen epidemiologischen Daten einerseits und
2. einen Vergleich der vertragsärztlichen Diagnosen aus den regionalen Routinedaten (vier KVen) mit einer bundesweiten Auswertung aus GKV-Routinedaten andererseits.

Die Arbeiten für diesen zweiten Berichtsteil sind weiter vorangeschritten und werden im Folgenden kurz zusammengefasst. Für diese zweite Vergleichsstrategie steht eine veröffentlichte Auswertung aus bundesweiten ambulanten Abrechnungsdaten der Gmünder Ersatzkasse (GEK) zur Verfügung [GEK Report „ambulant-ärztliche Versorgung 2008“]. Hier wird zunächst geprüft, inwieweit die Er-

gebnisse der GEK aus allen 17 KV-Bereichen mit Ergebnissen für dieselbe Kasse aus den vier dem BA zur Verfügung stehenden KV-Bereichen übereinstimmen.

Ausgewertet und analysiert sind bisher Prävalenzen, wie sie sich aus den vertragsärztlichen Diagnosen ergeben, auf der Ebene sehr grober Krankheitsgruppen, nämlich der Ebene der ICD-Kapitel. Die Ergebnisse aus den vier KVen und Deutschland insgesamt sind direkt vergleichbar, da sie sich auf die Versicherten derselben Kasse beziehen, einen gemeinsamen Zweijahreszeitraum 2006/07 abdecken und sämtlich auf die deutsche Wohnbevölkerung von 2005 alters- und geschlechtsadjustiert sind.

Die Versichertenstichprobe des BA ist hinsichtlich Übereinstimmung auf der Ebene von Diagnosekapiteln bei grob-orientierender epidemiologischer Durchsicht für ihre Subgruppe der ca. 100.000 GEK-Versicherten als „unauffällig“ einzuschätzen. Das Ausmaß der Abweichung zu den bundesweiten GEK-Report-Daten beträgt auf der Ebene der einzelnen Diagnosekapitel in 2006 maximal 0,0%-Punkte und minimal -1,9%-Punkte (Spannweite 1,9%-Punkte und Mittelwert -0,7%-Punkte) und in 2007 maximal -0,1%-Punkte und minimal -2,3%-Punkte (Spannweite 2,2%-Punkte und Mittelwert -0,9%-Punkte). Angesichts der – wie vorher schon dargestellt – nicht ganz vollständigen Abrechnungs- und Diagnosedaten in der BA-Stichprobe ist diese leichte Unterschätzung der bundesweiten GEK-Prävalenzen zu erwarten.

Die Veränderungsrate von 2006 zu 2007 ist in der Klumpenstichprobe der vier KVen auf der Ebene der ICD-Kapitel tendenziell etwas geringer ausgeprägt als in den bundesweiten GEK-Report-Daten. Für das Jahr 2008 liegt noch kein GEK-Report vor.

Als Zwischenfazit kann festgehalten werden, dass das Exempel der GEK Hinweise darauf liefert, dass die Veränderungsrate 2006/07 durch die regionale Zusammensetzung der BA-Stichprobe leicht unterschätzt wird.

2.8 Zusammenfassung der Vollständigkeits- und Repräsentativitätsprüfungen

Um die Ergebnisse aus den Vollständigkeits- und Repräsentativitätsprüfungen zusammenzufassen, zeigt Tabelle 20 einen Überblick über alle möglichen Problemfelder und deren geschätzte Verzerrung auf die diagnosebezogenen Veränderungsdaten. Wenn nur das Vorzeichen einer Verzerrung abschätzbar war, aber noch keine weitere Quantifizierung stattgefunden hat, ist als Eintrag < 0 bzw. > 0

vermerkt. Die quantifizierbaren Verzerrungen sind betragsmäßig angegeben, die abschließende Zeile fasst nur diese quantitativen Positionen zusammen, eine Zusammenfassung der qualitativen Angaben ist wegen der unterschiedlichen Vorzeichen nicht möglich.

Mögliches Problemfeld	Maximale geschätzte Verzerrung auf die diagnosebezogenen Veränderungsraten	
	2006/07	2007/08
1. Dokumentationslücken wg. fehlender Zusammenführbarkeit	Im Wesentlichen behoben (Ausschluss technisch unzulässiger Versichertenmeldungen, rd. 2,7–2,9 % der Versicherten)	
2. Verbliebene Dokumentationslücken		
a) Kassenart 1	- 0,3	+ 0,4
b) Kassenart 2	0*	- 0,4
3. Dokumentationslücken wegen		
1. Fehlender Fremdfälle	+ 0,2	> 0 (k.A.)
2. Sonstige	n.v.	< 0 (k.A.)
4. Dokumentationslücken bei unterjährig Versicherten (Wechsler)	> 0 (k.A.)	-,-
5. Unvollständig dokumentierte Versichertenzeiten	(k.A.)	(k.A.)
6. Fehlerhaft niedrige Versichertenzeiten (Ersatzkasse 1)	-,-	(k.A.)
7. Fusionsbedingte Probleme über die o. g. hinaus	> 0 (k.A.)	> 0 (k.A.)
8. Repräsentativität bzgl. Verteilung Stadt/Land	0*	0*
9. Repräsentativität bzgl. demographischer Struktur	+ 0,1	+ 0,1
10. Repräsentativität im Diagnosenmix	< 0 (k.A.)	n.v.
Untersuchte Effekte zusammen	0,0	+ 0,1

*Tabelle 20: Zusammenfassung Problemfelder
(Hinweis: 0*: keine nennenswerte Verzerrung,
-,- keine Verzerrung festgestellt,
k.A. geringfügige Verzerrung, noch nicht abgeschätzt
n.v.: nicht verfügbar)*

1. Da die Datenlieferungen der Versichertenstammdaten an die Datenstelle des Bewertungsausschusses nicht immer fehlerfrei sind, müssen je nach betrachtetem Jahr zwischen 2,7 % und 2,9 % aufgrund von Formalfehlern aus der Versichertenstichprobe ausgeschlossen werden (vgl. Kapitel 2.3.1).

2. Ein weiteres mögliches Problemfeld sind Schwankungen in der ärztlichen Inanspruchnahme (vgl. Kapitel 2.3.4). Im Fall der Kassen der Kassenart 1 könnten mögliche Verzerrungen in der Veränderungsrate auftreten, die sich auf bis zu -0,31 Prozentpunkte bezüglich der Rate 2006/07 und +0,40 Prozentpunkte bezüglich der Rate 2007/08 belaufen. Das Problem, welches für die beiden Kassen der Kassenart 2 besteht, führt hingegen zu gegensätzlichen Verzerrungen, die für die Veränderungsrate 2006/07 nicht nennenswert sind und für die Rate 2007/08 etwa -0,39 Prozentpunkte betragen. Beides zusammen genommen führt zu der Einschätzung, dass die Raten hierdurch nicht beeinflusst oder geringfügig verzerrt werden.
3. In der Stichprobe kann die Inanspruchnahme von Leistungen außerhalb der vier KVen nicht berücksichtigt werden, was die Veränderungsraten möglicherweise verzerren könnte. Die Untersuchung in Kapitel 2.5 zeigt, dass der Anteil der nicht beobachtbaren Diagnosen 2006 auf 2007 ansteigt, was einen Anstieg der Veränderungsrate 2006/07 von etwa 0,2 Prozentpunkten bewirken könnte. Zwischen 2007 und 2008 verlaufen die Entwicklungen bei den beiden Datenquellen nicht gleichartig. Es könnte zu einer leicht positiven Verzerrung kommen, die sich aber von Null nicht deutlich unterscheidet. Bezüglich sonstiger Dokumentationslücken (vgl. Kapitel 2.3.2) können für die morbiditätsbedingte Veränderungsrate 2006/07 keine Aussagen getroffen werden, da die C4-Daten für das Jahr 2006 nicht vorliegen. Für die Veränderungsrate 2007/08 ergibt sich eine geringfügige negative Verzerrung, deren Höhe durch die Art der Diagnosen bestimmt wird und noch nicht abgeschätzt ist.
4. Die Abnahme der unterjährig Versicherten von 2006 auf 2007 könnte eine geringfügige positive Verzerrung auf die Veränderung des Morbiditätsindex haben, was bisher aber noch nicht genau abgeschätzt werden konnte. Eine Abschätzung möglicher Verzerrungen in der Veränderungsrate 2007/08 ist aufgrund der Überlagerungen der Fehlerquellen bei den unterjährigen Versicherten derzeit nicht möglich (vgl. Kapitel 2.3.3).
5. Die Sonderuntersuchungen zu fehlenden Versichertenzeiten in Kapitel 2.3.3 ergaben, dass nicht alle Kassen die Versichertendatensätze stets vollständig übermittelt hatten. Die verzerrende Wirkung auf die Veränderungsrate konnte noch nicht ermittelt werden.
6. Die in Kapitel 2.3.3 durchgeführte Sonderuntersuchung zu Quartalsdatensätzen mit fälschlicherweise verkürzten Versichertenzeiten zeigte für die Ersatz-

kasse 1 von 2007 auf 2008 einen starken Anstieg der unterjährig Versicherten. Die sich daraus ergebenden Verzerrungen auf die Veränderungsrate konnten noch nicht abgeschätzt werden, dürften aber nur sehr klein sein.

7. Die Wirkung der fusionsbedingten Probleme auf die Veränderungsrate konnte noch nicht abgeschätzt werden. Es wird erwartet, dass beide Veränderungsraten leicht überschätzt werden. (vgl. Kap. 2.3.5).
8. Eine verzerrende Wirkung könnte auch von einer nicht-repräsentativen und über die Jahre hinweg schwankende Verteilung der Bevölkerung auf städtische und ländliche Regionen ausgehen (vgl. Kapitel 2.4.1). Sowohl die Abweichungen zwischen den betrachteten Jahren als auch die Abweichungen zwischen Versichertenstichprobe und KM6 betragen deutlich weniger als einen Prozentpunkt. Aus diesem Grund kann dieser mögliche Problembereich keine nennenswerte verzerrende Wirkung auf die Berechnung der Veränderungsraten haben.
9. Die Berechnungen in Kapitel 2.4.3 ergaben, dass über die betrachteten Jahre hinweg die Veränderungen der demographischen Morbiditätsindizes in Datengrundlage und GKV insgesamt recht ähnlich ausfielen. Der verzerrende Effekt wurde dort auf einen Promillepunkt beziffert.
10. Das Exempel der GEK liefert Hinweise darauf, dass die Veränderungsrate 2006/07 durch die regionale Zusammensetzung der BA-Stichprobe leicht unterschätzt wird (vgl. Kapitel 2.7). Für das Jahr 2008 liegt noch kein GEK-Report vor, sodass für die Veränderungsrate 2007/08 keine Aussage getroffen werden kann.

Fasst man die quantifizierten Effekte der möglichen Problemfelder zusammen, ergibt sich für die Veränderungsrate 2006/07 eine Verzerrung von 0,0 und für die Veränderungsrate 2007/08 eine Verzerrung von +0,1 Prozentpunkten.

Die lediglich qualifizierten Effekte weisen zum Teil in unterschiedliche Richtungen.

Korrekturfaktoren für die in Kapitel 1 bezifferten Veränderungsraten lassen sich daher nicht angeben.

2.9 Analysen zur Schätzgenauigkeit bei der Bestimmung von morbiditätsbedingten Veränderungsraten

Die Bestimmung von morbiditätsbedingten Veränderungsraten wird in der mathematischen Statistik als Durchführung einer „Punktschätzung“ verstanden. Die Raten basieren auf Stichproben von Versicherten und ihren Leistungen und Diagnosen; sie können daher unterschiedlichen Fehlern unterworfen sein:

1. Die Punktschätzung kann verzerrt sein. Gründe dafür können beispielsweise in fehlender Repräsentativität der Stichprobe für die Grundgesamtheit (hier: die gesetzlich Krankenversicherten insgesamt) liegen.
2. Eine Punktschätzung weist immer eine (Stichproben-bedingte) Schwankungsbreite auf. Der wahre Wert in der Grundgesamtheit kann höher oder niedriger ausfallen als das Stichprobenergebnis.

Die hier in Abschnitt 1 vorgelegten Untersuchungen beziehen sich auf den zuerst genannten möglichen Fehler (Verzerrungen). Der Frage der Höhe des Stichprobenfehlers (Schwankungsbreite) geht dieser Bericht nicht nach.

Zu diesem zweiten Aspekt liegt dem Bewertungsausschuss allerdings eine aktuelle Expertise von Prof. Schäfer (ISEG)¹⁵ vor, in der erste Abschätzungen für den Stichprobenfehler der morbiditätsbedingten Veränderungsraten gegeben werden.

Insofern ergänzen sich die beiden in etwa zeitgleich vorgelegten Berichte.

¹⁵ Das Gutachten „Expertise zur Eignung der Datengrundlage des Bewertungsausschusses zur Ermittlung der Veränderung der Morbiditätsstruktur gemäß § 87a Abs. 5 SGB V unter stichprobentheoretischen Gesichtspunkten“ ist unter dem Link www.gkv-spitzenverband.de/upload/prof_schaefer_expertise_stichprobenqualität2010_14081.pdf (Stand 22.07.2010) abrufbar.

3 Berechnung von Veränderungsdaten

Die Beauftragung des Bewertungsausschusses aus dem Beschluss der 184. Sitzung enthielt neben den Plausibilitäts- und Repräsentativitätsprüfungen des Kapitels 2 auch Aufträge zur Berechnung von Veränderungsdaten. Im folgenden Kapitel wird zuerst eine Einführung zum theoretischen Hintergrund der Berechnungen gegeben, anschließend werden Ergebnisse vorgestellt.

3.1 Hintergrund zur Modellierung

Das Ziel dieses Abschnittes ist es, die unterschiedlichen Möglichkeiten Veränderungsdaten zu berechnen darzustellen. Die Berechnung basiert auf einer Klassifikation. Klassifikation stellt in diesem Zusammenhang die Zuordnung von ICD-Kodes zu Risikokategorien (CC = condition category) dar. Die CC stellen die zentralen Bausteine der Klassifikation dar; sie werden im Folgenden als „diagnosebezogene Haupteffekte“ bezeichnet. Darüber hinaus gibt es in den Modellen gruppierte demographische Angaben der Versicherten (Alters- und Geschlechtsgruppe), sie werden im Folgenden als „demographische Haupteffekte“ bezeichnet. In einigen Ausgestaltungen gibt es darüber hinaus noch Kategorien, die Wechselwirkungen zwischen Alter und einem diagnosebezogenen Haupteffekt abbilden (K(H)CC, O(H)CC). Weitere mögliche Wechselwirkungsterme bilden das Zusammentreffen von bestimmten diagnosebezogenen Haupteffekten ab.

3.1.1 Darstellung der Varianten A, F, G und H

Die Varianten A, F, G und H bezeichnen Klassifikationsmodelle.

Variante A wurde von der Arbeitsebene des Bewertungsausschusses in Zusammenarbeit mit DxCG erstellt. Die Varianten F und G sind Weiterentwicklungen dieser Arbeiten durch das Institut des BA in Zusammenarbeit mit der Arbeitsebene des Bewertungsausschusses. Variante H ist ein Vorschlag des unparteiischen Vorsitzenden des Bewertungsausschusses auf Basis von Variante F und G. Bei der Weiterentwicklung wurde die Zuordnung von 340 ICD-Kodes geändert.

Varianten A und F enthalten jeweils nur Diagnosen mit dem Zusatzkennzeichen „G“ (gesichert). Die Klassifikation der Diagnosen mit Zusatzkennzeichen „G“ entspricht in den Varianten G und H der Variante F. Variante H enthält neben der Klassifikation der „G“-Diagnosen noch eine CC mit ausgewählten Diagnosen mit Zusatzkennzeichen „Z“. Variante G enthält über die Risikokategorien der Varian-

te H hinaus zwei CCs mit ausgewählten Diagnosen mit Zusatzkennzeichen „A“ (Ausschluss von).

Tabelle 21 gibt eine Übersicht über die Eckwerte der Varianten.

Variante	Beschreibung	Zahl der diagnosebezogenen Haupteffekte
A	Nur Diagnosen mit Zusatzkennzeichen „G“.	193
F	Nur Diagnosen mit Zusatzkennzeichen „G“.	195
G	Diagnosen mit Zusatzkennzeichen „G“ und ausgewählte Diagnosen mit Zusatzkennzeichen „A“ und „Z“.	198
H	Diagnosen mit Zusatzkennzeichen „G“ und ausgewählte Diagnosen mit Zusatzkennzeichen „Z“.	196

Tabelle 21: Übersicht zu den Variante A, F, G und H

Im Anhang 18 wird ein Überblick über die Erarbeitung der Varianten F und G gegeben.

3.1.2 Erläuterung eines „kompletten“, „reduzierten“ oder „komprimierten“ Modells

Für die diagnosebezogenen und die demographischen Haupteffekte (sowie für weitere Risikokategorien) werden mithilfe einer Regressionsrechnung jeweils die Koeffizienten dieser bestimmt. Die Koeffizienten werden in der weiteren Berichtslegung als Kostengewichte bezeichnet. Aus jedem Kostengewicht wird ein Relativgewicht berechnet, indem das entsprechende Kostengewicht durch den durchschnittlichen Leistungsbedarf aller Versicherten, auf denen die Kostengewichte berechnet wurden, geteilt wird. Die Kosten- bzw. Relativgewichte stellen den zusätzlichen Leistungsbedarf eines Versicherten aufgrund dieses Haupteffektes (bzw. dieser Risikokategorie) dar.

3.1.2.1 Komplette Modelle

Ausgangspunkt für die Erstellung aller Modelle sind jeweils die Varianten nach Abschnitt 3.1.1.

Beim kompletten Modell werden den diagnosebezogenen Haupteffekten der jeweiligen Variante noch 32 Alters- und Geschlechtskategorien und je diagnosebezogenem Haupteffekt eine weitere Risikokategorie für Versicherte unter 18

Jahre und eine für Versicherte über 64 Jahre hinzugefügt (K(H)CC und O(H)CC). Im Ergebnis umfasst ein komplettes Modell vor Beginn der Modellierung rund 650 Risikokategorien. Im Rahmen der Modellierung werden Risikokategorien, die eine zu geringe Besetzung aufweisen, nicht statistisch relevant oder negativ sind aus dem Modell ausgeschlossen. Am Ende der Modellierung verbleiben in dem Modell 200–300 Risikokategorien. Diese definieren das komplette Modell.

Der Modellierung liegt eine sogenannte Kalibrierungsstrategie zugrunde. Die Kalibrierungsstrategien für komplette Modelle finden sich in Anhang 18 (hierarchisierte Modelle) und Anhang 19 (nicht hierarchisierte Modelle).

3.1.2.2 Reduzierte Modelle

Beim reduzierten Modell werden den Risikokategorien der jeweiligen Variante die 32 Alters- und Geschlechtskategorien hinzugefügt und sechs Risikokategorien ausgeschlossen¹⁶. Auf diesen knapp 230 Risikokategorien wird eine Regressionsrechnung durchgeführt, ohne dass weitere Modellierungsschritte vorgenommen werden. Anhand der berechneten Relativgewichte werden die 30 Risikokategorien mit der höchsten ökonomischen Relevanz (berechnet aus Häufigkeit der Risikokategorie multipliziert mit dem Relativgewicht) und Risikokategorien mit einem Relativgewicht von mehr als 1,5 ausgewählt. Alle anderen Risikokategorien werden gemäß der Zuordnung zu ACC¹⁷ zusammengefasst; diese neu zusammengefassten Risikokategorien kürzen wir mit „RCC“ ab. Die verbliebenen Risikokategorien, die RCC und die Alters- und Geschlechtsvariablen umfassen das reduzierte Modell. Es enthält ca. 90 Risikokategorien.

Eine Zuordnung der Risikokategorien zu den ACC findet sich in Anhang 22.

Eine Modellausgestaltung weist eine Besonderheit auf, die in Abschnitt 3.3.3 erläutert wird.

Die Kalibrierungsstrategie zur Modellierung von reduzierten und komprimierten Modellen (siehe 3.1.2.3) findet sich in Anhang 20. Es soll an dieser Stelle auf

¹⁶ Ausgeschlossen werden die CC 166 (Schwere Symptome, abnorme Befunde); CC 167 (Schwindel, Unwohlsein und andere Allgemeinsymptome), CC 179 (verschiedene Eingriffe, Status nach operativen Eingriffen, Nachbehandlung, elektive Chirurgie), CC 182 (Rehabilitation), CC 183 (Vorsorgeuntersuchungen, Impfungen, Beratung zur Kontrazeption) und CC 184 (Erkrankungen in der Eigen- oder Familienanamnese)

¹⁷ Aggregated Condition Category (ACC) ist eine übergeordnete Kategorie aus der Original-Zuordnung von DxCG. Sie ordnet die Risikokategorien 30 Gruppen zu.

Schritt 5 hingewiesen werden. Ist das Kostengewicht einer RCC negativ, so wird folgendes Vorgehen gewählt:

1. Es werden die Korrelationskoeffizienten der RCC mit allen anderen RCC berechnet.
2. Die drei RCC mit dem jeweils höchsten Korrelationskoeffizienten aus der Korrelation mit der RCC mit negativem Kostengewicht werden für die weiteren Schritte herangezogen.
3. Tritt die RCC mit negativem Kostengewicht gemeinsam mit einer der drei am höchsten korrelierenden RCC bei einem Versicherten auf, wird die RCC mit negativem Kostengewicht unterdrückt (auf Null gesetzt).
4. Anschließend werden die Kostengewichte mittels der linearen Regression erneut berechnet.
5. Bleibt das Kostengewicht der RCC negativ, so wird das Kostengewicht dieser RCC auf Null gesetzt und erneut eine Regression durchgeführt.

In Abschnitt 3.2.1 (Tabelle 25) wird ein Überblick über die sich so ergebenden Beziehungen zwischen RCCs gegeben.

3.1.2.3 Komprimierte Modelle

Bis auf den Ausschluss der sechs Risikokategorien am Anfang des Algorithmus (Schritt 1 in Anhang 20) entspricht das komprimierte Modell dem reduzierten Modell. Im Gegensatz zum reduzierten Modell umfasst es aber den kompletten Diagnosesatz.

3.1.3 Erläuterung „hierarchisiert“ vs. „nicht hierarchisiert“

3.1.3.1 Allgemeine Darstellung

Bei der Hierarchisierung werden Gruppen von Risikokategorien in eine Reihenfolge gebracht. Die Gruppen umfassen jeweils Risikokategorien, die aus einem medizinischen Bereich (meistens aus einer ACC) kommen, und werden nach ökonomischer Schwere sortiert. Erhält ein Versicherter verschiedene Risikokategorien aus einem Hierarchiebaum, so wird nur die Risikokategorie in der höchsten Hierarchiestufe gewertet. In der Hierarchie darunter liegende Risikokategorien werden für diese Versicherten ausgeblendet (auf Null gesetzt). Abbildung 8 zeigt einen fiktiven Hierarchiebaum. Hat ein Versicherter die Risikokategorie A und B, so wird B auf Null gesetzt. Hat ein Versicherter die Risikokategorien B und C, so

hat die Hierarchisierung keine Wirkung, da beide Risikokategorien auf der gleichen Hierarchiestufe stehen.

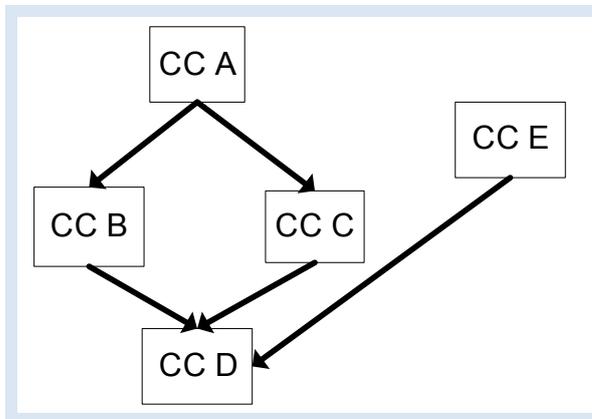


Abbildung 8: Beispiel eines Hierarchiebaums

Bei einem nicht hierarchischen Modell wird das oben beschriebene Verfahren nicht angewendet.

3.1.3.2 Hierarchisierung in komprimierten und reduzierten Modellen

Die Hierarchisierung im komprimierten und reduzierten Modell basiert auf der Hierarchisierung des kompletten Modells. Eine hierarchische Beziehung aus dem kompletten Modell bleibt erhalten, insofern die beteiligten Risikokategorien im reduzierten Modell (einzeln oder gemeinsam) bestehen bleiben. Hierarchische Beziehungen, die im kompletten Modell nicht bestanden hatten, werden nicht hinzugefügt.

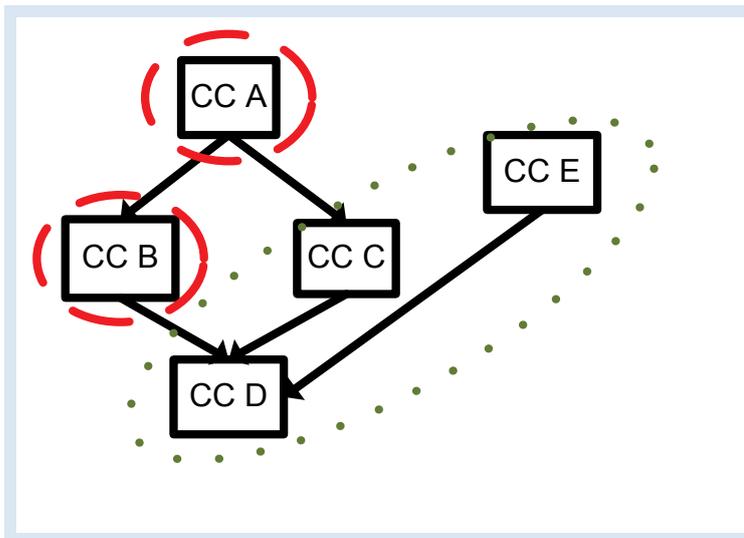


Abbildung 9: Erstes Beispiel für Hierarchisierung im reduzierten und komprimierten Modell

Bei den in Abbildung 9 und Abbildung 10 dargestellten Beispielen stellen die gestrichelt umrandeten CCs jeweils Risikokategorien dar, die im reduzierten bzw. komprimierten Modell einzeln stehen bleiben, die gepunktet umrandeten CCs stellen die Risikokategorien dar, die zu einer RCC zusammengefasst werden. Im Fall des 1. Beispiels würde es eine Hierarchisierung „CC A“ über „CC B“ geben, die RCC steht in keiner hierarchischen Beziehung zu den verbliebenen CC. Im 2. Beispiel steht die RCC über der „CC D“.

Im Anhang 7 findet sich die Hierarchietabelle für reduzierte und komprimierte Modelle.

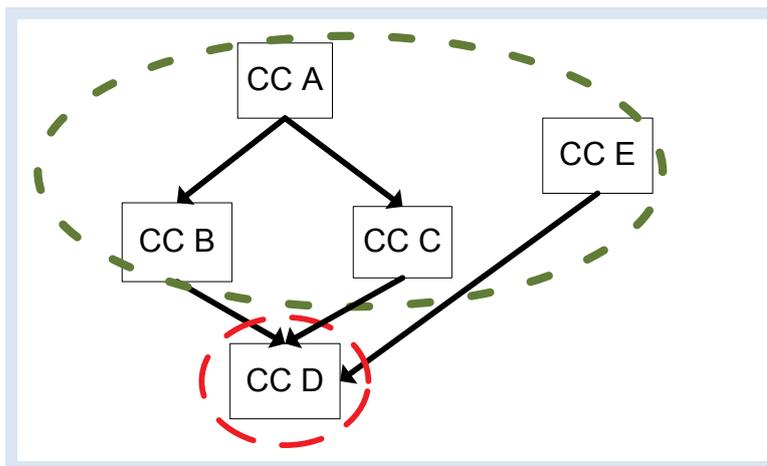


Abbildung 10: Zweites Beispiel für Hierarchisierung im reduzierten und komprimierten Modell

3.1.4 Erläuterung „zeitgleich“ vs. „prospektiv“

Auf Basis der Risikokategorien eines Modells werden für diese jeweils Relativ- bzw. Kostengewichte geschätzt. Anhand dieser Gewichte und der kodierten Krankheiten eines Versicherten wird anschließend der Leistungsbedarf je Versicherten mit unterschiedlichem Zeithorizont geschätzt. Der Zeithorizont der Schätzung muss sich auch in der Berechnung der Kostengewichte wieder finden.

Bei einem zeitgleichen Modell werden die Diagnosen eines Jahres verwendet, um den Leistungsbedarf desselben Jahres zu schätzen. Bei einem 1-Jahresprospektiven Modell werden die Diagnosen eines Jahres verwendet, um den Leistungsbedarf des Folgejahres (Jahr 2) zu schätzen. Bei einem 2-Jahresprospektiven Modell wird mit den Diagnosen eines Jahres der Leistungsbedarf zwei Jahre später geschätzt. Abbildung 11 zeigt eine Darstellung des Zusammenhangs von Diagnosen und geschätztem Leistungsbedarf je nach Zeithorizont.

	Jahr 1	Jahr 2	Jahr 3
zeitgleich	geschätzter LB Diagnosen		
1-Jahres-Prospektiv	Diagnosen	geschätzter LB	
2-Jahres-Prospektiv	Diagnosen		geschätzter LB

Abbildung 11: Darstellung des Zeithorizontes bei der Schätzung des Leistungsbedarfs

Wie im ersten Abschnitt erwähnt, muss der Zeitbezug der Schätzung des Leistungsbedarfs schon bei der Berechnung der Kosten- bzw. Relativgewichte berücksichtigt werden. Abbildung 12 zeigt, aus welchen Jahren der Leistungsbedarf und die Diagnosen, für den jeweiligen Zeitbezug, gewählt werden müssen.

	Jahr 1	Jahr 2	Jahr 3
zeitgleich	tatsächlicher LB Diagnosen		
1-Jahres-Prospektiv	Diagnosen	tatsächlicher LB	
2-Jahres-Prospektiv	Diagnosen		tatsächlicher LB

Abbildung 12: Darstellung des Zeithorizonts bei der Berechnung der Kosten-/Relativgewichte

3.1.5 Datengrundlage des Bewertungsausschusses

Die Berechnungen der Veränderungsraten werden auf der Datengrundlage des Bewertungsausschusses durchgeführt. Eine genaue Beschreibung dieser Datengrundlage findet sich im Abschnitt 2.1, in diesem Abschnitt soll nur ein kurzer Überblick gegeben werden.

3.1.5.1 Umfang der Datengrundlage

Die Datengrundlage basiert auf einer Vollerhebung aus den Kassenärztlichen Vereinigungen Bremen, Niedersachsen, Nordrhein und Thüringen. Es werden alle Versicherten nach Wohnortprinzip in die Datengrundlage aufgenommen. Die Versichertenstammdaten (Alter, Geschlecht, Zahl der Versichertentage) werden von den Krankenkassen geliefert, die Diagnose- und Leistungsdaten von den KVen. Anhand der pseudonymisierten Versichertennummer werden die Datensätze zusammengeführt.

Aus der Vollerhebung wird ein Drittel der Stichprobe gezogen (etwa 6 Mio. Versicherten je Jahr) und diese für Berechnungen geteilt in eine Lernstichprobe (3 Mio. Versicherte) und eine Validierungsstichprobe (etwa 2,8 Mio. Versicherte).

Die Lernstichprobe ist für zeitgleiche Modelle identisch mit der Kalibrierungsstichprobe, anhand derer die Modellbildung und die Berechnung von Relativgewichten erfolgt. Bei den prospektiven Modellen ist die Kalibrierungsstichprobe nur eine Teilmenge der Lernstichprobe, die wie folgt gebildet wird. Die Lernstichprobe des Jahres des zu verwendenden Leistungsbedarfes (Leistungsjahr) wird über die ID der Versicherten mit den Versicherten und ihren Diagnosen im Diagnosejahr verknüpft. Anschließend werden nur diejenigen Daten für die Kalibrierung ausgewählt, für welche Daten aus beiden Jahren (Diagnosejahr und Leistungsjahr) vorliegen. Die Kalibrierungsstichprobe der prospektiven Modelle enthält somit keine vor dem Leistungsjahr verstorbenen oder in die PKV gewechselten Versicherten, keine nach dem Diagnosejahr neugeborenen Versicherten und keine Versicherten, die im Übergang vom Diagnosejahr zum Leistungsjahr die Kasse oder KV gewechselt haben¹⁸. Unterjährig Versicherte im Diagnosejahr oder im Leistungsjahr werden vollständig von der Lernstichprobe in die Kalibrierungsstichprobe übernommen.

3.1.5.2 Ausschluss von Datensätzen

Versicherte

Wenn bei der Qualitätskontrolle der von den Krankenkassen gelieferten Daten bestimmte Auffälligkeiten auftreten, werden die Versicherten dieser Kranken-

¹⁸ In Ausnahmefällen kann es dazu kommen, dass die Kalibrierungsstichprobe für 2-Jahresprospektive Modelle Versicherte enthält, die die Kasse oder KV gewechselt haben, wenn diese anschließend mit derselben Versichertennummer wieder „zurückkehren“.

kassen aus der Stichprobe ausgeschlossen. Um Verzerrungen zwischen den Jahren zu vermeiden, erfolgt der Ausschluss jeweils in allen Jahren der Stichprobe. Eine genaue Erläuterung der Ausschlüsse und der dazugehörigen Statistik findet sich im Abschnitt 2.2.3.

Diagnosen

Diagnosen von nicht zur Diagnosekodierung verpflichteten Ärzten werden bei der Berechnung der Veränderungsraten nicht berücksichtigt (eine Auflistung der Arztgruppen findet sich in Anhang 8). Darüber hinaus werden Diagnosen, die keine Arztgruppenkennung tragen, bei den Berechnungen nicht berücksichtigt. Vereinzelt weitere Ausschlussgründe finden sich in Abschnitt 2.2.4.

3.1.5.3 Anpassen des Leistungsbedarfs an die aktuelle Beschlusslage des Bewertungsausschusses

Um eine Vergleichbarkeit zwischen den Ausgangswerten (sog. korrigierter Behandlungsbedarf gemäß § 87c Abs. 4 SGB V) und den Veränderungsraten herzustellen, wird hier (für die Berechnung von Veränderungsraten) der Leistungsbedarf an die Beschlusslage des Bewertungsausschusses für das Jahr 2009 angepasst (Beschlüsse des 8. EBA, 2008). Darüber hinaus erleichtert diese Vorgehensweise die Vergleichbarkeit von Kosten- und Relativgewichten zwischen den Jahren 2007 (EBM 2000 plus) und 2008 (EBM 2008).

Die Anpassungsschritte umfassen die folgenden Punkte:

- Die Trennung in Leistungen innerhalb der morbiditätsbedingten Gesamtvergütung (MGV) und außerhalb der morbiditätsbedingten Gesamtvergütung (EGV) wird nachvollzogen.
- Euro-Leistungen werden mittels Orientierungswert 2009, (3,5001 Cent) in Punkte umgerechnet (gilt für MGV und EGV).
- Die im 8. EBA beschlossenen Anpassungsfaktoren für die Bewertung einzelner Leistungen im EBM für das Jahr 2009 werden bei Punktzahlleistungen der MGV und EGV angewendet.
- Die im 8. EBA beschlossene Veränderungsrate der morbiditätsbedingten Leistungsmenge je Versicherten des Jahres 2009 gegenüber dem Jahr 2008 gemäß § 87c Abs. 4 SGB V wird auf Punktzahl- und Euroleistungen innerhalb der MGV angewendet (5,1 %).

- Abrechnungsdaten aus 2007 werden um EBM08-Effekt fortentwickelt (entfällt für Datenbasis 2008). Dieser ist innerhalb der MGV einheitlich 1,097; außerhalb der MGV segmentspezifisch (Stand 8. EBA).
- Die vom 8. EBA beschlossenen HVV-Quoten werden bei der herangezogenen Datengrundlage des Bewertungsausschusses aus vier KV-Bereichen nicht angewendet, um eine Verzerrung der angeforderten, sachlich-rechnerisch richtiggestellten Leistungsmengen durch unterschiedliche Honorarbegrenzungsregelungen in den KV-Bereichen zu vermeiden.

Die weiteren Verarbeitungen (Modellierungen und Berechnung von Veränderungsdaten) verwenden ausschließlich kalkulatorische Punktzahlen aus der MGV, brutto (d. h. vor Anwendung honorarwirksamer Mengenbegrenzungen) und angepasst, wie eben beschrieben.

3.2 Kennzahlen und Gütemaße der Modellierung

Dieser Abschnitt berichtet für die unterschiedlichen Kombinationen der Modellierungsmöglichkeiten, wie sie im Abschnitt 3.1 beschrieben wurden, Kennzahlen und Gütemaße.

Tabelle 22 gibt einen Überblick über die beauftragten Varianten. „Variante“ stellt dabei die Variante gemäß Abschnitt 3.1.1 dar, „Umfang“, ob das Modell „komplett“, „reduziert“ oder „komprimiert“ ist (vgl. Abschnitt 3.1.2), „Hierarchisierung“ bezieht sich darauf, ob das Modell hierarchisiert wurde (vgl. Abschnitt 0), „Zeitbezug LB“ bezieht sich auf das Jahr, aus dem der Leistungsbedarf in der Modellierung verwendet wurde, und „Zeithorizont der Modellierung“ verweist auf eine prospektive oder zeitgleiche Modellierung (vgl. Abschnitt 3.1.4). Die Spalte „bestellt von“ gibt als weitere Information an, wer das Modell beauftragt hat. In der Spalte „Abgrenzung MGV“ bedeutet „V404“ (MART InBA 4) und „V309“ (MART InBA 3), dass die Abgrenzung MGV zu EGV der Abgrenzung entspricht, die für das Jahr 2009 gültig war. Der Eintrag „X404“ (MART InBA 4) bedeutet, dass abweichend von der Abgrenzung MGV / EGV, wie sie im Jahr 2009 gültig war, noch Dialyse-Sachkosten in die EGV aufgenommen wurden und damit der Leistungsbedarf dieses Bereiches in den morbiditätsbedingten Veränderungsdaten nicht berücksichtigt wird. Die Modellausgestaltung „reduziert 130“ weist eine Besonderheit auf, die in Abschnitt 3.3.3 erläutert wird.

Während der Bearbeitung hat der unparteiische Vorsitzende neue Modellausgestaltungen beauftragt und für die anderen die Prioritäten geändert. Auf Wunsch

des unparteiischen Vorsitzenden sind diese Modelle nicht Gegenstand dieses und der folgenden Abschnitte.

Variante	Umfang	Hierarchisierung	Zeitbezug LB	Zeithorizont der Modellierung	Abgrenzung MGV	bestellt von
A	Reduziert	mit	2007	1 Jahr	V404	GKV
F	Reduziert	ohne	2007	1 Jahr	V404	GKV
F	Reduziert	ohne	2008	2 Jahre	V404	GKV
F	Reduziert	mit	2007	1 Jahr	V404	GKV
F	Reduziert-130	mit	2008	1 Jahr	V404	GKV
F	Komplett	ohne	2007	0 Jahre	V309	KBV
F	Komplett	ohne	2008	2 Jahre	V404	KBV
F	Komplett	mit	2008	2 Jahre	V404	EBA
G	Komplett	ohne	2007	0 Jahre	V309	KBV
G	Komplett	ohne	2008	0 Jahre	V404	BA
G	Komplett	ohne	2008	2 Jahre	V404	KBV
G	Komplett	ohne	2008	2 Jahre	X404	KBV

Tabelle 22: Ausgestaltung der beauftragten Varianten

3.2.1 Kennzahlen der Modellierung

Die Tabellen Kennzahlen (Tabelle 23 für reduzierte Modelle und Tabelle 24 für komplette Modelle) berichtet wichtige Informationen zu den Modellen. „N (Kalibrierungsstichprobe aus der Lernstichprobe)“ gibt die Zahl der Versicherten an, auf der die Relativgewichte modelliert wurden. Die Spalten „Zahl der verbliebenen Risikokategorien“ geben für die unterschiedlichen Arten von Risikokategorien an, wie viele in dem Modell verblieben sind. Die Spalte „Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz“ gibt den Anteil der auf die Alters- und Geschlechtsgruppen entfallenden ökonomischen Relevanz (berechnet als Relativgewicht mal absolute Häufigkeit) an der gesamten ökonomischen Relevanz an. Die Spalte „Kurzbeschreibung“ gibt eine kurze Zusammenfassung der Modellvarianten, die ausführliche Darstellung der beauftragten Varianten erfolgt in Tabelle 22.

Kurzbeschreibung	Zahl der verbleibenden Risikokategorien					Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz
	N (Kalibrierungsstichprobe aus der Lernstichprobe)	Demographische Haupteffekte	diagnosebezogene Haupteffekte	Zahl der RCC	Gesamt	
Var. A, reduziert, mit Hier., LB07, 1-J-prosp., V404	2.823.493	32	33	28	93	27,68 %
Var. F, reduziert, ohne Hier., LB07, 1-J-prosp., V404	2.823.493	32	32	28	92	27,76 %
Var. F, reduziert, ohne Hier., LB08, 2-J-prosp., V404	2.668.118	32	30	27	89	36,04 %
Var. F, reduziert, mit Hier., LB07, 1-J-prosp., V404	2.823.493	32	32	28	92	27,42 %
Var. F, reduziert 130, mit Hier., LB08, 1-J-prosp., V404	2.804.342	32	30	27	89	29,54 %

Tabelle 23: Kennzahlen der berechneten reduzierten Varianten

Kurzbeschreibung	Zahl der verbleibenden Risikokategorien					Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz
	N (Kalibrierungsstichprobe aus der Lernstichprobe)	Demographische Haupteffekte	diagnosebezogene Haupteffekte	Zahl der RCC	Gesamt	
Var. F, komplett, ohne Hier., LB07, zeitgl., V309	3.000.000	22	156	(111)	289	7,36 %
Var. F, komplett, ohne Hier., LB08, 2-J-prosp., V404	2.668.118	32	136	(86)	254	36,38 %

Kurzbeschreibung	Zahl der verbleibenden Risikokategorien					Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz
	N (Kalibrierungsstichprobe aus der Lernstichprobe)	Demographische Haupteffekte	diagnosebezogene Haupteffekte	Zahl der RCC	Gesamt	
Var. F, komplett, mit Hier., LB08, 2-J-prosp., V404	2.668.118	32	137	(93)	262	35,07 %
Var. G, komplett, ohne Hier., LB07, zeitgl., V309	3.000.000	26	156	(116)	298	8,05 %
Var. G, komplett, ohne Hier., LB08, zeitgl., V404	3.000.000	26	156	(116)	298	9,90 %
Var. G, komplett, ohne Hier., LB08, 2-J-prosp., V404	2.668.118	32	139	(81)	253	36,40 %
Var. G, komplett, ohne Hier., LB08, 2-J-prosp., X404	2.668.118	32	160	(121)	313	37,19 %

Tabelle 24: Kennzahlen der berechneten kompletten Varianten

Wie in Abschnitt 3.1.2.2 beschrieben werden bei reduzierten und komprimierten Modellen Hierarchien zwischen RCCs eingefügt, wenn RCCs in der Regression negative Kostengewichte erhalten. Wenn das Einfügen von Hierarchien die Negativität nicht beheben kann, werden die RCCs aus dem Modell ausgeschlossen (auf Null gesetzt). Tabelle 25 gibt einen Überblick über negative RCC in den berechneten Modellen. Die Spalte „Kurzbeschreibung“ gibt eine kurze Zusammenfassung der Modellvarianten, die ausführliche Darstellung der beauftragten Varianten erfolgt in Tabelle 22. Die Spalte „RCC mit negativen Kostengewichten“ gibt die RCC an, die nach Schritt 4 des Kalibrierungsvorganges (Anhang 20) negative Kostengewichte haben. Die Spalte „erfolgreich verwendete Beziehungen zwischen korrelierenden RCCs“ gibt einen Überblick über Hierarchien, die dazu führten, dass das Kostengewicht der RCC positiv wurde und die Spalte „Ausschluss“ enthält eine Auflistung der RCC, deren Kostengewichte auch nach Hierarchisierung negativ blieben und die daher ausgeschlossen wurden.

Kurzbeschreibung	RCCs mit negativen Kostengewichten	Erfolgreich verwendete Beziehungen zwischen korrelierenden RCCs	Ausschluss
Var. A, reduziert, mit Hier., LB07, 1-J-prosp., V404	RCC006 RCC024	RCC016 über RCC006 RCC020 über RCC006 RCC022 über RCC006	RCC024
Var. F, reduziert, ohne Hier., LB07, 1-J-prosp., V404	RCC006 RCC024	RCC016 über RCC006 RCC018 über RCC006 RCC020 über RCC006	RCC024
Var. F, reduziert, ohne Hier., LB08, 2-J-prosp., V404	RCC006 RCC024 RCC026	RCC016 über RCC006 RCC018 über RCC006 RCC020 über RCC006	RCC024 RCC028
Var. F, reduziert, mit Hier., LB07, 1-J-prosp., V404	RCC006 RCC024	RCC016 über RCC006 RCC018 über RCC006 RCC020 über RCC006	RCC024
Var. F, reduziert 130, mit Hier., LB08, 1-J-prosp., V404	RCC006 RCC024 RCC026	RCC016 über RCC006 RCC018 über RCC006 RCC022 über RCC006	RCC024 RCC028

Tabelle 25: Überblick über negative RCC in reduzierten Modellen

3.2.2 Gütemaße der Modellierung

Die Tabelle Gütemaße (Tabelle 26) gibt verschiedene Werte zur Beurteilung der Modelle an. Die Werte „relativer Prognosefehler“ werden berechnet aus dem „mean absolute prediction error“ (MAPE). Die Formel wird in Abbildung 13 dargestellt¹⁹. MAPE wird berechnet als Betrag der Differenz des tatsächlichen Leistungsbedarfs je Versicherten (y_i) zum geschätzten Leistungsbedarf je Versicherten (\hat{y}_i). Dieses wird über alle Versicherten $i=1, \dots, n$ der Stichprobe summiert und durch die Zahl der Versicherten n in der Stichprobe dividiert.

Der relative Prognosefehler ergibt sich anschließend durch die Relativierung am mittleren tatsächlichen Leistungsbedarf.

$$MAPE = \frac{\sum_i |y_i - \hat{y}_i|}{n}$$

Abbildung 13: Formel zur Berechnung des Prognosefehlers

¹⁹ Auf die Darstellung der in den Berechnungen vorgenommenen Gewichtungen mit den Versichertenzeiten wird hier der Übersichtlichkeit halber verzichtet.

Die Berechnung des relativen Prognosefehlers wird sowohl für die Stichprobe ausgeführt, auf der die Relativgewichte modelliert wurden (Kalibrierungsstichprobe), als auch auf einer weiteren Stichprobe (Validierungsstichprobe).

Bei prospektiven Modellen entspricht die Bildung der Validierungsstichprobe der Bildung der Kalibrierungsstichprobe (vgl. Abschnitt 3.1.5.1).

Der Wert „adjustiertes R^2 “ ist das Bestimmtheitsmaß der Modelle, das den Verlust von Freiheitsgraden bei der Aufnahme von Variablen ins Modell berücksichtigt. Die Formel zur Berechnung des adjustierten R^2 ist in Abbildung 14 angegeben, wobei p die Anzahl der Variablen im Modell darstellt.

$$R_{adj}^2 = 1 - \frac{\frac{1}{n-p-1} \sum_i (y_i - \hat{y}_i)^2}{\frac{1}{n-1} \sum_i (y_i - \bar{y})^2}$$

Abbildung 14: Formel zur Berechnung des adjustierten R^2

Kurzbeschreibung	Relativer Prognosefehler (MAPE / Mean_LB)		Predictive Ratio (geschätzt / tatsächlich)		Adjustiertes R^2
	in der Kalibrierungsstichprobe	in der Validierungsstichprobe	in der Kalibrierungsstichprobe	in der Validierungsstichprobe	
Var. A, reduziert, mit Hier., LB07, 1-J-prosp., V404	62,25 %	62,07 %	1,00000	0,99746	43,4 %
Var. F, reduziert, ohne Hier., LB07, 1-J-prosp., V404	60,59 %	60,27 %	1,00000	0,99641	52,1 %
Var. F, reduziert, ohne Hier., LB08, 2-J-prosp., V404	61,40 %	61,24 %	1,00000	0,99841	41,8 %
Var. F, reduziert, mit Hier., LB07, 1-J-prosp., V404	60,82 %	60,52 %	1,00000	0,99647	51,9 %
Var. F, reduziert 130, mit Hier., LB08, 1-J-prosp., V404	54,61 %	54,43 %	1,00000	0,99963	60,3 %

Kurzbeschreibung	Relativer Prognosefehler (MAPE / Mean_LB)		Predictive Ratio (geschätzt / tatsächlich)		Adjus- tiertes R ²
	in der Kalibrie- rungs- stich- probe	in der Validie- rungs- stichprobe	in der Kalibrier- ungs- stich- probe	in der Validie- rungsstichpro- be	
Var. F, komplett, ohne Hier., LB07, zeitgl., V309	48,85 %	48,95 %	1,00000	1,00000	57,0 %
Var. F, komplett, ohne Hier., LB08, 2-J-prosp., V404	60,95 %	60,80 %	1,00000	0,99901	42,7 %
Var. F, komplett, mit Hier., LB08, 2-J-prosp., V404	61,05 %	61,59 %	1,00000	0,97437	42,5 %
Var. G, komplett, ohne Hier., LB07, zeitgl., V309	48,72 %	48,83 %	1,00000	0,99996	57,0 %
Var. G, komplett, ohne Hier., LB08, zeitgl., V404	46,51 %	46,40 %	1,00000	0,99792	61,8 %
Var. G, komplett, ohne Hier., LB08, 2-J-prosp., V404	60,93 %	60,78 %	1,00000	0,99898	42,7 %
Var. G, komplett, ohne Hier., LB08, 2-J-prosp., X404	64,54 %	57,37 %	1,00000	1,00065	28,3 %

Tabelle 26: Gütemaße der beauftragten Varianten

3.3 Berechnung der Veränderungsrate

In diesem Abschnitt wird das Vorgehen bei der Berechnung der Veränderungsrate beschrieben. Im ersten Teil wird das allgemeine Vorgehen erläutert und im zweiten Abschnitt auf mögliche Informationsverluste bei der Berechnung von prospektiven Modellen eingegangen.

3.3.1 Allgemeines Vorgehen

Die Berechnung der Veränderungsrate der Morbidität findet bei jeder Variante nach dem gleichen Schema statt, welches im Folgenden vorgestellt wird. Weicht die Berechnung bei einer Variante von diesem Schema ab, so wird dieses in der Dokumentation der Variante dargelegt.

Das Schaubild in Anhang 23 stellt beispielhaft für ein zeitgleiches Modell für den Vergleich von 2006 und 2007 die einzelnen Berechnungsschritte grafisch dar.

In Anhang 24 werden jeweils die Formeln der einzelnen Schritte aufgezeigt.

Klassifizierung der Versicherten

Je Variante und Periode werden die Risikokategorien der Versicherten für die Bezugs- und die Folgeperiode ermittelt.

Die Klassifikation der Versicherten erfolgt grundsätzlich auf Jahresdatensätzen mit vier Quartalen.

Berechnung der Kostengewichte der Risikokategorien

Für alle Risikokategorien werden aus der Leistungsmenge der Kalibrierungsstichprobe des ausgewählten Abrechnungsjahres *und den Diagnosehäufigkeiten des Basisjahres* Kostengewichte berechnet. Die Kostengewichte werden taggenau ermittelt und für die weiteren Berechnungen fixiert.

Da das Schaubild in Anhang 23 ein zeitgleiches Modell darstellt, werden der Leistungsbedarf und die Diagnosen des Jahres 2007 zur Berechnung der Kostengewichte verwendet.

Berechnung des 100%-Wertes

Die mittlere Leistungsmenge (versichertenzeitgewichteter Durchschnitt über die Versicherten im Abrechnungsjahr) stellt den 100%-Wert dar. Dieser wird auch fixiert.

Berechnung der Relativgewichte

Die Kostengewichte der einzelnen Risikokategorien (Schritt 2) im Verhältnis zum 100%-Wert ergeben die einzelnen Relativgewichte. Sie bleiben für die weiteren Schritte fixiert.

Risikowert eines Versicherten

Die Risikokategorien eines Versicherten, die für die Bezugsperiode festgestellt wurden, werden mit den Relativgewichten bewertet. Die Summe der Relativgewichte eines Versicherten in der Bezugsperiode ist sein Risikowert.

Morbiditätsindex des Kollektivs in der Bezugsperiode

Der Morbiditätsindex ist der mit den Versichertentagen gewichtete Mittelwert der Risikowerte der Bezugsperiode. Dieser ist in der Kalibrierungsstichprobe in der Regel²⁰ gleich Eins.

Berechnung des Morbiditätsindex des Kollektivs in der Folgeperiode

Die Schritte „Risikowerte eines Versicherten“ und „Morbiditätsindex des Kollektivs in der Bezugsperiode“ werden mit den Risikokategorien, die für die Folgeperiode festgestellt wurden, wiederholt.

Veränderungsrate zwischen Bezugs- und Folgeperiode

Die Veränderungsrate ist der Morbiditätsindex der Folgeperiode im Verhältnis zum Morbiditätsindex der Bezugsperiode abzüglich Eins.

3.3.2 Informationsverluste bei prospektiven Modellen

Bei der Diskussion zu prospektiven Modellen, insbesondere bei Modellen, die eine 2-Jahres-Prospektion umfassen, gab es immer wieder Bedenken, dass es zu Informationsverlusten kommen kann. In der Diskussion wurden die Berechnung der Kosten- bzw. Relativgewichte und die Berechnung der Risikowerte der Versicherten als die zwei Berechnungsschritte, bei denen es zu möglichen Informationsverlusten kommen kann, benannt. Es wird daher im Folgenden auf diese beiden Schritte näher eingegangen.

Berechnung der Relativgewichte

Bei der Berechnung der Relativgewichte erfolgt die Zuweisung der Diagnosen nach Formel 1 aus Anhang 24. Die Variable A_{DIAG} entspricht jeweils dem Jahr 1. Der Jahresbezug des Leistungsbedarfs (A_{LB}) bei der Berechnung der Kostengewichte (Formel 2 aus dem Anhang 24) ist in einer einjährigen Prospektion das Jahr 2 und in einer zweijährigen Prospektion des Jahres 3.

Da bei Wechsel der Versichertennummer (vgl. Kapitel 2.1.4) zwischen Jahr 1 und Jahr 2 bzw. Jahr 3 die Diagnosen und der Leistungsbedarf eines Versicherten für die Berechnung der Kostengewichte nicht zusammengeführt werden können, können diese Versicherten nicht bei der Berechnung der prospektiven Kosten-

²⁰ Bei sog. erwartungstreuen Modellen ist dieser Wert in der Kalibrierungsstichprobe stets Eins.

gewichte berücksichtigt werden. Dieser Verlust von Informationen kann theoretisch zu Verzerrungen der Relativgewichte führen²¹.



Abbildung 15: Versichertenkollektiv im Jahr 1 und Jahr 3

Abbildung 15 und Abbildung 16 zeigen die möglichen Informationsverluste bei einer Zwei-Jahres Prospektion noch einmal schematisch auf. Die erste Abbildung stellt die Versichertenkollektive in den beiden Jahren dar. In der zweiten Abbildung wird gezeigt, dass nur die Versicherten, die in beiden Jahren identifizierbar sind, in die Berechnung der Relativgewichte eingehen können.

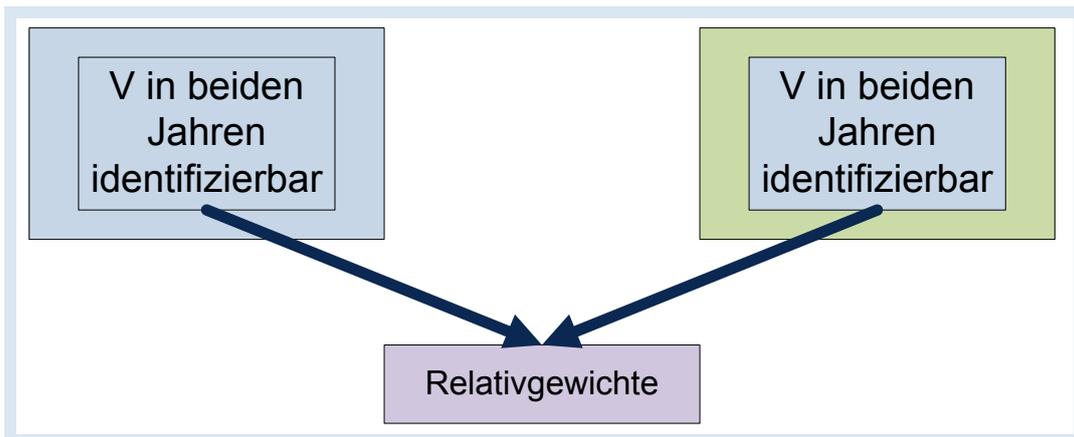


Abbildung 16: Versicherte zur Berechnung von Relativgewichten

Berechnung der Risikowerte der Versicherten

Bei der Berechnung der Risikowerte der Versicherten (Formel 5 in Anhang 24) entspricht die Jahresangabe A_M dem Jahr, dessen Risikokategorien zur Berech-

²¹ Dieser Informationsverlust wird auf den späteren Bearbeitungen mit der mittlerweile beschlossenen Geburtstagsstichprobe nicht mehr auftreten, da dort Pseudonyme der neuen, lebenslangen Versichertennummern vorgesehen sind und somit Versicherte auch bei Kassen- oder Regi-
onswechsel nachverfolgt werden können.

nung des Morbiditätsindex verwendet werden. Im Beispiel der Schaubilder wäre dieses Jahr 3. Es werden also auf allen Versicherten des Jahres 3 die Relativgewichte aufgesetzt und damit der geschätzte Leistungsbedarf (oder synonym der Risikowert der Versicherten) für das Jahr 5 berechnet. Wie in Abbildung 17 dargestellt, gehen bei diesem Rechenschritt keine Informationen verloren und es kann daher auch nicht zu Verzerrungen kommen.

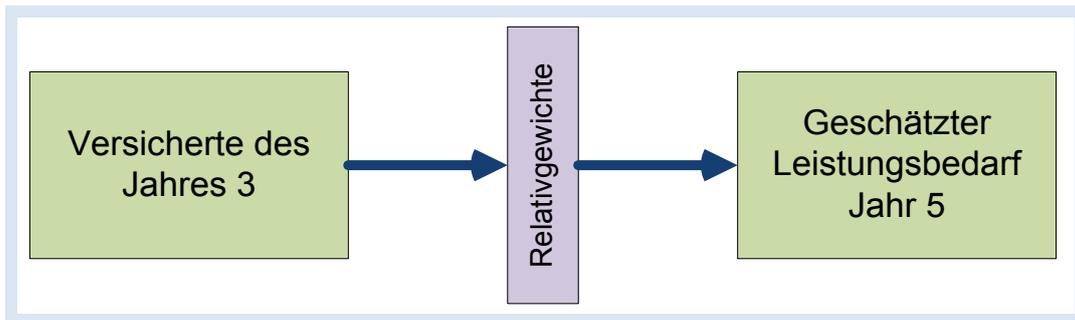


Abbildung 17: Berechnung der Risikowerte der Versicherten

3.3.3 Änderungen in der Klassifikation der CC 130

Aufgrund des Beschlusses der Kassenseite (V 2.0) wird für die Berechnung der Variante F reduziert130 (Kurzbeschreibung: Var. F, reduziert 130, mit Hier., LB08, 1-J-prosp., V404) eine andere Vorgehensweise bei der Klassifikation der Versicherten gewählt. Die Klassifikation entspricht im ersten Schritt der regulären Klassifikation der Variante F. In einem zweiten Schritt wird bei allen Versicherten, die aufgrund der Diagnosekodierung die Risikokategorie 130 aufweisen, diese auf Null zurückgesetzt. Es wird dann in einem separaten Schritt ermittelt, bei welchen Versicherten eine Gebührenordnungsposition des Bereichs Dialysesachkosten abgerechnet wurde, und bei diesen Versicherten für das jeweilige Jahr Risikokategorie 130 auf „1“ gesetzt. Die Risikokategorie 130 wird in dieser Variante also nicht durch Diagnosen, sondern durch die Abrechnung bestimmter Gebührenordnungspositionen getriggert. Eine Übersicht über die entsprechenden GOP steht in Anhang 25.

3.3.4 Ergebnisse der berechneten Varianten

Tabelle 27 gibt eine Übersicht über die Veränderungsdaten der berechneten Varianten.

Aus internen Bearbeitungsgründen wurden zum Teil morbiditätsbedingte Veränderungsdaten für weitere Zeitbezüge ebenfalls ausgerechnet.

Kurzbeschreibung	Zeitbezug Rate	Morbiditäts-Index			Veränderungsrate	
		2006	2007	2008	2006/2007	2007/2008
Var. A, reduziert, mit Hier., LB07, 1-J-prosp., V404	2006/07	1,02	1,04		2,08%	
Var. F, reduziert, ohne Hier., LB07, 1-J-prosp., V404	2006/07	1,02	1,04	1,08	2,30%	3,83%
Var. F, reduziert, ohne Hier., LB08, 2-J-prosp., V404	2007/08	1,02	1,04	1,08	2,11%	3,44%
Var. F, reduziert, mit Hier., LB07, 1-J-prosp., V404	2006/07	1,02	1,04	1,08	2,26%	3,81%
Var. F, reduziert 130;mit Hier., LB08, 1-J-prosp., V404	2007/08	0,99	1,01	1,03	1,82%	2,26%
Var. F, komplett, ohne Hier., LB07, zeitgl., V309	2007/08	0,97	1,00	1,05	2,91%	4,63%
Var. F, komplett, ohne Hier., LB08, 2-J-prosp., V404	2007/08	1,00	1,02	1,06	2,36%	3,67%
Var. F, komplett, mit Hier., LB08, 2-J-prosp., V404	2007/08	1,02	1,04	1,08	2,26%	3,59%
Var. G, komplett, ohne Hier., LB07, zeitgl., V309	2006/07 2007/08	0,97	1,00	1,04	2,88%	4,58%
Var. G, komplett, ohne Hier., LB08, zeitgl., V404	2007/08		0,96	1,00		4,44%
Var. G, komplett, ohne Hier., LB08, 2-J-prosp., V404	2007/08	1,02	1,05	1,08	2,37%	3,67%
Var. G, komplett, ohne Hier., LB08, 2-J-prosp., X404	2007/08	1,01	1,02	1,05	1,82%	2,68%

Tabelle 27: Übersicht über die Ergebnisse der berechneten Varianten

3.4 Sonderauswertungen

Die Beauftragung des Instituts des Bewertungsausschusses aus der 184. Sitzung des Bewertungsausschusses beinhaltet auch Auswertungen, die spezielle Aspekte des Themas „Repräsentativität der Datengrundlage und Berechnung von Veränderungsraten“ umfassen. Diese Auswertungen sollen in diesem Abschnitt vorgestellt werden.

Die ersten beiden Auswertungen betrachten dabei die Stabilität von Kostengewichten einerseits interregional und andererseits intertemporal. Im Zusammenhang mit der interregionalen Stabilität werden auch regional berechnete Veränderungsdaten berichtet. Diese Aspekte werden in den beiden Kapiteln 3.4.1 und 3.4.2 betrachtet.

3.4.1 Interregionale Stabilität von Relativgewichten

Die Auswertung der interregionalen Stabilität der Relativgewichte bezieht sich auf zwei Teilaspekte, einerseits werden die Relativgewichte selbst und andererseits die regional berechneten Veränderungsdaten betrachtet.

3.4.1.1 Vergleich regionaler Relativgewichte

Die Berechnung der Relativgewichte erfolgt mit Variante G, zeitgleich, nicht hierarchisiert, einmal als komplettes Modell und einmal als reduziertes Modell. Sie wird je Modell anhand der folgenden Schritte durchgeführt:

1. Berechnung von Relativgewichten auf der Datengrundlage des Bewertungsausschusses (4 KVen) erfolgt anhand der Formeln 1 bis 4 im Anhang 24.
2. Durchlaufen des Modellierungsalgorithmus und dabei Auswahl von Risikokategorien. Fixierung der Risikokategorien.
3. Auf Basis der fixierten Risikokategorien werden KV-spezifische Relativgewichte berechnet. Dazu wird das korrespondierende Kostengewicht durch einen KV-spezifischen 100%-Wert geteilt (Formel 3 und Formel 4 im Anhang 24)
4. Überprüfung der KV-spezifischen Relativgewichte auf Negativität und statistische Signifikanz.
5. Berechnung der prozentualen Abweichung der KV-spezifischen Relativgewichte von den Relativgewichten berechnet auf der 4 KVen-Datengrundlage.

Ergebnisse des kompletten Modells

Tabelle 28 gibt einen Überblick über die Kennzahlen der Berechnungen auf den einzelnen KVen und Tabelle 29 enthält einen Überblick über die Signifikanz und die Abweichungen der regionalen Relativgewichte von den Relativgewichten berechnet auf den 4 KVen gemeinsam. Es sei zur Bewertung der Ergebnisse in Bremen und Thüringen darauf hingewiesen, dass aufgrund der geringeren Einwohnerzahlen der beiden KVen auch die Anzahl der Versicherten für die Kalibrie-

ung entsprechend gering ist. Die Ergebnisse der Kalibrierung sind dementsprechend instabil.

	4 KVen	Bremen	Niedersachsen	Nordrhein	Thüringen
Anzahl Versicherte	3.000.000	96.882	1.159.702	1.389.559	352.825
mittlerer LB MGW (zehntel Punkte)	102.935	115.964	108.671	98.276	98.796
Anzahl Variablen	298	298	298	298	298
Anzahl negativer geschätzter Leistungsbedarfe	0	1.417	2	3	6.525

Tabelle 28: Kennzahlen zur Berechnung der regionalen Relativgewichte (komplettes Modell)

Kennzahl	4 KVen	Bremen	Nieder- sachsen	Nord- rhein	Thürin- gen
Anzahl der Relativgewichte im Modell	298	298	298	298	298
Davon					
statistisch insignifikant ($ t\text{-Wert} < 1,96$)	0	142	50	47	105
statistisch signifikant	298	156	248	251	193
davon					
negative Haupteffekte (AGE, CC)	0	6	1	0	0
negative Interaktionen (KCC, OCC)	77	24	54	56	38
positiv	221	126	193	195	155
davon mit einer Abweichung im Relativgewicht gegenüber der Lernmenge 4 KVen von:					
100 % und mehr	0	30	5	4	18
50 % bis unter 100 %	0	12	23	9	22
30 % bis unter 50 %	0	23	23	29	32
20 % bis unter 30 %	0	17	24	36	25
10 % bis unter 20 %	0	22	40	52	23
0 bis unter 10 %	0	22	78	65	35

Tabelle 29: Abweichungen der regionalen Relativgewichte von den Relativgewichten der vier KVen (komplettes Modell)

In Tabelle 68 im Anhang 21 wird eine leicht modifizierte Auswertung zu den Abweichungen der regionalen Relativgewichte dargestellt. Bei dieser Auswertung wurde beim 3. Schritt der oben genannten Abfolge nicht der KV-spezifische durchschnittliche Leistungsbedarf verwendet, sondern der durchschnittliche Leistungsbedarf über die 4 KVen hinweg.

Ergebnisse des reduzierten Modells

Tabelle 30 gibt einen Überblick über die Kennzahlen der Berechnungen auf den einzelnen KVen und Tabelle 31 enthält einen Überblick über die Signifikanz und die Abweichungen der regionalen Relativgewichte von den Relativgewichten berechnet auf den 4 KVen gemeinsam. Es sei zur Bewertung der Ergebnisse in Bremen und Thüringen darauf hingewiesen, dass aufgrund der geringeren Ein-

wohnerzahlen der beiden KVen auch die Anzahl der Versicherten für die Kalibrierung entsprechend gering ist.

	4 KVen	Bremen	Niedersachsen	Nordrhein	Thüringen
Anzahl Versicherte	3.000.000	96.882	1.159.702	1.389.559	352.825
mittlerer LB MGV (zehntel Punkte)	102.935	115.964	108.671	98.276	98.796
Anzahl Variablen	81	81	81	81	81
Anzahl negativer geschätzter Leistungsbedarfe	0	334	0	3	6.851

Tabelle 30: Kennzahlen zur Berechnung der regionalen Relativgewichte (reduziertes Modell)

Kennzahl	4 KVen	Bremen	Niedersachsen	Nordrhein	Thüringen
Anzahl der Relativgewichte im Modell	81	81	81	81	81
davon					
statistisch insignifikant ($ t\text{-Wert} < 1,96$)	0	16	1	2	7
statistisch signifikant	81	65	80	79	74
davon					
negative Haupteffekte (AGE, CC, RCC)	0	0	0	0	0
positiv	81	65	80	79	74
davon mit einer Abweichung im Relativgewicht gegenüber der Lernmenge 4 KVen von:					
100 % und mehr	0	1	1	0	0
50 % bis unter 100 %	0	4	0	0	3
30 % bis unter 50 %	0	12	8	6	14
20 % bis unter 30 %	0	10	7	12	16
10 % bis unter 20 %	0	19	16	25	18
0 bis unter 10 %	0	19	48	36	23

Tabelle 31: Abweichungen der regionalen Relativgewichte von den Relativgewichten der vier KVen (reduziertes Modell)

In Tabelle 69 in Anhang 21 wird eine leicht modifizierte Auswertung zu den Abweichungen der regionalen Relativgewichte dargestellt. Bei dieser Auswertung wurde beim 3. Schritt, der oben genannten Abfolge, nicht der KV-spezifische durchschnittliche Leistungsbedarf, sondern der durchschnittliche Leistungsbedarf über die 4 KVen hinweg verwendet.

3.4.1.2 Regionale Veränderungsraten

Zur Berechnung der regionalen Veränderungsraten werden auf der 4-KVen-Datengrundlage Relativgewichte der Variante G, komplett, nicht hierarchisiert und zeitgleich berechnet. Auf Grundlage dieser wird dann jeweils auf einem nach KVen abgegrenzten Teil der Grouper-Schnittstelle eine Veränderungsrate je KV berechnet. Tabelle 32 gibt einen Überblick über die regionalen Veränderungsraten und die regionalen Morbiditätsindizes.

Region	LB für Relativgew. aus	Index			Rate	
		2006	2007	2008	2006/2007	2007/2008
4 KVen	2007	0,9709	0,9989 ²²	1,0446	2,88%	4,58%
Bremen		0,9586	0,9810	1,0126	2,33%	3,23%
Niedersachsen		0,9703	0,9994	1,0427	3,00%	4,33%
Nordrhein		0,9726	0,9999	1,0448	2,80%	4,49%
Thüringen		0,9718	0,9977	1,0589	2,66%	6,14%
4 KVen	2008		0,9575	1,0000		4,44%

Tabelle 32: Regionale Veränderungsraten

3.4.2 Intertemporale Stabilität der Relativgewichte

Um die intertemporale Stabilität der Relativgewichte zu vergleichen, werden auf den Jahren 2007 und 2008 Relativgewichte der Variante G, komplett, nicht hierarchisiert und zeitgleich berechnet.

Die Berechnung erfolgt anhand der folgenden Schritte:

²² Der Wert ist nicht 1, da die Berechnungen der Relativgewichte und die Berechnungen des Index auf unterschiedlichen Stichproben erfolgten.

1. Die Berechnung von Relativgewichten für das Jahr 2007 auf der Datengrundlage des Bewertungsausschusses erfolgt anhand der Formeln 1 bis 4 im Anhang 24.
2. Durchlaufen der Kalibrierungsstrategie und dabei Auswahl von Risikokategorien. Fixierung der Risikokategorien.
3. Auf der Basis der fixierten Risikokategorien werden für das Jahr 2008 Kostengewichte berechnet. Die Berechnung der Relativgewichte 2008 erfolgt in zwei unterschiedlichen Versionen, einmal mit dem durchschnittlichen Leistungsbedarf 2007 und einmal mit dem durchschnittlichen Leistungsbedarf 2008. (Formel 3 und Formel 4 im Anhang 24)
4. Überprüfung der KV-spezifischen Relativgewichte auf Negativität und statistische Signifikanz.
5. Berechnung der prozentualen Abweichung der Relativgewichte 2008 von den Relativgewichten 2007.

Tabelle 33 gibt einen Überblick über die Kennzahlen der Berechnung zur Einschätzung der intertemporalen Stabilität der Relativgewichte.

	4 KVen, 2007	4 KVen, 2008
Anzahl Versicherte	3.000.000	3.000.000
mittlerer LB MGV	102.935	103.370
Anzahl Variablen	298	298
adj. R2	57,0%	61,8%
Anzahl negativer geschätzter Leistungsbedarfe	0	32

Tabelle 33: Kennzahlen zur Berechnung der intertemporalen Stabilität der Relativgewichte

Tabelle 34 zeigt die Ergebnisse bei einer Berechnung der Relativgewichte mit dem durchschnittlichen Leistungsbedarf 2008.

Kennzahl	4 KVen, 2007	4 KVen, 2008
Anzahl der Relativgewichte im Modell	298	298
davon		
statistisch insignifikant ($ t\text{-Wert} < 1,96$)	0	27
statistisch signifikant	298	271
davon		
negative Haupteffekte (AGE, CC)	0	3
negative Interaktionen (KCC, OCC)	77	67
positiv	221	201
davon mit einer Abweichung im Relativgewicht gegenüber der Lernmenge 4 KVen von:		
100 % und mehr	0	10
50 % bis unter 100 %	0	20
30 % bis unter 50 %	0	30
20 % bis unter 30 %	0	32
10 % bis unter 20 %	0	48
0 bis unter 10 %	0	61

Tabelle 34: Intertemporale Stabilität von Relativgewichten (Berechnung der Relativgewichte 2008 mit dem durchschnittlichen Leistungsbedarf 2008)

Tabelle 35 zeigt die Ergebnisse bei einer Berechnung der Relativgewichte mit dem durchschnittlichen Leistungsbedarf 2007.

Kennzahl	4 KVen, 2007	4 KVen, 2008
Anzahl der Relativgewichte im Modell	298	298
davon		
statistisch insignifikant ($ t\text{-Wert} < 1,96$)	0	27
statistisch signifikant	298	271
davon		
negative Haupteffekte (AGE, CC)	0	3
negative Interaktionen (KCC, OCC)	77	67
positiv	221	201
davon mit einer Abweichung im Relativgewicht gegenüber der Lernmenge 4 KVen von:		
100 % und mehr	0	10
50 % bis unter 100 %	0	20
30 % bis unter 50 %	0	29
20 % bis unter 30 %	0	31
10 % bis unter 20 %	0	53
0 bis unter 10 %	0	58

Tabelle 35: Intertemporale Stabilität von Relativgewichten (Berechnung der Relativgewichte 2008 mit dem durchschnittlichen Leistungsbedarf 2007)

4 Das beschlossene Modell

Die berichteten Ergebnisse der Kapitel 2 und 3 dieses Berichts beruhen auf einem Auftrag des Bewertungsausschusses an das Institut des Bewertungsausschusses aus der 184. Sitzung des Bewertungsausschusses.

Ausgehend von diesem Auftrag hat der 15. Erweiterte Bewertungsausschuss²³ in seinem Beschluss zur Berechnung der morbiditätsbedingten Gesamtvergütung 2010 eine morbiditätsbedingte Veränderungsrate beschlossen, welche sich aus einer diagnosenbezogenen Veränderungsrate und einer Veränderungsrate nach Alter und Geschlecht zusammensetzt. Der diagnosenbezogene Teil der morbiditätsbedingten Veränderungsrate 2009 auf 2010 wurde mit einem Wert von 2,5034% beschlossen. Sie geht mit 61% in die morbiditätsbedingte Veränderungsrate ein. Die Veränderungsrate nach Alter und Geschlecht wurde mit einer Höhe von 0,3423% beschlossen. Sie geht mit 39% in die Gesamtrate ein. (Beschluss des 15. EBA, Teil B, Nr 2.3.1 bis 2.3.3).

Das Ziel der folgenden Abschnitte ist es, die Modelle zu den beschlossenen Veränderungsraten darzustellen. Der Abschnitt 4.3 zum Modell zur Berechnung des diagnosebezogenen Anteils der Veränderungsrate enthält darüber hinaus wichtige Kennzahlen und Gütemaße zur Beurteilung des Modells.

Es wird dabei versucht, das Kapitel so zu gestalten, dass es auch unabhängig von den vorherigen Kapiteln gelesen werden kann. Im Unterschied zu vorherigen Kapiteln werden Ergebnisse, die einen etwas höheren Platzbedarf haben, daher auch nicht im Anhang, sondern direkt im Fließtext dargestellt.

4.1 Abgrenzung des Leistungsbedarfs zur Berechnung der Relativgewichte im beschlossenen Modell

Der Leistungsbedarf des Jahres 2008, welcher zur Berechnung der Relativgewichte herangezogen wurde, wurde mit den folgenden Schritten an die Beschlusslage des Bewertungsausschusses für das Jahr 2009 angepasst:

- Die Trennung in Leistungen innerhalb der morbiditätsbedingten Gesamtvergütung (MGV) und außerhalb der morbiditätsbedingten Gesamtvergütung (EGV) wird nachvollzogen. In Entsprechung zur Beschlusslage des Jahres 2010

²³ Der Beschluss ist unter <http://www.institut-des-bewertungsausschusses.de/ba/beschluesse.php> veröffentlicht.

werden die Leistungen der Dialysesachkosten außerhalb der morbiditätsbedingten Gesamtvergütung zugeordnet.

- Euro-Leistungen werden mittels Orientierungswert (2009, 3,5001 Cent) in Punkte umgerechnet.
- Die im 8. EBA beschlossenen Anpassungsfaktoren für die Bewertung einzelner Leistungen im EBM für das Jahr 2009 werden bei Punktzahlleistungen der MGV angewendet.
- Die im 8. EBA beschlossene Veränderungsrate der morbiditätsbedingten Leistungsmenge je Versicherten des Jahres 2009 gegenüber dem Jahr 2008 gemäß § 87c Abs. 4 SGB V wird auf Punktzahl- und Euroleistungen innerhalb der MGV angewendet (5,1 %).
- Die vom 8. EBA beschlossenen HVV-Quoten werden bei der herangezogenen Datengrundlage des Bewertungsausschusses aus vier KV-Bereichen nicht angewendet, um eine Verzerrung der angeforderten, sachlich-rechnerisch richtiggestellten Leistungsmengen durch unterschiedliche Honorarbegrenzungsregelungen in den KV-Bereichen zu vermeiden.

4.2 Darstellung der Veränderungsrate nach Alter und Geschlecht

Die Veränderungsrate nach Alter und Geschlecht wurde in der Differenzierung der KM-6 berechnet. Die Kostengewichte basieren auf einem 2-Jahresprospektiven Modell mit dem Diagnosejahr 2006 und dem Leistungsjahr 2008. Im Unterschied zur Bildung der Kalibrierungstichproben eines prospektiven Modells mit diagnosebezogenen Risikokategorien wie in Abschnitt 3.1.5.1 beschrieben, wird hier die Kalibrierungstichprobe aus der Lernstichprobe 2008 gebildet, indem das Alter der Versicherten in 2008 minus zwei gerechnet wird. Versicherte mit negativem Alter werden anschließend ausgeschlossen.

Eine Darstellung der Relativgewichte erfolgt in Tabelle 36. Die Zahl der Versicherten je Alters- und Geschlechtsgruppe wurde aus den KM-6-Daten des Jahres 2007 und 2008 berechnet.

Risikoklasse	Relativgewicht
Frauen bis unter 15 Jahre	0,623
Frauen 15 bis unter 20 Jahre	0,814
Frauen 20 bis unter 25 Jahre	0,806

Risikoklasse	Relativgewicht
Frauen 25 bis unter 30 Jahre	0,907
Frauen 30 bis unter 35 Jahre	0,974
Frauen 35 bis unter 40 Jahre	1,002
Frauen 40 bis unter 45 Jahre	1,064
Frauen 45 bis unter 50 Jahre	1,147
Frauen 50 bis unter 55 Jahre	1,245
Frauen 55 bis unter 60 Jahre	1,338
Frauen 60 bis unter 65 Jahre	1,438
Frauen 65 bis unter 70 Jahre	1,551
Frauen 70 bis unter 75 Jahre	1,642
Frauen 75 bis unter 80 Jahre	1,645
Frauen 80 bis unter 85 Jahre	1,638
Frauen 85 bis unter 90 Jahre	1,664
Frauen 90 und mehr Jahre	1,770
Männer bis unter 15 Jahre	0,649
Männer 15 bis unter 20 Jahre	0,436
Männer 20 bis unter 25 Jahre	0,414
Männer 25 bis unter 30 Jahre	0,462
Männer 30 bis unter 35 Jahre	0,535
Männer 35 bis unter 40 Jahre	0,599
Männer 40 bis unter 45 Jahre	0,659
Männer 45 bis unter 50 Jahre	0,751
Männer 50 bis unter 55 Jahre	0,886
Männer 55 bis unter 60 Jahre	1,062
Männer 60 bis unter 65 Jahre	1,231
Männer 65 bis unter 70 Jahre	1,401
Männer 70 bis unter 75 Jahre	1,587
Frauen 75 bis unter 80 Jahre	1,705
Männer 80 bis unter 85 Jahre	1,743
Männer 85 bis unter 90 Jahre	1,768
Männer 90 und mehr Jahre	1,898

Tabelle 36: Relativgewichte für die Veränderungsrate nach Alter und Geschlecht

4.3 Darstellung des beschlossenen Modells zur Berechnung des diagnosenbezogenen Anteils

Der diagnosenbezogene Teil der morbiditätsbedingten Veränderungsrate wurde mit der Variante H als komprimiertes, hierarchisiertes und 2-Jahres-prospektives Modell berechnet.

Variante H beschreibt die Klassifikationsvariante, die den Ausgangspunkt für die Erstellung des Modells darstellt. Variante H wurde durch die unabhängigen Mitglieder des Erweiterten Bewertungsausschusses, auf Basis von aktuellen Arbeitsständen des Bewertungsausschusses, eingebracht. Die Variante umfasst 196 Risikokategorien²⁴ mit Diagnosen mit dem Zusatzkennzeichen „G“ (gesicherte Diagnose) und 1 Risikokategorie mit Diagnosen mit dem Zusatzkennzeichen „Z“ (Zustand nach Diagnose).

Den 196 Risikokategorien werden noch 32 Alters- und Geschlechtskategorien hinzugefügt und auf diesen knapp 230 Risikokategorien wird eine Regressionsrechnung durchgeführt, ohne dass weitere Modellierungsschritte vorgenommen werden. Ergebnis dieser Regressionsrechnung ist für jede Risikokategorie ein Kostengewicht (*der entsprechende Regressionskoeffizient*). Diese Kostengewichte werden durch den durchschnittlichen Leistungsbedarf aller Versicherten in der Berechnungsstichprobe geteilt, um so die Relativgewichte²⁵ zu erhalten. Eine ausführliche Darstellung der Rechenschritte findet sich in Anhang 24. Anhand der berechneten Relativgewichte werden die 30 Risikokategorien mit der höchsten ökonomischen Relevanz (berechnet aus der Besetzungshäufigkeit der Risikokategorie multipliziert mit dem Relativgewicht) und Risikokategorien mit einem Relativgewicht von mehr als 1,5 ausgewählt. Diese Risikokategorien werden im Folgenden als „TCC“²⁶ bezeichnet. Alle anderen Risikokategorien werden gemäß der Zuordnung zu ACC²⁷ zusammengefasst; diese neu zusammengefassten Risikoklassen werden mit „RCC“ abgekürzt. Die TCCs, die RCCs und die Alters- und Ge-

²⁴ Im folgenden Text werden die erklärenden Variablen, die im beschlossenen Modell in der beschlossenen Ausgestaltung enthalten sind, als Risikoklassen bezeichnet. Alle anderen erklärenden Variablen werden als Risikokategorien bezeichnet.

²⁵ Die Kosten- bzw. Relativgewichte stellen den zusätzlichen Leistungsbedarf einer Person aufgrund dieser zusätzlichen Risikokategorie dar.

²⁶ Die Nummerierung der TCC entspricht jeweils der ihr zugrunde liegenden (H)CC, d. h. wenn z. B. die „CC 6“ zu den TOP 30 CC gehört, wird die daraus entstehende TCC als „TCC 6“ bezeichnet.

²⁷ Aggregated Condition Category (ACC) ist eine übergeordnete Kategorie aus der Original-Zuordnung von DxCG. Sie ordnet die Risikokategorien 30 Gruppen zu.

schlechtskategorien bilden das beschlossene Modell. Es enthält 91 Risikoklassen und wird als komprimiertes Modell bezeichnet.

Die Kalibrierungsstrategie zur Modellierung des komprimierten Modells findet sich in Anhang 20. Es soll an dieser Stelle auf Schritt 5 hingewiesen werden. Ist das Kostengewicht einer RCC negativ, so wird folgendes Vorgehen gewählt:

1. Es werden die Korrelationskoeffizienten der RCC mit allen anderen RCC berechnet.
2. Die drei RCC mit dem jeweils höchsten Korrelationskoeffizienten aus der Korrelation mit der RCC mit negativem Kostengewicht werden für die weiteren Schritte herangezogen.
3. Tritt die RCC mit negativem Kostengewicht gemeinsam mit einer der drei am höchsten korrelierenden RCC bei einem Versicherten auf, wird die RCC mit negativem Kostengewicht unterdrückt (auf Null gesetzt).
4. Anschließend werden die Kostengewichte mittels der linearen Regression erneut berechnet.
5. Bleibt das Kostengewicht der RCC negativ, so wird das Kostengewicht dieser RCC auf Null gesetzt und erneut eine Regression durchgeführt.

Der Anhang 26 gibt einen Überblick über die Korrelationskoeffizienten. Im beschlossenen Modell wurde die RCC 24 auf Null gesetzt, da das zuvor beschriebene Verfahren die Negativität der Kostengewichte nicht beheben konnte. Bei der RCC 28 konnte die Negativität der Kostengewichte durch die Unterdrückung der RCC 28 bei Auftreten der RCC 5, RCC 13 oder RCC 21 behoben werden.

Am Ende des Modellierungsvorganges steht die Hierarchisierung mit einer abschließend durchgeführten Regression. Das Institut und der GKV-Spitzenverband hatten hierfür eine Hierarchisierungstabelle abgestimmt, die für alle komprimierten und reduzierten Modelle, die im Rahmen des Auftrags aus der 184. Sitzung des Bewertungsausschusses berechnet wurden, angewendet wurde. Diese Hierarchisierungstabelle enthält die hierarchischen Beziehungen zwischen den Risikoklassen. Abschnitt 4.3.1 gibt eine Übersicht über die hierarchischen Beziehungen, die in dem beschlossenen Modell angewendet wurden.

Die Berechnung der Kosten- bzw. Relativgewichte basiert auf einem 2-Jahresprospektiven Modell. Dieses bedeutet, dass mithilfe der Diagnosen des Jahres 2006 und dem Leistungsbedarf des Jahres 2008 die Kosten- bzw. Relativgewichte

berechnet wurden. Die Berechnung der Veränderungsrate basiert auf den Morbiditätsindizes der Jahre 2007 und 2008 (vgl. Anhang 24).

4.3.1 Klassifikation und Hierarchisierung im beschlossenen Modell

In Abschnitt 4.2 wurde beschrieben, dass beim komprimierten Modell bestimmte Risikokategorien des ursprünglichen Modells aufgrund ihrer ökonomischen Relevanz und / oder ihres Relativgewichts ausgewählt werden und diese als Risikokategorien stehen bleiben. Alle anderen ehemaligen Risikokategorien werden in RCC zusammengefasst. Eine Darstellung der Zuordnung für jeden ICD-Kode auf TCC und RCC für die der Beschlussfassung zugrunde liegenden Jahre 2006, 2007 und 2008 erfolgt in einem separaten Dokument.

Darüber hinaus wurde in Abschnitt 4.2 beschrieben, dass die Hierarchisierung für komprimierte Modelle allgemein erarbeitet wurde. In den folgenden Schaubildern (Abbildung 18 bis Abbildung 23) wird ein Überblick über die angewendeten Hierarchien im beschlossenen Modell gegeben.

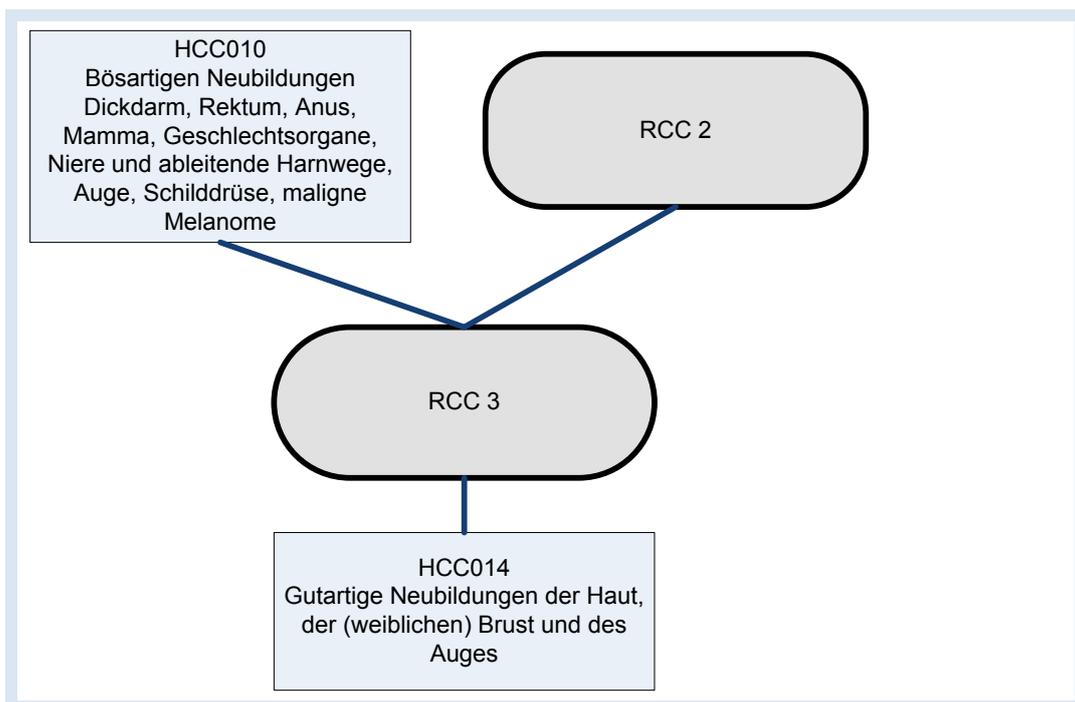


Abbildung 18: Hierarchisierung innerhalb der ACC 2 und ACC 3 im beschlossenen Modell

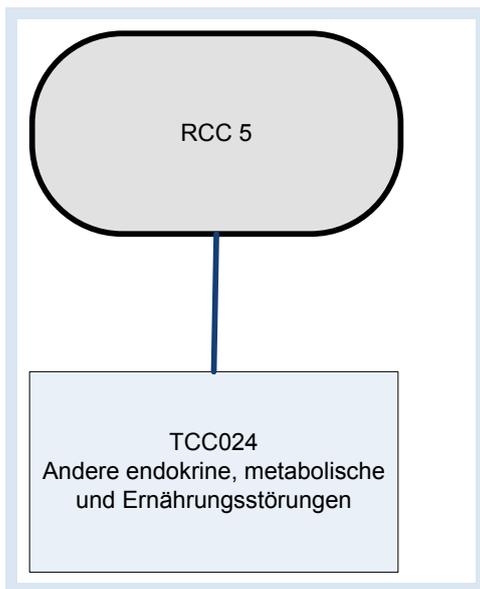


Abbildung 19: Hierarchisierung innerhalb der ACC 5 im beschlossenen Modell

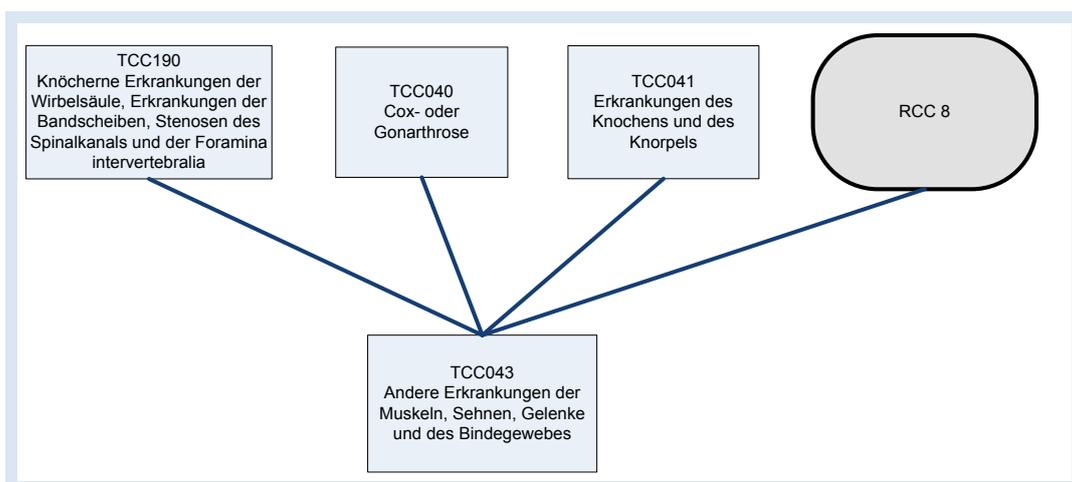


Abbildung 20: Hierarchisierung innerhalb der ACC 8 im beschlossenen Modell

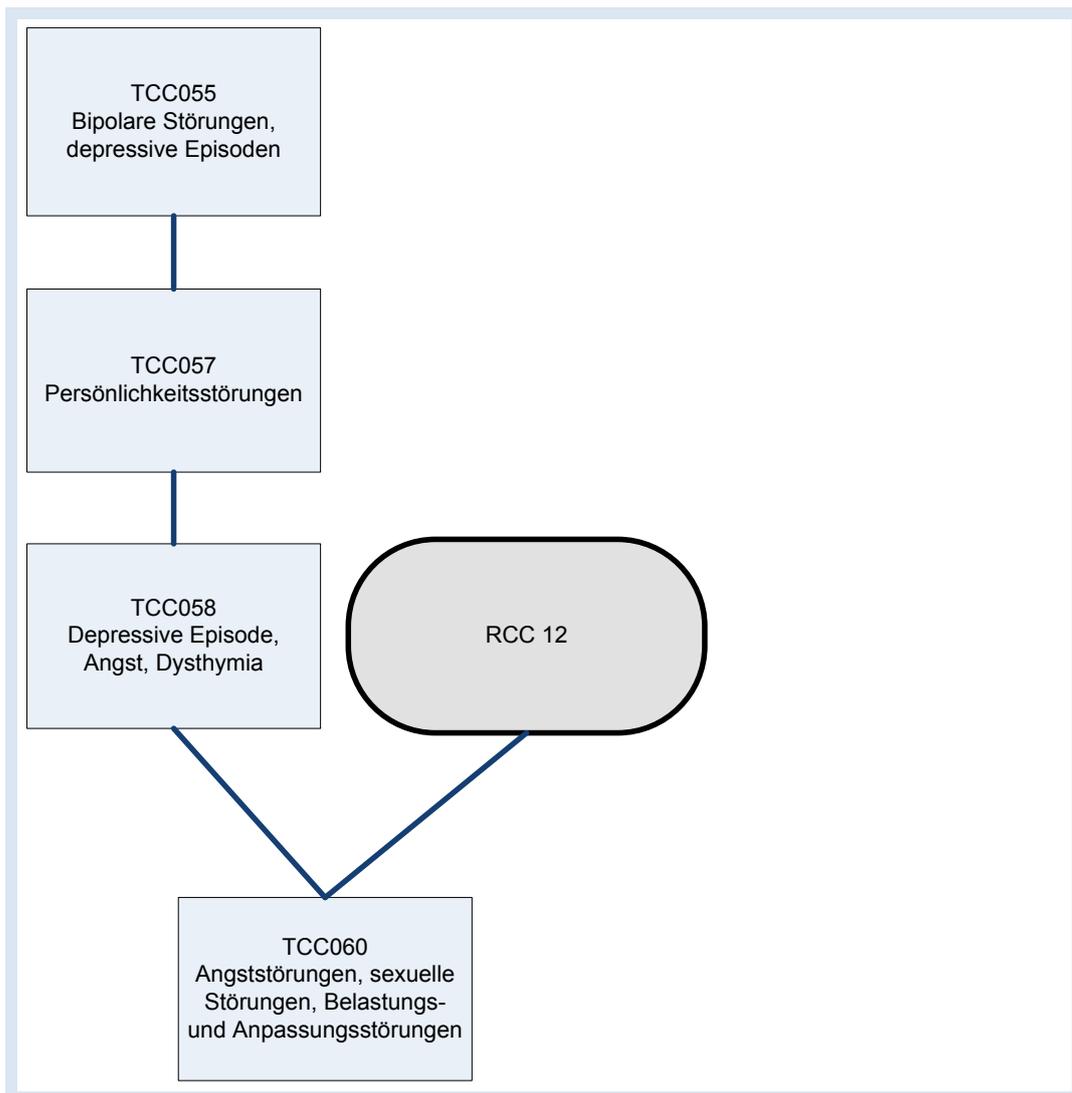


Abbildung 21: Hierarchisierung innerhalb der ACC 12 im beschlossenen Modell

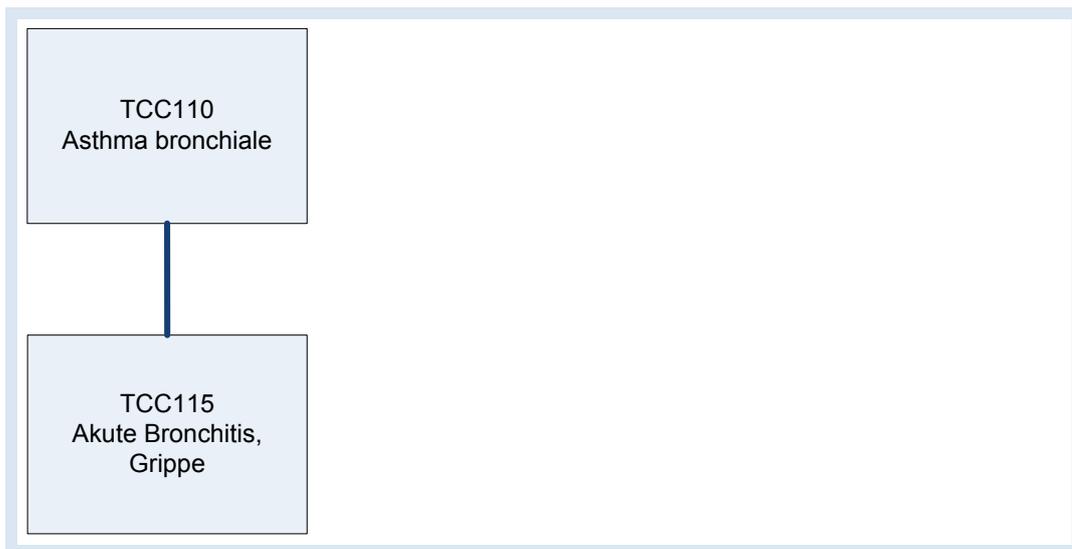


Abbildung 22: Hierarchisierung innerhalb der ACC 19 im beschlossenen Modell

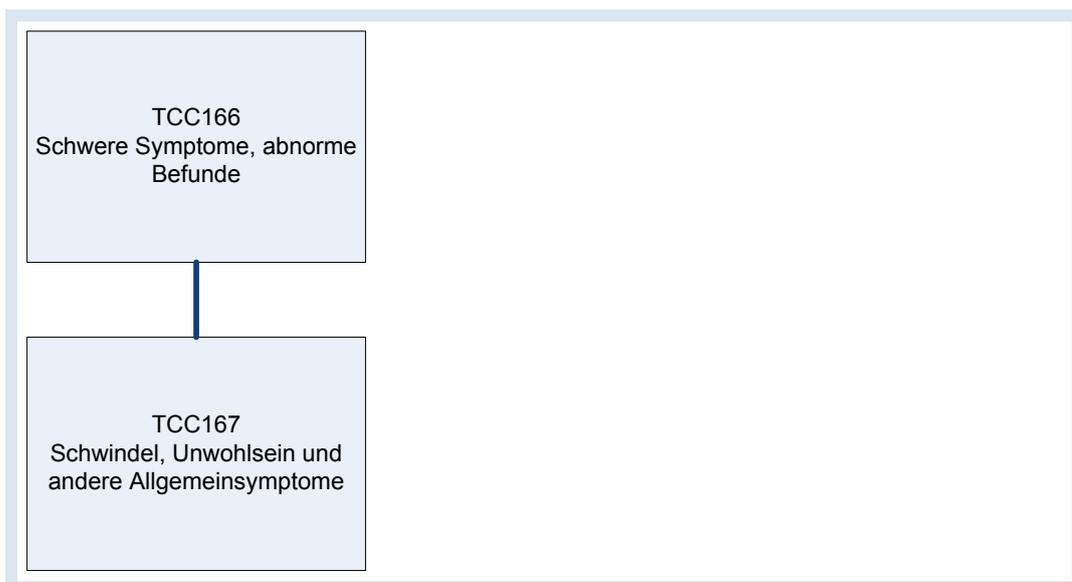


Abbildung 23: Hierarchisierung innerhalb der ACC 27 im beschlossenen Modell

4.3.2 Kennzahlen

Dieser Abschnitt berichtet wichtige Kennzahlen zu der beschlossenen Ausgestaltung. Die Kennzahlen werden in einer ersten Übersicht (Tabelle 37) nur für die beschlossene Ausgestaltung dargestellt und in einer zweiten Übersicht (Tabelle 38) für weitere Ausgestaltungen, die das Institut im Rahmen des Auftrags aus der

184. Sitzung des Bewertungsausschusses berechnet hatte. Die beschlossene Ausgestaltung steht in dieser Tabelle in der untersten Zeile. Die weiteren Ausgestaltungen werden in dieser Darstellung nur mit einer Kurzbeschreibung dargestellt, die ausführliche Beschreibung der Ausgestaltung findet sich in Abschnitt 3.2 (Tabelle 22).

Die Spalte „Kurzbeschreibung“ gibt eine kurze Zusammenfassung der Modellvarianten, die ausführliche Darstellung der beauftragten Ausgestaltungen erfolgt in Tabelle 22. Die Spalte „N (Kalibrierungstichprobe aus der Lernstichprobe)“ gibt die Zahl der Versicherten an, auf der die Relativgewichte modelliert wurden. Die Spalten „Zahl der Risikoklassen“ geben die Zahl der im beschlossenen Modell enthaltenen Risikoklassen des jeweiligen Typs an. Die Spalte „Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz“ gibt den Anteil des auf die Alters- und Geschlechtsklassen entfallenden erwarteten Leistungsbedarfs am gesamten Leistungsbedarf an.

Kurzbeschreibung	N (Kalibrierungstichprobe aus der Lernstichprobe)	Zahl der Risikoklassen				Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz
		Alters- und Geschlechtsklassen	TCC	RCC	Gesamt	
Var. H, komprimiert, mit Hier., LB08, 2-J-prosp., X404	2.668.118	32	32	28	92	35,00 %

Tabelle 37: Kennzahlen zum beschlossenen Modell

Zum Vergleich werden in der folgenden Tabelle (Tabelle 38) die Kennzahlen für die weiteren Modelle, die im Rahmen der Beauftragung aus der 184. Sitzung des Bewertungsausschusses berechnet wurden, dargestellt. Die Spalte „diagnosebezogene Haupteffekte“ berichtet bei reduzierten und komprimierten Modellen die Zahl der TCC und bei kompletten Modellen die Zahl der CC bzw. HCC, die im Modell verbleiben.

Kurzbeschreibung	Zahl der Risikokategorien					Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz
	N (Kalibrierungsstichprobe aus der Lernstichprobe)	Demographische Haupteffekte	Diagnosebezogene Haupteffekte (+RCC)	Wechselwirkungen	Gesamt	
Var. A, reduziert, mit Hier., LB07, 1-J-prosp., V404	2.823.493	32	33 (+28)	-	93	27,68 %
Var. F, reduziert, ohne Hier., LB07, 1-J-prosp., V404	2.823.493	32	32 (+28)	-	92	27,76 %
Var. F, reduziert, ohne Hier., LB08, 2-J-prosp., V404	2.668.118	32	30 (+27)	-	89	36,04 %
Var. F, reduziert, mit Hier., LB07, 1-J-prosp., V404	2.823.493	32	32 (+28)	-	92	27,42 %
Var. F, reduziert 130, mit Hier., LB08, 1-J-prosp., V404	2.804.342	32	30 (+27)	-	89	29,54 %
Var. F, komplett, ohne Hier., LB07, zeitgl., V309	3.000.000	22	156	111	289	7,36 %
Var. F, komplett, ohne Hier., LB08, 2-J-prosp., V404	2.668.118	32	136	86	254	36,38 %
Var. F, komplett, mit Hier., LB08, 2-J-prosp., V404	2.668.118	32	137	93	262	35,07 %
Var. G, komplett, ohne Hier., LB07, zeitgl., V309	3.000.000	26	156	116	298	8,05 %
Var. G, komplett, ohne Hier., LB08, zeitgl., V404	3.000.000	26	156	116	298	9,90 %
Var. G, komplett, ohne Hier., LB08, 2-J-prosp., V404	2.668.118	32	139	81	253	36,40 %

Kurzbeschreibung	Zahl der Risikokategorien					Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz
	N (Kalibrierungsstichprobe aus der Lernstichprobe)	Demographische Haupteffekte	Diagnosebezogene Haupteffekte (+RCC)	Wechselwirkungen	Gesamt	
Var. G, komplett, ohne Hier., LB08, 2-J-prosp., X404	2.668.118	32	160	121	313	37,19 %

Tabelle 38: Kennzahlen der beschlossenen Ausgestaltung im Vergleich mit weiteren Ausgestaltungen

4.3.3 Gütemaße

Die Tabellen „Gütemaße“ geben verschiedene Werte zur Beurteilung der Modelle an. Wie im vorherigen Abschnitt wird die Tabelle einmal nur für die beschlossene Ausgestaltung (Tabelle 39) und einmal als Überblick mit weiteren bearbeiteten Ausgestaltungen dargestellt (Tabelle 40). Die beschlossene Ausgestaltung steht in der zweiten Tabelle in der untersten Spalte. Die weiteren Ausgestaltungen werden in dieser Darstellung nur mit einer Kurzbeschreibung dargestellt, die ausführliche Beschreibung der Ausgestaltung findet sich in Abschnitt 3.2 (Tabelle 22).

Der „relativer Prognosefehler“ wird berechnet aus dem „mean absolute prediction error“ (MAPE). Die Formel wird in Abbildung 24 dargestellt²⁸. MAPE wird berechnet als Betrag der Differenz des tatsächlichen Leistungsbedarfs je Versicherten (y_i) zum geschätzten Leistungsbedarf je Versicherten (\hat{y}_i). Dieses wird über alle Versicherten $i=1, \dots, n$ der Stichprobe summiert und durch die Zahl der Versicherten n in der Stichprobe dividiert.

Der relative Prognosefehler ergibt sich anschließend durch die Relativierung am mittleren tatsächlichen Leistungsbedarf (\bar{y}).

²⁸ Auf die Darstellung der in den Berechnungen vorgenommenen Gewichtungen mit den Versichertenzeiten wird hier der Übersichtlichkeit halber verzichtet.

$$MAPE = \frac{\sum_i |y_i - \hat{y}_i|}{n}$$

Abbildung 24: Formel zur Berechnung des mittleren absoluten Prognosefehlers

Die Berechnung des relativen Prognosefehlers wird sowohl für die Stichprobe ausgeführt, auf der die Relativgewichte modelliert wurden (Kalibrierungsstichprobe), als auch auf einer weiteren Stichprobe (Validierungsstichprobe).

Bei prospektiven Modellen erfolgt die Bildung der Validierungsstichprobe entsprechend der Bildung der Kalibrierungsstichprobe (vgl. Abschnitt 3.1.5.1).

Für die Berechnung der Predictive Ratios wurde sowohl der durchschnittliche tatsächliche Leistungsbedarf 2008 als auch der durchschnittliche geschätzte Leistungsbedarf 2008 berechnet. Der geschätzte Leistungsbedarf 2008 wurde mit Diagnosen des Jahres 2006 berechnet. Die Abgrenzung des Leistungsbedarfs ist bei beiden Berechnungen gleich.

Der Wert „adjustiertes R^2 “ ist das Bestimmtheitsmaß der Modelle, das den Verlust von Freiheitsgraden bei der Aufnahme von Variablen ins Modell berücksichtigt. Die entsprechende Formel ist in Abbildung 25 angegeben, wobei p die Anzahl der Variablen im Modell darstellt.

$$R_{adj}^2 = 1 - \frac{\frac{1}{n-p-1} \sum_i (y_i - \hat{y}_i)^2}{\frac{1}{n-1} \sum_i (y_i - \bar{y})^2}$$

Abbildung 25: Formel zur Berechnung des adjustierten R^2

Kurzbeschreibung	Relativer Prognosefehler (MAPE / Mean_LB)		Predictive Ratio (geschätzt / tatsächlich)		Adjus- tiertes R ²
	in der Kalibrie- rungsstich- probe	in der Validie- rungs- stichprobe	in der Kalibrie- rungsstich- probe	in der Vali- dierungssti- chprobe	
Var. H, komprimiert mit Hier., LB08, 2-J-prosp., X404	58,13 %	58,10 %	1,00000	1,00047	27,0 %

Tabelle 39: Gütemaße der beschlossenen Ausgestaltung

Kurzbeschreibung	Relativer Prognosefehler (MAPE / Mean_LB)		Predictive Ratio (geschätzt / tatsächlich)		Adjus- tiertes R ²
	in der Kalibrie- rungsstich- probe	in der Validie- rungs- stichprobe	in der Kalibrie- rungsstich- probe	in der Vali- dierungs- stichprobe	
Var. A, reduziert, mit Hier., LB07, 1-J-prosp., V404	62,25 %	62,07 %	1,00000	0,99746	43,4 %
Var. F, reduziert, ohne Hier., LB07, 1-J-prosp., V404	60,59 %	60,27 %	1,00000	0,99641	52,1 %
Var. F, reduziert, ohne Hier., LB08, 2-J-prosp., V404	61,40 %	61,24 %	1,00000	0,99841	41,8 %
Var. F, reduziert, mit Hier., LB07, 1-J-prosp., V404	60,82 %	60,52 %	1,00000	0,99647	51,9 %
Var. F, reduziert 130, mit Hier., LB08, 1-J-prosp., V404	54,61 %	54,43 %	1,00000	0,99963	60,3 %
Var. F, komplett, ohne Hier., LB07, zeitgl., V309	48,85 %	48,95 %	1,00000	1,00000	57,0 %
Var. F, komplett, ohne Hier., LB08, 2-J-prosp., V404	60,95 %	60,80 %	1,00000	0,99901	42,7 %
Var. F, komplett, mit Hier., LB08, 2-J-prosp., V404	61,05 %	61,59 %	1,00000	0,97437	42,5 %

Kurzbeschreibung	Relativer Prognosefehler (MAPE / Mean_LB)		Predictive Ratio (geschätzt / tatsächlich)		Adjus- tiertes R ²
	in der Kalibrie- rungsstich- probe	in der Validie- rungs- stichprobe	in der Kalibrie- rungsstich- probe	in der Vali- dierungs- stichprobe	
Var. G, komplett, ohne Hier., LB07, zeitgl., V309	48,72 %	48,83 %	1,00000	0,99996	57,0 %
Var. G, komplett, ohne Hier., LB08, zeitgl., V404	46,51 %	46,40 %	1,00000	0,99792	61,8 %
Var. G, komplett, ohne Hier., LB08, 2-J-prosp., V404	60,93 %	60,78 %	1,00000	0,99898	42,7 %
Var. G, komplett, ohne Hier., LB08, 2-J-prosp., X404	64,54 %	57,37 %	1,00000	1,00065	28,3 %
Var. H, komprimiert, mit Hier., LB08, 2-J-prosp., X404	58,12 %	58,10 %	1,00000	1,00047	27,0 %

Tabelle 40: Gütemaße der beschlossenen Ausgestaltung im Vergleich mit weiteren beauftragten Ausgestaltungen

Während die Predictive Ratios ein Maß für die Erwartungstreue der Modelle sind, gibt der relative absolute Prognosefehler (MAPE / Mean_LB) die mittlere absolute Schätzungenauigkeit auf Versichertenebene an *in Beziehung zum mittleren Leistungsbedarf*. Die Verteilung des Prognosefehlers je Versicherten (PE in Abbildung 26) in der Validierungsstichprobe ist für das beschlossene Modell in Abbildung 27 und Abbildung 28 in Form eines Histogramms dargestellt, wobei auf der waagerechten Achse die Prognosefehler der Versicherten und auf der senkrechten Achse die relative Häufigkeit der Versicherten mit einem entsprechenden Prognosefehler abgetragen wird.

$$PE_i = y_i - \hat{y}_i$$

Abbildung 26: Formel zur Berechnung des Prognosefehlers

Aus Darstellungsgründen musste die Validierungsstichprobe hierfür zusätzlich getrimmt werden, d. h., es wurden Versicherte mit Extremwerten an den Verteilungsenden ausgeschlossen. In Abbildung 27 wurden jeweils die Beobachtungen ausgeschlossen, welche 0,1 % der Versichertenjahre am oberen Ende der Vertei-

lung beinhalten und die Beobachtungen, welche 0,1 % der Versichertenjahre am unteren Ende der Verteilung beinhalten.

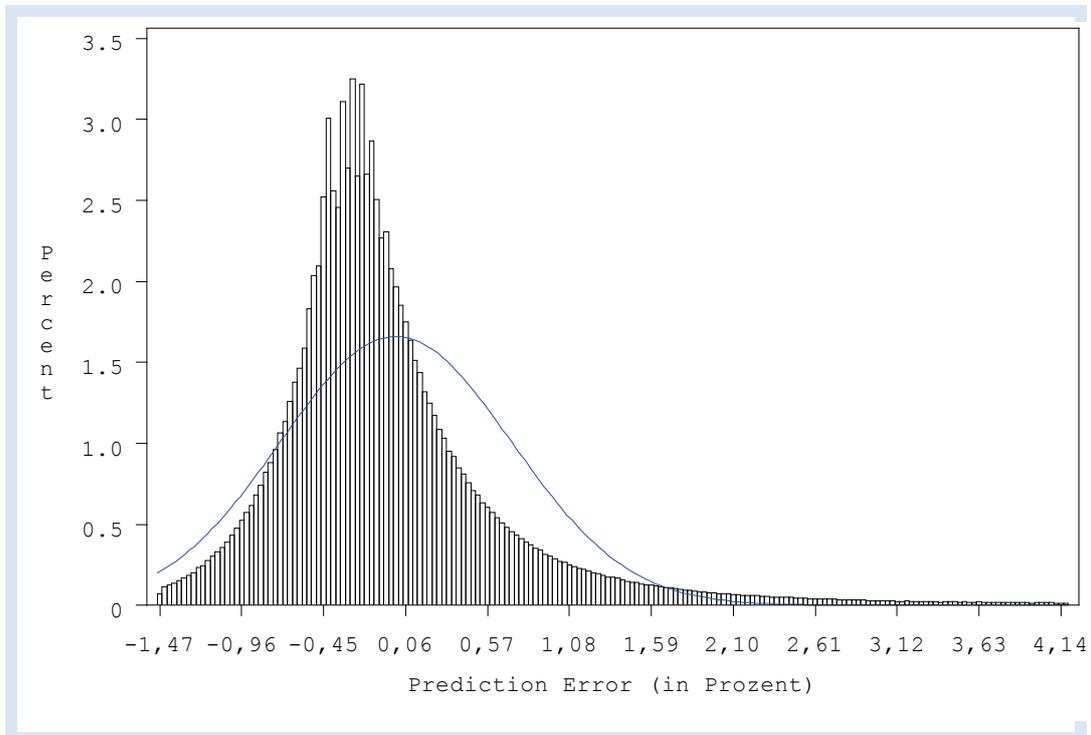


Abbildung 27: Histogramm der Prognosefehler je Versicherten in der Validierungsstichprobe

In Abbildung 28 wurden alle Beobachtungen ausgeschlossen, die außerhalb des Bereichs [unteres Quartil - $1,5 \cdot$ Interquartilsabstand; oberes Quartil + $1,5 \cdot$ Interquartilsabstand] liegen. Die Quartile und der Interquartilsabstand wurden dabei jeweils gewichtet mit den Versichertenjahren berechnet.

In beiden Grafiken ist außerdem eine Normalverteilungskurve mit $\mu=0$ und der jeweils aus den Daten geschätzten Standardabweichung eingefügt.

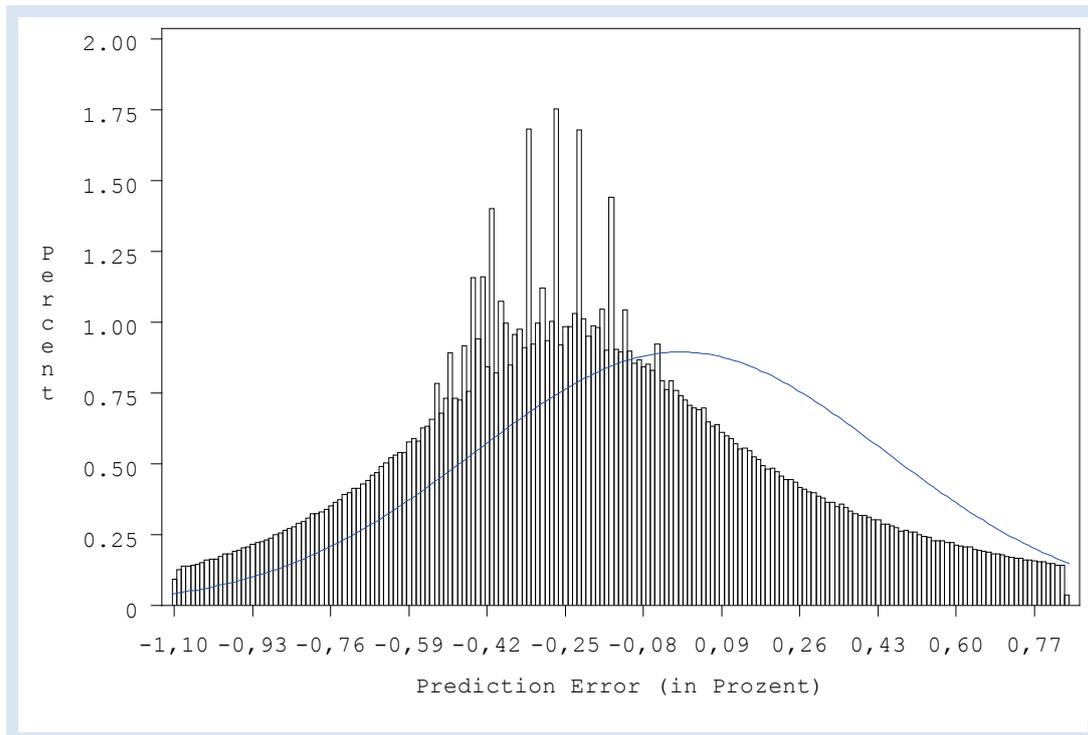


Abbildung 28: Histogramm der Prognosefehler je Versicherten in der Validierungsstichprobe

4.3.4 Predictive Ratios

In Abschnitt 4.3.3 wurden die Predictive Ratios berechnet für die Gesamtheit der Kalibrierungs- und der Validierungsstichprobe dargestellt. In diesem Abschnitt werden nun Predictive Ratios für einzelne Untergruppen dargestellt. Diese Auswertungen werden nur für die beschlossene Ausgestaltung vorgenommen.

Im ersten Teil werden die Versicherten einmal nach der Höhe der für sie geschätzten und einmal nach ihren tatsächlich abgerechneten Punkten unterteilt und die Predictive Ratios für jede Gruppe einzeln dargestellt. Im zweiten Teil werden die Predictive Ratios für jede Risikogruppe einzeln dargestellt.

4.3.4.1 Predictive Ratios getrennt nach der Höhe des geschätzten und tatsächlichen Leistungsbedarfs

Die Darstellung des Leistungsbedarfs in Tabelle 41 und Tabelle 42 erfolgt in Punkten. Die Darstellung von „N“ erfolgt in Zahl der Versicherten.

In Tabelle 41 wurden die Versicherten anhand ihres erwarteten Leistungsbedarfs in vierundzwanzig Gruppen unterteilt. Die Gruppenbildung erfolgte dabei mit

Hilfe von Perzentilen, anhand derer die Versicherten in jeweils gleich große Gruppen unterteilt wurden. Als Gruppengrenzen wurden nur gerundete Werte verwendet.

Klassen des erwarteten Leistungsbedarfs	N	Ø geschätzter LB	Ø tatsächlicher LB	Predictive Ratio
<= 2.650	113.168	2.007,74	2.722,24	0,73753
> 2.650 –3.300	143.538	2.968,45	3.056,78	0,97110
> 3.300 –4.000	108.741	3.696,55	3.973,03	0,93041
> 4.000 –4.500	138.014	4.250,67	3.934,33	1,08041
> 4.500 –5.100	123.948	4.798,13	4.617,74	1,03906
> 5.100 –5.750	130.202	5.417,36	5.285,02	1,02504
> 5.750 –6.400	127.115	6.083,08	5.938,42	1,02436
> 6.400 –7.100	132.445	6.749,40	6.529,49	1,03368
> 7.100 –7.800	128.246	7.444,22	7.257,85	1,02568
> 7.800 –8.500	120.936	8.146,61	7.972,41	1,02185
> 8.500 –9.250	120.825	8.868,99	8.680,59	1,02170
> 9.250 –10.100	125.943	9.669,69	9.627,07	1,00443
> 10.100 –11.000	119.866	10.541,21	10.489,63	1,00492
> 11.000 –12.000	119.833	11.489,86	11.492,79	0,99974
> 12.000 –13.200	125.776	12.584,30	12.651,71	0,99467
> 13.200 –14.600	125.517	13.878,91	14.110,07	0,98362
> 14.600 –16.300	125.656	15.415,37	15.643,55	0,98541
> 16.300 –18.500	124.518	17.346,44	17.521,32	0,99002
> 18.500 –22.000	128.514	20.083,95	20.167,31	0,99587
> 22.000 –26.100	75.905	23.782,21	23.761,64	1,00087
> 26.100 –29.100	26.273	27.437,48	27.173,99	1,00970
> 29.100 –32.300	13.457	30.486,51	30.058,03	1,01426
> 32.300 –51.900	10.954	36.895,05	36.142,95	1,02081
>= 51.900	2.653	74.498,68	74.676,97	0,99761

Tabelle 41: Vergleich des durchschnittlichen tatsächlichen und geschätzten Leistungsbedarfs in Abhängigkeit von der Höhe der geschätzten Punkte.

In Abbildung 29 wird der Vergleich des durchschnittlichen tatsächlichen und geschätzten Leistungsbedarfs grafisch dargestellt.



Abbildung 29: Vergleich des tatsächlichen und geschätzten durchschnittlichen Leistungsbedarfs in Abhängigkeit von der Höhe der geschätzten Punkte

In Abbildung 30 ist derselbe Vergleich nur für die unteren fünf Kostenklassen abgebildet.

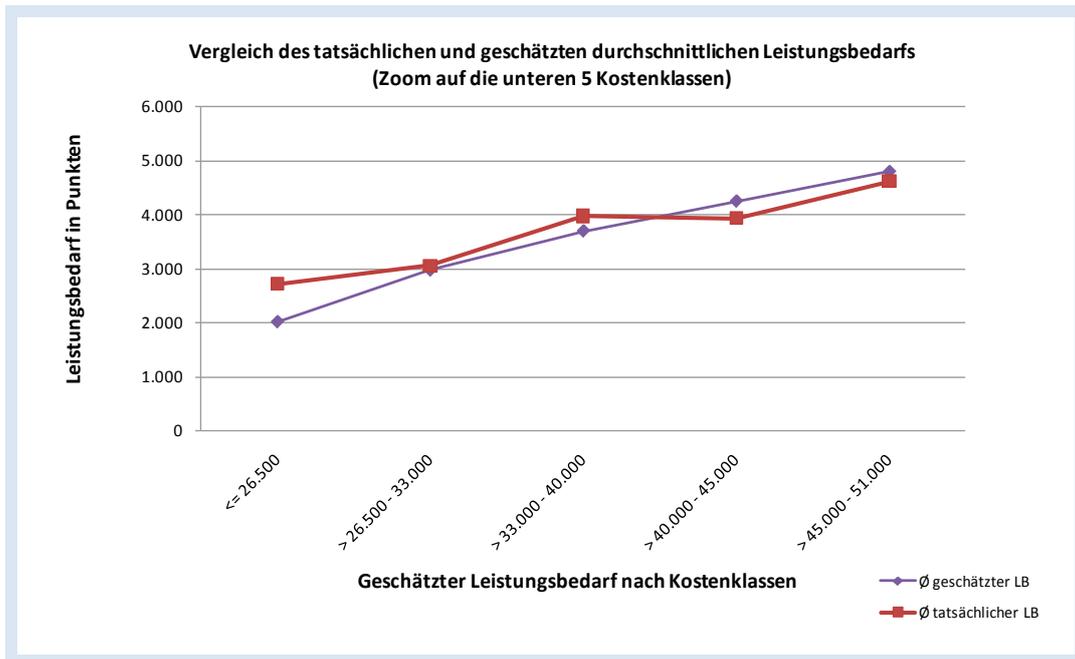


Abbildung 30: Vergleich des tatsächlichen und geschätzten durchschnittlichen Leistungsbedarfs in Abhängigkeit von der Höhe der geschätzten Punkte mit Zoom auf die 5 unteren Kostenklassen

Für die Darstellung der Predictive Ratios in der folgenden Tabelle, wurden die Versicherten anhand ihres tatsächlichen Leistungsbedarfs in vierundzwanzig Gruppen unterteilt. Als Gruppengrenzen wurden dieselben Werte wie in Tabelle 41 verwendet.

Klassen des geschätzten Leistungsbedarfs	N	Ø geschätzter LB	Ø tatsächlicher LB	Predictive Ratio
<= 2.650	616.779	5.375,58	965,41	5,56817
> 2.650 –3.300	112.165	6.585,74	2.982,53	2,20811
> 3.300 –4.000	115.040	7.097,43	3.650,61	1,94417
> 4.000 –4.500	81.536	7.418,55	4.245,48	1,74740
> 4.500 –5.100	95.154	7.844,12	4.798,64	1,63466
> 5.100 –5.750	96.263	8.205,52	5.425,01	1,51254
> 5.750 –6.400	93.506	8.672,80	6.075,76	1,42744
> 6.400 –7.100	99.498	9.284,10	6.745,64	1,37631
> 7.100 –7.800	91.948	9.797,27	7.444,28	1,31608
> 7.800 –8.500	83.580	10.240,52	8.145,07	1,25727
> 8.500 –9.250	82.349	10.688,89	8.870,41	1,20500

Klassen des geschätzten Leistungsbedarfs	N	Ø geschätzter LB	Ø tatsächlicher LB	Predictive Ratio
> 9.250 –10.100	85.739	11.188,47	9.666,98	1,15739
> 10.100 –11.000	81.059	11.686,30	10.540,35	1,10872
> 11.000 –12.000	80.654	12.155,30	11.490,58	1,05785
> 12.000 –13.200	84.799	12.706,92	12.581,75	1,00995
> 13.200 –14.600	83.513	13.287,97	13.879,48	0,95738
> 14.600 –16.300	85.520	13.972,22	15.419,36	0,90615
> 16.300 –18.500	87.908	14.665,91	17.354,59	0,84507
> 18.500 –22.000	100.035	15.624,00	20.131,28	0,77611
> 22.000 –26.100	77.357	16.646,91	23.889,61	0,69683
> 26.100 –29.100	37.886	17.417,84	27.513,88	0,63306
> 29.100 –32.300	29.363	17.854,45	30.620,23	0,58309
> 32.300 –51.900	72.454	18.589,38	39.457,78	0,47112
>= 51.900	37.938	20.370,10	78.968,47	0,25795

Tabelle 42: Vergleich des durchschnittlichen tatsächlichen und geschätzten Leistungsbedarfs in Abhängigkeit von der Höhe des tatsächlichen Leistungsbedarfs.

In Abbildung 31 wird der Vergleich des durchschnittlichen tatsächlichen und geschätzten Leistungsbedarfs nun in Abhängigkeit von der Höhe der tatsächlichen Punkte grafisch dargestellt.

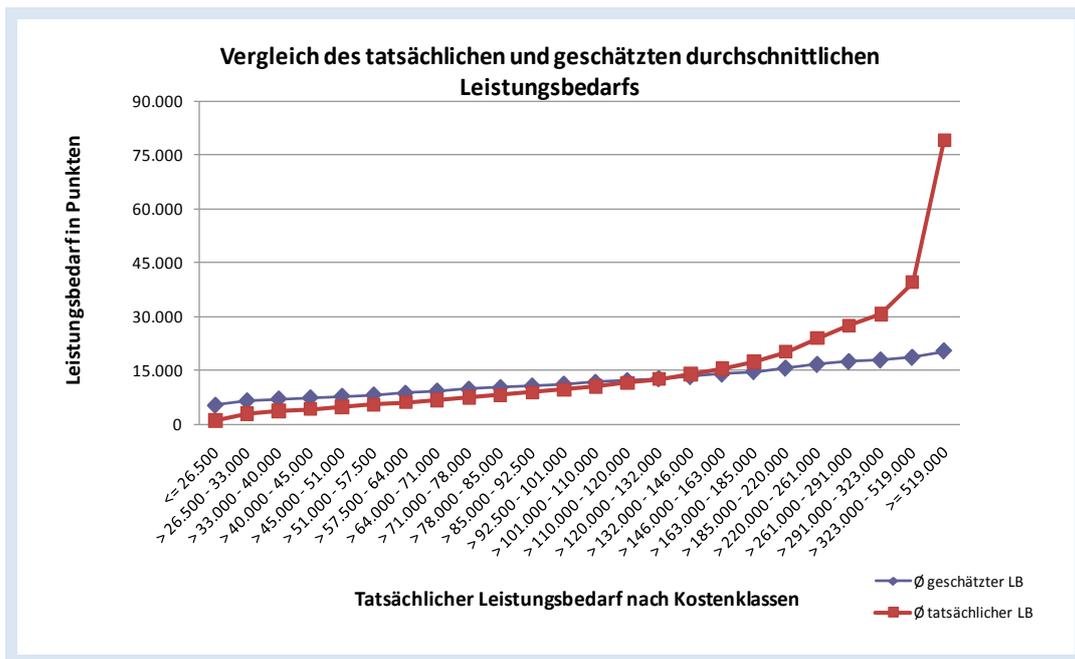


Abbildung 31: Vergleich des tatsächlichen und geschätzten durchschnittlichen Leistungsbedarfs in Abhängigkeit von der Höhe der tatsächlichen Punkte

4.3.4.2 Predictive Ratios je Risikoklasse

In den folgenden Tabellen werden die Predictive Ratios je Risikogruppe der beschlossenen Variante in der Validierungsstichprobe dargestellt. Die ersten zwei Spalten beschreiben die Risikogruppe. In der dritten und vierten Spalte wird der durchschnittliche Leistungsbedarf aller Versicherten mit dieser Risikokategorie berechnet, einmal tatsächlich und einmal geschätzt. Die Darstellung erfolgt in Punkten. In der fünften und sechsten Spalte werden je Risikokategorie die Predictive Ratios und der relative Prognosefehler dargestellt. In Tabelle 43 werden die Werte für die TCC dargestellt. Abbildung 32 gibt einen grafischen Überblick über die Predictive Ratios der TCC.

TCC	Beschreibung	N	Ø geschätzter LB (Standardabweichung)	Ø tatsächlicher LB (Standardabweichung)	Predictive Ratio (geschätzt / tatsächlich)	Relativer Prognosefehler (MAPE / Mean_LB)
6	Andere Infektionskrankheiten	541.943	11.410,19 (6.649,12)	11.398,25 (13.379,14)	1,00105	55,41 %

TCC	Beschreibung	N	Ø geschätzter LB (Standard-abweichung)	Ø tatsächlicher LB (Standard-abweichung)	Predictive Ratio (geschätzt / tatsächlich)	Relativer Prognosefehler (MAPE / Mean_LB)
10	Bösartige Neubildungen Dickdarm, Rektum, Anus, Mamma, Geschlechtsorgane, Niere und ableitende Harnwege, Auge, Schilddrüse, maligne Melanome	64.027	21.318,23 (7.438,25)	21.295,22 (17.107,96)	1,00108	45,98 %
14	Gutartige Neubildungen der Haut, der (weiblichen) Brust und des Auges	164.411	13.290,49 (6.103,23)	13.285,64 (14.096,28)	1,00036	54,43 %
19	Diabetes mellitus ohne oder mit nicht näher bezeichneten Komplikationen, Schwangerschaftsdiabetes	199.313	17.299,70 (7.805,19)	17.330,27 (13.953,00)	0,99824	43,42 %
24	Andere endokrine, metabolische und Ernährungsstörungen	770.542	14.443,78 (6.365,58)	14.436,96 (13.406,07)	1,00047	49,98 %
36	Andere gastrointestinale Erkrankungen	632.479	14.066,10 (7.222,19)	14.050,29 (14.191,36)	1,00113	51,52 %
38	Entzündliche Systemerkrankungen des rheumatoiden Formenkreises	82.251	18.848,70 (7.384,31)	18.835,89 (15.422,10)	1,00068	47,51 %
40	Cox- oder Gonarthrose	211.938	17.761,97 (7.136,64)	17.784,93 (14.078,84)	0,99871	45,67 %
41	Erkrankungen des Knochens und des Knorpels	201.641	17.694,90 (7.377,56)	17.692,86 (15.006,35)	1,00012	47,59 %

TCC	Beschreibung	N	Ø geschätzter LB (Standardabweichung)	Ø tatsächlicher LB (Standardabweichung)	Predictive Ratio (geschätzt / tatsächlich)	Relativer Prognosefehler (MAPE / Mean_LB)
43	Andere Erkrankungen der Muskeln, Sehnen, Gelenke und des Bindegewebes	542.601	11.495,75 (5.661,26)	11.474,44 (12.695,97)	1,00186	56,51 %
55	Bipolare Störungen, depressive Episoden	67.665	21.234,28 (6.819,38)	21.359,63 (20.164,38)	0,99413	56,27 %
57	Persönlichkeitsstörungen	50.839	18.388,24 (6.149,14)	18.209,87 (20.613,32)	1,00980	65,88 %
58	Depressive Episode, Angst, Dysthymia	144.576	17.849,52 (6.403,43)	17.799,28 (16.931,50)	1,00282	54,41 %
60	Angststörungen, sexuelle Störungen, Belastungs- und Anpassungsstörungen	206.158	13.479,34 (5.682,68)	13.480,65 (13.593,11)	0,99990	53,57 %
76	Radikulopathien, andere neurologische Erkrankungen und Folgen von Verletzungen des Nervensystems	367.395	15.883,37 (7.135,66)	15.871,85 (14.990,68)	1,00073	51,31 %
84	Koronare Herzkrankheit / andere chronisch-ischämische Erkrankungen des Herzens	155.317	18.701,85 (7.932,95)	18.655,99 (14.372,41)	1,00246	42,97 %
91	Essentielle Hypertonie	598.228	15.609,97 (7.227,04)	15.571,65 (13.393,65)	1,00246	46,14 %
110	Asthma bronchiale	140.905	14.100,71 (6.916,44)	14.066,86 (14.388,94)	1,00241	51,78 %
115	Akute Bronchitis, Grippe	377.956	11.159,98 (6.320,08)	11.124,41 (12.616,57)	1,00320	54,86 %
124	Andere Erkrankungen des Auges	662.221	13.968,21 (7.248,63)	13.980,87 (13.938,81)	0,99909	49,60 %

TCC	Beschreibung	N	Ø geschätzter LB (Standardabweichung)	Ø tatsächlicher LB (Standardabweichung)	Predictive Ratio (geschätzt / tatsächlich)	Relativer Prognosefehler (MAPE / Mean_LB)
127	Andere Erkrankungen HNO Bereich	955.766	11.502,78 (6.562,43)	11.508,88 (13.167,23)	0,99947	54,87 %
129	Terminale Niereninsuffizienz	3.819	65.115,07 (18.186,65)	66.817,19 (41.464,56)	0,97453	47,29 %
130	Abhängigkeit von Dialyse	1.885	80.429,02 (14.034,22)	81.642,85 (38.521,43)	0,98513	33,93 %
139	Störungen der Menstruation, klimakterische Störungen, andere Erkrankungen der weiblichen Geschlechtsorgane	542.157	13.622,43 (6.111,88)	13.612,07 (14.390,87)	1,00076	54,18 %
153	Verbrennungen 1. und 2. Grades, andere Erkrankungen der Haut und Hautanhangsgebilde	582.408	13.064,80 (6.948,17)	13.059,00 (13.992,24)	1,00044	52,88 %
162	Extremitätenfrakturen, andere Verletzungen	397.033	11.877,83 (7.131,48)	11.892,13 (13.389,16)	0,99880	54,18 %
166	Schwere Symptome, abnorme Befunde	502.953	13.069,76 (6.895,36)	13.056,77 (14.140,86)	1,00099	53,67 %
167	Schwindel, Unwohlsein und andere Allgemeinsymptome	440.943	13.854,43 (6.784,74)	13.826,55 (14.143,52)	1,00202	53,17 %
179	Verschiedene Eingriffe, Status nach operativen Eingriffen, Nachbehandlung, elektive Chirurgie	203.419	17.493,07 (7.804,61)	17.465,55 (14.921,73)	1,00158	45,82 %

TCC	Beschreibung	N	Ø geschätzter LB (Standardabweichung)	Ø tatsächlicher LB (Standardabweichung)	Predictive Ratio (geschätzt / tatsächlich)	Relativer Prognosefehler (MAPE / Mean_LB)
183	Vorsorgeuntersuchungen, Impfungen, Beratung zur Kontrazeption	1.185.915	12.544,62 (6.502,42)	12.521,28 (13.288,73)	1,00186	52,67 %
190	Knöchernerkrankungen der Wirbelsäule, Erkrankungen der Bandscheiben, Stenosen des Spinalkanals und der Foramina intervertebralia	224.426	17.511,74 (7.134,39)	17.519,01 (15.453,41)	0,99958	49,88 %
194	Migräne und Kopfschmerzen weiterer Genesen	207.841	13.806,48 (6.686,14)	13.796,28 (15.001,24)	1,00074	55,66 %

Tabelle 43: Predictive Ratios je TCC

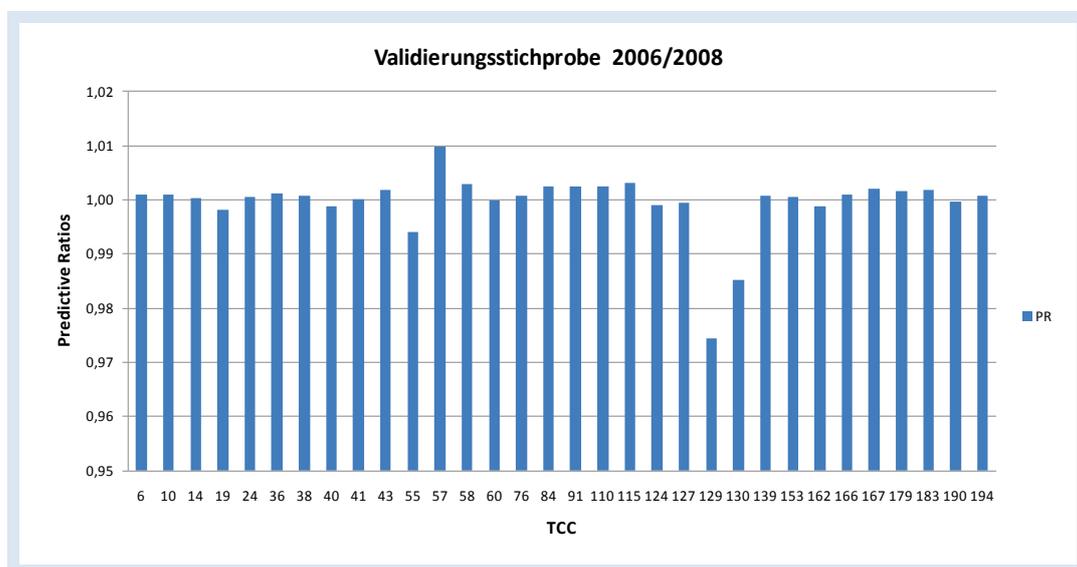


Abbildung 32: grafische Darstellung der Predictive Ratios je TCC

In Abbildung 33 und Tabelle 44 werden die Predictive Ratios für die RCC der beschlossenen Ausgestaltung dargestellt.

RCC	Beschreibung	N	Ø geschätzter LB (Standard-abweichung)	Ø tatsächlicher LB (Standard-abweichung)	Predictive Ratio (geschätzt / tatsächlich)	Relativer Prognosefehler (MAPE / Mean_LB)
1	Infektiöse und parasitäre Krankheiten; nach Variante 15. EBA	15.997	17.754,98 (8.863,75)	17.859,88 (17.354,41)	0,99413	55,91 %
2	Bösartige Neubildungen; nach Variante 15. EBA	28.096	23.177,91 (7.658,36)	22.974,00 (20.971,65)	1,00888	54,97 %
3	Sonstige Neubildungen und Krankheiten der Mamma [Brustdrüse]; nach Variante 15. EBA	160.793	15.307,74 (6.588,62)	15.321,56 (14.554,18)	0,99910	50,75 %
4	Diabetes mellitus; nach Variante 15. EBA	83.193	19.472,78 (9.017,08)	19.560,70 (15.769,79)	0,99551	42,17 %
5	Sonstige endokrine, Ernährungs- und Stoffwechselkrankheiten; nach Variante 15. EBA	42.914	19.293,51 (12.931,49)	19.375,65 (20.900,12)	0,99576	49,24 %
6	Krankheiten der Leber; nach Variante 15. EBA	151.234	15.752,86 (7.454,64)	15.688,51 (14.175,07)	1,00410	49,37 %
7	Sonstige Krankheiten des Verdauungssystems; nach Variante 15. EBA	88.585	16.601,18 (8.126,10)	16.556,71 (15.561,37)	1,00269	50,07 %
8	Krankheiten des Muskel-Skelettsystems und des Bindegewebes; nach Variante 15. EBA	143.683	15.098,91 (7.431,58)	15.093,77 (14.954,56)	1,00034	52,12 %

RCC	Beschreibung	N	Ø geschätzter LB (Standardabweichung)	Ø tatsächlicher LB (Standardabweichung)	Predictive Ratio (geschätzt / tatsächlich)	Relativer Prognosefehler (MAPE / Mean_LB)
9	Krankheiten des Blutes und der blutbildenden Organe sowie bestimmte Störungen mit Beteiligung des Immunsystems; nach Variante 15. EBA	133.799	16.695,78 (9.869,58)	16.790,22 (17.913,24)	0,99438	51,39 %
10	Störungen des Bewusstseins und neuropsychologische Symptome; nach Variante 15. EBA	74.406	18.630,37 (7.730,11)	18.545,39 (14.115,36)	1,00458	45,92 %
11	Psychische und Verhaltensstörungen durch Alkohol-, Drogen und sonstige Substanzen; nach Variante 15. EBA	99.461	13.597,80 (6.937,91)	13.571,08 (14.708,76)	1,00197	57,88 %
12	Sonstige psychische Krankheiten; nach Variante 15. EBA	75.816	19.051,88 (6.965,91)	18.871,26 (19.124,17)	1,00957	58,78 %
13	Entwicklungs- und Lernstörungen, Intelligenzminderung, Ticstörungen und komplexe Fehlbildungssyndrome; nach Variante 15. EBA	93.335	10.425,54 (5.240,09)	10.368,90 (13.253,80)	1,00546	61,85 %
14	Sonstige neurologische Krankheiten; nach Variante 15. EBA	94.910	19.307,73 (8.685,70)	19.364,40 (16.187,10)	0,99707	46,06 %

RCC	Beschreibung	N	Ø geschätzter LB (Standard-abweichung)	Ø tatsächlicher LB (Standard-abweichung)	Predictive Ratio (geschätzt / tatsächlich)	Relativer Prognosefehler (MAPE / Mean_LB)
15	Herzstillstand, Ateminsuffizienz, Tod und Vorhandensein eines Tracheostomas; nach Variante 15. EBA	13.775	19.421,94 (8.435,70)	19.264,45 (15.952,32)	1,00818	46,45 %
16	Krankheiten des Herz-Kreislaufsystems; nach Variante 15. EBA	343.838	17.207,72 (7.814,37)	17.172,79 (14.565,31)	1,00203	46,07 %
17	Zerebrovaskuläre Krankheiten; nach Variante 15. EBA	101.237	19.194,09 (7.884,94)	19.142,47 (14.470,78)	1,00270	43,66 %
18	Sonstige Blutgefäß-Krankheiten; nach Variante 15. EBA	399.043	15.935,14 (7.413,38)	15.913,22 (14.598,62)	1,00138	49,07 %
19	Krankheiten der unteren Atemwege; nach Variante 15. EBA	186.015	15.809,68 (7.714,07)	15.792,60 (14.740,89)	1,00108	48,90 %
20	Augen-Krankheiten; nach Variante 15. EBA	284.390	17.702,12 (7.469,43)	17.734,57 (14.112,58)	0,99817	43,49 %
21	HNO- und stomatologische Krankheiten; nach Variante 15. EBA	185.972	15.233,34 (7.635,59)	15.243,56 (14.364,15)	0,99933	48,60 %
22	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation; nach Variante 15. EBA	345.542	16.850,71 (8.390,40)	16.839,16 (15.737,95)	1,00069	48,59 %

RCC	Beschreibung	N	Ø geschätzter LB (Standard-abweichung)	Ø tatsächlicher LB (Standard-abweichung)	Predictive Ratio (geschätzt / tatsächlich)	Relativer Prognosefehler (MAPE / Mean_LB)
23	Krankheiten der Genitalorgane; nach Variante 15. EBA	235.131	15.421,30 (7.029,40)	15.400,96 (14.287,41)	1,00132	49,26 %
25	Krankheiten der Haut und der Unterhaut; nach Variante 15. EBA	115.375	13.676,99 (7.705,27)	13.619,42 (14.538,61)	1,00423	52,53 %
26	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen; nach Variante 15. EBA	319.629	14.651,44 (7.599,96)	14.636,13 (14.973,12)	1,00105	51,64 %
28	Krankheiten und Zustände des Neugeborenen; nach Variante 15. EBA	3.930	7.220,30 (2.455,50)	7.167,50 (4.438,46)	1,00737	45,71 %
29	Krankheiten und Zustände nach sonstigen Transplantationen und nach besonderen medizinischen Maßnahmen; nach Variante 15. EBA	8.543	22.449,96 (10.703,32)	22.560,53 (19.760,54)	0,99510	47,84 %

RCC	Beschreibung	N	Ø geschätzter LB (Standardabweichung)	Ø tatsächlicher LB (Standardabweichung)	Predictive Ratio (geschätzt / tatsächlich)	Relativer Prognosefehler (MAPE / Mean_LB)
30	Untersuchungen, prophylaktische Maßnahmen, anamnestische Angaben und Zustände nach weiteren medizinischen Maßnahmen; nach Variante 15. EBA	160.723	18.426,40 (7.784,29)	18.369,57 (16.043,12)	1,00309	47,12 %

Tabelle 44: Predictive Ratios je RCC

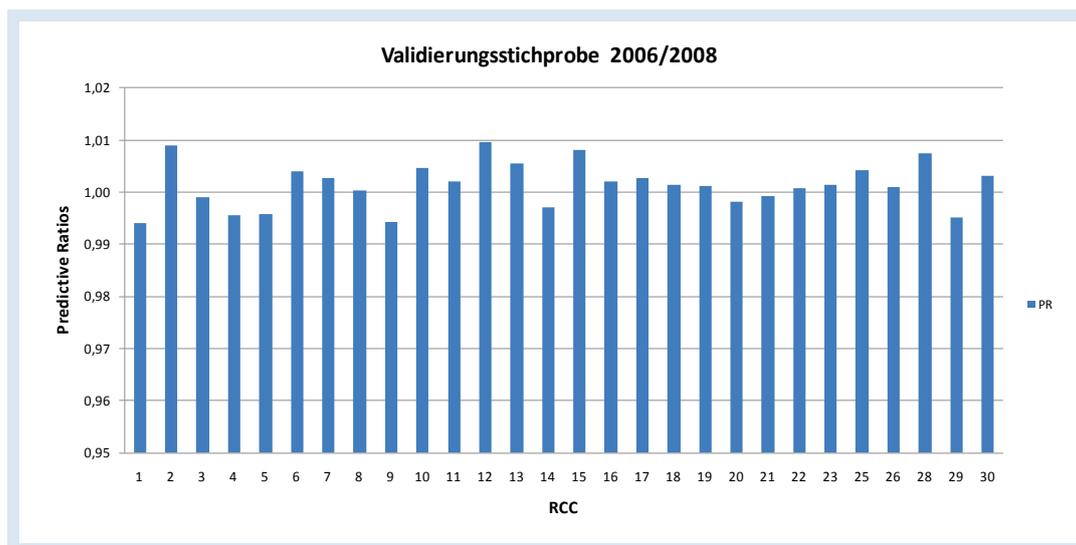


Abbildung 33: grafische Darstellung der Predictive Ratios je RCC

In Tabelle 45 werden die Predictive Ratios für die Alters- und Geschlechtsgruppen der beschlossenen Ausgestaltung dargestellt.

Alters- und Geschlechtsklasse	Be- schreibung	N	Ø geschätzter LB (Standard- abweichung)	Ø tatsächlicher LB (Standard- abweichung)	Predictive Ratio (ge- schätzt / tatsäch- lich)	Relativer Prognosef- ehler (MAPE / Mean-LB)
1	Weiblich, 0 bis 5 Jahre	62.197	6.533,67 (2.734,16)	6.538,42 (6.134,23)	0,99927	52,58 %
2	Weiblich, 6 bis 12 Jahre	87.765	5.578,19 (2.757,22)	5.503,48 (8.566,73)	1,01357	67,12 %
3	Weiblich, 13 bis 17 Jahre	69.059	7.758,15 (3.175,62)	7.767,86 (10.154,28)	0,99875	60,33 %
4	Weiblich, 18 bis 24 Jahre	93.442	8.035,54 (3.698,68)	7.949,45 (10.520,81)	1,01083	65,11 %
5	Weiblich, 25 bis 34 Jahre	138.926	9.436,59 (4.287,83)	9.394,10 (13.505,48)	1,00452	68,81 %
6	Weiblich, 35 bis 44 Jahre	209.276	10.162,29 (4.871,75)	10.208,74 (13.979,31)	0,99545	66,86 %
7	Weiblich, 45 bis 54 Jahre	193.227	11.729,80 (5.740,71)	11.742,54 (14.462,93)	0,99891	61,82 %
8	Weiblich, 55 bis 59 Jahre	83.747	13.176,97 (6.439,09)	13.129,66 (13.860,03)	1,00360	54,97 %
9	Weiblich, 60 bis 64 Jahre	73.148	14.182,43 (6.740,62)	14.195,59 (13.307,70)	0,99907	49,93 %
10	Weiblich, 65 bis 69 Jahre	98.518	15.145,99 (6.984,98)	15.027,21 (12.850,83)	1,00790	46,87 %
11	Weiblich, 70 bis 74 Jahre	76.249	16.016,66 (7.384,46)	15.964,81 (12.849,76)	1,00325	44,25 %
12	Weiblich, 75 bis 79 Jahre	62.571	16.023,68 (7.440,30)	16.071,35 (12.280,10)	0,99703	43,18 %
13	Weiblich, 80 bis 84 Jahre	49.608	15.949,12 (7.450,76)	15.883,23 (11.590,36)	1,00415	43,54 %
14	Weiblich, 85 bis 89 Jahre	24.380	16.347,03 (6.980,47)	16.513,29 (11.729,45)	0,98993	44,85 %
15	Weiblich, 90 bis 94 Jahre	8.969	17.262,10 (6.163,67)	17.080,79 (10.994,43)	1,01061	44,97 %
16	Weiblich, 95 Jahre und älter	2.397	17.777,81 (5.529,11)	17.793,84 (10.658,37)	0,99910	42,87 %
17	Männlich, 0 bis 5 Jahre	65.009	7.317,83 (2.941,60)	7.326,79 (7.079,26)	0,99878	52,96 %
18	Männlich, 6 bis 12 Jahre	92.003	6.203,88 (2.975,36)	6.210,00 (9.851,56)	0,99901	70,69 %

Alters- und Geschlechtsklasse	Be-schreibung	N	Ø geschätzter LB (Standard-abweichung)	Ø tatsächlicher LB (Standard-abweichung)	Predictive Ratio (ge-schätzt / tatsäch-lich)	Relativer Prognosefehler (MAPE / Mean-LB)
19	Männlich, 13 bis 17 Jahre	70.849	4.595,34 (2.718,56)	4.604,44 (6.982,22)	0,99802	72,36 %
20	Männlich, 18 bis 24 Jahre	95.796	4.143,99 (2.727,00)	4.153,66 (6.570,98)	0,99767	80,60 %
21	Männlich, 25 bis 34 Jahre	129.746	4.997,00 (3.279,06)	5.007,86 (8.811,32)	0,99783	81,23 %
22	Männlich, 35 bis 44 Jahre	193.398	6.222,91 (3.981,00)	6.227,94 (10.068,43)	0,99919	76,51 %
23	Männlich, 45 bis 54 Jahre	177.003	7.983,00 (4.998,10)	7.999,91 (11.163,11)	0,99789	68,71 %
24	Männlich, 55 bis 59 Jahre	74.034	10.414,91 (5.972,86)	10.432,03 (12.162,05)	0,99836	60,05 %
25	Männlich, 60 bis 64 Jahre	64.108	12.061,58 (6.393,91)	12.054,88 (12.369,63)	1,00056	53,87 %
26	Männlich, 65 bis 69 Jahre	84.578	13.689,26 (6.893,65)	13.665,09 (13.160,13)	1,00177	49,69 %
27	Männlich, 70 bis 74 Jahre	60.102	15.459,83 (7.267,69)	15.475,72 (13.572,31)	0,99897	46,49 %
28	Männlich, 75 bis 79 Jahre	40.782	16.617,90 (7.751,25)	16.635,66 (13.499,09)	0,99893	44,23 %
29	Männlich, 80 bis 84 Jahre	21.104	16.961,60 (7.696,60)	17.123,18 (13.419,55)	0,99056	43,63 %
30	Männlich, 85 bis 89 Jahre	7.646	17.297,67 (7.382,21)	17.265,84 (12.499,48)	1,00184	44,36 %
31	Männlich, 90 bis 94 Jahre	2.029	18.038,55 (6.564,33)	18.305,33 (12.590,23)	0,98543	45,05 %
32	Männlich, 95 Jahre und älter	377	18.642,67 (5.782,90)	18.457,68 (11.180,24)	1,01002	42,80 %

Tabelle 45: Predictive Ratios für Alters- und Geschlechtsgruppen

4.3.5 Prävalenzen der Risikoklassen

In Tabelle 46 wird für jede Risikoklasse der beschlossenen Ausgestaltung dargestellt, wie viele Versicherte jeder Risikokategorie in den unterschiedlichen Stich-

proben zur Berechnung der diagnosenbezogenen Veränderungsrate zugeordnet wurden.

Risiko- klasse	Stichprobe zur Berechnung der Kostengewichte (Leistungsbedarf 2008, Diagnosen 2006) (insgesamt: 2.668.118)		Stichprobe zur Berechnung des Morbi-Index 2007 (insgesamt: 3.000.000)		Stichprobe zur Berechnung des Morbi-Index 2008 (insgesamt: 3.000.000)	
	N	Anteil in Prozent	N	Anteil in Prozent	N	Anteil in Prozent
AGE1	65.944	2,47 %	74.008	2,47 %	73.329	2,44 %
AGE2	93.006	3,49 %	101.660	3,39 %	99.900	3,33 %
AGE3	72.747	2,73 %	80.352	2,68 %	77.481	2,58 %
AGE4	98.981	3,71 %	124.839	4,16 %	125.405	4,18 %
AGE5	147.965	5,55 %	175.453	5,85 %	174.834	5,83 %
AGE6	222.210	8,33 %	240.056	8,00 %	231.003	7,70 %
AGE7	205.410	7,70 %	227.528	7,58 %	234.059	7,80 %
AGE8	88.351	3,31 %	98.324	3,28 %	99.439	3,31 %
AGE9	77.579	2,91 %	81.962	2,73 %	81.940	2,73 %
AGE10	104.588	3,92 %	105.674	3,52 %	102.156	3,41 %
AGE11	81.581	3,06 %	89.726	2,99 %	97.109	3,24 %
AGE12	66.494	2,49 %	70.128	2,34 %	69.689	2,32 %
AGE13	52.602	1,97 %	59.324	1,98 %	60.443	2,01 %
AGE14	26.351	0,99 %	37.097	1,24 %	40.725	1,36 %
AGE15	9.738	0,36 %	12.996	0,43 %	11.905	0,40 %
AGE16	2.483	0,09 %	5.326	0,18 %	5.628	0,19 %
AGE17	69.150	2,59 %	77.631	2,59 %	76.764	2,56 %
AGE18	97.737	3,66 %	106.201	3,54 %	104.766	3,49 %
AGE19	76.085	2,85 %	84.797	2,83 %	82.210	2,74 %
AGE20	101.591	3,81 %	126.858	4,23 %	129.308	4,31 %
AGE21	136.810	5,13 %	166.487	5,55 %	165.094	5,50 %
AGE22	206.247	7,73 %	223.598	7,45 %	213.546	7,12 %
AGE23	187.887	7,04 %	209.389	6,98 %	215.307	7,18 %
AGE24	78.300	2,93 %	87.292	2,91 %	88.177	2,94 %
AGE25	67.818	2,54 %	71.688	2,39 %	71.472	2,38 %
AGE26	89.480	3,35 %	92.353	3,08 %	88.622	2,95 %
AGE27	63.722	2,39 %	73.702	2,46 %	79.921	2,66 %

Risiko- klasse	Stichprobe zur Berechnung der Kostengewichte (Leistungsbedarf 2008, Diagnosen 2006) (insgesamt: 2.668.118)		Stichprobe zur Berechnung des Morbi-Index 2007 (insgesamt: 3.000.000)		Stichprobe zur Berechnung des Morbi-Index 2008 (insgesamt: 3.000.000)	
	N	Anteil in Prozent	N	Anteil in Prozent	N	Anteil in Prozent
AGE28	44.163	1,66 %	50.065	1,67 %	49.976	1,67 %
AGE29	22.360	0,84 %	28.814	0,96 %	31.865	1,06 %
AGE30	8.148	0,31 %	12.479	0,42 %	13.889	0,46 %
AGE31	2.188	0,08 %	3.241	0,11 %	2.982	0,10 %
AGE32	402	0,02 %	952	0,03 %	1.056	0,04 %
TCC006	576.442	21,60 %	640.366	21,35 %	652.642	21,75 %
TCC010	69.165	2,59 %	83.975	2,80 %	91.728	3,06 %
TCC014	174.088	6,52 %	198.108	6,60 %	219.845	7,33 %
TCC019	211.290	7,92 %	242.486	8,08 %	253.321	8,44 %
TCC024	818.222	30,67 %	906.195	30,21 %	936.956	31,23 %
TCC036	670.382	25,13 %	764.878	25,50 %	793.231	26,44 %
TCC038	87.587	3,28 %	96.796	3,23 %	101.838	3,39 %
TCC040	224.240	8,40 %	251.853	8,40 %	264.867	8,83 %
TCC041	214.663	8,05 %	241.497	8,05 %	257.298	8,58 %
TCC043	575.438	21,57 %	631.084	21,04 %	634.480	21,15 %
TCC055	71.898	2,69 %	80.748	2,69 %	85.033	2,83 %
TCC057	54.090	2,03 %	62.616	2,09 %	67.102	2,24 %
TCC058	152.865	5,73 %	171.469	5,72 %	175.674	5,86 %
TCC060	219.356	8,22 %	235.783	7,86 %	233.207	7,77 %
TCC076	390.894	14,65 %	441.378	14,71 %	463.895	15,46 %
TCC084	165.598	6,21 %	185.562	6,19 %	188.754	6,29 %
TCC091	635.298	23,81 %	712.015	23,73 %	747.594	24,92 %
TCC110	150.265	5,63 %	171.881	5,73 %	177.714	5,92 %
TCC115	401.267	15,04 %	463.218	15,44 %	453.263	15,11 %
TCC124	702.781	26,34 %	761.050	25,37 %	778.228	25,94 %
TCC127	1.015.292	38,05 %	1.133.546	37,78 %	1.153.126	38,44 %
TCC129	3.968	0,15 %	6.015	0,20 %	6.725	0,22 %
TCC130	2.007	0,08 %	2.950	0,10 %	3.658	0,12 %
TCC139	574.485	21,53 %	630.834	21,03 %	643.483	21,45 %
TCC153	618.145	23,17 %	675.426	22,51 %	703.098	23,44 %

Risiko- klasse	Stichprobe zur Berechnung der Kostengewichte (Leistungsbedarf 2008, Diagnosen 2006) (insgesamt: 2.668.118)		Stichprobe zur Berechnung des Morbi-Index 2007 (insgesamt: 3.000.000)		Stichprobe zur Berechnung des Morbi-Index 2008 (insgesamt: 3.000.000)	
	N	Anteil in Prozent	N	Anteil in Prozent	N	Anteil in Prozent
TCC162	421.012	15,78 %	457.791	15,26 %	468.945	15,63 %
TCC166	535.077	20,05 %	597.764	19,93 %	604.398	20,15 %
TCC167	468.014	17,54 %	529.525	17,65 %	549.242	18,31 %
TCC179	216.151	8,10 %	252.808	8,43 %	269.484	8,98 %
TCC183	1.259.665	47,21 %	1.448.705	48,29 %	1.483.666	49,46 %
TCC190	238.105	8,92 %	270.623	9,02 %	291.786	9,73 %
TCC194	220.295	8,26 %	237.195	7,91 %	239.498	7,98 %
RCC001	16.877	0,63 %	19.126	0,64 %	18.571	0,62 %
RCC002	29.793	1,12 %	43.623	1,45 %	46.842	1,56 %
RCC003	169.935	6,37 %	188.751	6,29 %	203.668	6,79 %
RCC004	88.556	3,32 %	102.121	3,40 %	107.650	3,59 %
RCC005	45.484	1,70 %	57.176	1,91 %	59.401	1,98 %
RCC006	160.629	6,02 %	176.363	5,88 %	183.622	6,12 %
RCC007	94.049	3,52 %	105.391	3,51 %	110.153	3,67 %
RCC008	153.227	5,74 %	173.199	5,77 %	184.697	6,16 %
RCC009	142.547	5,34 %	164.968	5,50 %	174.975	5,83 %
RCC010	79.249	2,97 %	101.786	3,39 %	102.847	3,43 %
RCC011	104.952	3,93 %	126.969	4,23 %	140.174	4,67 %
RCC012	80.043	3,00 %	89.496	2,98 %	91.635	3,05 %
RCC013	99.676	3,74 %	110.319	3,68 %	114.362	3,81 %
RCC014	101.089	3,79 %	118.937	3,96 %	126.967	4,23 %
RCC015	14.693	0,55 %	19.073	0,64 %	19.508	0,65 %
RCC016	366.386	13,73 %	413.060	13,77 %	423.772	14,13 %
RCC017	107.401	4,03 %	126.784	4,23 %	131.954	4,40 %
RCC018	423.290	15,86 %	467.427	15,58 %	482.656	16,09 %
RCC019	197.851	7,42 %	227.189	7,57 %	231.126	7,70 %
RCC020	302.171	11,33 %	330.526	11,02 %	343.622	11,45 %
RCC021	196.564	7,37 %	216.070	7,20 %	224.144	7,47 %
RCC022	366.961	13,75 %	420.622	14,02 %	441.062	14,70 %
RCC023	249.723	9,36 %	274.029	9,13 %	284.956	9,50 %

Risiko- klasse	Stichprobe zur Berechnung der Kostengewichte (Leistungsbedarf 2008, Diagnosen 2006) (insgesamt: 2.668.118)		Stichprobe zur Berechnung des Morbi-Index 2007 (insgesamt: 3.000.000)		Stichprobe zur Berechnung des Morbi-Index 2008 (insgesamt: 3.000.000)	
	N	Anteil in Prozent	N	Anteil in Prozent	N	Anteil in Prozent
RCC025	122.536	4,59 %	135.967	4,53 %	139.356	4,65 %
RCC026	339.608	12,73 %	378.298	12,61 %	398.560	13,29 %
RCC028	4.283	0,16 %	4.166	0,14 %	3.146	0,10 %
RCC029	9.259	0,35 %	12.860	0,43 %	13.811	0,46 %
RCC030	170.984	6,41 %	207.756	6,93 %	239.254	7,98 %

Tabelle 46: Prävalenzen je Risikoklasse in der Lernstichprobe und den Validierungstichproben

4.3.6 Relativgewichte

Tabelle 47 zeigt die Kostengewichte und Relativgewichte der Risikoklassen im beschlossenen Modell. Die Kostengewichte werden in Punkten (nicht Zehntelpunkten) dargestellt. Ein Kostengewicht von 1.000 Punkten für die Risikoklasse A bedeutet, dass der Versicherte aufgrund der Risikoklasse A einen zusätzlichen Leistungsbedarf von 1.000 Punkten hat. Die Relativgewichte werden ermittelt, indem die Kostengewichte durch den durchschnittlichen Leistungsbedarf der Kalibrierungstichprobe geteilt werden. Sie stellen ebenfalls eine Zunahme des Leistungsbedarfes eines Versicherten bei Hinzukommen einer Risikoklasse dar, jedoch ausgedrückt in Anteilen am mittleren Leistungsbedarf. Die Spalte „Zahl der Versichertenjahre“ berichtet je Risikoklasse die Summe der Versichertenjahre der Versicherten in der Stichprobe zur Berechnung des Morbiditätsindex 2008 (Lernstichprobe), die die jeweilige Risikogruppe haben. Die RCC 24 und die RCC 28 wurden in der beschlossenen Ausgestaltung auf Null gesetzt (vgl. Kap.4.2). Die Zahl der Versichertenjahre je Risikoklasse in 2008 wird daher in Klammern berichtet.

Risikoklasse	Relativgewicht	Zahl der Versichertenjahre je Risikoklasse in 2008
AGE01	0,179	64.533
AGE02	0,188	96.251
AGE03	0,379	74.568
AGE04	0,300	114.054
AGE05	0,387	162.022

Risikoklasse	Relativgewicht	Zahl der Versichertenjahre je Risikoklasse in 2008
AGE06	0,405	221.357
AGE07	0,434	226.834
AGE08	0,437	97.015
AGE09	0,450	80.252
AGE10	0,467	100.448
AGE11	0,451	95.805
AGE12	0,385	68.400
AGE13	0,355	58.571
AGE14	0,382	38.533
AGE15	0,537	10.761
AGE16	0,654	4.751
AGE17	0,214	67.664
AGE18	0,220	100.945
AGE19	0,152	78.571
AGE20	0,152	114.589
AGE21	0,219	151.401
AGE22	0,267	203.924
AGE23	0,329	208.617
AGE24	0,408	85.982
AGE25	0,448	69.898
AGE26	0,495	86.996
AGE27	0,529	78.270
AGE28	0,509	48.465
AGE29	0,465	30.423
AGE30	0,471	12.863
AGE31	0,581	2.616
AGE32	0,738	852
TCC006	0,060	634.373
TCC010	0,402	88.725
TCC014	0,088	215.608
TCC019	0,106	246.795
TCC024	0,081	917.867
TCC036	0,087	771.045
TCC038	0,234	100.090

Risikoklasse	Relativgewicht	Zahl der Versichertenjahre je Risikoklasse in 2008
TCC040	0,141	260.369
TCC041	0,138	252.223
TCC043	0,093	617.856
TCC055	0,617	83.017
TCC057	0,595	65.309
TCC058	0,341	171.114
TCC060	0,123	226.950
TCC076	0,127	455.053
TCC084	0,123	183.421
TCC091	0,111	731.532
TCC110	0,204	172.966
TCC115	0,048	440.931
TCC124	0,098	764.275
TCC127	0,084	1.122.076
TCC129	3,142	6.278
TCC130	3,518	3.440
TCC139	0,093	627.975
TCC153	0,091	685.120
TCC162	0,065	457.541
TCC166	0,087	587.140
TCC167	0,085	535.147
TCC179	0,093	264.165
TCC183	0,060	1.447.467
TCC190	0,199	286.814
TCC194	0,091	233.882
RCC001	0,262	18.072
RCC002	0,634	43.036
RCC003	0,098	200.151
RCC004	0,126	105.065
RCC005	0,194	56.314
RCC006	0,009	179.593
RCC007	0,100	107.217
RCC008	0,086	180.575
RCC009	0,098	169.806

Risikoklasse	Relativgewicht	Zahl der Versichertenjahre je Risikoklasse in 2008
RCC010	0,175	97.159
RCC011	0,023	136.161
RCC012	0,383	89.432
RCC013	0,142	111.138
RCC014	0,248	123.059
RCC015	0,171	18.088
RCC016	0,123	412.001
RCC017	0,121	127.552
RCC018	0,065	471.193
RCC019	0,090	224.369
RCC020	0,098	339.231
RCC021	0,080	219.835
RCC022	0,125	429.050
RCC023	0,081	279.173
RCC024	0,000	(54.113)
RCC025	0,078	134.785
RCC026	0,083	389.398
RCC027	0,000	0
RCC028	0,052	2.001
RCC029	0,319	12.910
RCC030	0,135	233.881

Tabelle 47: Kosten- und Relativgewichte je Risikoklasse in der beschlossenen Ausgestaltung

4.3.7 Vergleich des beschlossenen Modells einerseits mit einem kompletten Modell und andererseits mit einem komprimierten Modell mit einer anderen Leistungsabgrenzung

Die Abgrenzung des Leistungsbedarfs und die Wahl eines komprimierten Modells stellten zwei wesentliche Aspekte der Beschlussfassung des Erweiterten Bewertungsausschusses dar. Um diese isoliert betrachten zu können, wurde das beschlossene Modell jeweils mit einem Modell verglichen, welches sich nur in einem dieser beiden Aspekte unterscheidet. Im ersten Teil des folgenden Abschnittes wurde ein komplettes Modell gerechnet und im zweiten Teil wurde der Leistungsbedarf anders abgegrenzt.

4.3.7.1 Vergleich mit einem kompletten Modell

Als Vergleichsmodell wurde ein komplettes Modell auf Variante H mit Hierarchisierung, 2 Jahre prospektiv gerechnet. Die Abgrenzung des Leistungsbedarfs entspricht der Beschlussfassung des 15. EBA.

Tabelle 48 zeigt die Informationen über Kennzahlen zu den Modellen.

Kurzbeschreibung	N (Kalibrierungsstichprobe aus der Lernstichprobe)	Zahl der Risikoklassen				Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz
		Demographische Haupteffekte	diagnosebezogene Haupteffekte (+RCC)	Wechselwirkungen	Gesamt	
Var. H, komprimiert, mit Hier., LB 2008, 2-J-prosp., X404	2.668.118	32	32 (+28)	-	92	35,00 %
Var. H, komplett, mit Hier., LB 2008, 2-J-prosp., X404	2.668.118	32	148	133	313	35,72 %

Tabelle 48: Kennzahlen des beschlossenen Modells im Vergleich mit einem kompletten Modell der gleichen Ausgestaltung

Tabelle 49 zeigt verschiedene Werte zur Beurteilung der Modelle an.

Kurzbeschreibung	Relativer Prognosefehler (MAPE / Mean_LB)		Predictive Ratio (geschätzt / tatsächlich)		Adjustiertes R ²
	in der Kalibrierungsstichprobe	in der Validierungsstichprobe	in der Kalibrierungsstichprobe	in der Validierungsstichprobe	
Var. H, komprimiert, mit Hier., LB 2008, 2-J-prosp., X404	58,13%	58,10%	1,00000	1,00047	27,0%
Var. H, komplett, mit Hier., LB 2008, 2-J-prosp., X404	57,72%	57,83%	1,00000	0,99615	27,6%

Tabelle 49: Gütemaße des beschlossenen Modells im Vergleich mit einem kompletten Modell der gleichen Ausgestaltung

4.3.7.2 Vergleich mit anderer Leistungsabgrenzung

Als Vergleichsmodell wurde ein komprimiertes Modell auf Variante H mit Hierarchisierung, 2 Jahres prospektiv gerechnet. Der Leistungsbedarf des Vergleichsmodell enthält Dialysesachkosten.

Tabelle 50 zeigt die Informationen über Kennzahlen zu den Modellen.

Kurzbeschreibung	N (Kalibrierungsstichprobe aus der Lernstichprobe)	Zahl der Risikoklassen				Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz
		Demographische Haupteffekte	Diagnosebezogene Haupteffekte (+RCC)	Wechselwirkungen	Gesamt	
Var. H, komprimiert, mit Hier., LB 2008, 2-J-prosp., X404	2.668.118	32	32 (+28)	-	92	35,00%
Var. H, komprimiert, mit Hier., LB 2008, 2-J-prosp., V404	2.668.118	32	30 (+28)	-	90	33,87%

Tabelle 50: Kennzahlen des beschlossenen Modells im Vergleich mit einem komprimierten Modell anderer Abgrenzung des Leistungsbedarfs

Tabelle 51 zeigt verschiedene Werte zur Beurteilung der Modelle an.

Kurzbeschreibung	Relativer Prognosefehler (MAPE / Mean_LB)		Predictive Ratio (geschätzt / tatsächlich)		Adjustiertes R ²
	in der Kalibrierungsstichprobe	in der Validierungsstichprobe	in der Kalibrierungsstichprobe	in der Validierungsstichprobe	
Var. H, komprimiert, mit Hier., LB 2008, 2-J-prosp., X404	58,13%	58,10%	1,00000	1,00047	27,0%
Var. H, komprimiert, mit Hier., LB 2008, 2-J-prosp., V404	61,40%	61,25%	1,00000	0,99833	41,7%

Tabelle 51: Gütemaße des beschlossenen Modells im Vergleich mit einem komprimierten Modell anderer Abgrenzung des Leistungsbedarfs

5 Ausblick auf die zukünftige Datenlieferungen

Der vorliegende Bericht hat in Kapitel 2 Ergebnisse zu verschiedenen Analysen der Datengrundlage dargestellt und berichtet, inwieweit vorhandene Schwachstellen einen Einfluss auf die Berechnung von morbiditätsbedingten Veränderungsrate haben können. In Kapitel 3 wurden die Ergebnisse der beauftragten morbiditätsbedingten Veränderungsrate berichtet und unterschiedliche Überprüfungen zum Klassifikationsmodell vorgenommen.

Das Ziel dieses Abschnittes ist es einen Ausblick zu geben, wie schon geplante Änderungen in den Datenlieferungen an den Bewertungsausschuss berichtete Schwachstellen der Datengrundlage beheben können und welche Schwachstellen bleiben und daher systematisch angegangen werden sollen.

Für den Herbst 2009 ist geplant, dass dem Bewertungsausschuss erstmalig Daten aus einer bundesweiten Versichertenstichprobe (ca. 13 % der Versicherten) zur Verfügung gestellt werden. Aufgrund der bundesweiten Datenerhebung wird es keine Verzerrungen aufgrund fehlender Diagnosen sowie fehlendem Leistungsbedarf aus dem Fremdkassenzahlungsausgleich (vgl. Abschnitt 2.5) mehr geben.

Da mit dieser Datenlieferung auch die lebenslange Versichertennummer in die Daten aufgenommen wird und die Stichprobenziehung aufgrund von Geburtstagen vorgenommen wird, können Versicherte auch nach Umzug, Statuswechsel oder Versicherungswechsel zu einer Person zusammengeführt werden. Mögliche Verzerrungen der morbiditätsbedingten Veränderungsrate aufgrund von Unterjährigkeit werden daher deutlich verringert.

Schwachstellen der Datengrundlage, die mit Fusionen und unvollständigen Datenlieferungen zusammenhängen, können durch die bundesweite Versichertenstichprobe zunächst nicht behoben werden. Hier wird das Institut des Bewertungsausschusses allerdings versuchen, mit der Arbeitsebene des Bewertungsausschusses Verbesserungen an den Prüfroutinen zu besprechen, um Probleme frühzeitig erkennen zu können.

Aus Sicht des Instituts ist es darüber hinaus sinnvoll, stärker und frühzeitiger als bisher in die ETL-Prozesse bei der Datenstelle des Bewertungsausschusses involviert zu werden. Ein ETL-Prozess umfasst alle Schritte von der Annahme der Daten von den Lieferanten über notwendige Transformationen bis zur Ablage in einem geeigneten Datenbankschema.

Das Ziel dabei muss es sein, Inkonsistenzen oder Lücken in den gelieferten Daten frühzeitiger als bisher zu erkennen und umgehend Rückmeldung zum Korrekturbedarf zu geben. Im Rahmen der Qualitätssicherung der Grouper-Schnittstellen der Jahre 2006 bis 2008 wurden Inplausibilitäten entdeckt, die nicht in der eigentlichen Erstellung der Schnittstellen begründet waren. Vielmehr waren es Probleme im zugrunde liegenden Datenkörper. Eine strengere und umfassendere Prüfung in frühen Phasen der Datenbewirtschaftung würde helfen, viele Fehler am Ende der Prozesskette zu vermeiden.

Anhang 1 Technische Aspekte der Datengrundlage

In diesem Abschnitt sollen die technischen Aspekte der Datengrundlage näher ausgeführt werden. Insbesondere werden in diesem Abschnitt diejenigen Sachverhalte beschrieben, die für eine korrekte Pseudonymisierung der Krankenversicherenummer eines Versicherten notwendig sind. Weiterhin sollen alle Aspekte benannt und erläutert werden, die eine korrekte Zuordnung von Versichertenstammdaten zu Krankenkassen und zu ärztlichen Leistungsdaten erlauben. Insbesondere wird das Thema Kassenfusion behandelt.

Identifikation von Krankenkassen

KVK-IK der Krankenversichertenkarte:

In den ambulanten, ärztlichen Leistungsdaten wird dieses Merkmal, welches auf der Krankenversichertenkarte, in der Arztpraxis und Rahmen der Datenverarbeitung der Abrechnungsdaten der Kassenärztlichen Vereinigung verwendet wird, siebenstellig geführt.

Sofern ein KVK-IK im Zuge der Datenerhebung in einer Arztpraxis nicht eindeutig bestimmbar ist, wird ersatzweise eine Vertragskassennummer oder ein Krankenkassenname erhoben. In der Kassenärztlichen Vereinigung wird das fehlende KVK-IK durch ein Abrechnungs-IK ersetzt.

Das KVK-IK im Datenbestand der Krankenkasse ist immer neunstellig inklusive dem Präfix „10“, ansonsten weist dieses Merkmal die gleiche Ausprägung wie die Daten der Krankenversichertenkarte aus.

Das KVK-IK ist nicht immer aktuell: ggf. hatte die Kassenseite ein anderes Institutionskennzeichen an die Datenstelle übermittelt, als auf der Krankenversichertenkarte vermerkt/verwendet war (sog. veraltetes IK, Geschäftsstellen-IK, IK vor Fusion usw.). Um diesen Fehler auszugleichen wurde die Satzart 108 definiert, die einem (veraltetem oder kasseninternen) Institutionskennzeichen ein bekanntes Institutionskennzeichen (in Relation zu Satzart 092) zuweist.

Betriebsnummer:

Ordnungsnummer einer Krankenkasse, über die die sozialversicherungspflichtigen Abgaben der Arbeitgeber und Arbeitnehmer abgewickelt werden. Jeder Betriebsnummer können mehrere Institutionskennzeichen der Krankenversichertenkarte zugeordnet sein und umgekehrt.

Kassenfusion

Die Kassenfusion wird in dieser Ausarbeitung daraufhin eingeschränkt, dass eine Krankenkasse die Versicherten der fusionierten Kasse aufnimmt. Die Krankenversichertendaten der Versicherten der aufgenommenen Kasse ändern sich. Deshalb erhalten diese eine neue Krankenversichertenkarte. Im Zuge der Datenlieferung an die Datenstelle des Bewertungsausschusses kommt es zu einer Übermittlung von zusätzlichen Versichertenstammdaten bei der aufnehmenden Kasse. Die Verpflichtung zur Datenlieferung der aufgenommenen Kasse erlischt.

Die Beziehung zwischen dem Institutionskennzeichen der fusionierten Kasse und des Rechtsnachfolgers wird als separate Satzart 110 übermittelt.

Identifikation von Versicherten

Ein Versicherter wird stets durch die Merkmale Versichertennummer und KVK-IK der Kasse eindeutig identifiziert. Die Versichertennummer zu einer Person kann sich jedoch ändern, wenn sich persönliche Merkmale verändern (z.B. bei Namenswechsel, Wechsel des Versichertenstatus, Teilnahme an einem strukturierten Behandlungsprogramm (DMP) usw.). In der Regel ändert dieser Wechsel jedoch nicht das Institutionskennzeichen auf der Krankenversichertenkarte.

Die Versichertennummer ist nach der Spezifikation der Daten der Krankenversichertenkarte mindestens 7-stellig und maximal 12-stellig.

Diese mögliche Varianz bei der Stellenzahl der Krankenversichertennummer wird zur Vereinfachung der Datenverarbeitung für die Datenlieferung an den Bewertungsausschuss so behandelt, dass kürzere Versichertennummern von vorn auf 12 Stellen aufzufüllen sind.

Sollten alphabetische Zeichen Bestandteil der Versichertennummer sein, so werden diese aus der Zeichenkette entfernt (mündliche Vereinbarung AG IT). Die verkürzte Versichertennummer wird auch in diesem Fall mit führenden Nullen aufgefüllt.

Achtung: In der Arztpraxis ist nicht immer bekannt, dass ein Patient die Kasse gewechselt hat oder eine Kassenfusion stattgefunden hat, sofern der Versicherte noch mit der alten Krankenversichertenkarte den Arzt aufsucht. Mitunter erreichen deshalb alte, und damit ungültige Abrechnungsinformationen, die Kassenärztliche Vereinigung.

Die Versichertennummer in der Satzart 090 entspricht per Definition der auf 12 Stellen normalisierten Versichertennummer. Sofern dieser Sachverhalt bei der

Datenerhebung unberücksichtigt bleibt, kann hierin schon ein Grund für ein nicht funktionierendes Matching liegen.

In der Datenlieferung an den Bewertungsausschuss werden jeder Kombination aus (Pseudonym(Krankenversicherternummer), KVK-IK) ein Paar (PersonID, Betriebsnummer) in der Satzart 090 zugewiesen. Die PersonID und B_KassenNr²⁹ ist ein Merkmal, das einen Versicherten über den Zeitraum der Datenlieferung hinweg eindeutig kennzeichnet. Sofern ein Versicherter mehrere Versichertennummern bei der gleichen Kasse aufweist, kann es zu einem Paar (PersonID, Betriebsnummer) mehrere Datensätze in der Satzart 090 geben.

Pseudonymisierung der Versicherten

In der Datenstelle des Bewertungsausschusses werden keine Krankenversicherternummern geführt, sondern Pseudonyme dieser Krankenversicherternummern zu einem Versicherten. Damit die Zusammenführung von Versichertenstammdaten mit den ärztlichen Abrechnungsdaten möglich ist, müssen die Schritte der Normalisierung der Versichertennummer auf der Kassenseite wie seitens der Kassenärztlichen Vereinigungen gleichartig umgesetzt werden. Nur so ist sichergestellt, dass die im Anschluss hieran berechneten Pseudonyme, bei der Verwendung gleicher Schlüsselworte, ein gleiches Pseudonym ergeben.

Insofern gibt die Satzart 090 richtigerweise die Abbildung (pseudonym(Krankenversicherternummer), KVK-IK) → (PersonID, Betriebsnummer) wieder. In einem gleichartig gehaltenen zweiten Pseudonymisierungsschritt entstehen diejenigen Pseudonyme, die an die Datenstelle des Bewertungsausschusses übermittelt werden müssen. Bei der Verwendung gleicher Schlüsselworte werden auch hier gleiche Pseudonyme erzeugt.

Zusammenfassend lässt sich der hier implementierte Pseudonymisierungsvorgang wie folgt beschreiben:

Stufe 1: Normalisierung und Pseudonymisierung mit Schlüsselwort dieser Stufe

Stufe 2: Pseudonymisierung mit Schlüsselwort dieser Stufe

Die Schlüsselworte der beiden Stufen der Pseudonymisierung sind voneinander verschieden. Die Beteiligten an den jeweiligen Stufen der Pseudonymisierung kennen nur die Schlüsselworte ihrer Stufe.

Abbildung 34 verdeutlicht abschließend den Pseudonymisierungsvorgang.

²⁹ Aufgrund der kassenspezifischen PersonID ist eine Längsschnittfähigkeit von Personen bei Kassenwechsel nicht gegeben.

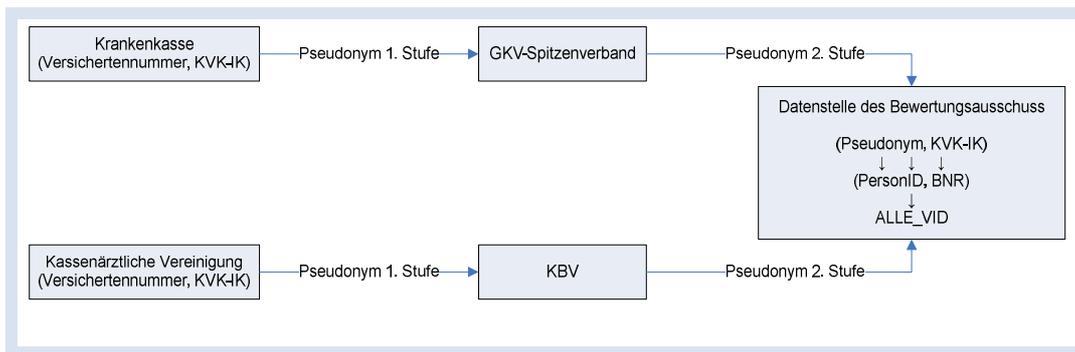


Abbildung 34: Pseudonymisierungsvorgang

Zusammenführung von Versichertenstammdaten und ärztlichen Abrechnungsdaten in der Datenstelle (Matching)

Idealerweise wird für jedes Paar (Krankenversichertennummer, KVK-IK) ein Wertepaar ((Pseudonym (Krankenversichertennummer), KVK-IK) sowohl in den Versichertenstammdaten der Krankenkassen als auch in den ambulanten, ärztlichen Abrechnungsdaten berechnet.

Hiermit sind die Voraussetzungen geschaffen, die Versichertenstammdaten mit den Leistungsdaten der Kassenärztlichen Vereinigungen zusammenzuführen. In der Terminologie der Datenstellen spricht man hier von „Matching“.

Grundsätzlich muss an dieser Stelle darauf hingewiesen werden, dass nicht alle Daten gematcht werden können:

- Es gibt Versicherte, zu denen keine Falldaten vorliegen (no show = ein Versicherter nimmt in diesem Quartal keine ärztlichen Leistungen in Anspruch)
- Es gibt Falldaten, zu denen es keine Versichertenstammdaten gibt (no matching = ein Versicherter, zu dem es keine Versichertenstammdaten gibt, nimmt ärztliche Leistungen in der Erhebungsregion in Anspruch)
- Es gibt Abrechnungsdaten mit KVK-IKs von historischen Kassen der Kassenart 2: Bei der Lieferung der Abrechnungsdaten wurden von der KV Nordrhein zwei KVK-IKs geliefert, die veraltet sind. Es handelt sich um die IKs zweier ehemaliger Kassen. Diese Daten konnten nicht mit den Versichertenstammdaten gematcht werden, da die Lieferung dieser Kassen für diese Versicherten das aktuelle, gültige IK aufweist. Die betreffenden Abrechnungsdatensätze sind somit nicht für die Jahre 2006 bis 2008 in der Grouper-Schnittstelle enthalten.

- In der Datenlieferung der Ersatzkasse 1 konnten 1344 VIDS nicht gemacht werden. Im Bezug auf die Grouper-Schnittstelle bedeutet dieses, dass pro Jahr ca. 200 Versicherte nicht gemacht werden konnten.

Damit das Matching erfolgreich sein kann, müssen folgende Kriterien erfüllt sein:

- Die Versicherten wohnen im PLZ-Bereich der vier beteiligten Kassenärztlichen Vereinigungen (Bremen, Niedersachsen, Nordrhein, Thüringen). Um diese Zuordnung prüfen zu können gibt es eine Postleitzahlen-Stammdatei, die eine Zuordnung zwischen PLZ und Kassenärztlicher Vereinigung erlaubt.
- Die Krankenkasse hat ihren Sitz in dem PLZ-Bereich der vier Kassenärztlichen Vereinigungen bzw. rechnet per Wohnort-Prinzip Versicherte aus dem PLZ-Bereich der vier Kassenärztlichen Vereinigungen ab.
- Es werden nur Kassen und Leistungsdaten von Versicherten der gesetzlichen Krankenversicherung berücksichtigt.
- Das Institutionskennzeichen der KVK ist gültig bzw. zu jeder Betriebsnummer gibt es eine Liste von in der Vergangenheit gültigen Institutionskennzeichen. Das im Abrechnungsfall enthaltene Institutionskennzeichen muss aus dieser Liste sein. Gleichzeitig muss die pseudonymisierte Versichertennummer zu den Versichertennummern derjenigen Betriebsnummer gehören, auf die das KVK-IK verweist.

Um Versicherte eindeutig zu kennzeichnen, wird in der Datenstelle des Bewertungsausschusses das Attribut „ALLE_VID“ geführt. Dieses ordnet jedem Paar (PersonID, Betriebsnummer) einen Wert zu. Das Attribut ist ein ganzzahliges Attribut und wird bei jeder Datenlieferung an die Datenstelle für neue Versicherte fortgeschrieben. Die Attributwerte von „ALLE_VID“ werden in eine Sequenz fortgeschrieben.

Schritt 1: Fortschreibung des Merkmals „ALLE_VID“

Wenn neue Versichertenstammdaten der Satzart 090 geladen werden, wird anhand der Attribute (PersonID, Betriebskassennummer) geprüft, ob diese Informationen bereits existieren. Ist diese Merkmals-Kombination nicht vorhanden, dann wird der Datensatz neu aufgenommen und ein Wert für „ALLE_VID“ für diesen Datensatz bestimmt.

Schritt 2: Aktualisierung der Versichertenstammdaten Satzart 091

Zu jedem Datensatz aus Satzart 090 wird ein Datensatz in Satzart 091 aufgenommen (siehe Datensatzbeschreibung zu Satzart 091). In Satzart 091 werden

die Versichertenstammdaten geführt. Zu jedem Datensatz aus Satzart 091 muss mindestens ein Datensatz in Satzart 090 existieren. Falls diese Bedingung nicht zutrifft, ist dieses als Fehler auszuweisen.

Über die Satzart 091 lassen sich generell alle Versicherten eines Jahres / Quartals feststellen.

Schritt 3: Verknüpfung zwischen Satzart 090 und Satzart 092

- Die von der KBV gelieferten Leistungsdaten und die von den Krankenkassen übermittelten Versichertenpseudonyme müssen zusammengeführt werden. Dazu sind als gemeinsame Attribute das KVK-IK und die (zweistufig) pseudonymisierte Versichertennummer in den genannten Satzarten vorhanden.
- Sofern ein Datenpaar (KVK-IK, pseudonymisierte Versichertennummer) aus Satzart 092 in Satzart 090 gefunden werden kann, gibt es zu den Leistungsdaten Versichertenstammdaten.
- Falls die Kombination (KVK-IK, Pseudonym) aus Satzart 092 ein veraltetes IK aufweist, wird dieses über die Satzart 108 in das KVK-IK umgeschlüsselt, so dass mit diesem IK das Matching zwischen Satzart 090 und Satzart 092 durchgeführt werden kann
- Das Matching zwischen Satzart 092 und Satzart 090 wird dadurch materialisiert, dass der aus Satzart 090 stammende Attributwert des Attributes „ALLE_VID“ in den Datensatz aus Satzart 092 (und den assoziierten Datensätzen für abgerechnete ambulante ärztliche Leistungen sowie den dokumentierten Diagnosen) nachgetragen wird.
- Ausschließlich Datensätze, die in Satzart 092 in diesem zusätzlichen Attribut einen Eintrag aufweisen, gelten als „gematcht“.

Nachdem das Matching durchgeführt wurde, müssen zu dem Paar (B_KASSNR, ALLE_VID) noch die aktuell gültigen Krankenkassenangaben ermittelt werden. Das Schaubild in Abbildung 35 zeigt diesen Prozess.

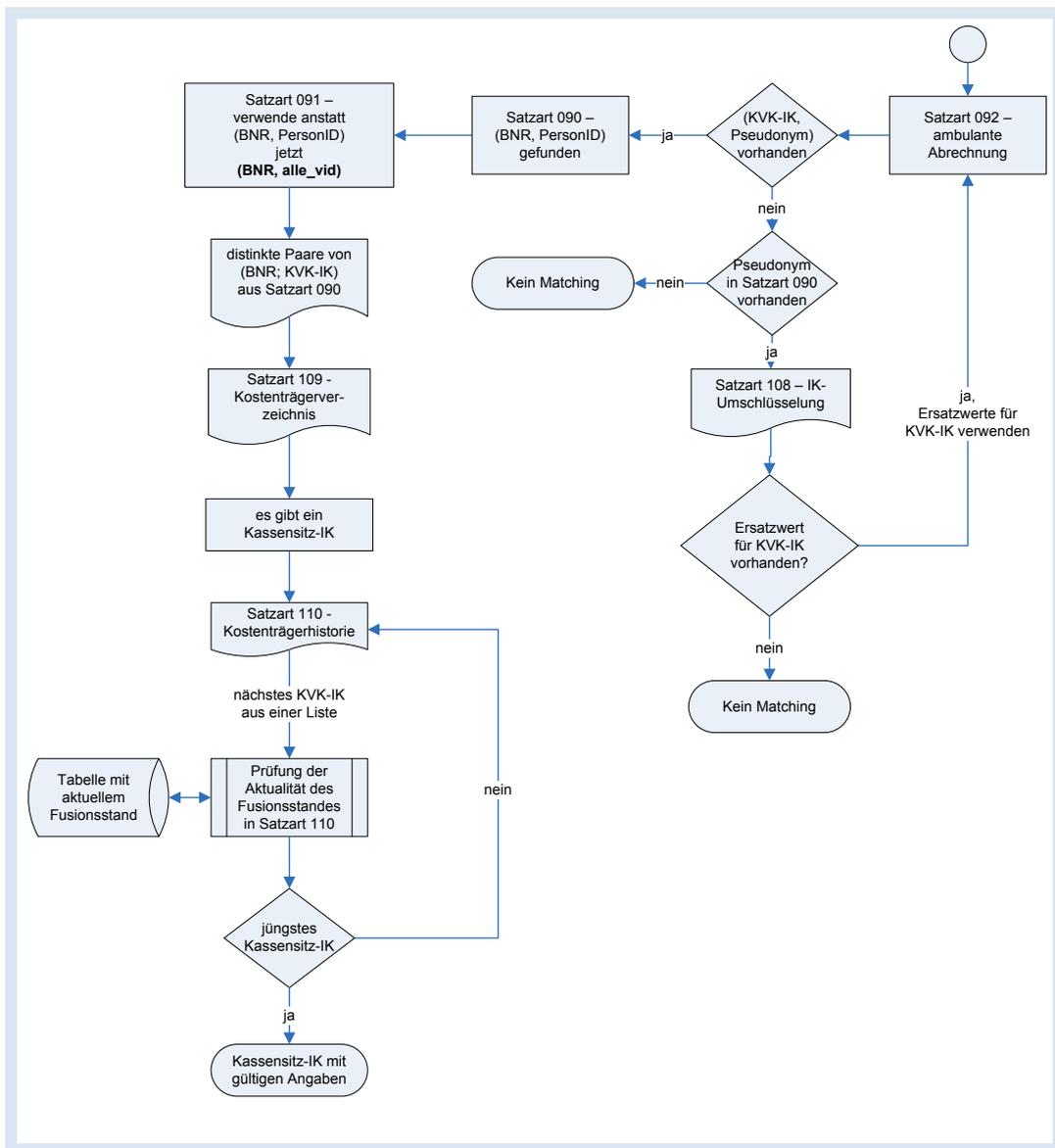


Abbildung 35: Matching und Bestimmung aktueller Krankenkassenangaben

Anhang 2 Datensatzbeschreibungen der Sonderauswertung zum Fremdkassenzahlungsausgleich (Lieferung der Kassenseite)

Abfrage zur Zahl der Diagnosen				
Jahr	WOP_KV	Arztsitz-KV	Zahl der Diagnosen-Datensätze auf Basis aller Arztgruppen	davon Zahl der Datensätze von Ärzten ohne Kodierverpflichtung (Anlage 1)
2006	Bremen	Bremen		
2006	Bremen	Niedersachsen		
2006	Bremen	Nordrhein		
2006	Bremen	Thüringen		
2006	Bremen	alle anderen		
2006	Niedersachsen	Bremen		
2006	Niedersachsen	Niedersachsen		
2006	Niedersachsen	Nordrhein		
2006	Niedersachsen	Thüringen		
2006	Niedersachsen	alle anderen		
2006	Nordrhein	Bremen		
2006	Nordrhein	Niedersachsen		
2006	Nordrhein	Nordrhein		
2006	Nordrhein	Thüringen		
2006	Nordrhein	alle anderen		
2006	Thüringen	Bremen		
2006	Thüringen	Niedersachsen		
2006	Thüringen	Nordrhein		
2006	Thüringen	Thüringen		
2006	Thüringen	alle anderen		
2007	Bremen	Bremen		
2007	Bremen	Niedersachsen		
2007	Bremen	Nordrhein		
2007	Bremen	Thüringen		
2007	Bremen	alle anderen		

Abfrage zur Zahl der Diagnosen				
Jahr	WOP_KV	Arztsitz-KV	Zahl der Diagnosen-Datensätze auf Basis aller Arztgruppen	davon Zahl der Datensätze von Ärzten ohne Kodierverpflichtung (Anlage 1)
2007	Niedersachsen	Bremen		
2007	Niedersachsen	Niedersachsen		
2007	Niedersachsen	Nordrhein		
2007	Niedersachsen	Thüringen		
2007	Niedersachsen	alle anderen		
2007	Nordrhein	Bremen		
2007	Nordrhein	Niedersachsen		
2007	Nordrhein	Nordrhein		
2007	Nordrhein	Thüringen		
2007	Nordrhein	alle anderen		
2007	Thüringen	Bremen		
2007	Thüringen	Niedersachsen		
2007	Thüringen	Nordrhein		
2007	Thüringen	Thüringen		
2007	Thüringen	alle anderen		
2008	Bremen	Bremen		
2008	Bremen	Niedersachsen		
2008	Bremen	Nordrhein		
2008	Bremen	Thüringen		
2008	Bremen	alle anderen		
2008	Niedersachsen	Bremen		
2008	Niedersachsen	Niedersachsen		
2008	Niedersachsen	Nordrhein		
2008	Niedersachsen	Thüringen		
2008	Niedersachsen	alle anderen		

Abfrage zur Zahl der Diagnosen				
Jahr	WOP_KV	Arztsitz-KV	Zahl der Diagnosen-Datensätze auf Basis aller Arztgruppen	davon Zahl der Datensätze von Ärzten ohne Kodierverpflichtung (Anlage 1)
2008	Nordrhein	Bremen		
2008	Nordrhein	Niedersachsen		
2008	Nordrhein	Nordrhein		
2008	Nordrhein	Thüringen		
2008	Nordrhein	alle anderen		
2008	Thüringen	Bremen		
2008	Thüringen	Niedersachsen		
2008	Thüringen	Nordrhein		
2008	Thüringen	Thüringen		
2008	Thüringen	alle anderen		

Tabelle 52: Datensatzbeschreibung zur „Abfrage zur Zahl der Diagnosen“

Anhang 3 Datensatzbeschreibungen der Sonderauswertung zum Fremdkassenzahlungsausgleich (Lieferung der Kassenärztlichen Bundesvereinigung)

Abfrage der C4-Daten						
Jahr	WOP_KV	Arztsitz-KV	Summe des Leistungsbedarfs in Punkte (LB_Pkt) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Punkten (LB_Pkt) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)	Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)
2007	Bremen	Bremen				
2007	Bremen	Niedersachsen				
2007	Bremen	Nordrhein				
2007	Bremen	Thüringen				
2007	Bremen	alle anderen				
2007	Niedersachsen	Bremen				
2007	Niedersachsen	Niedersachsen				
2007	Niedersachsen	Nordrhein				
2007	Niedersachsen	Thüringen				
2007	Niedersachsen	alle anderen				
2007	Nordrhein	Bremen				
2007	Nordrhein	Niedersachsen				
2007	Nordrhein	Nordrhein				
2007	Nordrhein	Thüringen				
2007	Nordrhein	alle anderen				

Abfrage der C4-Daten						
Jahr	WOP_KV	Arztsitz-KV	Summe des Leistungsbedarfs in Punkte (LB_Pkt) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Punkten (LB_Pkt) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)	Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)
2007	Thüringen	Bremen				
2007	Thüringen	Niedersachsen				
2007	Thüringen	Nordrhein				
2007	Thüringen	Thüringen				
2007	Thüringen	alle anderen				
2008	Bremen	Bremen				
2008	Bremen	Niedersachsen				
2008	Bremen	Nordrhein				
2008	Bremen	Thüringen				
2008	Bremen	alle anderen				
2008	Niedersachsen	Bremen				
2008	Niedersachsen	Niedersachsen				
2008	Niedersachsen	Nordrhein				
2008	Niedersachsen	Thüringen				
2008	Niedersachsen	alle anderen				

Abfrage der C4-Daten						
Jahr	WOP_KV	Arztsitz-KV	Summe des Leistungsbedarfs in Punkte (LB_Pkt) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Punkten (LB_Pkt) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)	Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)
2008	Nordrhein	Bremen				
2008	Nordrhein	Niedersachsen				
2008	Nordrhein	Nordrhein				
2008	Nordrhein	Thüringen				
2008	Nordrhein	alle anderen				
2008	Thüringen	Bremen				
2008	Thüringen	Niedersachsen				
2008	Thüringen	Nordrhein				
2008	Thüringen	Thüringen				
2008	Thüringen	alle anderen				

Tabelle 53: Datensatzbeschreibung zur „Abfrage der C4-Daten“

Anlage 3 – Abfrage der Daten der Ersatzkassen						
Jahr	WOP_KV	Arztsitz-KV	Summe des Leistungsbedarf in Punkte (LB_Pkt) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Punkten (LB_Pkt) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)	Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)
2006	Bremen	Bremen				
2006	Bremen	Niedersachsen				
2006	Bremen	Nordrhein				
2006	Bremen	Thüringen				
2006	Bremen	alle anderen				
2006	Niedersachsen	Bremen				
2006	Niedersachsen	Niedersachsen				
2006	Niedersachsen	Nordrhein				
2006	Niedersachsen	Thüringen				
2006	Niedersachsen	alle anderen				
2006	Nordrhein	Bremen				
2006	Nordrhein	Niedersachsen				
2006	Nordrhein	Nordrhein				
2006	Nordrhein	Thüringen				
2006	Nordrhein	alle anderen				

Anlage 3 – Abfrage der Daten der Ersatzkassen						
Jahr	WOP_KV	Arztsitz-KV	Summe des Leistungsbedarf in Punkte (LB_Pkt) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Punkten (LB_Pkt) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)	Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)
2006	Thüringen	Bremen				
2006	Thüringen	Niedersachsen				
2006	Thüringen	Nordrhein				
2006	Thüringen	Thüringen				
2006	Thüringen	alle anderen				
2007	Bremen	Bremen				
2007	Bremen	Niedersachsen				
2007	Bremen	Nordrhein				
2007	Bremen	Thüringen				
2007	Bremen	alle anderen				
2007	Niedersachsen	Bremen				
2007	Niedersachsen	Niedersachsen				
2007	Niedersachsen	Nordrhein				
2007	Niedersachsen	Thüringen				
2007	Niedersachsen	alle anderen				

Anlage 3 – Abfrage der Daten der Ersatzkassen						
Jahr	WOP_KV	Arztsitz-KV	Summe des Leistungsbedarf in Punkte (LB_Pkt) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Punkten (LB_Pkt) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)	Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)
2007	Nordrhein	Bremen				
2007	Nordrhein	Niedersachsen				
2007	Nordrhein	Nordrhein				
2007	Nordrhein	Thüringen				
2007	Nordrhein	alle anderen				
2007	Thüringen	Bremen				
2007	Thüringen	Niedersachsen				
2007	Thüringen	Nordrhein				
2007	Thüringen	Thüringen				
2007	Thüringen	alle anderen				
2008	Bremen	Bremen				
2008	Bremen	Niedersachsen				
2008	Bremen	Nordrhein				
2008	Bremen	Thüringen				
2008	Bremen	alle anderen				

Anlage 3 – Abfrage der Daten der Ersatzkassen						
Jahr	WOP_KV	Arztstz-KV	Summe des Leistungsbedarf in Punkte (LB_Pkt) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Punkten (LB_Pkt) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)	Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV aller Arztgruppen	davon Summe des Leistungsbedarfs in Euro (LB_Eur) vor HVV von Ärzten ohne Kodierverpflichtung (Anlage 1)
2008	Niedersachsen	Bremen				
2008	Niedersachsen	Niedersachsen				
2008	Niedersachsen	Nordrhein				
2008	Niedersachsen	Thüringen				
2008	Niedersachsen	alle anderen				
2008	Nordrhein	Bremen				
2008	Nordrhein	Niedersachsen				
2008	Nordrhein	Nordrhein				
2008	Nordrhein	Thüringen				
2008	Nordrhein	alle anderen				
2008	Thüringen	Bremen				
2008	Thüringen	Niedersachsen				
2008	Thüringen	Nordrhein				
2008	Thüringen	Thüringen				
2008	Thüringen	alle anderen				

Tabelle 54: Datensatzbeschreibung zur „Abfrage der Daten der Ersatzkassen“

Anhang 4 Umfang der bilateralen Prüfungen zwischen Datengeber und Datenstelle

Diese Prüfungen wurden von der AG IT erarbeitet.

Satzart 90:

1. Prüfung Primärschlüssel
2. Vergleich frühere Lieferung
3. Anzahl Datensätze
4. Physische Person
5. Versichertenkarte je Person
6. Anzahl KV-Wechsler
7. KV-Rate KV-Wechsler

Satzart 91:

8. Prüfung Primärschlüssel
9. Vergleich frühere Lieferung
10. Anzahl Datensätze je Quartal
11. Verknüpfbarkeit Satzart 90 und Satzart 91
12. Versichertentage
13. Geschlecht
14. Verteilung nach Geschlecht
15. Verschiedene Geschlechter
16. Alter
17. Verschiedene Angaben zum Geburtsdatum
18. Verteilung nach Alter
19. RSA-Gruppe
20. Versichertenstatus
21. Nie im PLZ-Bereich
22. PLZ und Geburtstag

- 23. Verstorbene
- 24. Mehrfach Verstorbene
- 25. Verschiedene Angaben zu verstorben

Satzart 108:

- 26. Primärschlüsseleigenschaft
- 27. Lücken im Quartalsverlauf
- 28. Verknüpfbarkeit Satzart 90 und Satzart 92 vor/nach Umschlüsselung

Satzart 109:

- 29. Primärschlüsseleigenschaft
- 30. Vollständigkeit (alle KVK-IK aus Satzart 90 auch in Satzart 109)
- 31. Rückwärtsgerichtete Gültigkeitszeiträume
- 32. Kontinuität der Gültigkeitszeiträume

Satzart 110:

- 33. Prüfung Primärschlüssel
- 34. Überschneidungsfreiheit
- 35. Rückwärtsgerichtete Gültigkeitszeiträume
- 36. Vollständigkeit
- 37. Zeitliche Lücken
- 38. Rechtsnachfolger nicht angegeben
- 39. Rechtsnachfolger nicht vorhanden

Satzart 92:

- 40. Anzahl der Datensätze je Abrechnungsjahr
- 41. Anzahl der Datensätze je Abrechnungsquartal und KV
- 42. Anzahl ambulanter Datensätze je Abrechnungsjahr und KV

43. Anzahl ambulanter Datensätze je Abrechnungsquartal und KV
44. Anzahl stationärer Datensätze je Abrechnungsjahr und KV
45. Anzahl stationärer Datensätze je Abrechnungsquartal und KV
46. Amb. Punktzahl nach Abrechnungsjahr und KV
47. Amb. Punktzahl nach Abrechnungsquartal und KV
48. Stat. Punktzahl nach Abrechnungsjahr und KV
49. Stat. Punktzahl nach Abrechnungsquartal und KV
50. Amb. Kosten nach Abrechnungsjahr und KV
51. Amb. Kosten nach Abrechnungsquartal und KV
52. Stat. Kosten nach Abrechnungsjahr und KV
53. Stat. Kosten nach Abrechnungsquartal und KV
54. Summe Arzt-Patientenkontakte nach Abrechnungsjahr und KV
55. Summe Arzt-Patientenkontakte nach Abrechnungsquartal und KV
56. Fälle ohne Patientenbezug nach Abrechnungsjahr und KV

Satzart 93:

57. Anzahl Diagnosen je Abrechnungsjahr und KV
58. Anzahl Diagnosen je Abrechnungsquartal und KV
59. Anzahl Diagnosen mit Angabe Diagnosesicherheit je Abrechnungsjahr
60. Anzahl Diagnosen mit Angabe Diagnosesicherheit je Abrechnungsquartal
61. Anzahl Diagnosen mit Angabe Seitenlokation je Abrechnungsjahr
62. Anzahl Diagnosen mit Angabe Seitenlokation je Abrechnungsquartal
63. Diagnoseangaben ohne Patientenbezug
64. Die 50 häufigsten Diagnosen im Abrechnungsjahr
65. Die 50 häufigsten Diagnosen im Abrechnungsquartal

Satzart 100:

66. Anzahl der Datensätze je Abrechnungsjahr und KV

- 67. Anzahl der Datensätze je Abrechnungsquartal und KV
- 68. Punktzahl nach Abrechnungsjahr und KV
- 69. Punktzahl nach Abrechnungsquartal und KV
- 70. Kosten nach Abrechnungsjahr und KV
- 71. Kosten nach Abrechnungsquartal und KV
- 72. GNR ohne Fallbezug nach KV

Satzart 105:

- 73. Anzahl der Datensätze je Abrechnungsjahr und KV
- 74. Anzahl der Datensätze je Abrechnungsquartal und KV
- 75. Anzahl ambulanter Datensätze je Abrechnungsjahr und KV
- 76. Anzahl ambulanter Datensätze je Abrechnungsquartal und KV
- 77. Anzahl stationärer Datensätze je Abrechnungsjahr und KV
- 78. Anzahl stationärer Datensätze je Abrechnungsquartal und KV

Satzart 101:

- 79. Anzahl der Datensätze je Abrechnungsjahr und KV
- 80. Anzahl der Datensätze je Abrechnungsquartal und KV
- 81. Anzahl unterschiedliche Praxispseudonyme je Abrechnungsjahr und KV
- 82. Anzahl unterschiedliche Praxispseudonyme je Abrechnungsquartal und KV
- 83. Anzahl unterschiedliche Praxispseudonyme je Fachgruppe je Abrechnungsjahr und KV
- 84. Anzahl unterschiedliche Praxispseudonyme je Fachgruppe je Abrechnungsquartal und KV
- 85. Anzahl unterschiedliche Praxispseudonyme je Facharzt/Hausarztkennung je Abrechnungsjahr und KV
- 86. Anzahl unterschiedliche Praxispseudonyme je Facharzt/Hausarztkennung je Abrechnungsquartal und KV

- 87. Anzahl Ärzte je Abrechnungsjahr und KV
- 88. Anzahl Ärzte je Abrechnungsquartal und KV
- 89. Anzahl Ärzte je Fachgruppe Abrechnungsjahr und KV
- 90. Anzahl Ärzte je Fachgruppe, Abrechnungsquartal und KV

Satzart 102:

- 91. Anzahl der Datensätze je Abrechnungsjahr und KV
- 92. Anzahl der Datensätze je Abrechnungsquartal und KV
- 93. Anzahl unterschiedliche Arzt pseudonyme je Abrechnungsjahr und KV
- 94. Anzahl unterschiedliche Arzt pseudonyme je Abrechnungsquartal und KV
- 95. Anzahl Arzt pseudonyme je Fachgruppe je Abrechnungsjahr und KV
- 96. Anzahl Arzt pseudonyme je Fachgruppe je Abrechnungsquartal und KV
- 97. Anzahl unterschiedliche Praxis pseudonyme je Abrechnungsjahr und KV
- 98. Anzahl unterschiedliche Praxis pseudonyme je Abrechnungsquartal und KV
- 99. Anzahl unterschiedlicher Praxis pseudonyme je Fachgruppe, Abrechnungsjahr und KV
- 100. Anzahl unterschiedlicher Praxis pseudonyme je Fachgruppe, Abrechnungsquartal und KV

Anhang 5 Umfang der Plausibilitätsprüfungen von Datenstelle und AG IT

Diese Auswertungen wurden von der AG IT erarbeitet.

Auswertungen auf dem zusammengeführten Datenbestand:

1. Anzahl Fälle nach KV-Bereich und Quartal
2. Anzahl Fälle nach Kassenart/Kasse und Quartal
3. Anzahl Fälle nach Geschlecht und Quartal
4. Anzahl Fälle nach Versichertenstatus und Quartal
5. Anzahl Fälle nach Geburtsjahrgang und Quartal
6. Kosten (gesamt) nach KV-Bereich und Quartal
7. durchschnittliche Kosten je Versicherten nach KV-Bereich und Quartal (Minimum und Maximum)
8. Punkte (gesamt) nach KV-Bereich und Quartal
9. durchschnittliche Punkte je Versicherten nach KV-Bereich und Quartal (Minimum und Maximum)
10. Gesamtleistungsbedarf nach KV-Bereich und Quartal
11. Gesamtleistungsbedarf nach Kassenart/Kasse und Quartal
12. Gesamtleistungsbedarf nach Geschlecht und Quartal
13. Gesamtleistungsbedarf nach Geburtsjahrgang und Quartal
14. durchschnittlicher Leistungsbedarf je Versicherten nach KV-Bereich und Quartal
15. durchschnittlicher Leistungsbedarf je Versicherten nach Kassenart/Kasse und Quartal
16. durchschnittlicher Leistungsbedarf je Versicherten nach Geschlecht und Quartal
17. durchschnittlicher Leistungsbedarf je Versicherten nach Geburtsjahrgang der Versicherten und Quartal
18. Anzahl Versicherte nach Geschlecht und Quartal
19. Anzahl Versicherte nach Versichertenstatus und Quartal

20. Anzahl Versicherte mit Kennzeichen „verstorben“ nach Kassenart/Kasse und Quartal
21. durchschnittliche Anzahl der Versichertentage nach Kassenart/Kasse und Quartal (Minimum und Maximum)
22. durchschnittliche Anzahl der Versichertentage nach Geschlecht und Quartal (Minimum und Maximum)
23. durchschnittliche Anzahl der Arzt-/Patientenkontakte je Versicherten nach KV-Bereich und Quartal (Minimum und Maximum)
24. durchschnittliche Anzahl der Arzt-/Patientenkontakte je Versicherten nach Kassenart/ Kasse und Quartal (Minimum und Maximum)
25. durchschnittliche Anzahl der Arzt-/Patientenkontakte je Versicherten nach Geschlecht und Quartal (Minimum und Maximum)
26. durchschnittliche Anzahl der Arzt-/Patientenkontakte je Versicherten nach Geburtsjahrgang und Quartal (Minimum und Maximum)
27. Die häufigsten 50 Diagnosen (je Kennzeichen Diagnosesicherheit) mit Anzahlen nach KV-Bereich und Quartal
28. Die häufigsten 50 Diagnosen (je Kennzeichen Diagnosesicherheit) mit Anzahlen nach Geschlecht und Quartal
29. Die häufigsten 50 Gebührennummern mit Anzahlen nach KV-Bereich und Quartal
30. Die häufigsten 50 Gebührennummern mit Anzahlen nach Geschlecht und Quartal
31. Fälle ohne Leistungsbedarf
32. Fälle ohne Diagnoseangaben
33. Anzahl der Versicherten ohne Leistungsanspruchnahme nach KV-Bereich und Quartal
34. Anzahl der Versicherten ohne Leistungsanspruchnahme nach Kassenart/Kasse und Quartal
35. Anzahl Fälle ohne Versichertentage (nach Satzart 91) nach KV-Bereich und Quartal (ggf. letzten Datensatz verwenden)
36. Anzahl Fälle ohne Versichertentage nach Kassenart/Kasse und Quartal

Nicht-gematchter Datenbestand:

37. Anzahl Fälle nach KV-Bereich und Quartal
38. Anzahl Fälle nach Kassenart/Kasse und Quartal
39. Kosten (gesamt) nach KV-Bereich und Quartal
40. durchschnittliche Kosten je Versicherten nach KV-Bereich und Quartal (Minimum und Maximum)
41. Punkte (gesamt) nach KV-Bereich und Quartal
42. durchschnittliche Punkte je Versicherten nach KV-Bereich und Quartal (Minimum und Maximum)
43. Gesamtleistungsbedarf nach KV-Bereich und Quartal
44. Gesamtleistungsbedarf nach Kassenart/Kasse und Quartal
45. durchschnittlicher Leistungsbedarf je Versicherten nach KV-Bereich und Quartal
46. durchschnittlicher Leistungsbedarf je Versicherten nach Kassenart/Kasse und Quartal
47. durchschnittliche Anzahl der Arzt-/Patientenkontakte je Versicherten nach KV-Bereich und Quartal (Minimum und Maximum)
48. durchschnittliche Anzahl der Arzt-/Patientenkontakte je Versicherten nach Kassenart/Kasse und Quartal (Minimum und Maximum)
49. Die häufigsten 50 Gebührennummern mit Anzahlen nach KV-Bereich und Quartal
50. Fälle ohne Leistungsbedarf
51. Anzahl der Versicherten ohne Leistungsanspruchnahme nach KV-Bereich und Quartal
52. Anzahl der Versicherten ohne Leistungsanspruchnahme nach Kassenart/Kasse und Quartal

Anhang 6 Simulation zu Unterjährigkeit von Versicherten

A 6.1 Hintergrund und mögliche Auswirkungen

In der Datengrundlage des BA gibt es eine Reihe von Versicherten, die durch einen Wechsel der Personen ID innerhalb eines Jahres (z. B. durch Kassenwechsel) als zwei oder mehr Versicherte weiterverarbeitet werden. Diese Aufspaltung einer physischen Person in zwei oder mehrere Versicherte in der Datengrundlage kann dann problematisch werden, wenn die „geteilten“ Versicherten nicht alle Diagnosen des „kompletten“ Versicherten mitnehmen.

Dadurch kann es zu Verzerrungen der geschätzten Kosten- und Relativgewichte kommen (**Fehler 1**). Außerdem wird der anschließend mit den Kostengewichten geschätzte Leistungsbedarf \hat{y} durch die Gewichtung der Kostengewichte β mit den jeweiligen Versicherungszeitanteilen (*eligf* bzw. $(1-\textit{eligf})$) der „geteilten“ Versicherten zu gering ausfallen (**Fehler 2**).

$$\begin{aligned}\hat{y} &= \beta_1 + \beta_2 \\ \hat{y}_1 &= \textit{eligf} * \beta_1 + \textit{eligf} * \beta_2 \\ \hat{y}_2 &= (1 - \textit{eligf}) * \beta_1 + (1 - \textit{eligf}) * \beta_2 \\ \Rightarrow \hat{y} &= \hat{y}_1 + \hat{y}_2\end{aligned}$$

Beide „geteilten“ Versicherten haben dieselben Diagnosen wie der „komplette“ Versicherte

$$\begin{aligned}\hat{y} &= \beta_1 + \beta_2 \\ \hat{y}_1 &= \textit{eligf} * \beta_1 \\ \hat{y}_2 &= (1 - \textit{eligf}) * \beta_2 \\ \Rightarrow \hat{y} &> \hat{y}_1 + \hat{y}_2\end{aligned}$$

Die „geteilten“ Versicherten haben jeweils nur eine Diagnose des „kompletten“ Versicherten (**Fehler 2**)

A 6.2 Unterjährigkeit in der Datengrundlage (3-Mio-Stichprobe)

Jahr	Ø Anzahl Versichertentage	Minimale Anzahl Versichertentage	Maximale Anzahl Versichertentage
2007	344.47	1	365
2008	347.51	1	366

Tabelle 55: Durchschnittliche Anzahl Versichertentage

Perzentile	Versichertentage 2007	Versichertentage 2008
100 % (Maximum)	365	366
75 %	365	366
50 % (median)	365	366
25 %	365	365
10 %	302	335
5 %	181	184
1 %	59	60
0 % (Minimum)	1	1

Tabelle 56: Perzentile der Anzahl Versichertentage

A 6.3 Simulation der Auswirkungen von Unterjährigkeit auf die Regressionsergebnisse (Fehler 1)

A 6.3.1 Zeitgleiches Modell

Durchführung

Vergleich der Kostengewichte eines Modells inklusive aller Versicherten der 3-Mio-Stichprobe 2007 (Stichprobe „3-Mio-2007“) mit einem Modell, welches nur ganzjährig Versicherte (Versichertentage=365), Neugeborene und Verstorbene aus der 3-Mio-Stichprobe 2007 verwendet (Stichprobe „3-Mio-2007 ganzjährig“)

Verwendet wurde ein Modell der Variante F Version 0.2.2 zeitgleich mit Anwendung der kompletten Kalibrierungsstrategie ohne Hierarchisierung.

Die berechneten Kostengewichte werden zur Abschätzung des Fehlers 2 in Teil 3 weiterverwendet.

Ergebnisse

Stichprobe	N	Adjustiertes R ²	MAPE in Punkten (in %)	Standardabweichung des MAPE in Punkten
„3-Mio-2007“	3.000.000	56,82 %	5.025,48 (48,92 %)	18.314,39
„3-Mio-2007 ganzjährig“	2.657.475	57,23 %	5.069,04 (48,35 %)	19.026,54

Tabelle 57: Vergleich eines zeitgleichen Modells der Stichprobe „3-Mio-2007“ mit einem zeitgleichen Modell der Stichprobe „3-Mio-2007 ganzjährig“

Die Kostengewichte der beiden Modelle wurden zusätzlich mit einem t-Test für verbundene Stichproben miteinander verglichen. Zwei Stichproben heißen verbunden (auch: abhängig), wenn je zwei Messwerte (x_i , y_i) einander paarweise zugeordnet werden können und die Paare untereinander unabhängig sind. Für jedes $i=1, \dots, n$ sind die beiden Zufallsvariablen X_i und Y_i dagegen voneinander abhängig; die gemeinsame (bivariate) Verteilung von (X_i, Y_i) kann für jedes Paar verschieden sein. Der Test vergleicht die zwei Stichproben anhand des Mittelwerts der Differenzen zwischen den Messwertepaaren dieser zwei Stichproben.³⁰

Die Nullhypothese, dass die Differenz der Kostengewichte der beiden Modelle im Mittel gleich Null ist, konnte nicht abgelehnt werden.

Der in Tabelle 58 dargestellte Test umfasst nur den Vergleich der Kostengewichte von Risikokategorien, die in mindestens einem Modell enthalten sind. Der Einschluss von Risikokategorien, die in beiden Modellen ausgeschlossen wurden, würde dazu führen, dass eine Ablehnung der Nullhypothese unwahrscheinlicher wird und das Testergebnis somit verfälschen.

³⁰ Vgl. z.B. Mosler, Schmid (2008): Wahrscheinlichkeitsrechnung und schließende Statistik

Differenz der Kostengewichte (KG)	N	Mittelwert	t-Wert	PR > t
„KG 3-Mio-2007“ – „KG 3-Mio-2007 ganzjährig“	288	27,01	0,73	0,4659

Tabelle 58: Ergebnisse des t-Tests zum Vergleich der Kostengewichte

Abbildung 36 zeigt zusätzlich die Differenzen zwischen den Kostengewichten der einzelnen Alters- und Geschlechtsgruppen und CCs, wobei eine positive Differenz bedeutet, dass das Kostengewicht der Stichprobe „3-Mio-2007 ganzjährig“ größer ist als in der Stichprobe „3-Mio-2007“.

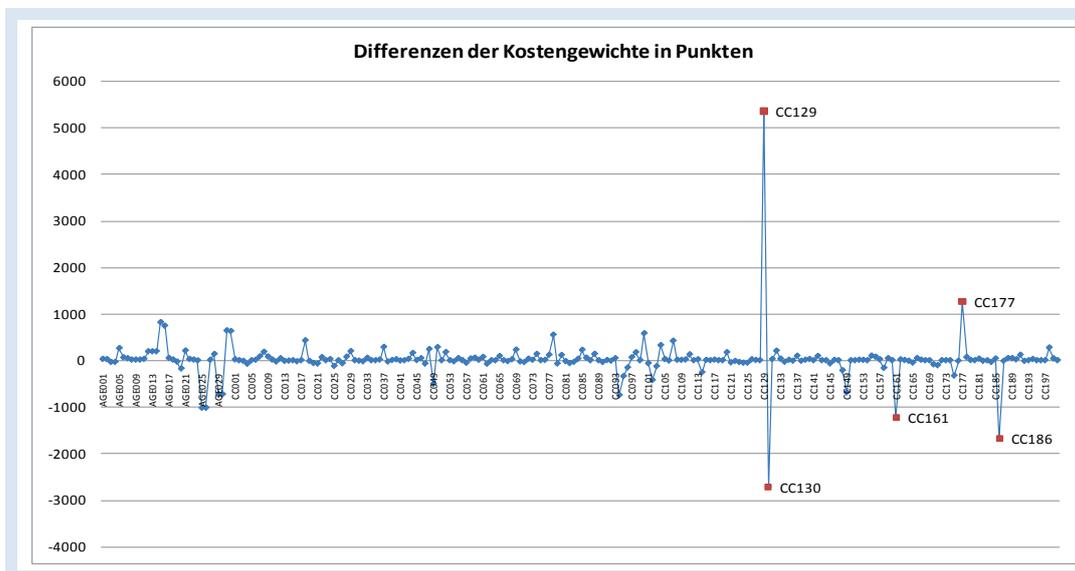


Abbildung 36: Differenzen der Kostengewichte zwischen den zeitgleichen Modellen in Punkten

Tabelle 59 enthält zusätzliche Informationen zu den Risikogruppen mit auffälligen Abweichungen im Kostengewicht zwischen den Modellen.

Risikokategorie	„KG 3-Mio-2007“ in Punkten	„KG 3-Mio-2007 ganzjährig“ in Punkten	„Häufigkeit 3-Mio-2007“	„Häufigkeit 3-Mio-2007 ganzjährig“
CC129 Terminale Niereninsuffizienz	286.889,57	281.524,18	5.962	5.866
CC130 Abhängigkeit von Dialyse	315.939,10	318.660,47	2.903	2.874
CC161 Traumatische Amputationen, Infektion des Stumpfes	0,00	1.227,43	3.088	2.923
CC177 Extremitätenverlust, Phantomschmerz	8.798,21	7.532,79	2.672	2.608
CC186 Multiples Myelom/Plasmozytom und andere chronische Leukämien	11.872,11	13.550,13	2.204	2.139

Tabelle 59: Risikokategorien mit auffälligen Abweichungen im Kostengewicht zwischen den beiden zeitgleichen Modellen

A 6.3.2 Prospektives Modell

Durchführung

Vergleich der Kostengewichte eines Modells, welches alle Versicherten der 3-Mio-Stichprobe 2008 verwendet, die auch in der Datengrundlage 2006 mit mindestens einem Versichertentag vorhanden sind (Stichprobe „3-Mio-2006/2008 prosp“) mit einem Modell, welches nur die Versicherten der 3-Mio-Stichprobe 2008 verwendet, die in 2006 ganzjährig versichert (Versichertentage=365) oder neugeboren waren (Stichprobe „3-Mio-2006/2008 prosp ganzjährig“).

Verwendet wurde ein Modell der Variante H Version 0.3.0 2-Jahres-prospektiv mit Anwendung der kompletten Kalibrierungsstrategie ohne Hierarchisierung.

Ergebnisse

Stichprobe	N	Adjustiertes R ²	MAPE in Punkten (in %)	Standardabweichung des MAPE in Punkten
„3 Mio-2006/2008 prosp.“	2.668.118	28,30%	5.731,13 (56,44 %)	8.875,23
„3 Mio-2006/2008 prosp. ganzjährig“	2.450.429	28,45 %	5.797,94 (56,82 %)	8.927,02

Tabelle 60: Vergleich eines prospektiven Modells der Stichprobe „3-Mio-2006/2008 prosp“ mit einem prospektiven Modell der Stichprobe „3-Mio-2006/2008 prosp ganzjährig“

Die Kostengewichte der beiden Modelle wurden zusätzlich mit einem t-Test für verbundene Stichproben miteinander verglichen. Zwei Stichproben heißen verbunden (auch: abhängig), wenn je zwei Messwerte (x_i , y_i) einander paarweise zugeordnet werden können und die Paare untereinander unabhängig sind. Für jedes $i=1, \dots, n$ sind die beiden Zufallsvariablen X_i und Y_i dagegen voneinander abhängig; die gemeinsame (bivariate) Verteilung von (X_i , Y_i) kann für jedes Paar verschieden sein. Der Test vergleicht die zwei Stichproben anhand des Mittelwerts der Differenzen zwischen den Messwertepaaren dieser zwei Stichproben.

Die Nullhypothese, dass die Differenz der Kostengewichte der beiden Modelle im Mittel gleich Null ist, konnte nicht abgelehnt werden.

Der in Tabelle 61 dargestellte Test umfasst nur den Vergleich der Kostengewichte von Risikokategorien, die in mindestens einem Modell enthalten sind. Der Einschluss von Risikokategorien, die in beiden Modellen ausgeschlossen wurden, würde dazu führen, dass eine Ablehnung der Nullhypothese unwahrscheinlicher wird und das Testergebnis somit verfälschen.

Differenz der Kostengewichte (KG)	N	Mittelwert	t-Wert	PR > t
„KG 3Mio-2006/2008 prosp.“ – „KG 3Mio-2006/2008 prosp. ganzjährig“	316	88,876	0,36	0,7162

Tabelle 61: Ergebnisse des t-Tests zum Vergleich der Kostengewichte

Abbildung 37 zeigt zusätzlich die Differenzen zwischen den Kostengewichten der einzelnen Alters- und Geschlechtsgruppen und CCs, wobei eine positive Differenz

bedeutet, dass das Kostengewicht der Stichprobe „3-Mio-2006/2008 prosp ganz-jährig“ größer ist als in der Stichprobe „3-Mio-2006/2008 prosp“.

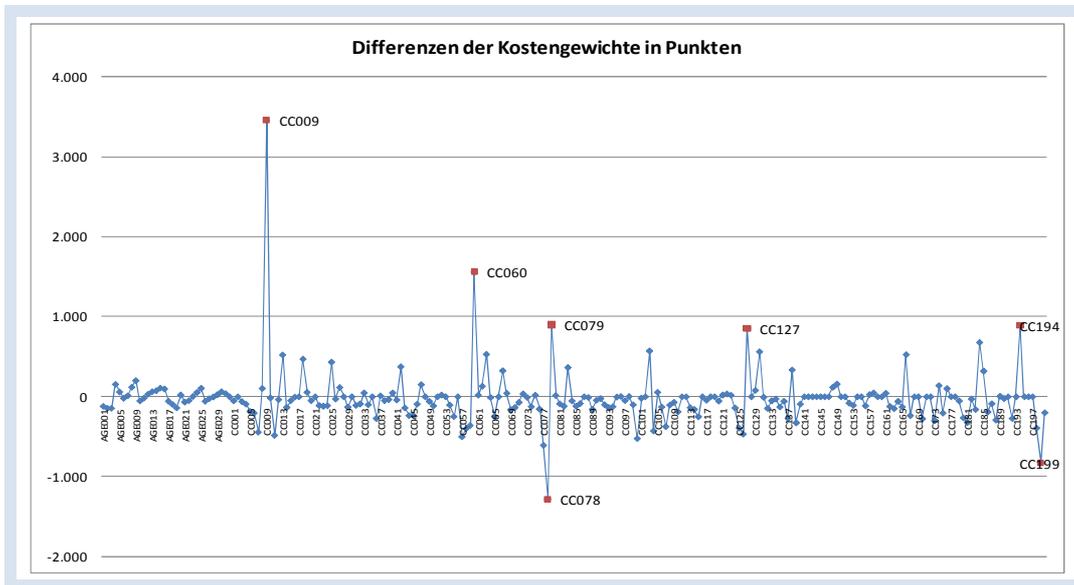


Abbildung 37: Differenzen der Kostengewichte zwischen den prospektiven Modellen in Punkten

Tabelle 62 enthält zusätzliche Informationen zu den Risikogruppen mit auffälligen Abweichungen im Kostengewicht zwischen den Modellen.

Risikokategorie	„KG 3-Mio-2006/2008 prosp“ in Punkten	„KG 3-Mio-2006/2008 prosp ganz-jährig“ in Punkten	„Häufigkeit 3-Mio-2006/2008 prosp“	„Häufigkeit 3-Mio-2006/2008 prosp ganz-jährig“
CC009 Bösartige Neubildungen Kopf-Hals-Region, ZNS, peripheres Nervensystem, Mediastinum, Lunge, Knochen, Gelenke, Bindegewebe, Uterus, Adnexe	3.460,96	0,00	8.051	7.798
CC060 Angststörungen, sexuelle Störungen, Belastungs- und Anpassungsstörungen	1.559,23	0,00	349.132	328.495
CC078 Atemstillstand	3.501,48	4.783,58	160	147
CC079 Herz-Kreislauf-Versagen, Schock, akute und chronische respiratorische Insuffizienz	899,40	0,00	13.804	13.327

Risikokategorie	„KG 3-Mio-2006/2008 prosp“ in Punkten	„KG 3-Mio-2006/2008 prosp ganz-jährig“ in Punkten	„Häufigkeit 3-Mio-2006/2008 prosp“	„Häufigkeit 3-Mio-2006/2008 prosp ganz-jährig“
CC127 Andere Erkrankungen HNO Bereich	851,01	0,00	1.015.278	948.569
CC194 Migräne und Kopfschmerzen weiterer Genesen	887,49	0,00	220.295	205.250
CC199 Umschriebene Entwicklungsstörung	0,00	825,32	57.553	53.318

Tabelle 62: Risikokategorien mit auffälligen Abweichungen im Kostengewicht zwischen den beiden prospektiven Modellen

A 6.3.3 Zusammenfassung

Die hinsichtlich der Kostengewichts-Differenzen auffälligen Risikokategorien unterscheiden sich für die zeitgleichen und prospektiven Modelle deutlich voneinander.

Um dies zu veranschaulichen, stellt Abbildung 38 die Differenzen der Kostengewichte der CCs der zeitgleichen und der prospektiven Modelle gemeinsam in einer Grafik dar.

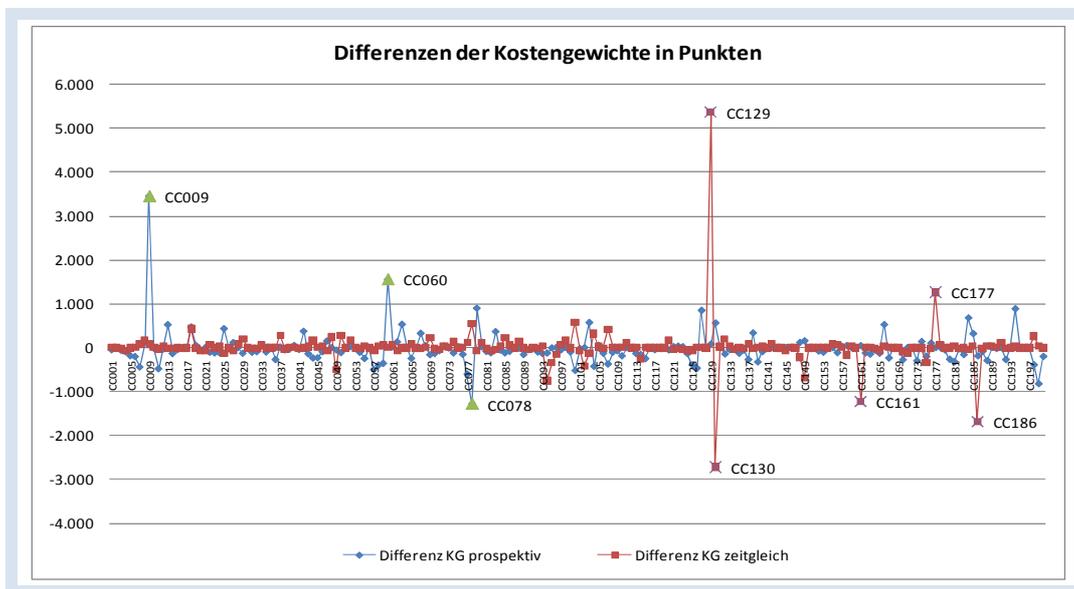


Abbildung 38: Differenzen der Kostengewichte zwischen den zeitgleichen und den prospektiven Modellen

Insgesamt zeigt sich, dass einzelne Kostengewichte deutliche Unterschiede zwischen den Stichproben aufweisen. Ein Einfluss auf die Veränderungsrate ist daher nicht auszuschließen.

A 6.4 Simulation der Auswirkungen von Unterjährigkeit auf die Berechnung des geschätzten LB / Morbiditätsindex (Fehler 2)

Durchführung

Für die Simulation wurden 500.000 ganzjährig Versicherte aus dem Jahr 2007 zufällig ausgewählt (Stichprobe „Simulation“) und in jeweils 2 Versicherte zerteilt (Stichprobe „Simulation Split“). Die Zerteilung der Versicherten erfolgte dabei jeweils im Übergang zwischen den vier Quartalen. Ein Versicherter wurde also in zwei Versicherte zerteilt, von denen entweder einer ein Quartal und einer drei Quartale versichert war oder von denen beide je zwei Quartale versichert waren.

Für beide Stichproben wird dann der geschätzte Leistungsbedarf je Versicherten berechnet. In Tabelle 63 anhand von Relativgewichten, die auf Basis der 3-Mio-Stichprobe 2007 („3-Mio-2007“) mit einem zeitgleichen kompletten Modell Variante F ohne Hierarchisierung berechnet wurden. In Tabelle 64 anhand von Relativgewichten, die auf Basis der 3-Mio-Stichprobe 2007 ohne unterjährig Versicherte („3-Mio-2007 ganzjährig“) mit einem zeitgleichen kompletten Modell Variante F ohne Hierarchisierung berechnet wurden.

Dieser geschätzte Leistungsbedarf wurde dann jeweils mit dem tatsächlichen Leistungsbedarf in der Stichprobe verglichen.

Ergebnisse

Stichprobe	Ø tatsächlicher LB in Punkten	Ø geschätzter LB in Punkten	PR (geschätzt / tatsächlich)
„3Mio-2007“	10.272,12	10.272,12	1,00000
„3-Mio-2007 ganzjährig“	10.483,40	10.574,87	1,00872
„Simulation“ (500 Tsd. ganzjährig Versicherte)	10.400,34	10.512,63	1,01080
„Simulation Split“	10.400,34	8.344,95	0,80237

Tabelle 63: Vergleich des geschätzten und tatsächlichen LB in den Stichproben mit Kostengewichten der Stichprobe „3-Mio-2007“

Stichprobe	Ø tatsächlicher LB in Punkten	Ø geschätzter LB in Punkten	PR (geschätzt / tatsächlich)
„3Mio-2007“	10.272,12	10.180,15	0,99105
„3-Mio-2007 ganzjährig“	10.483,40	10.483,40	1,00000
„Simulation“ (500 Tsd. ganzjährig Versicherte)	10.400,34	10.421,08	1,00199
„Simulation Split“	10400,34	8.265,30	0,79471

Tabelle 64: Vergleich des geschätzten und tatsächlichen LB in den Stichproben mit Kostengewichten der Stichprobe „3-Mio-2007 ganzjährig“

Anhang 7 Hierarchietabelle für reduzierte und komprimierte Modelle

Diese Hierarchietabelle wurde in der AG Grouperanpassung erarbeitet und konsentiert.

HCC / RCC über	HCC / RCC unter
HCC010	HCC014
HCC010	RCC003
HCC022	HCC024
HCC040	HCC043
HCC041	HCC043
HCC055	HCC057
HCC055	HCC058
HCC055	HCC059
HCC055	HCC060
HCC057	HCC058
HCC057	HCC059
HCC057	HCC060
HCC058	HCC059
HCC058	HCC060
HCC059	HCC060
HCC110	HCC115
HCC129	HCC131
HCC130	HCC131
HCC166	HCC167
HCC190	HCC043
RCC002	HCC014
RCC002	RCC003
RCC005	HCC024
RCC008	HCC043
RCC012	HCC059
RCC012	HCC060

Anhang 8 Nicht zur Diagnosekodierung verpflichtete Arztgruppen

Bezeichnung	Code
FA Laboratoriumsmedizin	110
FA Mikrobiologie und Infektionsepidemiologie; obsolet	111
TG Mikrobiologie; obsolet	112
FA Pathologie	170
TG Neuropathologie; obsolet	171
FA Pharmakologie und Toxikologie	180
TG Klinische Pharmakologie; obsolet	181
FA Radiologie (vor 1980); obsolet	190
TG Neuroradiologie; obsolet	192
FA Radiologische Diagnostik; obsolet	194
FA Diagnostische Radiologie; obsolet	195
SP Kinderradiologie	196
SP Neuroradiologie	197
FA Strahlentherapie und Radiologische Diagnostik; obsolet	199
FA Rechtsmedizin	240
FA Hygiene und Umweltmedizin	250
FA Neuropathologie	271
FA Klinische Pharmakologie	281
FA Anatomie	301
FA Biochemie	302
FA Transfusionsmedizin	303

Bezeichnung	Code
Medizinische Informatik	306
Pathologische Physiologie	307
FA Physiologie	308
FA Physiotherapie; neue Bundesländer; obsolet	309
FA Sozialhygiene; neue Bundesländer; obsolet	311
Physiologische Chemie	317
SP Echokardiologie herznaher Gefäße; obsolet	321
Biomathematik	350
Biophysik	351
Geschichte der Medizin	352
Industrietoxikologie	353
Klinische Strahlenphysik	354
Medizinische Wissenschaftsinformation	355
Medizinische Physik und Biophysik	356
FA Experimentelle und diagnostische Mikrobiologie; neue Bundesländer; obsolet	357
Fachbiologie der Medizin	358
Fachzahnarzt für Mikrobiologie	359
Fachzahnarzt für theoretisch-experimentelle Medizin	361
Fachwissenschaftler Chemie und Labordiagnostik	371
Fachwissenschaftler Genetik	372
Fachwissenschaftler Immunologie	373
Fachwissenschaftler Zytologie/Histologie	374

Bezeichnung	Code
FA Arbeitshygiene; neue Bundesländer; obsolet	500
FA Blutspende- und Transfusionsmedizin; kammerindividuell; obsolet	501
FA Hygiene; kammerindividuell; obsolet	502
FA Pathologische Anatomie; kammerindividuell; obsolet	506
FA Mikrobiologie, Virologie und Infektionsepidemiologie	532
FA Radiologie	535

Anhang 9 Ausschlussgründe und ihre Häufigkeiten

Zurzeit gelangen die folgenden Delete-Flags zum Einsatz:

- *D*: Es liegen Doubletten, d. h. mehrfach gelieferte Kombinationen von PERSONID und B_KASSNR vor.
- *N*: Die Kosten/Punktzahlen sind negativ.
- *G*: Der Versicherte bezieht Leistungen vor seiner Geburt.
- *4*: Es liegen mehr als vier Versichertenpseudonyme in Satzart 090 und mindestens ein Versichertentag in Satzart 091 vor.
- *S*: Die Kasse, bei der eine Person versichert ist, hat im aktuellen Quartal eine „Showquote“ von 0. Dies bedeutet, dass **alle** Versicherten der betreffenden Kasse keine ärztlichen Leistungen in Anspruch genommen haben und entspricht damit einem unrealistischen Sachverhalt.
- *UBS*: Versichertenstammdatensatz in Satzart 091 fehlt temporär (sog. Überspringer-Flag).
- *AGE*: Das Lebensalter ist negativ.
- *SEX*: Keine eindeutigen Angaben zum Geschlecht des Versicherten.
- *BIRTH*: Keine eindeutigen Angaben zum Geburtsdatum des Versicherten.
- *DUR*: Die Anzahl der dokumentierten Versichertentage ist zu niedrig (≤ 0 im Jahr) oder zu hoch (mehr als im Quartal möglich).
- *BEK*: Person_ID nicht immer eindeutig

	Verteilung der Ausschlussgründe in %		
	2006	2007	2008
D	16,6496	14,5687	14,2757
N	–	–	–
G	–	0,0017	–
4	2,8527	2,6637	2,7556
S	57,3212	61,5589	61,5376
UBS	19,7007	17,9534	17,9952
AGE	0,0006	0,0006	0,0006
SEX	1,0402	1,0491	1,2063
BIRTH	1,0384	1,0798	1,1910
DUR	–	–	–
BEK	0,1164	0,0880	0,1182
SEX, BIRTH	0,0208	0,0148	0,0196
S, UBS	0,0861	0,0744	0,0741
S, BIRTH	0,0095	0,0114	0,0141
S, SEX	0,0048	0,0062	0,0067
S, SEX, BIRTH	0,0006	0,0011	0,0012
4, UBS	0,0434	0,0414	0,0441
4, S	0,0077	0,0074	0,0080
4, BIRTH	0,0012	0,0011	0,0012
4, SEX	0,0012	0,0011	0,0018
4, BEK	0,0006	0,0006	0,0006
D, S	0,2875	0,2640	0,1904
D, 4	0,7663	0,5654	0,5058
D, 4, S	0,0154	0,0148	0,0159
D, BIRTH	0,0053	0,0051	0,0055
D, SEX	0,0065	0,0062	0,0067
D, SEX, BIRTH	0,0042	0,0040	0,0043
D, 4, BIRTH	0,0012	0,0011	0,0012
D, 4, SEX	0,0006	0,0006	0,0006
D, 4, SEX, BIRTH	0,0012	0,0011	0,0012
D, BEK	0,0018	0,0006	0,0012

	Verteilung der Ausschlussgründe in %		
	2006	2007	2008
UBS, BIRTH	0,0048	0,0040	0,0049
UBS, SEX	0,0089	0,0085	0,0092
G, BIRTH	0,0006	0,0011	0,0012
Gesamt in Personen	168.334	176.145	163.305
Gesamt in % der Brutto-Personenzahl	2,74	2,90	2,73

*Tabelle 65: Verteilung der Ausschlussgründe
(Quelle: Daten des BA; eigene Berechnungen und Darstellung)*

Anhang 10 Unterjährig Versicherte (Versichertenzeit < 356 Tage)

Anzahl Versicherte	2006	2007	2008
Absolut			
Gesamt	5.968.440	5.890.005	5.823.029
(in VJ)	(5.602.963,7)	(5.559.555,6)	(5.529.294,7)
• Ganzjährig Versicherte	5.172.183	5.165.956	5.183.062
• Unterjährig Versicherte	796.257	724.049	639.967
(in VJ)	(431.301,3)	(394.085,7)	(357.865,3)
Prozentual			
• Ganzjährig Versicherte	86,659 %	87,707 %	89,010 %
• Unterjährig Versicherte	13,341 %	12,293 %	10,990 %
(in VJ)	(7,226 %)	(6,691 %)	(6,146 %)
1. Geburten	0,704 %	0,697 %	0,714 %
2. Sterbefälle	0,957 %	0,945 %	1,014 %
3. Sterbefälle im 1. Lebensj.	0,001 %	0,002 %	0,001 %
4. Sonstige	11,678 %	10,650 %	9,261 %

Tabelle 66: Verteilung der unterjährig Versicherten (Versichertenzeit < 356 Tage)
(Quelle: Daten des BA; eigene Berechnungen und Darstellung)

Anhang 11 Illustration des Problems fehlender Versichertendaten

In den Daten der Krankenkassen ist zu beobachten, dass für einzelne Kassen die Versichertenstammdaten unvollständig übermittelt werden. Als Beispiel sei hier eine Kasse aufgeführt. Abbildung 39 zeigt die quartalsweise übermittelte Anzahl der Versichertenstammdatensätzen.

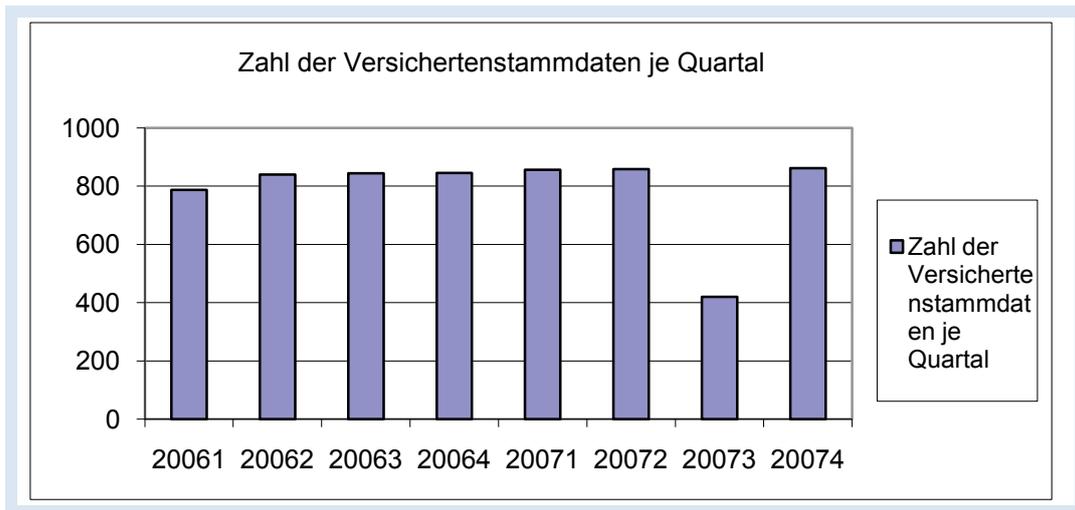


Abbildung 39: Anzahl der quartalsweise übermittelten Versichertenstammdatensätzen am Beispiel einer Kasse

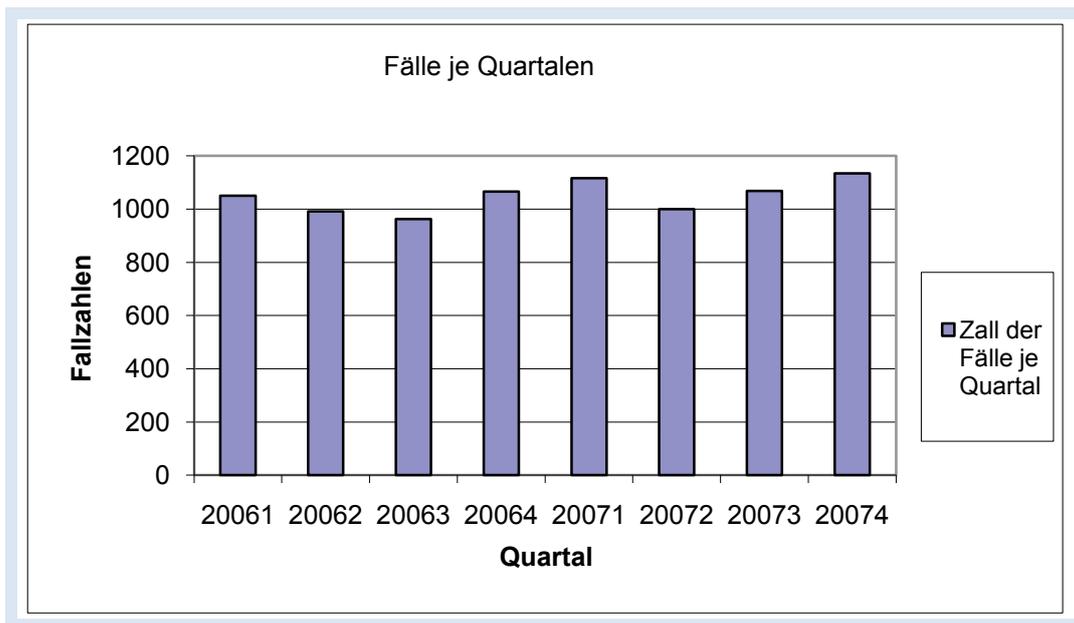


Abbildung 40: Zahl der Abrechnungsfälle je Quartal am Beispiel einer Kasse

Während die Versichertenstammdaten offensichtlich unvollständig übermittelt wurden, zeigt die Zahl der quartalsweise ermittelten Abrechnungsfälle keine Lücken (Abbildung 40). Bei der weiteren Nutzung der Daten ist davon auszugehen, dass die Versichertenzeiten dieser Krankenkasse unterrepräsentiert sind.

Das Beispiel steht für eine Vielzahl – oft kleinerer – Kassen, die Befunde setzen sich auch ins Jahr 2008 fort.

Anhang 12 Verarbeitung von unterjährigen Versicherten

Die Modellierung des geschätzten Leistungsbedarfs erfolgt in mehreren Schritten, wobei bei unterjährig Versicherten (zu den Ursachen der Unterjährigkeit vgl. Kapitel 2.1.4) Verzerrungen entstehen können, die im Folgenden erläutert werden.

IDNO	Q/J	DIAG	LB	VZ	
4711	I/2007	ICD1	}	GOP1	90
		ICD2		GOP2	
		ICD3		GOP3	
		ICD4		GOP4	
4711	II/2007	ICD5	}	GOP5	91
		ICD6		GOP6	
		ICD7		GOP7	
0815	III/2007	ICD8	}	GOP8	90
		ICD9		GOP9	
		ICD10		GOP10	
0815	IV/2007	ICD11	}	GOP11	91
		ICD12		GOP12	
		ICD13		GOP13	
		ICD14		GOP14	

Abbildung 41: Unterjährige Versicherte

In Abbildung 41 ist der Versicherte mit der Versichertennummer 4711 lediglich die Quartale I/2007 und II/2007 versichert. Er weist in den ersten beiden Quartalen die Diagnosen 1 bis 7 mit den entsprechenden GOPs auf. Der Versicherte ist ein Kassenwechsler, der im dritten und vierten Quartal 2007 bei einer anderen Kasse die Versichertennummer 0815 zugewiesen bekommt. In den entsprechenden Quartalen weist er die Diagnosen 8 bis 14 auf. Bei der Verarbeitung im Grouper bekommt dieser fiktive Versicherte aufgrund seiner Diagnosen in den ersten beiden Quartalen die CC007 zugewiesen und die Diagnosen der zweiten Quartale des Jahres 2007 führen in die CC117.

Für sich allein unproblematisch ist die geteilte Versicherungszeit – der Versicherte erhält unter seinen beiden Versichertennummern jeweils hälftige Jahresbeträge; bei z. B. chronischen Erkrankungen sind die hälftigen Zuschläge gleich groß und summieren sich korrekterweise zum Jahresbetrag.

Der Sachverhalt ist dagegen problematisch, wenn sich die Risikoklassen zwischen den Halbjahren unterscheiden.

Ein Versicherter, der ganzjährig versichert ist und in der ersten Jahreshälfte Diagnosen aufweist, die in die CC007 führen und in der zweiten Jahreshälfte Diagnosen, die die CC117 triggern, bekommt beide Risikozuschläge voll „angerechnet“.

Der aufgrund der technischen Gegebenheiten geteilte Versicherte (zwei Versichertennummern aufgrund von Kassenwechsel) hingegen erhält CC007 im ersten Halbjahr ($\frac{1}{2}$ Risikozuschlag) und CC117 im zweiten Halbjahr (ebenfalls nur ein halber Risikozuschlag).

Jede Veränderung im Anteil der technisch aufgeteilten Personen kann daher zu leichten Verzerrungen in der Veränderungsrate führen, in prospektiver Modellierung tendenziell schwächer als in zeitgleicher.

Anhang 13 Aufteilung der Altersgruppen in der KM6

Diese Aufteilung der Altersgruppen entspricht der Aufteilung der durch das BMG gewählten Altersgruppenaufteilung in den KM6-Daten.

- Bis unter 15 Jahre
- 15 bis unter 20 Jahre
- 20 bis unter 25 Jahre
- 25 bis unter 30 Jahre
- 30 bis unter 35 Jahre
- 35 bis unter 40 Jahre
- 40 bis unter 45 Jahre
- 45 bis unter 50 Jahre
- 50 bis unter 55 Jahre
- 55 bis unter 60 Jahre
- 60 bis unter 65 Jahre
- 65 bis unter 70 Jahre
- 70 bis unter 75 Jahre
- 75 bis unter 80 Jahre
- 80 bis unter 85 Jahre
- 85 bis unter 90 Jahre
- 90 und mehr Jahre

Anhang 14 Vergleich der Altersstruktur in Versichertenstichprobe und KM6 (2006 und 2007)

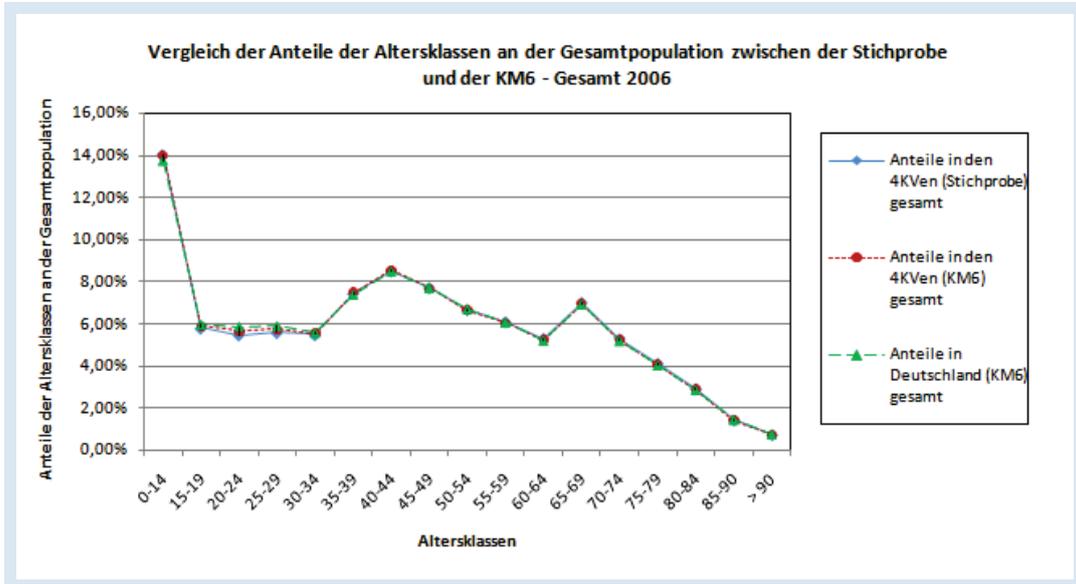


Abbildung 42: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Gesamte Bevölkerung – 2006
 (Quelle: Daten des statistischen Bundesamtes und des BA; eigene Berechnungen und Darstellung)

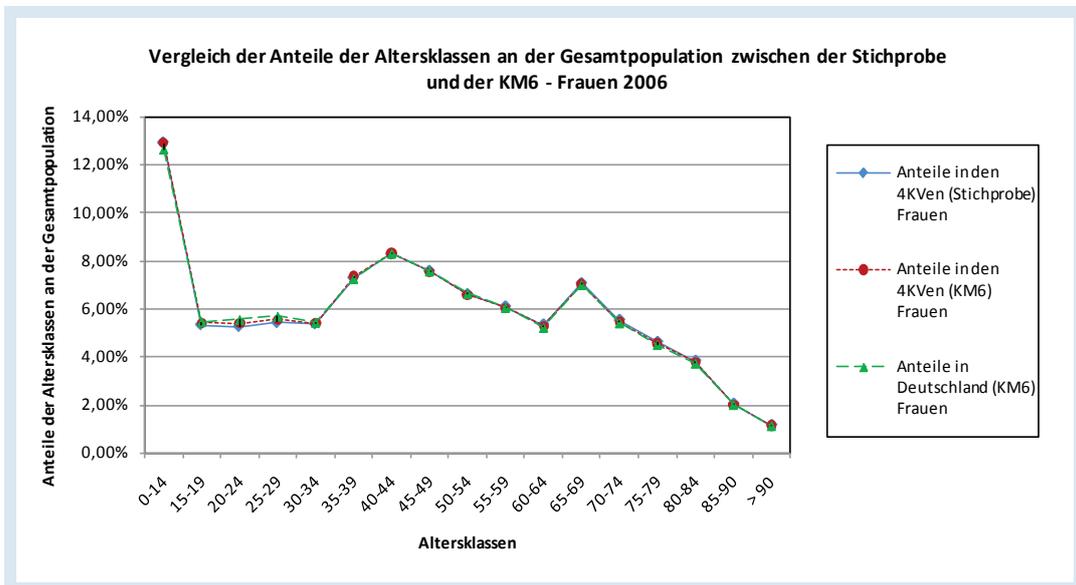


Abbildung 43: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Frauen – 2006
 (Quelle: Daten des statistischen Bundesamtes und des BA; eigene Berechnungen und Darstellung)

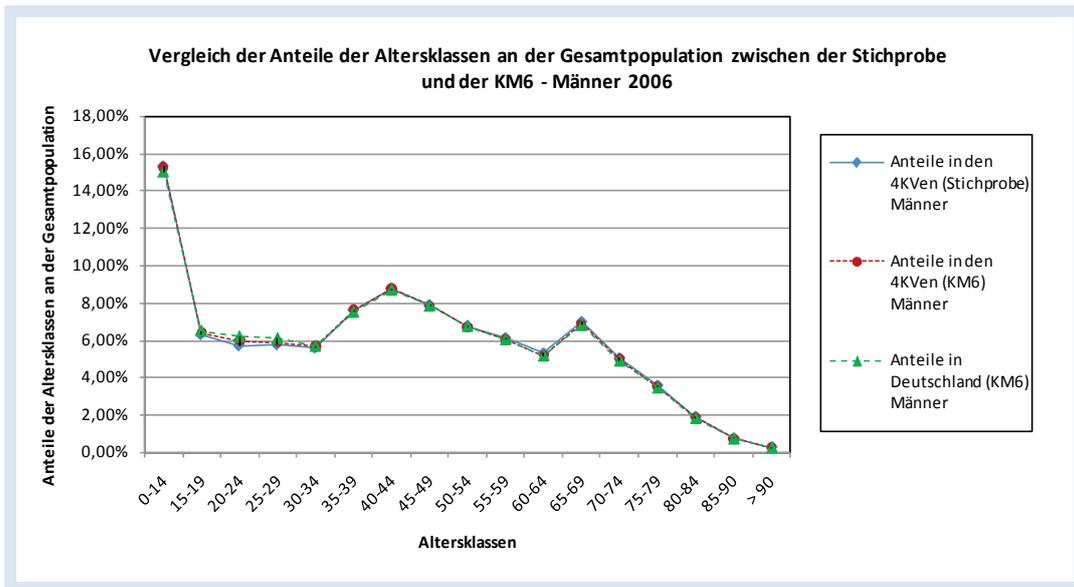


Abbildung 44: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Männer – 2006
(Quelle: Daten des statistischen Bundesamtes und des BA; eigene Berechnungen und Darstellung)

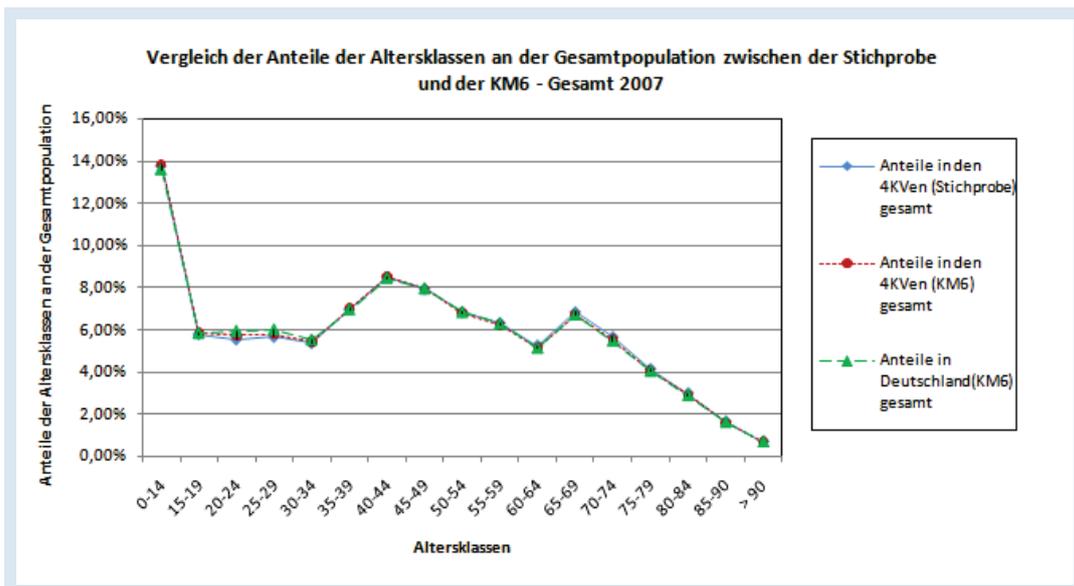


Abbildung 45: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Gesamte Bevölkerung – 2007
(Quelle: Daten des statistischen Bundesamtes und des BA; eigene Berechnungen und Darstellung)

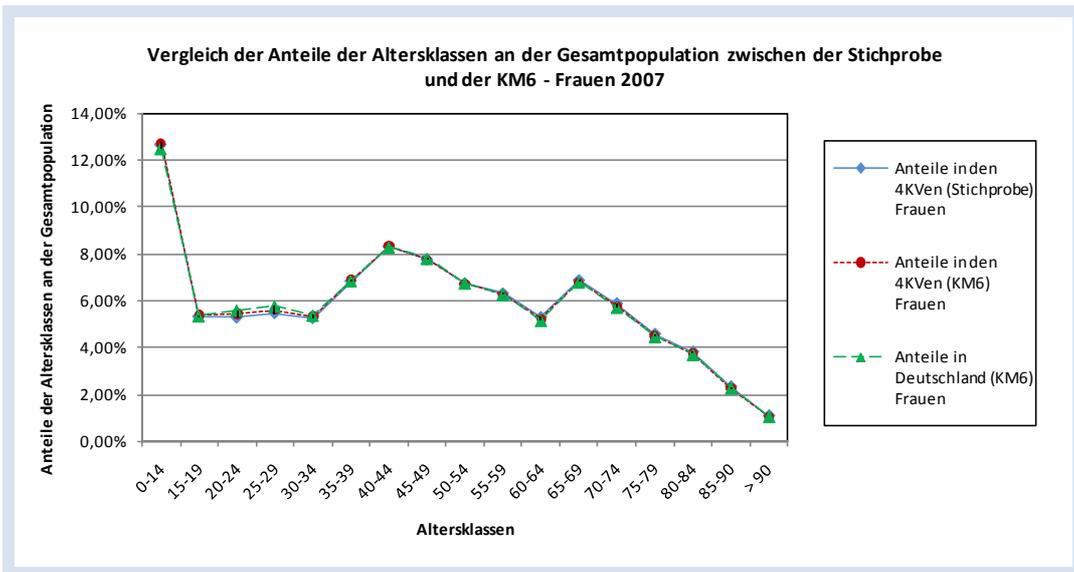


Abbildung 46: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Frauen – 2007
(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

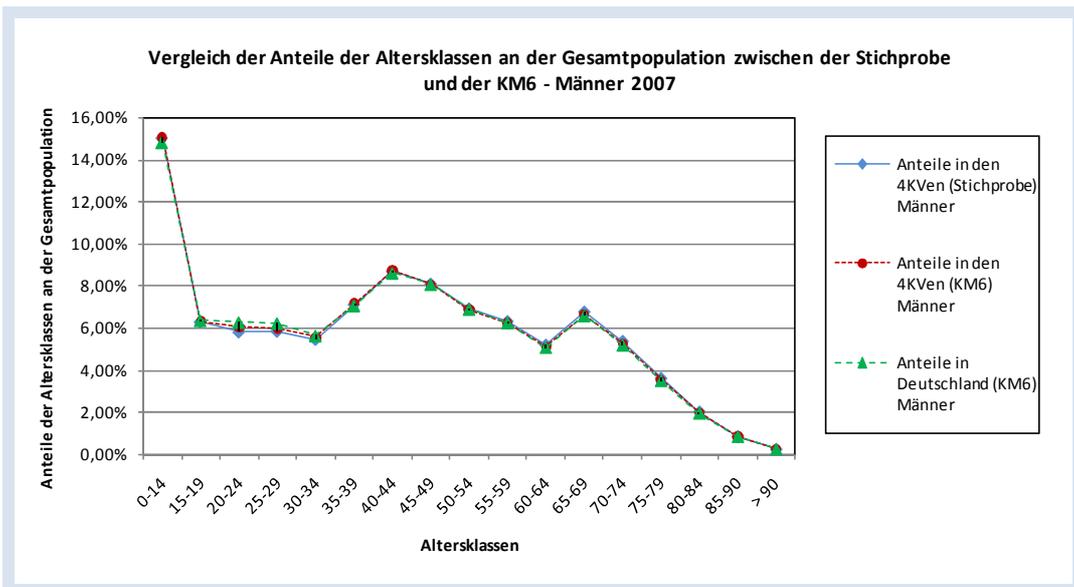


Abbildung 47: Vergleich der Altersstruktur in Versichertenstichprobe und KM6 – Männer – 2007
(Quelle: KM6 2006, 2007, 2008 und Daten des BA; eigene Berechnungen und Darstellung)

Anhang 15 Externer Vergleich der Relativgewichte bei einer Klassifikation nach Alter, Geschlecht, EM-Status

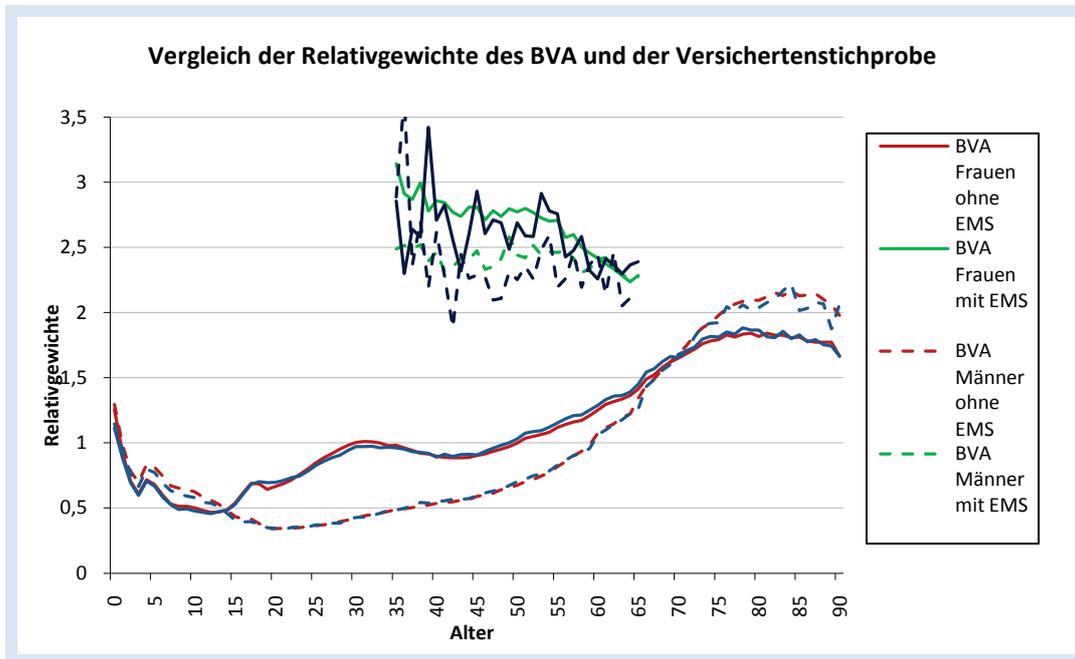


Abbildung 48: Vergleich der Relativgewichte der Versichertenstichprobe und des BVA nach EM-Status

(Quelle: Daten des BVA, Daten des BA; eigene Berechnungen und Darstellung)

Anhang 16 Weitere Auswertungen zum Kapitel „Fremdkassen-zahlungsausgleich“

WOP_KV	Anteil des Leistungsbedarfs außerhalb der Datengrundlage (alle Kassen)	
	2007	2008
Bremen	2,29 %	2,64 %
Niedersachsen	4,54 %	4,77 %
Nordrhein	3,22 %	3,31 %
Thüringen	5,44 %	5,14 %

Tabelle 67: Anteil des Leistungsbedarfs außerhalb der Datengrundlage am WOP-Leistungsbedarf je KV (Quelle: Sonderabfrage durch die KBV)

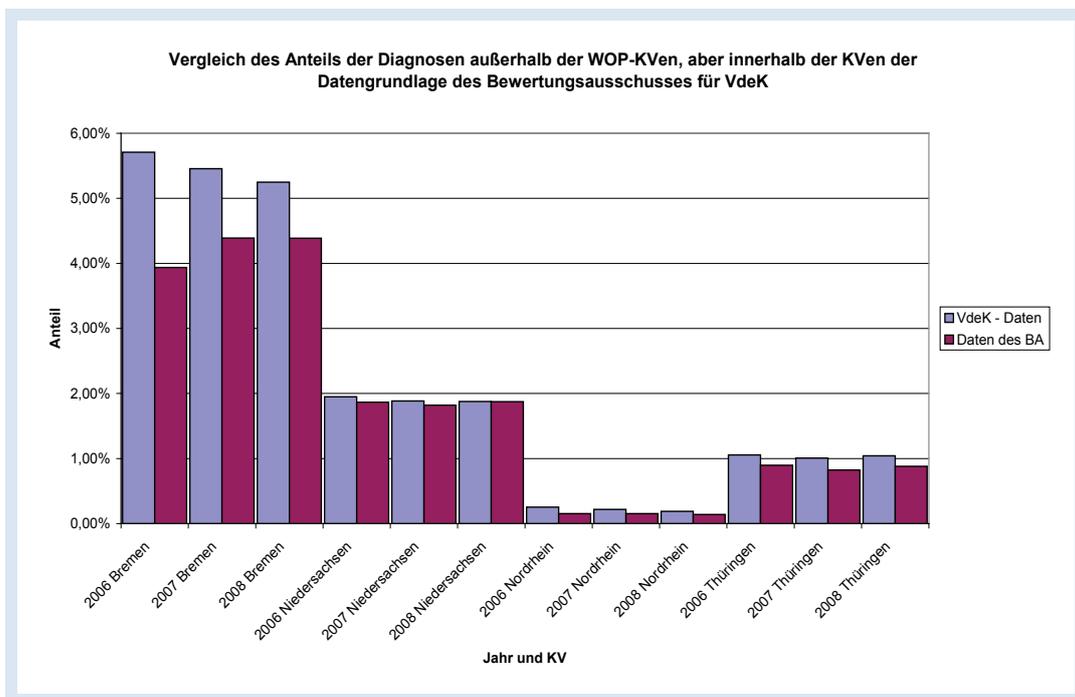
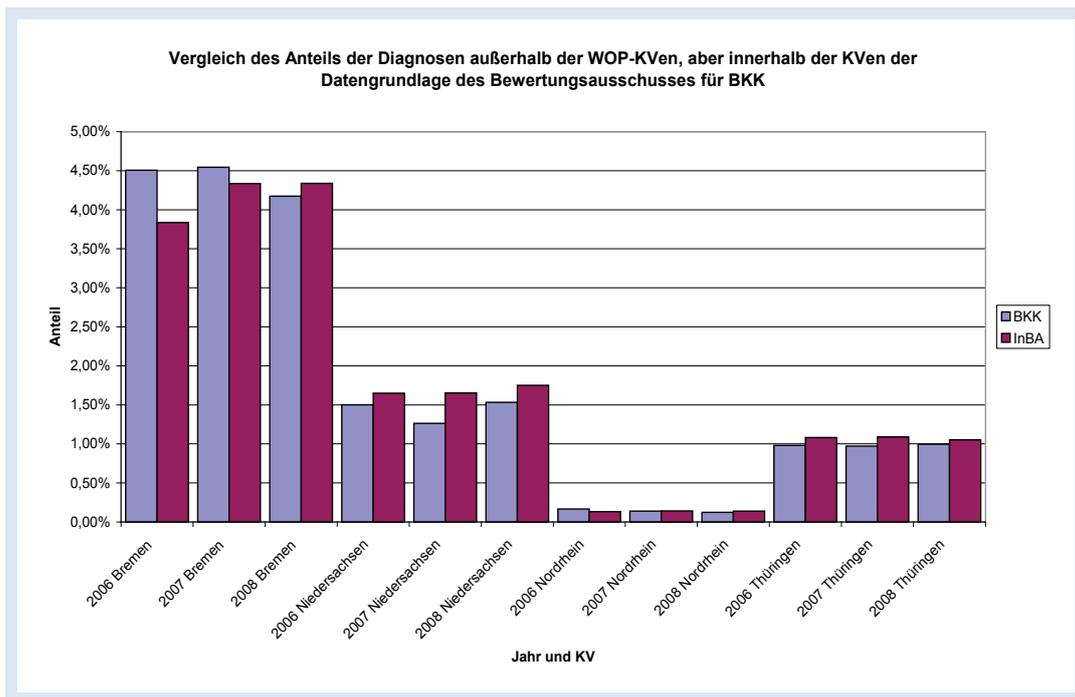


Abbildung 49: Vergleich des Anteils der Diagnosen aus den drei nicht-WOP-KV der Datengrundlage zur Summe der WOP-KV für vdek (gemäß Datengrundlage BA) (Quelle: Daten des Bewertungsausschusses und Sonderabfrage durch die Ersatzkassen)



*Abbildung 50: Vergleich des Anteils der Diagnosen aus den drei nicht-WOP-KV der Datengrundlage zur Summe der WOP-KV für BKK (gemäß Datengrundlage BA)
(Quelle: Daten des Bewertungsausschusses und Sonderabfrage durch die Betriebskrankenkassen)*

Anhang 17 Überblick über die Erarbeitung der Varianten F und G

Der Bewertungsausschuss hatte im Jahr 2004 ein Klassifikationssystem erworben, welches die komplette Versorgung von Versicherten abbildet. Um eine größere Passgenauigkeit für die ambulante deutsche Versorgung herzustellen, wurden durch das Institut des Bewertungsausschusses Änderungsvorschläge für die Klassifikation erarbeitet, die dann mit der Arbeitsebene des Bewertungsausschusses (AG Grouperanpassung und AG medizinische Grouperanpassung) konsentiert wurden.

Die Auswahl der Risikokategorien, welche auf Passgenauigkeit für die ambulante Versorgung in Deutschland überprüft wurden, erfolgte aufgrund der ökonomischen Relevanz der Risikokategorien. Bei der Überprüfung wurden siebzehn Risikokategorien und damit ca. 3.000 ICD-Kodes überprüft. Dieses entspricht ca. 40 % der ökonomischen Relevanz und 20 % aller ICD-Kodes.

Der Konsentierungsprozess führte zu den folgenden Anpassungen:

- 1 Kode erhielt eine Änderung beim oberen Alter
- 15 Kodes erhielten Änderungen beim unteren Alter
- 296 Kodes mit Änderung der Zuordnung zur DxGruppe (Änderung der Zuordnung zur DxGruppe bzw. Überführung in neu geschaffene DxGruppen)
- Es kam zu keiner Löschung oder Zusammenführung von DxGruppen
- Die Variante F arbeitet mit 763, die Variante G mit 766 durch Kodes angesteuerte („nicht leere“) DxGruppen
- 340 Kodes mit Änderung der Zuordnung zur CC
- Die Variante F arbeitet mit 195, die Variante G mit 198 durch Kodes angesteuerte („nicht leere“) CC.

Anhang 18 Kalibrierungsstrategie für hierarchische komplette Modelle

Diese Kalibrierungsstrategie wurde von der KBV erarbeitet.

Regression ohne Intercept

Nicht modifiziertes R^2 wird ausgerechnet

1. Schätzung des vollen Regressionsmodells: Kitchen Sink

```
proc reg data = regress;
```

```
    model depend = AG001-AG032
```

```
        HCC001-HCC200
```

```
        KHCC001-KHCC200
```

```
        OHCC001-OHCC200
```

```
        ZHCC001-ZHCC029 / noint;
```

```
run;
```

2. Ausschluss von Interaktionen (KHCC's, OHCC's, ZHCC's) mit einer Prävalenz < 500

Erneute Schätzung des Regressionsmodells.

3. Ausschluss von Interaktionen (KHCC's, OHCC's, ZHCC's) mit $Pr \geq 0.05$, stepwise regression.

Erneute Schätzung des Regressionsmodells.

4. Zusammenfassen bzw. Ausschluss von AG-Gruppen mit negativen und/oder nicht signifikanten³¹ Kostengewichten. Die AG-Gruppen sollen möglichst nicht ausgeschlossen, sondern zu größeren AG-Gruppen zusammengefasst werden.

Zusammenfassung nur von benachbarten AG-Gruppen.

Zusammenfassung/Ausschluss möglichst symmetrisch für beide Geschlechtsgruppen

Aufnahme von NoHCC in die Regressionsgleichung, wenn mindestens eine AG-Gruppe ausgeschlossen wurde.

Erneute Schätzung des Regressionsmodells.

³¹ Signifikanz: $Pr < 0.05$

Schritt 4 wird bei Bedarf mehrmals wiederholt.

5. Ausschluss von einzelnen HCC's mit negativen β -Kostengewichten, die in der jeweils niedrigsten Hierarchiestufe liegen oder keiner Hierarchisierung hinterlegen. Es ist zuerst aber zu überprüfen, ob die negativen Kostengewichte der HCCs in der jeweils niedrigsten Hierarchiestufe durch das Zusammenfassen (constraints) mit HCCs in der jeweils höheren (nächstniedrigeren) Hierarchiestufe sich korrigieren lassen.
Beim Ausschluss von einer HCC wird die entsprechende KHCC und OHCC auch ausgeschlossen.

Erneute Schätzung des Regressionsmodells.

Schritt 5 wird bei Bedarf mehrmals wiederholt.

6. Zusammenfassen von einzelnen HCCs mit negativen oder nicht signifikanten Kostengewichten, die nicht in der jeweils niedrigsten Hierarchiestufe liegen. Falls trotz Zusammenfassens der HCCs negative oder nicht signifikante Kostengewichte im Modell bleiben, werden diese HCCs ausgeschlossen.
Beim Ausschluss von einer HCC wird die entsprechende KHCC und OHCC auch ausgeschlossen.

Erneute Schätzung des Regressionsmodells.

Schritt 6 wird bei Bedarf mehrmals wiederholt.

7. Zusammenfassen von HCCs bei Verletzung der Hierarchien.
Nur diejenigen Verletzungen werden korrigiert, bei denen die Differenz der Kostengewichte kleiner als ca. 5 % des durchschnittlichen Leistungsbedarfs beträgt. Andere Verletzungen bleiben im Modell – keine strenge Monotonie.

Erneute Schätzung des Regressionsmodells.

Schritt 7 wird bei Bedarf mehrmals wiederholt.

8. Zusammenfassen bzw. Ausschluss von Interaktionen (KHCCs, OHCCs) mit nicht signifikanten Kostengewichten und/oder aufgrund negativer Nettoeffekte ($\text{Net Kids} = \beta\text{CC} + \beta\text{KCC} \geq 0$, $\text{Net Old} = \beta\text{CC} + \beta\text{OCC} \geq 0$)

Erneute Schätzung des Regressionsmodells.

Schritt 8 wird bei Bedarf mehrmals wiederholt.

9. Zusammenfassen bzw. Ausschluss von Interaktionen (KHCCs, OHCCs) bei Verletzung der Hierarchien von Nettoeffekten
Falls die Hierarchieverletzungen der HCCs nicht korrigiert wurden (sind im Modell geblieben), werden die entsprechenden Hierarchieverletzungen von Nettoeffekten auch vernachlässigt, somit werden die entsprechenden KHCCs

und OHCCs nicht zusammengefasst oder ausgeschlossen und bleiben im Modell.

Erneute Schätzung des Regressionsmodells.

Schritt 9 wird bei Bedarf mehrmals wiederholt.

10. Erneute Überprüfung von allen Koeffizienten auf die Signifikanz. Bei Bedarf Wiederholung der Schritte 4-6 und 8.

Erneute Schätzung des Regressionsmodells.

Schritt 10 wird bei Bedarf mehrmals wiederholt.

11. Erneute Überprüfung der hierarchischen Struktur der HCC-Kostengewichte und der Nettoeffekte. Bei Bedarf Wiederholung der Schritte 7 und 9. Erneute Schätzung des Regressionsmodells.

Schritt 11 wird bei Bedarf mehrmals wiederholt.

12. Ausschluss von Krankheitsinteraktionen (ZHCC´s) mit nicht signifikanten Kostengewichten

Erneute Schätzung des Regressionsmodells.

Schritt 12 wird bei Bedarf mehrmals wiederholt.

13. Ausschluss von Krankheitsinteraktionen (ZHCC´s) falls das Kostengewicht negativ ist und sein Absolutbetrag größer als das kleinste HCC-Kostengewicht in der Definition der jeweiligen Interaktion ist.

Erneute Schätzung des Regressionsmodells.

Schritt 13 wird bei Bedarf mehrmals wiederholt.

14. Ausschluss von NoHCC falls das Kostengewicht negativ und/oder nicht signifikant ist.

Erneute Schätzung des Regressionsmodells.

Schritte 10 – 13 werden noch mal wiederholt.

Zusammenfassen von Kostengewichten:

- Zusammenfassung erfolgt mittels „constraints“. Als Folge der Zusammenfassung erhalten die zusammengefassten Variablen (AG-Gruppen, HCCs, KHCCs, OHCCs) den gleichen Betrag für die Kostengewichte.
- Zusammenfassung nur von benachbarten AG-Gruppen.
- Zusammenfassung nur von benachbarten HCCs innerhalb eines Hierarchiebildes.

Falls mehrere Hierarchieverletzungen innerhalb eines Hierarchiebildes existieren, entscheidet man nach der Ähnlichkeit der Kostengewichte welche und wie viele HCCs zusammengefasst werden. Dabei können mehrere constraints mit jeweils zwei oder mehr HCCs durchgeführt werden.

Anhang 19 Kalibrierungsstrategie für nicht hierarchische komplette Modelle³²

Diese Kalibrierungsstrategie wurde von der KBV erarbeitet.

Regression ohne Interzept

Nicht modifiziertes R^2

1. Kitchen Sink: ADEPEND = AG001-AG032 CC001-CC200 KCC001-KCC200 OCC001-OCC200 (ohne ZCCs)
2. Ausschluss von Interaktionen (KCCs / OCCs) mit einer Häufigkeit < 500
3. Ausschluss von Interaktionen (KCCs / OCCs) mit $Pr \geq 0,05$ (stepwise regression)
Schätzung des Regressionsmodells
4. Zusammenfassen bzw. Ausschluss von AG-Gruppen mit negativen und / oder nicht signifikanten Kostengewichten ($Pr \geq 0,05$)
symmetrisch für beide Geschlechtsgruppen
Aufnahme von NOCC, falls AG-Gruppen ausgeschlossen wurden
Schätzung des Regressionsmodells
Schritt bei Bedarf wiederholen
5. Ausschluss von CCs mit negativen Kostengewichten, beginnend mit dem niedrigsten Kostengewicht je ACC
beim Ausschluss einer CC wird die entsprechende KCC und OCC auch ausgeschlossen
ggf. NOCC neu erzeugen (künstliche Aufstockung)
Schätzung des Regressionsmodells
Schritt bei Bedarf wiederholen
6. Ausschluss von CCs mit nicht signifikanten Kostengewichten ($Pr \geq 0,05$)
beim Ausschluss einer CC wird die entsprechende KCC und OCC auch ausgeschlossen
ggf. NOCC neu erzeugen (künstliche Aufstockung)
Schätzung des Regressionsmodells
Schritt bei Bedarf wiederholen

³² Die KBV verwendet hier den allgemeinen Begriff „Koeffizienten“ statt dem im Bericht verwendeten Begriff „Kostengewichte“.

7. Ausschluss von Interaktionen (KCCs / OCCs) mit nicht signifikanten Kostengewichten ($Pr \geq 0,05$) und / oder negativen Nettoeffekten ($Net\ Kids = \beta_{CC} + \beta_{KCC} \geq 0$; $Net\ Old = \beta_{CC} + \beta_{OCC} \geq 0$)
ggf. NOCC neu erzeugen (künstliche Aufstockung)
Schätzung des Regressionsmodells
Schritt bei Bedarf wiederholen
8. Überprüfung von allen Kostengewichten auf Signifikanz
bei Bedarf Wiederholung der Schritte 4-7
Schätzung des Regressionsmodells
Schritt bei Bedarf wiederholen
9. ggf. Ausschluss von NOCC, falls das Kostengewicht negativ und / oder nicht signifikant ($Pr \geq 0,05$) ist
Schätzung des Regressionsmodells
Schritt 8 wiederholen

Anhang 20 Kalibrierungsstrategie für reduzierte und komprimierte Modelle

Diese Kalibrierungsstrategie wurde vom GKV-Spitzenverband erarbeitet.

Regression ohne Interzept

1. Ausschluss von Risikoklassen
CCs der ACC27: CC166, CC167
CCs der ACC30 (außer CC180 und CC181): CC179, CC182, CC183, CC184
2. Regression mit 32 AG-Gruppen und 200 CCs
3. Reduktion der verbliebenen Risikoklassen
Aufnahme der
 - Top 30 CCs mit höchster ökonomischer Relevanz (Häufigkeit mal Relativgewicht aus Schritt 2)
 - der CCs mit Relativgewicht $\geq 1,5$ in das Modell

Die nicht ausgewählten unter den verbliebenen CCs werden entsprechend ihrer zugehörigen ACC-Gruppen zusammengefasst (Bildung von RCCs: eine RCC entspricht in Ihrer Zusammensetzung einer ACC abzüglich der CCs, die aufgrund hoher ökonomischer Relevanz bzw. einem Relativgewicht $> 1,5$ separat im Modell stehenbleiben)
4. Regression mit 32 AG-Gruppen und den Risikogruppen des reduzierten Modells (CCs, RCCs)
5. bei negativen Kosten- bzw. Relativgewichten von RCCs:
Jeweils die 3 RCCs mit der größten Korrelation zur RCC mit negativem Kosten- / Relativgewicht hierarchisch über die RCC mit negativem Kosten- / Relativgewicht stellen
erneute Regression
Hat diese Strategie keinen Erfolg, wird die RCC ausgeschlossen.
erneute Regression
6. Falls Alters-Geschlechtsgruppen negative und / oder insignifikante Kosten- bzw. Relativgewichte haben, diese zusammenfassen (analog komplettes Modell der KBV)
erneute Regression
7. ggf. Hierarchisierung

Anhang 21 Berechnung der regionalen Relativgewichte mit dem durchschnittlichen Leistungsbedarf der 4 KVen

Komplettes Modell:

Kennzahl	4 KVen	Bremen	Nieder- sachsen	Nordrhein	Thüringen
Anzahl der Relativgewichte im Modell	298	298	298	298	298
davon					
statistisch insignifikant (t-Wert < 1,96)	0	142	50	47	105
statistisch signifikant	298	156	248	251	193
davon					
negative Haupteffekte (AGE, CC)	0	6	1	0	0
negative Interaktionen (KCC, OCC)	77	24	54	56	38
positiv	221	126	193	195	155
davon mit einer Abweichung im Relativgewicht gegenüber der Lernmenge 4 KVen von:					
100 % und mehr	0	33	7	4	18
50 % bis unter 100 %	0	19	24	10	18
30 % bis unter 50 %	0	24	29	28	37
20 % bis unter 30 %	0	13	27	33	23
10 % bis unter 20 %	0	13	57	48	28
0 bis unter 10 %	0	24	49	72	31

Tabelle 68: Abweichungen der regionalen Relativgewichte von den Relativgewichten der vier KVen berechnet mit dem durchschnittlichen Leistungsbedarf der 4 KVen (komplettes Modell)

Reduziertes Modell:

Kennzahl	4 KVen	Bremen	Nieder- sachsen	Nordrhein	Thüringen
Anzahl der Relativgewichte im Modell	81	81	81	81	81
davon					
statistisch insignifikant (t-Wert < 1,96)	0	16	1	2	7
statistisch signifikant	81	65	80	79	74
davon					
negative Haupteffekte (AGE, CC, RCC)	0	0	0	0	0
positiv	81	65	80	79	74
davon mit einer Abweichung im Relativgewicht gegenüber der Lernmenge 4 KVen von:					
100 % und mehr	0	2	1	0	0
50 % bis unter 100 %	0	6	1	0	1
30 % bis unter 50 %	0	13	7	1	15
20 % bis unter 30 %	0	12	11	13	13
10 % bis unter 20 %	0	14	22	24	27
0 bis unter 10 %	0	18	38	41	18

Tabelle 69: Abweichungen der regionalen Relativgewichte von den Relativgewichten der vier KVen berechnet mit dem durchschnittlichen Leistungsbedarf der 4 KVen (reduziertes Modell)

Anhang 22 Übersicht der Zuordnung von Risikokategorien (CC) zu ACC

CC	ACC	Label der ACC
CC001	ACC001	Infektiöse und parasitäre Krankheiten
CC002	ACC001	Infektiöse und parasitäre Krankheiten
CC003	ACC001	Infektiöse und parasitäre Krankheiten
CC004	ACC001	Infektiöse und parasitäre Krankheiten
CC005	ACC001	Infektiöse und parasitäre Krankheiten
CC006	ACC001	Infektiöse und parasitäre Krankheiten
CC007	ACC002	Bösartige Neubildungen
CC008	ACC002	Bösartige Neubildungen
CC009	ACC002	Bösartige Neubildungen
CC010	ACC002	Bösartige Neubildungen
CC011	ACC003	Sonstige Neubildungen und Krankheiten der Mamma [Brustdrüse]
CC012	ACC003	Sonstige Neubildungen und Krankheiten der Mamma [Brustdrüse]
CC013	ACC003	Sonstige Neubildungen und Krankheiten der Mamma [Brustdrüse]
CC014	ACC003	Sonstige Neubildungen und Krankheiten der Mamma [Brustdrüse]
CC015	ACC004	Diabetes mellitus
CC016	ACC004	Diabetes mellitus
CC017	ACC004	Diabetes mellitus
CC018	ACC004	Diabetes mellitus
CC019	ACC004	Diabetes mellitus
CC020	ACC004	Diabetes mellitus
CC021	ACC005	Sonstige endokrine, Ernährungs- und Stoffwechselkrankheiten
CC022	ACC005	Sonstige endokrine, Ernährungs- und Stoffwechselkrankheiten
CC023	ACC005	Sonstige endokrine, Ernährungs- und Stoffwechselkrankheiten
CC024	ACC005	Sonstige endokrine, Ernährungs- und Stoffwechselkrankheiten
CC025	ACC006	Krankheiten der Leber
CC026	ACC006	Krankheiten der Leber
CC027	ACC006	Krankheiten der Leber
CC028	ACC006	Krankheiten der Leber
CC029	ACC006	Krankheiten der Leber
CC030	ACC006	Krankheiten der Leber
CC031	ACC007	Sonstige Krankheiten des Verdauungssystems
CC032	ACC007	Sonstige Krankheiten des Verdauungssystems

CC	ACC	Label der ACC
CC033	ACC007	Sonstige Krankheiten des Verdauungssystems
CC034	ACC007	Sonstige Krankheiten des Verdauungssystems
CC035	ACC007	Sonstige Krankheiten des Verdauungssystems
CC036	ACC007	Sonstige Krankheiten des Verdauungssystems
CC037	ACC008	Krankheiten des Muskel-Skelett-Systems und des Bindegewebes
CC038	ACC008	Krankheiten des Muskel-Skelett-Systems und des Bindegewebes
CC039	ACC008	Krankheiten des Muskel-Skelett-Systems und des Bindegewebes
CC040	ACC008	Krankheiten des Muskel-Skelett-Systems und des Bindegewebes
CC041	ACC008	Krankheiten des Muskel-Skelett-Systems und des Bindegewebes
CC042	ACC008	Krankheiten des Muskel-Skelett-Systems und des Bindegewebes
CC043	ACC008	Krankheiten des Muskel-Skelett-Systems und des Bindegewebes
CC044	ACC009	Krankheiten des Blutes und der blutbildenden Organe sowie bestimmte Störungen mit Beteiligung des Immunsystems
CC045	ACC009	Krankheiten des Blutes und der blutbildenden Organe sowie bestimmte Störungen mit Beteiligung des Immunsystems
CC046	ACC009	Krankheiten des Blutes und der blutbildenden Organe sowie bestimmte Störungen mit Beteiligung des Immunsystems
CC047	ACC009	Krankheiten des Blutes und der blutbildenden Organe sowie bestimmte Störungen mit Beteiligung des Immunsystems
CC048	ACC010	Störungen des Bewusstseins und neuropsychologische Symptome
CC049	ACC010	Störungen des Bewusstseins und neuropsychologische Symptome
CC050	ACC010	Störungen des Bewusstseins und neuropsychologische Symptome
CC051	ACC011	Psychische und Verhaltensstörungen durch Alkohol-, Drogen und sonstige Substanzen
CC052	ACC011	Psychische und Verhaltensstörungen durch Alkohol-, Drogen und sonstige Substanzen
CC053	ACC011	Psychische und Verhaltensstörungen durch Alkohol-, Drogen und sonstige Substanzen
CC054	ACC012	Sonstige psychische Krankheiten
CC055	ACC012	Sonstige psychische Krankheiten
CC057	ACC012	Sonstige psychische Krankheiten
CC058	ACC012	Sonstige psychische Krankheiten
CC059	ACC012	Sonstige psychische Krankheiten
CC060	ACC012	Sonstige psychische Krankheiten
CC061	ACC013	Entwicklungs- und Lernstörungen, Intelligenzminderung, Ticstörungen und komplexe Fehlbildungs-Syndrome

CC	ACC	Label der ACC
CC062	ACC013	Entwicklungs- und Lernstörungen, Intelligenzminderung, Ticstörungen und komplexe Fehlbildungs-Syndrome
CC063	ACC013	Entwicklungs- und Lernstörungen, Intelligenzminderung, Ticstörungen und komplexe Fehlbildungs-Syndrome
CC064	ACC013	Entwicklungs- und Lernstörungen, Intelligenzminderung, Ticstörungen und komplexe Fehlbildungs-Syndrome
CC065	ACC013	Entwicklungs- und Lernstörungen, Intelligenzminderung, Ticstörungen und komplexe Fehlbildungs-Syndrome
CC067	ACC014	Sonstige neurologische Krankheiten
CC068	ACC014	Sonstige neurologische Krankheiten
CC069	ACC014	Sonstige neurologische Krankheiten
CC070	ACC014	Sonstige neurologische Krankheiten
CC071	ACC014	Sonstige neurologische Krankheiten
CC072	ACC014	Sonstige neurologische Krankheiten
CC073	ACC014	Sonstige neurologische Krankheiten
CC074	ACC014	Sonstige neurologische Krankheiten
CC075	ACC014	Sonstige neurologische Krankheiten
CC076	ACC014	Sonstige neurologische Krankheiten
CC077	ACC015	Herzstillstand, Ateminsuffizienz, Tod und Vorhandensein eines Tracheostomas
CC078	ACC015	Herzstillstand, Ateminsuffizienz, Tod und Vorhandensein eines Tracheostomas
CC079	ACC015	Herzstillstand, Ateminsuffizienz, Tod und Vorhandensein eines Tracheostomas
CC080	ACC016	Krankheiten des Herz-Kreislaufsystems
CC081	ACC016	Krankheiten des Herz-Kreislaufsystems
CC082	ACC016	Krankheiten des Herz-Kreislaufsystems
CC083	ACC016	Krankheiten des Herz-Kreislaufsystems
CC084	ACC016	Krankheiten des Herz-Kreislaufsystems
CC085	ACC016	Krankheiten des Herz-Kreislaufsystems
CC086	ACC016	Krankheiten des Herz-Kreislaufsystems
CC087	ACC016	Krankheiten des Herz-Kreislaufsystems
CC088	ACC016	Krankheiten des Herz-Kreislaufsystems
CC089	ACC016	Krankheiten des Herz-Kreislaufsystems
CC090	ACC016	Krankheiten des Herz-Kreislaufsystems
CC091	ACC016	Krankheiten des Herz-Kreislaufsystems

CC	ACC	Label der ACC
CC092	ACC016	Krankheiten des Herz-Kreislaufsystems
CC093	ACC016	Krankheiten des Herz-Kreislaufsystems
CC094	ACC016	Krankheiten des Herz-Kreislaufsystems
CC095	ACC017	Zerebrovaskuläre Krankheiten
CC096	ACC017	Zerebrovaskuläre Krankheiten
CC097	ACC017	Zerebrovaskuläre Krankheiten
CC098	ACC017	Zerebrovaskuläre Krankheiten
CC099	ACC017	Zerebrovaskuläre Krankheiten
CC100	ACC017	Zerebrovaskuläre Krankheiten
CC101	ACC017	Zerebrovaskuläre Krankheiten
CC102	ACC017	Zerebrovaskuläre Krankheiten
CC103	ACC017	Zerebrovaskuläre Krankheiten
CC104	ACC018	Sonstige Blutgefäß-Krankheiten
CC105	ACC018	Sonstige Blutgefäß-Krankheiten
CC106	ACC018	Sonstige Blutgefäß-Krankheiten
CC107	ACC019	Krankheiten der unteren Atemwege
CC108	ACC019	Krankheiten der unteren Atemwege
CC109	ACC019	Krankheiten der unteren Atemwege
CC110	ACC019	Krankheiten der unteren Atemwege
CC111	ACC019	Krankheiten der unteren Atemwege
CC112	ACC019	Krankheiten der unteren Atemwege
CC113	ACC019	Krankheiten der unteren Atemwege
CC114	ACC019	Krankheiten der unteren Atemwege
CC115	ACC019	Krankheiten der unteren Atemwege
CC116	ACC020	Augen-Krankheiten
CC117	ACC020	Augen-Krankheiten
CC118	ACC020	Augen-Krankheiten
CC119	ACC020	Augen-Krankheiten
CC120	ACC020	Augen-Krankheiten
CC121	ACC020	Augen-Krankheiten
CC122	ACC020	Augen-Krankheiten
CC123	ACC020	Augen-Krankheiten
CC124	ACC020	Augen-Krankheiten
CC125	ACC021	HNO- und stomatologische Krankheiten

CC	ACC	Label der ACC
CC126	ACC021	HNO- und stomatologische Krankheiten
CC127	ACC021	HNO- und stomatologische Krankheiten
CC128	ACC022	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation
CC129	ACC022	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation
CC130	ACC022	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation
CC131	ACC022	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation
CC132	ACC022	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation
CC133	ACC022	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation
CC134	ACC022	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation
CC135	ACC022	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation
CC136	ACC022	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation
CC137	ACC023	Krankheiten der Genitalorgane
CC138	ACC023	Krankheiten der Genitalorgane
CC139	ACC023	Krankheiten der Genitalorgane
CC140	ACC023	Krankheiten der Genitalorgane
CC141	ACC024	Krankheiten und Zustände bei Schwangerschaft, Geburt und Wochenbett
CC142	ACC024	Krankheiten und Zustände bei Schwangerschaft, Geburt und Wochenbett
CC143	ACC024	Krankheiten und Zustände bei Schwangerschaft, Geburt und Wochenbett
CC144	ACC024	Krankheiten und Zustände bei Schwangerschaft, Geburt und Wochenbett
CC145	ACC024	Krankheiten und Zustände bei Schwangerschaft, Geburt und Wochenbett
CC146	ACC024	Krankheiten und Zustände bei Schwangerschaft, Geburt und Wochenbett
CC147	ACC024	Krankheiten und Zustände bei Schwangerschaft, Geburt und Wochenbett
CC148	ACC025	Krankheiten der Haut und der Unterhaut

CC	ACC	Label der ACC
CC149	ACC025	Krankheiten der Haut und der Unterhaut
CC150	ACC025	Krankheiten der Haut und der Unterhaut
CC151	ACC025	Krankheiten der Haut und der Unterhaut
CC152	ACC025	Krankheiten der Haut und der Unterhaut
CC153	ACC025	Krankheiten der Haut und der Unterhaut
CC154	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC155	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC156	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC157	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC158	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC159	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC160	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC161	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC162	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC163	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC164	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen
CC165	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen

CC	ACC	Label der ACC
CC166	ACC027	Symptome, Krankheitszeichen und anderenorts nicht klassifizierte Störungen und Zustände
CC167	ACC027	Symptome, Krankheitszeichen und anderenorts nicht klassifizierte Störungen und Zustände
CC168	ACC028	Krankheiten und Zustände des Neugeborenen
CC169	ACC028	Krankheiten und Zustände des Neugeborenen
CC170	ACC028	Krankheiten und Zustände des Neugeborenen
CC171	ACC028	Krankheiten und Zustände des Neugeborenen
CC172	ACC028	Krankheiten und Zustände des Neugeborenen
CC173	ACC029	Krankheiten und Zustände nach sonstigen Transplantationen und nach besonderen medizinischen Maßnahmen
CC174	ACC029	Krankheiten und Zustände nach sonstigen Transplantationen und nach besonderen medizinischen Maßnahmen
CC175	ACC029	Krankheiten und Zustände nach sonstigen Transplantationen und nach besonderen medizinischen Maßnahmen
CC176	ACC029	Krankheiten und Zustände nach sonstigen Transplantationen und nach besonderen medizinischen Maßnahmen
CC177	ACC029	Krankheiten und Zustände nach sonstigen Transplantationen und nach besonderen medizinischen Maßnahmen
CC178	ACC029	Krankheiten und Zustände nach sonstigen Transplantationen und nach besonderen medizinischen Maßnahmen
CC179	ACC030	Untersuchungen, prophylaktische Maßnahmen, anamnestische Angaben und Zustände nach weiteren medizinischen Maßnahmen
CC180	ACC030	Untersuchungen, prophylaktische Maßnahmen, anamnestische Angaben und Zustände nach weiteren medizinischen Maßnahmen
CC181	ACC030	Untersuchungen, prophylaktische Maßnahmen, anamnestische Angaben und Zustände nach weiteren medizinischen Maßnahmen
CC182	ACC030	Untersuchungen, prophylaktische Maßnahmen, anamnestische Angaben und Zustände nach weiteren medizinischen Maßnahmen
CC183	ACC030	Untersuchungen, prophylaktische Maßnahmen, anamnestische Angaben und Zustände nach weiteren medizinischen Maßnahmen
CC184	ACC030	Untersuchungen, prophylaktische Maßnahmen, anamnestische Angaben und Zustände nach weiteren medizinischen Maßnahmen
CC185	ACC002	Bösartige Neubildungen
CC186	ACC002	Bösartige Neubildungen
CC187	ACC002	Bösartige Neubildungen
CC188	ACC003	Sonstige Neubildungen und Krankheiten der Mamma [Brustdrüse]
CC189	ACC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funk-

CC	ACC	Label der ACC
		tionseinschränkungen
CC190	ACC008	Krankheiten des Muskel-Skelett-Systems und des Bindegewebes
CC191	ACC016	Krankheiten des Herz-Kreislaufsystems
CC192	ACC021	HNO- und stomatologische Krankheiten
CC193	ACC006	Krankheiten der Leber
CC194	ACC014	Sonstige neurologische Krankheiten
CC195	ACC018	Sonstige Blutgefäß-Krankheiten
CC196	ACC017	Zerebrovaskuläre Krankheiten
CC197	ACC011	Psychische und Verhaltensstörungen durch Alkohol-, Drogen und sonstige Substanzen
CC198	ACC012	Sonstige psychische Krankheiten
CC199	ACC013	Entwicklungs- und Lernstörungen, Intelligenzminderung, Ticstörungen und komplexe Fehlbildungs-Syndrome
CC200	ACC030	Untersuchungen, prophylaktische Maßnahmen, anamnestische Angaben und Zustände nach weiteren medizinischen Maßnahmen

Anhang 23 Schaubild zur Berechnung von Veränderungsraten

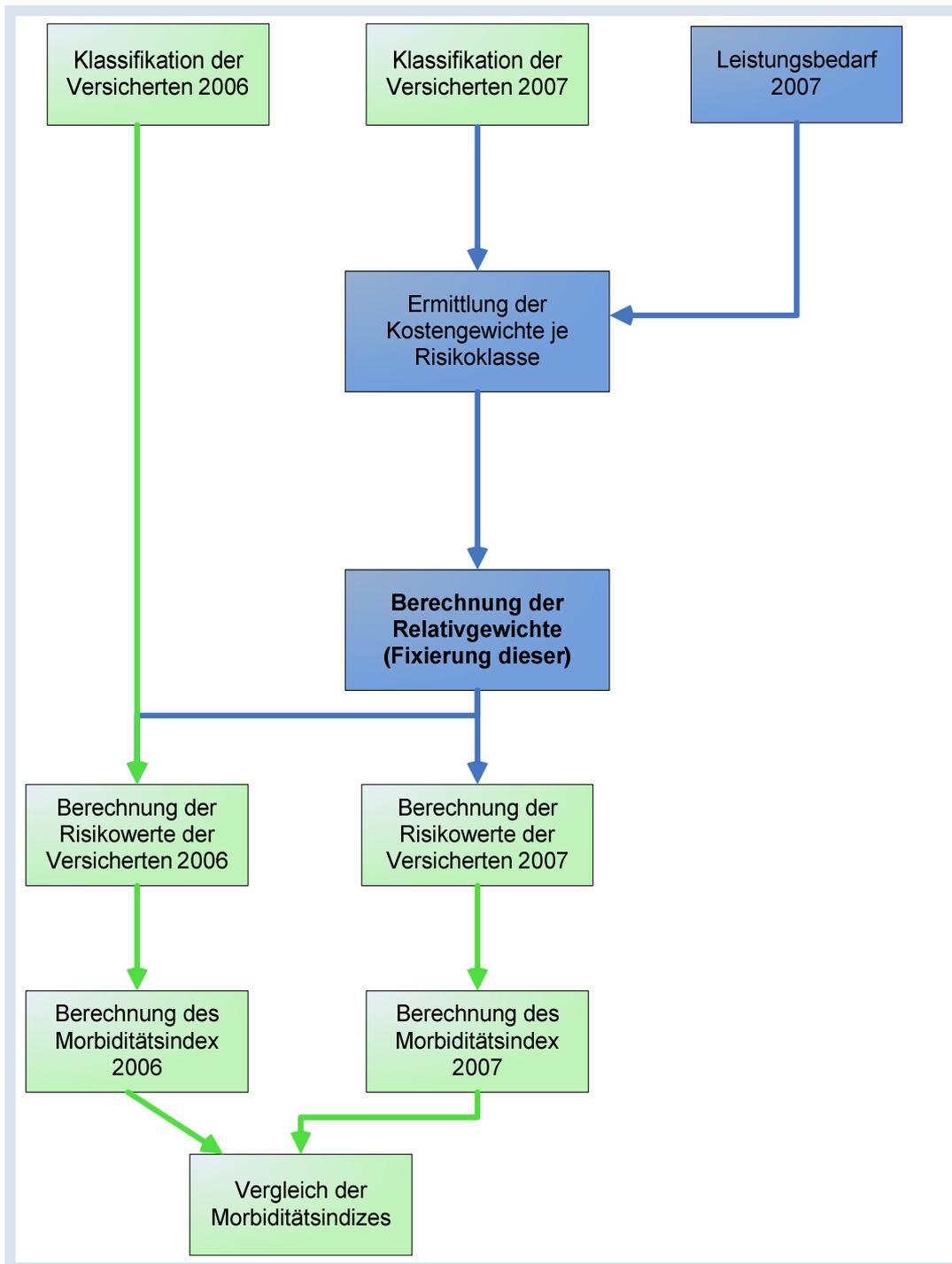


Abbildung 51: Schaubild zur Berechnung von Veränderungsraten am Beispiel einer zeitgleichen Modellierung

Anhang 24 Formeln zur Berechnung der Veränderungsrate

Klassifizierung der Versicherten (Formel 1)

$$ICD_1 \rightarrow RK_1$$

$$ICD_2 \rightarrow RK_1$$

$$ICD_3 \rightarrow RK_2$$

$$\vdots \quad \quad \quad \vdots$$

$$ICD_x \rightarrow RK_{m-z}$$

$$\vdots \quad \quad \quad \vdots$$

$$Merkmal_x \rightarrow RK_{m-1}$$

$$Merkmal_y \rightarrow RK_m$$

wenn $D_{i;A_{Diag}} = ICD_1$ dann

$$RK_{1;i;A_{Diag}} = 1$$

u.s.w

D_i = Diagnose / Merkmal des Versicherten i mit $i = 1, \dots, n$

ICD_1 = erste ICD des ICD Kataloges u.s.w.

A_{Diag} = Jahr der Diagnosenstellung / der Merkmalsausprägung

RK_j = Risikoklasse j mit $j=1, \dots, m$

Berechnung der Kostengewichte (Formel 2)

Das Kostengewicht KG_j der Risikoklasse j entspricht dem $\hat{\beta}_j$ aus der Regressionsrechnung.

$\hat{\beta}_j$ wird berechnet als Schätzung gewichteter kleinster Quadrate auf Basis annualisierter Leistungsbedarfe. Die Annualisierung der Leistungsbedarfe erfolgt durch einfache Hochrechnung der Leistungsbedarfe auf das Kalenderjahr:

$$ALB = \sum_{i=1}^n (LB_{i,A_{LB}} * \frac{365}{VT_{i,A_{LB}}})$$

ALB = Annualisierter Leistungsbedarf

A_{LB} = Jahresbezug des Leistungsbedarfs

Die Gewichtung erfolgt mit dem Anteil der Versichertentage am Kalenderjahr.

Berechnung des 100 % Wertes (Formel 3)

Der Durchschnittsbetrag ergibt sich als gewichteter Mittelwert annualisierter Individualbeträge.

$$DLB = \frac{\sum_{i=1}^n (LB_{i,ALB} * \frac{365}{VT_{i,ALB}}) * \frac{VT_{i,ALB}}{365}}{\sum_{i=1}^n \frac{VT_{i,ALB}}{365}}$$

DLB = Durchschnittlicher Leistungsbedarf über alle Versicherten

$VT_{j;A=t}$ = Versichertentage des Versicherten i im Jahr t

Berechnung des Relativgewichtes (Formel 4)

$$RG_j = \frac{KG_j}{DLB}$$

Risikowert des Versicherten i (Formel 5)

$$RW_{i,A_M} = \sum_{j=1}^m (RK_{i,j,A_M} * RG_j)$$

A_M = Jahresbezug des Morbiditätsindex

Morbiditätsindex (Formel 6)

$$MI_{A_M=t} = \frac{\sum_{i=1}^n (RW_{i,A_M} * \frac{VT_{i,A_M}}{365})}{\sum_{i=1}^n \frac{VT_{i,A_M}}{365}}$$

Veränderungsrate (Formel 7)

$$VR = \frac{MI_{A_M=t+1}}{MI_{A_M=t}}$$

Anhang 25 Triggerung der CC 130 durch die folgenden GOPs:

Abrechnungsjahr	GOP
2006 und 2007	40800; 40800F, 40802; 40802F, 40802W; 40804, 40804F, 40810, 40812, 40820, 40820F, 40800A, 40800U, 40800W, 40802A, 40802U, 40804A, 40804U, 40804W
2008	40800; 40800A, 40800F, 40800U, 40800W, 40801, 40802, 40802A, 40802F, 40802U, 40802W, 40803, 40804, 40804A, 40804F, 40804U, 40804W, 40805, 40806, 40807, 40808, 40810, 40811, 40812, 40813, 40820, 40820F, 40821, 40822

Tabelle 70: GOPs, die CC130 triggern

Anhang 26 Korrelationskoeffizienten der RCC 24 und RCC 28 im beschlossenen Modell

Korrelationskoeffizienten der RCC 28		Korrelationskoeffizienten der RCC 24	
Phi-Koeffizient	RCC	Phi-Koeffizient	RCC
1,0000	28	1,0000	24
0,0357	13	0,0727	9
0,0187	21	0,0345	23
0,0164	5	0,0259	12
0,0063	8	0,0109	22
0,0041	9	0,0106	18
0,0026	25	0,0094	11
0,0016	10	0,0059	5
0,0004	1	0,0038	3
0,0003	15	0,0005	1
-0,0006	29	-0,0006	26
-0,0010	23	-0,0028	15
-0,0038	3	-0,0037	7
-0,0042	7	-0,0040	29
-0,0042	19	-0,0050	25
-0,0043	14	-0,0063	8
-0,0046	2	-0,0064	30
-0,0056	17	-0,0068	28
-0,0059	16	-0,0086	2
-0,0068	26	-0,0088	14
-0,0068	24	-0,0135	13
-0,0078	12	-0,0144	10
-0,0082	30	-0,0184	4
-0,0090	4	-0,0193	6
-0,0097	11	-0,0194	21
-0,0117	6	-0,0198	19
-0,0123	22	-0,0222	17
-0,0141	20	-0,0327	16
-0,0184	18	-0,0398	20

Tabelle 71: Korrelationskoeffizienten der RCC 24 und RCC 28 im beschlossenen Modell

Anhang 27 Extra Anhang MPE

Kurzbeschreibung	Relativer Prognosefehler (MPE / Mean_LB)	
	in der Kalibrierungsstichprobe	in der Validierungsstichprobe
Var. A, reduziert, mit Hier., LB07, 1-J-prosp., V404	0,00%	0,25%
Var. F, reduziert, ohne Hier.; LB07, 1-J-prosp., V404	0,00%	0,36%
Var. F, reduziert, ohne Hier.; LB08, 2-J-prosp., V404	0,00%	0,16%
Var. F, reduziert, mit Hier., LB07, 1-J-prosp., V404	0,00%	0,35%
Var. F, reduziert 130;mit Hier., LB08, 1-J-prosp., V404	0,00%	0,37%
Var. F, komplett, ohne Hier.; LB07, zeitgl., V309	0,00%	0,00%
Var. F, komplett, ohne Hier.; LB08, 2-J-prosp., V404	0,00%	0,10%
Var. F, komplett, mit Hier., LB08, 2-J-prosp., V404	0,00%	2,56%
Var. G, komplett, ohne Hier.; LB07, zeitgl., V309	0,00%	0,00%
Var. G, komplett, ohne Hier.; LB08, zeitgl., V404	0,00%	0,21%
Var. G, komplett, ohne Hier.; LB08, 2-J-prosp., V404	0,00%	0,10%
Var. G, komplett, ohne Hier.; LB08, 2-J-prosp., X404	0,00%	-0,06%
Var. H, komprimiert mit Hier., LB08, 2-J-prosp., X404	0,00%	0,29%

Tabelle 72: Prognosefehler der beschlossenen Ausgestaltung im Vergleich mit weiteren beauftragten Ausgestaltungen

TCC	Beschreibung	Relativer Prognosefehler (MPE / Mean_LB)
6	Andere Infektionskrankheiten	-0,10%
10	Bösartige Neubildungen Dickdarm, Rektum, Anus, Mamma, Geschlechtsorgane, Niere und ableitende Harnwege, Auge, Schilddrüse, maligne Melanome	-0,11%
14	Gutartige Neubildungen der Haut, der (weiblichen) Brust und des Auges	0,05%
19	Diabetes mellitus ohne oder mit nicht näher bezeichneten Komplikationen, Schwangerschaftsdiabetes	0,18%
24	Andere endokrine, metabolische und Ernährungsstörungen	-0,05%
36	Andere gastrointestinale Erkrankungen	-0,11%
38	Entzündliche Systemerkrankungen des rheumatoiden Formenkreises	-0,06%
40	Cox- oder Gonarthrose	0,13%
41	Erkrankungen des Knochens und des Knorpels	-0,01%
43	Andere Erkrankungen der Muskeln, Sehnen, Gelenke und des Bindegewebes	-0,19%
55	Bipolare Störungen, depressive Episoden	0,59%
57	Persönlichkeitsstörungen	-0,98%
58	Depressive Episode, Angst, Dysthymia	-0,28%
60	Angststörungen, sexuelle Störungen, Belastungs- und Anpassungsstörungen	0,01%
76	Radikulopathien, andere neurologische Erkrankungen und Folgen von Verletzungen des Nervensystems	-0,07%
84	Koronare Herzkrankheit / andere chronisch-ischämische Erkrankungen des Herzens	-0,25%
91	Essentielle Hypertonie	-0,25%
110	Asthma bronchiale	-0,24%
115	Akute Bronchitis, Grippe	-0,32%
124	Andere Erkrankungen des Auges	0,09%
127	Andere Erkrankungen HNO Bereich	0,05%
129	Terminale Niereninsuffizienz	2,55%
130	Abhängigkeit von Dialyse	1,49%
139	Störungen der Menstruation, klimakterische Störungen, andere Erkrankungen der weiblichen Geschlechtsorgane	-0,08%

TCC	Beschreibung	Relativer Prognosefehler (MPE / Mean_LB)
153	Verbrennungen 1. und 2. Grades, andere Erkrankungen der Haut und Hautanhangsgebilde	-0,04%
162	Extremitätenfrakturen, andere Verletzungen	0,12%
166	Schwere Symptome, abnorme Befunde	-0,10%
167	Schwindel, Unwohlsein und andere Allgemeinsymptome	-0,20%
179	Verschiedene Eingriffe, Status nach operativen Eingriffen, Nachbehandlung, elektive Chirurgie	-0,16%
183	Vorsorgeuntersuchungen, Impfungen, Beratung zur Kontrazeption	-0,19%
190	Knöchernen Erkrankungen der Wirbelsäule, Erkrankungen der Bandscheiben, Stenosen des Spinalkanals und der Foramina intervertebralia	0,04%
194	Migräne und Kopfschmerzen weiterer Genesen	-0,08%

Tabelle 73: Predictive Ratios je TCC

RCC	Beschreibung	Relativer Prognosefehler (MPE / Mean_LB)
1	Infektiöse und parasitäre Krankheiten; nach Variante 15. EBA	0,59%
2	Bösartige Neubildungen; nach Variante 15. EBA	-0,89%
3	Sonstige Neubildungen und Krankheiten der Mamma [Brustdrüse]; nach Variante 15. EBA	0,10%
4	Diabetes mellitus; nach Variante 15. EBA	0,45%
5	Sonstige endokrine, Ernährungs- und Stoffwechselkrankheiten; nach Variante 15. EBA	0,43%
6	Krankheiten der Leber; nach Variante 15. EBA	-0,41%
7	Sonstige Krankheiten des Verdauungssystems; nach Variante 15. EBA	-0,26%
8	Krankheiten des Muskel-Skelett-Systems und des Bindegewebes; nach Variante 15. EBA	-0,04%
9	Krankheiten des Blutes und der blutbildenden Organe sowie bestimmte Störungen mit Beteiligung des Immunsystems; nach Variante 15. EBA	0,57%

RCC	Beschreibung	Relativer Prognosefehler (MPE / Mean_LB)
10	Störungen des Bewusstseins und neuropsychologische Symptome; nach Variante 15. EBA	-0,46%
11	Psychische und Verhaltensstörungen durch Alkohol-, Drogen und sonstige Substanzen; nach Variante 15. EBA	-0,20%
12	Sonstige psychische Krankheiten; nach Variante 15. EBA	-0,96%
13	Entwicklungs- und Lernstörungen, Intelligenzminderung, Ticstörungen und komplexe Fehlbildungs-Syndrome; nach Variante 15. EBA	-0,55%
14	Sonstige neurologische Krankheiten; nach Variante 15. EBA	0,29%
15	Herzstillstand, Ateminsuffizienz, Tod und Vorhandensein eines Tracheostomas; nach Variante 15. EBA	-0,83%
16	Krankheiten des Herz-Kreislaufsystems; nach Variante 15. EBA	-0,20%
17	Zerebrovaskuläre Krankheiten; nach Variante 15. EBA	-0,27%
18	Sonstige Blutgefäß-Krankheiten; nach Variante 15. EBA	-0,14%
19	Krankheiten der unteren Atemwege; nach Variante 15. EBA	-0,11%
20	Augen-Krankheiten; nach Variante 15. EBA	0,19%
21	HNO- und stomatologische Krankheiten; nach Variante 15. EBA	0,07%
22	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation; nach Variante 15. EBA	-0,07%
23	Krankheiten der Genitalorgane; nach Variante 15. EBA	-0,13%
25	Krankheiten der Haut und der Unterhaut; nach Variante 15. EBA	-0,42%
26	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen; nach Variante 15. EBA	-0,10%
29	Krankheiten und Zustände nach sonstigen Transplantationen und nach besonderen medizinischen Maßnahmen; nach Variante 15. EBA	0,49%
30	Untersuchungen, prophylaktische Maßnahmen, anamnestiche Angaben und Zustände nach weiteren medizinischen Maßnahmen; nach Variante 15. EBA	-0,31%

Tabelle 74: Predictive Ratios je RCC

Alters- und Geschlechts-klasse	Beschreibung	Relativer Prognosefehler (MPE / Mean_LB)
1	Weiblich, 0 bis 5 Jahre	0,06%
2	Weiblich, 6 bis 12 Jahre	-1,36%
3	Weiblich, 13 bis 17 Jahre	0,13%
4	Weiblich, 18 bis 24 Jahre	-1,08%
5	Weiblich, 25 bis 34 Jahre	-0,46%
6	Weiblich, 35 bis 44 Jahre	0,45%
7	Weiblich, 45 bis 54 Jahre	0,12%
8	Weiblich, 55 bis 59 Jahre	-0,36%
9	Weiblich, 60 bis 64 Jahre	0,10%
10	Weiblich, 65 bis 69 Jahre	-0,79%
11	Weiblich, 70 bis 74 Jahre	-0,33%
12	Weiblich, 75 bis 79 Jahre	0,30%
13	Weiblich, 80 bis 84 Jahre	-0,41%
14	Weiblich, 85 bis 89 Jahre	0,99%
15	Weiblich, 90 bis 94 Jahre	-1,06%
16	Weiblich, 95 Jahre und älter	0,09%
17	Männlich, 0 bis 5 Jahre	0,12%
18	Männlich, 6 bis 12 Jahre	0,10%
19	Männlich, 13 bis 17 Jahre	0,20%
20	Männlich, 18 bis 24 Jahre	0,23%
21	Männlich, 25 bis 34 Jahre	0,22%
22	Männlich, 35 bis 44 Jahre	0,08%
23	Männlich, 45 bis 54 Jahre	0,21%
24	Männlich, 55 bis 59 Jahre	0,16%
25	Männlich, 60 bis 64 Jahre	-0,06%
26	Männlich, 65 bis 69 Jahre	-0,18%
27	Männlich, 70 bis 74 Jahre	0,09%
28	Männlich, 75 bis 79 Jahre	0,12%
29	Männlich, 80 bis 84 Jahre	0,94%
30	Männlich, 85 bis 89 Jahre	-0,18%
31	Männlich, 90 bis 94 Jahre	1,47%
32	Männlich, 95 Jahre und älter	-1,04%

Tabelle 75: Predictive Ratios für Alters- und Geschlechtsgruppen