

# Populationsbildung auf Grundlage von Abrechnungsdaten der vertragsärztlichen Versorgung – Kurzversion

Mangiapane S, Riens B, Augustin J

**Zi** Zentralinstitut für die kassenärztliche Versorgung in Deutschland

Herbert-Lewin-Platz 3 - 10623 Berlin - Tel. (030) 4005-2450 - E-Mail: [zi@zi-berlin.de](mailto:zi@zi-berlin.de)

## Hintergrund

Abrechnungsdaten, welche im Rahmen der medizinischen Versorgung gesetzlich krankenversicherter Patienten routinemäßig an verschiedenen Stellen des deutschen Gesundheitssystems erfasst werden, gehören zu den wichtigsten Datenquellen der Versorgungsforschung. Da diese Daten primär zu Abrechnungszwecken erhoben werden, werden sie bei ihrer Verwendung in der Versorgungsforschung auch als Sekundärdaten bezeichnet (i.S. eines sekundären Einsatzzweckes).

Je nach Versorgungsbereich und Leistungsart werden Abrechnungsdaten an verschiedenen Stellen im Gesundheitswesen generiert. Die Daten der vertragsärztlichen Versorgung stellen insofern eine Besonderheit dar, als dass diese Daten im Rahmen der kollektivvertraglichen Abrechnungen nicht direkt an die Krankenkassen, sondern zunächst an die zuständige Kassenärztliche Vereinigung (KV) übermittelt werden. Mit den zusammengeführten pseudonymisierten vertragsärztlichen Abrechnungsdaten aller 17 KVen steht dem Zentralinstitut für die kassenärztliche Versorgung (ZI) ein Datenkörper zur Verfügung, der gegenüber den Daten der einzelnen Krankenkassen den Vorteil hat, die vertragsärztliche Versorgung bundesweit und Krankenkassen übergreifend darzustellen.

Ein Nachteil der Daten besteht darin, dass die sogenannten Versichertenstammdaten nicht enthalten sind. Das heißt, es liegen weder Informationen über den Versicherungsbeginn, noch über das Versicherungsende eines Versicherten bei einer bestimmten Krankenkasse (Versicherungszeiten) vor. Darüber hinaus ist die in den Daten enthaltene pseudonymisierte Versichertennummer allein nicht eindeutig, da dieselbe Versichertennummer bei verschiedenen Kassen vorkommen kann. Die Folge ist, dass bei alleiniger Betrachtung der Versichertennummer fälschlicherweise davon ausgegangen wird, dass es sich dabei um dieselbe Person handelt. Dass es dennoch Möglichkeiten gibt, die vertragsärztlichen Abrechnungsdaten der KVen personenbezogen zu analysieren und welche methodischen Konsequenzen damit verbunden sind, soll im Folgenden dargestellt werden.

## Methodik

### Datenherkunft und Dateninhalte

Abrechnungsdaten der vertragsärztlichen Versorgung gemäß § 295 SGB V stellen eine der Hauptdatenquellen für die im Rahmen des Versorgungsatlas durchgeführten Sekundärdatenanalysen dar. Sie werden von den an der vertragsärztlichen Versorgung teilnehmenden Ärzten zu Abrechnungszwecken generiert und

quartalsweise an die zuständige Kassenärztliche Vereinigung übermittelt, die auf Basis dieser Daten die Honorarabrechnung durchführt.

Die für Versorgungsanalysen relevanten Teile der Abrechnungsdaten enthalten für alle gesetzlich krankenversicherten Personen, die im jeweiligen Betrachtungsjahr mindestens einen Arztkontakt hatten, folgende Angaben:

- Patienten charakterisierende Angaben (pseudonymisierte Versichertennummer, Institutskennezeichen (IK) der Krankenkasse, Alter und Geschlecht)
- Arzt charakterisierende Merkmale (pseudonymisierte lebenslange Arztnummer, pseudonymisierte Betriebsstättennummer, Facharztgruppe, KV)
- Abgerechnete Leistungen (abgerechnete Gebührenordnungspositionen gemäß Einheitlichem Bewertungsmaßstab, sowie gemäß OPS (Operationen- und Prozedurenschlüssel) kodierte Schlüssel)
- Diagnosen (gemäß ICD 10 (Internationale Klassifikation von Krankheiten))

### Bildung von Patientenentitäten

Wie bereits einleitend erwähnt, enthalten die Daten, die von den Vertragsärzten an die Kassenärztlichen Vereinigungen übermittelt werden, im Gegensatz zu den Daten, die den gesetzlichen Krankenkassen vorliegen, keine Versichertenstammdaten. Das heißt, die Daten enthalten

- weder ein den Patienten eindeutig kennzeichnendes Merkmal,
- noch sind Versicherungszeiten verfügbar, aus denen hervorgeht, über welchen Zeitraum hinweg ein Patient bei einer Krankenkasse versichert war.

Während die fehlenden Versicherungszeiten derzeit ein für die Datenanalysen unlösbares Problem darstellen, kann die Kennzeichnung einer Patientenentität über eine Kombination verschiedener Merkmale versucht werden. So wird zwar dieselbe Versichertennummer bei verschiedenen Kassen vorkommen, sodass diese für die Kennzeichnung eines Patienten ungeeignet ist. Dieselbe Kombination aus Versichertennummer (pseudonymisiert), der IK einer

Krankenkasse und dem Geburtsdatum eines Patienten sollte hingegen einmalig sein, sodass über diese Kombination die Bildung einer Patientenentität vorgenommen werden kann.

Eine Patientenentität =

Versichertennummer (pseudonymisiert) +  
Krankenversicherungskarten-IK + Geburtsdatum

Grundsätzlich sind bei der beschriebenen Vorgehensweise einige Aspekte zu beachten, die die Zusammensetzung der gebildeten Populationen beeinflussen.

*a) Verwendung der Versichertennummer als Bestandteil der Patientenentität*

Ändert sich die Versichertennummer eines Patienten, wird durch den beschriebenen Algorithmus ein „neuer Patient“ erzeugt. Dies hat zur Folge, dass sich einerseits die Patientenzahl für das Beobachtungsjahr erhöht und andererseits der entsprechende Patient im Längsschnitt nicht weiter verfolgbar ist.

*b) Verwendung der Krankenversicherungskarten-IK als Bestandteil der Patientenentität*

Da jede Krankenkasse über eigene IKs verfügt, führt jeder Krankenkassenwechsel, den ein Patient vornimmt, zur Bildung einer neuen Patientenentität in den Abrechnungsdaten. Gleiches gilt für Kassenfusionen, bei denen entweder die IKs einer der fusionierenden Kassen beibehalten werden oder bei denen vollständig neue IKs zum Einsatz kommen. Darüber hinaus beinhaltet die Krankenversicherungskarten-IK auch Informationen über die KV, in deren Zuständigkeitsbereich der Wohnort des Versicherten gehört, sodass jeder Wohnortwechsel eines Patienten, der mit einem Wechsel der zuständigen KV verbunden ist, zu einer Änderung der Krankenversicherungskarten-IK führt. Somit sind auch Fusionen von KVen (wie z.B. in Baden-Württemberg und Rheinland-Pfalz, die ursprünglich in mehrere KV-Regionalstellen unterteilt waren) mit der Vergabe neuer IKs verbunden. Alle genannten Faktoren haben zur Folge, dass sich einerseits die Patientenzahl für das Beobachtungsjahr erhöht und andererseits der entsprechende Patient nicht weiter längsschnittlich verfolgbar ist.

### c) Verwendung des Geburtsdatums als Bestandteil der Patientenentität

Auch die Einbeziehung des Geburtsdatums in die Bildung von Patientenentitäten kann zur Folge haben, dass sich die Patientenzahl für das Beobachtungsjahr erhöht und Patienten ab einem bestimmten Zeitpunkt nicht weiter längsschnittlich verfolgbar sind. Dies ist immer dann der Fall, wenn für denselben Patienten (i.S.v. selber Versichertennummer und IK) fälschlicherweise mehrere Geburtsdaten vorliegen (z.B. bei fehlerhafter Datumsübertragung).

#### **Patientenausschluss**

Eine Mindestanforderung für die Bearbeitung von Fragestellungen der Versorgungsforschung besteht darin, dass die zur Verfügung stehenden Daten Angaben zum Alter und Geschlecht der zu untersuchenden Personen enthalten. Liegt darüber hinaus -wie im Versorgungsatlas- ein Fokus auf der Untersuchung regionaler Unterschiede, so muss zudem die regionale Zuordnung eines Patienten möglich sein.

Bezüglich der Angaben zum Geschlecht eines Patienten ist bekannt, dass es einige Praxisverwaltungssysteme gibt (z.B. in Laborpraxen, die keinen direkten Patientenkontakt haben), die für dieses Attribut einen Default-Wert einstellen. Weicht diese Default-Einstellung von der korrekten Geschlechtsangabe ab, existieren in den Daten für einen Patienten parallel zwei Geschlechtsangaben. Das ZI prüft im Rahmen von Plausibilitätsprüfungen, ob für betreffende Patienten im Beobachtungszeitraum Diagnosen kodiert wurden, die gemäß der Klassifikationsdateien des Deutschen Instituts für medizinische Dokumentation und Information (DIMDI) geschlechtsspezifisch sind. Sind diese spezifischen Diagnosen beobachtbar, wird die sich daraus ergebende falsche Geschlechtsangabe für diese Patienten korrigiert. Patienten mit zwei Geschlechtsangaben für die keine geschlechtsspezifischen Diagnosen kodiert wurden, müssen hingegen für Versorgungsanalysen ausgeschlossen werden.

Weiterhin finden Patienten mit ungültigem Geburtsdatum (z.B. 01.01.1800) bzw. unplausiblen Geburtsdatum (Alter > 109 Jahre), Patienten ohne Wohnortangabe und Patienten mit einer unplausibel hohen Anzahl an Arztkon-

takten (> 365 Tage pro Jahr) keine Berücksichtigung.

#### **Definition der Verdünnungsquote**

Die Folgen, welche die aufgeführten Faktoren auf die Populationsbildung haben, lassen sich in erster Linie daran erkennen, dass pro Jahr mehr Patienten beobachtbar sind, als Versicherte gemäß der amtlichen GKV-Statistik KM 6 zu erwarten sind (Bei der KM 6-Statistik handelt es sich um die einzige allgemein verfügbare Statistik, in der die Anzahl aller zum 1.7. eines Jahres gesetzlich krankenversicherter Personen, gegliedert nach Altersgruppe, Wohnort (KV), Versicherten-Status und Kassenart enthalten ist. Sie wird jährlich vom Bundesministerium für Gesundheit (BMG) veröffentlicht (1).).

Der Patientenüberschuss lässt sich aus dem Verhältnis der Anzahl der innerhalb eines Jahres beobachteten Patienten und der Anzahl GKV-Versicherter zum 1.7. eines Jahres quantifizieren und wird im Folgenden als Verdünnungsquote [%] bezeichnet. Auch wenn die Bezeichnung Verdünnungsquote zunächst kontraintuitiv erscheint (sie quantifiziert ja eigentlich einen Patientenüberschuss), so hat sich die Begrifflichkeit doch in der internen Diskussion durchgesetzt, da infolge dieses Überschusses Auswertungsergebnisse (zumindest bei Verdünnungsquoten > 100 %) in der Regel „verdünnt“ werden, also niedriger sind als sie bei korrekter Erfassung des Nenners tatsächlich ausfallen würden. Allerdings gibt es dabei auch Ausnahmen. So wird ein Ergebnis im Falle einer Verdünnungsquote von über 100 % immer dann zu niedrig ausfallen, wenn pro Patient aufsummierte Ereignisse dargestellt werden (z.B. Anzahl Leistungen pro Patient oder Leistungsbedarf pro Patient), weil sich in diesem Fall nur der Nenner und nicht der Zähler verändert. Bei der Bestimmung von Anteilen (z.B. Prävalenzen) sind im Falle von Verdünnungsquoten von über 100 % hingegen zwei Szenarien mit unterschiedlichen Folgen denkbar, die am Beispiel der Ermittlung der Diabetes-Jahresprävalenz verdeutlicht werden soll:

Szenario 1: Für einen Patienten wird im ersten Halbjahr des Jahres 2008 die Diagnose E 11 (Typ-II-Diabetes) kodiert. Der Patient wechselt im zweiten Halbjahr die Krankenkasse (erscheint somit als „neuer Patient“ in den Daten) und er-

hält als „neuer Patient“ keinen erneuten Diagnoseeintrag E11. Bei der Bestimmung der Diabetes-Jahresprävalenz (Anteil Typ-II-Diabetiker an allen beobachtbaren Patienten eines Jahres) würde somit der Nenner um eine Person erhöht werden, während der Zähler gleich bliebe. Szenario 1 führt somit zu einer Verdünnung der Jahresprävalenz.

Szenario 2: Für einen Patienten wird im ersten Halbjahr des Jahres 2008 die Diagnose E 11 kodiert. Der Patient wechselt im zweiten Halbjahr die Krankenkasse (erscheint somit als „neuer Patient“ in den Daten) und erhält als „neuer Patient“ ebenfalls einen Diagnoseeintrag E11. Bei der Bestimmung der Diabetes-Jahresprävalenz würden somit sowohl der Nenner als auch der Zähler um eine Person erhöht werden. Szenario 2 führt somit trotz einer Verdünnungsquote von 100 % nicht zu einer Verdünnung der Jahresprävalenz.

Da davon auszugehen ist, dass im Falle chronischer Erkrankungen wie Diabetes, Asthma etc. grundsätzlich in mehreren Quartalen eines Jahres Diagnoseeinträge für einen Patienten zu finden sind, sollte angenommen werden, dass der in Szenario 1 beschriebene Verdünnungseffekt eher bei akuten Erkrankungen (einmalige Diagnosestellung) auftreten wird.

Zusammenfassend heißt dies, dass bei Verdünnungsquoten über 100 % immer dann eine Unterschätzung eines Untersuchungsergebnisses zu erwarten ist, wenn pro Patient aufsummierte Ereignisse oder Prävalenzen seltener Ereignisse dargestellt werden.

## Ergebnisse

### Anzahl Patienten im Abrechnungsjahr 2007 und 2008

Die bundesweiten vertragsärztlichen Abrechnungsdaten umfassen in ihrer ursprünglichen Fassung  $N = 77.276.648$  Patientenentitäten für das Abrechnungsjahr 2007 und  $N = 79.219.078$  Patientenentitäten für das Jahr 2008 (statt Patientenentitäten wird im Weiteren die Bezeichnung Patienten verwendet).  $N = 3.161.944$  Patienten (4 %) des Jahres 2007 erfüllten mindestens eines der Ausschlusskriterien, so dass die Untersuchungspopulation des Jahres 2007  $N = 74.114.704$  Patienten umfasst. Da die Daten des Jahres 2008 deutlich mehr Patienten ent-

halten als die Daten des Jahres 2007, mussten für das Jahr 2008 zwar absolut betrachtet mehr Patienten ausgeschlossen werden ( $n = 3.198.318$ ), relativ betrachtet handelt es dabei aber um den gleichen Ausschlussanteil wie im Jahr 2007 (4 %). Insgesamt reduzierte sich in 2008 die Ausgangspopulation demnach von  $N = 79.219.078$  auf  $N = 76.020.760$  Patienten.

### Verdünnungsquote 2007

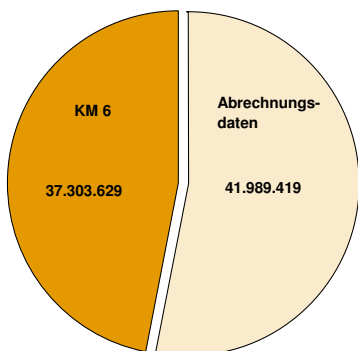
Die Auswertung zur Verdünnungsquote zeigt für das Jahr 2007, dass es sowohl eine regionale, als auch eine geschlechts- und altersspezifische Abhängigkeit gibt. Deutlich wird auch, dass eine Abhängigkeit von der Krankenkassenart besteht.

Laut KM 6-Statistik waren im Jahr 2007  $N = 70.314.011$  Personen gesetzlich krankenversichert. Dem gegenüber steht eine Patientenpopulation von insgesamt  $N = 74.114.704$  in den Abrechnungsdaten, so dass sich für das Jahr 2007 eine Verdünnungsquote von 105,4 % ergibt.

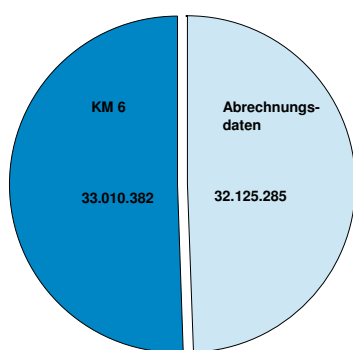
### Verdünnungsquote 2007 nach Geschlecht

Abbildung 1 zeigt die Anzahl weiblicher Versicherter (KM 6) und die Anzahl weiblicher Patienten (Abrechnungsdaten) auf Basis der Untersuchungspopulation. Der Abbildung ist zu entnehmen, dass nach der KM 6-Statistik  $N = 37.303.629$  weibliche GKV-Versicherte gemeldet sind, in den Abrechnungsdaten jedoch  $N = 41.989.419$  weibliche Patienten auftreten. Daraus ergibt sich eine Verdünnungsquote von 112,6 %.

Wie Abbildung 2 zeigt, liegt demgegenüber die Anzahl männlicher Patienten in den Abrechnungsdaten mit  $N = 32.125.285$  deutlich unter der Anzahl männlicher GKV-Versicherter ( $N = 33.010.382$ ), sodass sich für die männlichen Patienten eine Verdünnungsquote von 97,3 % ergibt.



**Abbildung 1: Gegenüberstellung Anzahl weiblicher Versicherter und Patienten für 2007**



**Abbildung 2: Gegenüberstellung Anzahl männlicher Versicherter und Patienten für 2007**

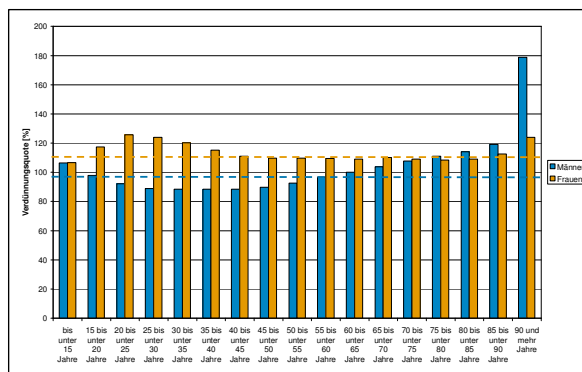
### Verdünnungsquote 2007 nach Altersklassen und Geschlecht

In *Abbildung 3* ist die Verdünnungsquote alters- und geschlechtsspezifisch für das Jahr 2007 dargestellt. Bei den Frauen zeigt sich, dass die Verdünnungsquote im überwiegenden Teil der Altersklassen unter dem Mittelwert der Frauen (112,5 %) liegt. Neben einer erhöhten Verdünnungsquote in der Altersklasse *90 und mehr Jahre*, fallen insbesondere die Klassen von *15 bis unter 40 Jahre* mit einer relativ hohen Verdünnungsquote auf. Mit 125,8 % liegt die Altersklasse *20 bis unter 25 Jahre* noch über der Altersklasse *90 und mehr Jahre* (124,0 %).

Im Vergleich zu den Frauen zeigt sich bei den Männern ein etwas anderer Zusammenhang zwischen der Höhe der Verdünnungsquote und dem Alter. So liegen die Verdünnungsquoten in den Altersgruppen *15 bis unter 60 Jahre* unter dem Mittelwert der Männer (97,3 %). Neben einer erhöhten Verdünnungsquote bei den unter 15jährigen zeigt sich eine stetig ansteigende erhöhte Verdünnungsquote in den Altersgruppen

ab 55 Jahren, die ihr Maximum mit 178,8 % in der Altersgruppe *90 und mehr Jahre* findet.

Im direkten Altersgruppenvergleich zeigt sich darüber hinaus, dass die Verdünnungsquote der Männer in den Altersklassen von *15 bis 65 Jahre* unter der der Frauen liegt, später jedoch (ab 75 Jahren) die der Frauen übersteigt.



**Abbildung 3: Verdünnungsquote 2007 dargestellt nach Alter und Geschlecht in %.**

Gestrichelte Linien kennzeichnen den geschlechtsspezifischen Mittelwert über alle Altersklassen

### Verdünnungsquote 2007 nach Region

*Abbildung 4* zeigt die Verdünnungsquote auf Basis der KVen insgesamt (flächenhafte Darstellung), sowie differenziert nach Geschlecht. Es wird deutlich, dass die KVen Rheinland-Pfalz, Hamburg, Baden-Württemberg und Mecklenburg-Vorpommern die höchsten Verdünnungsquoten aller KVen aufweisen. Mit einer Verdünnungsquote von 127,8 % nimmt Rheinland-Pfalz die Spitzenposition unter den KVen ein. Am unteren Ende der Skala liegt Sachsen als einzige KV mit einer Verdünnungsquote < 100 % (98,1 %).

Betrachtet man die geschlechtsspezifische Verdünnungsquote der KVen, fällt auf, dass Rheinland-Pfalz sowohl bei den Frauen (137,2 %), als auch bei den Männern (117,1 %) die höchste Verdünnungsquote aufweist. Für Sachsen sind demgegenüber für Männer und Frauen die niedrigsten Verdünnungsquoten zu beobachten (90,8 bzw. 104,5 %). Grundsätzlich zeigt sich in allen KVen eine höhere Verdünnungsquote bei den Frauen als bei den Männern, wobei einzig in Bremen die jeweiligen Verdünnungsquoten nahezu gleich sind (Frauen 109,1 und Männer 107,3 %). Während die Verdünnungsquote der Frauen in allen KVen



über 100 % liegt, ist dies bei Männern nur in den KVen Rheinland-Pfalz (117,1 %) und Bremen (109,1 %) der Fall. In Hamburg, Baden-Württemberg und Bayern liegen die Verdünnungsquoten männlicher Patienten nur geringfügig über 100 % (100,2 - 100,5 %).

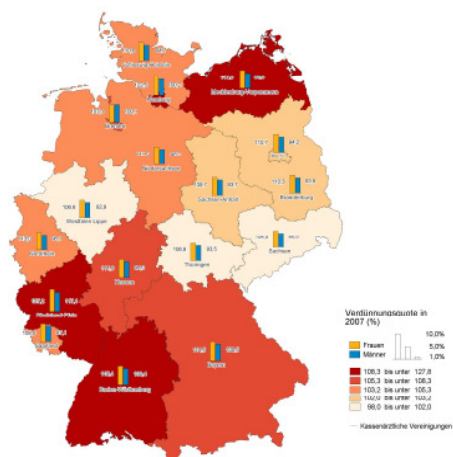


Abbildung 4: Verdünnungsquote 2007 auf KV-Ebene und nach Geschlecht in %.

### Verdünnungsquote 2008

Laut KM 6-Statistik waren im Jahr 2008  $N = 70.243.851$  Personen gesetzlich krankenversichert. Dem gegenüber steht eine Patientenpopulation von insgesamt  $N = 76.020.760$ , so dass sich für das Jahr 2008 eine Verdünnungsquote von 108,2 % ergibt.

### Verdünnungsquote 2008 nach Geschlecht

Abbildung 5 zeigt die Anzahl weiblicher Versicherter (KM 6) und die Anzahl weiblicher Patienten (Abrechnungsdaten) im Jahr 2008 auf Basis der Untersuchungspopulation. Insgesamt werden in den Abrechnungsdaten  $n = 5.915.165$  Frauen mehr beobachtet als nach KM 6-Statistik für das Jahr 2008 gemeldet sind. Daraus resultiert für die Frauen im Jahr 2008 eine Verdünnungsquote von 115,9 %. Demgegenüber unterscheidet sich das Verhältnis von Patienten und Versicherten für das Jahr 2008 bei den Männern nur geringfügig. So stehen  $N = 33.007.651$  GKV-Versicherte  $N = 32.869.395$  Patienten aus den Abrechnungsdaten gegenüber. Daraus ergibt sich eine Verdünnungsquote von 99,6 %, das heißt einer relativ geringen Abweichung von Patienten zu Versicherten (Abbildung 6).

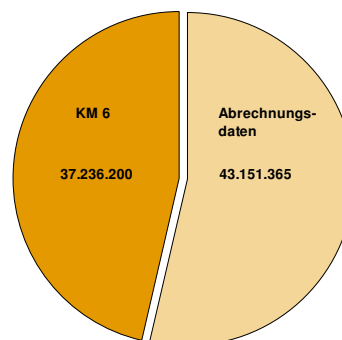


Abbildung 5: Gegenüberstellung Anzahl weiblicher Versicherter und Patienten für 2008

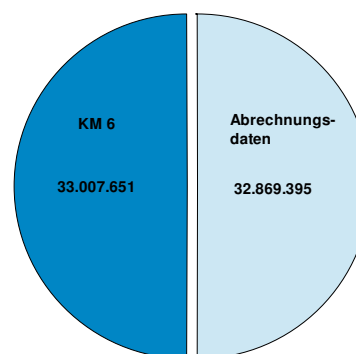


Abbildung 6: Gegenüberstellung Anzahl männlicher Versicherter und Patienten für 2008

### Verdünnungsquote 2008 nach Altersklassen und Geschlecht

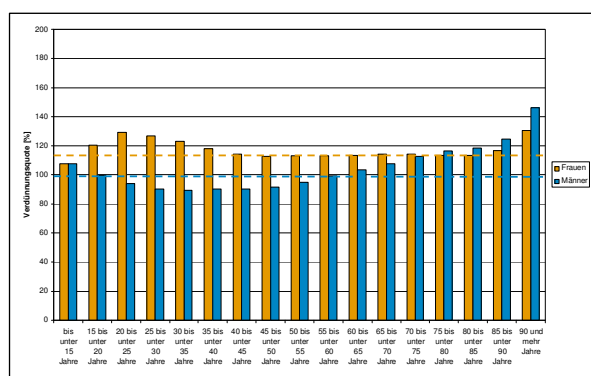
In *Abbildung 7* ist die Verdünnungsquote nach Alter für Frauen und Männer differenziert dargestellt. Das Mittel über alle Klassen liegt bei 115,9 % (Frauen) bzw. 99,6 % (Männer).

Bei den Frauen zeigt sich, dass die Verdünnungsquote in den Altersklassen *15 Jahre bis unter 40 Jahre* sowie in den Altersklassen ab 85 Jahren über dem Mittelwert der Frauen liegen (Maximum 126 % in der Altersklasse *90 und mehr Jahre*). Bis auf die Altersklasse *bis unter 15 Jahre* weisen alle sonstigen Altersklassen bei den Frauen Verdünnungsquoten auf, die nahe dem geschlechtsspezifischen Mittel liegen.

Wie bereits im Jahre 2007 ersichtlich war, zeigt sich für Männer auch im Jahr 2008 ein Zusammenhang zwischen der Verdünnungsquote und dem Alter, der nahezu konträr zur Verdünnungsquote der Frauen verläuft. So liegen die Verdünnungsquoten in den Altersgruppen *15 bis unter 60 Jahre* auf bzw. unter dem Mittelwert der Männer. Neben einer erhöhten Verdünnungs-

quote bei den unter 15jährigen zeigt sich eine stetig ansteigende Verdünnungsquote in den Altersgruppen ab 60 Jahren, die ihr Maximum mit 146 % in der Altersklasse 90 und mehr Jahre findet.

Im direkten Altersgruppenvergleich zeigt sich darüber hinaus, dass die Verdünnungsquote der Männer in den Altersgruppen *bis unter 75 Jahren* unter der der Frauen liegt, ab der Altersgruppe *75 und mehr Jahre* die der Frauen aber übersteigt. Dies wird ganz besonders in der Altersgruppe *90 und mehr Jahre* deutlich (146 % vs. 126 %)



**Abbildung 7: Verdünnungsquote 2008 dargestellt nach Alter und Geschlecht in %.**

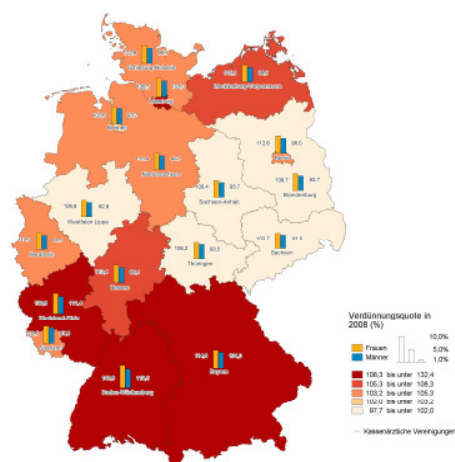
Gestrichelte Linien kennzeichnen den geschlechtsspezifischen Mittelwert über alle Altersklassen

### Verdünnungsquote 2008 nach Region

Abbildung 8 veranschaulicht die Verdünnungsquote auf Basis der KVen gesamt, sowie differenziert nach Männern und Frauen. Deutlich wird, dass die KVen Baden-Württemberg, Rheinland-Pfalz, Hamburg und Bayern im nationalen Vergleich besonders hohe Verdünnungsquoten aufweisen. So hat Baden-Württemberg mit 132,4 %, gefolgt von Rheinland-Pfalz mit 121,7 % die höchste Verdünnungsquote. Thüringen und Sachsen haben demgegenüber mit 100,2 % bzw. 97,7 % die niedrigste Verdünnungsquote aller KVen.

Betrachtet man die Verdünnungsquoten geschlechtsspezifisch, zeigt sich ein ähnliches Bild. Baden-Württemberg, Rheinland-Pfalz und Hamburg weisen sowohl bei den Männern (119,5; 111,4 bzw. 103 %), als auch bei den Frauen (143,8; 130,8 bzw. 126,5 %) die höchsten Verdünnungsquoten auf. Besonders

niedrige Verdünnungsquoten, insbesondere bei den Männern, zeigen die KVen Westfalen-Lippe (106,6 % (Frauen), 92,9 % (Männer)), Thüringen (106,2 % (Männer) 93,5 % (Frauen)) und Sachsen (103,7 % (Frauen), 91 % (Männer)). Grundsätzlich zeigt sich in allen KVen eine höhere Verdünnungsquote bei den Frauen als bei den Männern. Während die Verdünnungsquote der Frauen in allen KVen über 100 % liegt, ist dies bei den Männern nur in den KVen Bayerns (101 %), Hamburg (103 %), Rheinland-Pfalz (111,4 %) und Baden Württemberg (119,5 %) der Fall.



**Abbildung 8: Verdünnungsquote 2008 auf KV-Ebene und nach Geschlecht in %**

### Diskussion

Die Verdünnungsquoten der Jahre 2007 (105,4 %) und 2008 (108,2 %) veranschaulichen, dass die Anzahl der in den Abrechnungsdaten zu beobachtenden Patienten deutlich über der Anzahl gemäß KM 6-Statistik gemeldeter Versicherter liegt. Dabei zeigt sich, dass sich die Verdünnungsquote sowohl zwischen Männern und Frauen, als auch zwischen den verschiedenen Altersgruppen und im Ergebnis auch zwischen den einzelnen KVen unterscheidet.

So liegt die Verdünnungsquote der Frauen sowohl im Jahr 2007 als auch im Jahr 2008 mit 112,6 % bzw. 115,9 % deutlich über der Quote der Männer (97,3 % und 99,6 %). Eine Erklärung dafür könnte sein, dass Frauen grundsätzlich eher und zudem häufiger vertragsärztliche Leistungen in Anspruch nehmen (2,3) und somit für Frauen die Wahrscheinlichkeit höher ist, dass

Veränderungen in den Parametern Versichertennummer etc. zu der Verdünnungsproblematik führen (es ist zu erwarten, dass sich diese Parameter für Männer gleichermaßen ändern, da sie aber aufgrund geringerer Arztkontakte seltener in den Daten zu beobachten sind, fällt dies nicht derart ins Gewicht wie bei den Frauen). Bei Betrachtung der verschiedenen Altersgruppen zeigt sich, dass die Verdünnungsquote sowohl bei den Männern als auch bei den Frauen in der jüngsten Altersgruppe (bis unter 15 Jahre) und in der ältesten Altersgruppe (90 und mehr Jahre) besonders erhöht ist. Eine mögliche Erklärung dafür ist, dass im Falle unlesbarer oder fehlender Angaben zum Geburtsdatum die Defaultwerte „01.01.1900“ oder „01.01.2000“ als Angaben in den Daten übermittelt werden, die zur Erhöhung der Patientenzahl in der jüngsten bzw. ältesten Altersgruppe beitragen. Die insbesondere bei den Männern ab einem Alter von 75 Jahren deutlich zunehmenden und die der Frauen übersteigenden Verdünnungsquoten geben Hinweise darauf, dass der Erklärungsansatz, dass die Anzahl der Arztkontakte die Höhe der Verdünnungsquote beeinflusst, plausibel ist. So zeigen Auswertungen der BARMER GEK (2), dass die Arztkontaktraten von Männern ab dem 70. Lebensjahr die Kontaktraten der Frauen übersteigen, so dass in diesen Altersgruppen das Verdünnungsquotenproblem auch für die Männer in den Daten sichtbar wird.

Bei den Frauen ist darüber hinaus in den Altersgruppen 15 bis unter 40 Jahre eine besonders hohe Verdünnungsquote zu beobachten. Ein Grund dafür könnten erneut die im Vergleich zu Männern in dieser Altersklasse bei Frauen höheren Arztkontaktraten sein (v.a. gynäkologische Arztkontakte). Darüber hinaus ist die Kassenwechsellrate in dieser Altersgruppe bei Frauen höher als bei Männern (persönliche Angaben des Statistischen Bundesamtes auf Grundlage von Daten des Bundesversicherungsamtes des Jahres 2007). Da ein Kassenwechsel mit einer neuen Kombinationen von Versicherungsnummer, Krankenversicherungskarten-IK und Geburtsdatum einhergeht, kann dies zur Bildung neuer Patientenentitäten in dieser Altersgruppe führen.

Deutlich wird, dass sich die Faktoren Alter, Geschlecht und Krankenkassenart je nach regio-

ner Verteilung auf die Höhe der Verdünnungsquoten auf KV-Ebene auswirken und daher deutliche Unterschiede zwischen den KVen zu beobachten sind (Spanne der regionalen Verdünnungsquote: 98,1 % - 127,8 % (2007), bzw. 97,7 % - 132,4 % (2008)). So sind Regionen mit höherem Frauenanteil oder mit vielen Versicherten von Krankenkassen, die an Fusionen beteiligt sind, häufiger betroffen als Regionen, die entsprechend niedrigere Anteile aufweisen. Die Tatsache, dass die Verdünnungsquoten einiger KVen im Jahr 2008 deutlich von denen des Jahres 2007 abweichen, ist u.a. durch die Fusionen von Krankenkassen mit hohem regionalen Marktanteil (z.B. Baden-Württemberg (+ 18,2 %), Hamburg (+ 2,9 %)), bzw. der Fusion von KV-Regionalstellen zu erklären, bzw. hat ihre Ursache in einer Umstellung der Versichertennummersystematik innerhalb regional stark verteilter Krankenkassen. Dies erklärt auch den starken Abfall der Verdünnungsquoten in Rheinland-Pfalz und Bremen. Während Patienten in Rheinland-Pfalz im Jahr 2007 im Vergleich zum Jahr 2008 überdurchschnittlich stark von Krankenkassenfusionen bzw. der Fusion von KV-Regionalstellen betroffen waren, konnte für Patienten in Bremen im Jahr 2007 im Vergleich zum Jahr 2008 eine auffällige Häufung von neuen Versichertennummern bei männlichen Disease-Management-Programm-Teilnehmern in den Abrechnungsdaten beobachtet werden, die den insbesondere bei Männern zu beobachtenden starken Abfall der Verdünnungsquote im Jahr 2008 erklären könnten.

Darüber hinaus wird die Höhe der Verdünnungsquote auch durch die Art der Vergleichspopulation beeinflusst. So handelt es sich bei der aus den Abrechnungsdaten gebildeten Population um eine sogenannte Längsschnittpopulation, die angibt, wie viele Personen, an mindestens einem Tag eines Jahres vertragsärztliche Leistungen in Anspruch genommen haben. Die als Vergleichspopulation herangezogene GKV-Population gemäß KM 6-Statistik hingegen stellt eine Stichtagspopulation dar, die Angaben darüber macht, wie viele Personen am 1.7. eines Kalenderjahres gesetzlich krankenversichert sind. Diese Population setzt sich (ebenso wie die aus den Abrechnungsdaten gebildete Population) im Laufe eines Kalenderjahres ständig neu zusammen, da Versiche-



rungsverhältnisse unterjährig zu unterschiedlichen Zeitpunkten enden oder neu beginnen. Würde man statt der Stichtagspopulation hingegen eine Population zum Vergleich hinzuziehen, die alle Personen erfasst, die mindestens an einen Tag eines Kalenderjahres gesetzlich krankenversichert sind, so würde dies mehr Versicherte enthalten und somit der Quotient aus Patienten in den Abrechnungsdaten und gesetzlich Krankenversicherten spürbar geringer werden. So waren zum 1.7.2007 gemäß KM 6-Statistik beispielsweise  $N = 70.314.011$  Personen gesetzlich krankenversichert. Demgegenüber sind in die Berechnungen des BVA zum morbiditätsorientierten Risikostrukturausgleich im selben Jahr  $N = 74.070.004$  Versicherte einbezogen worden und somit genau die Personen, die im Jahr 2007 mindestens einen Tag gesetzlich krankenversichert waren (4) und nahezu identisch mit der Anzahl der in den Abrechnungsdaten enthaltenen Patienten ist ( $N = 74.114.704$ ). Somit wird deutlich, dass die KM 6-Statistik als Bezugspopulation zu einer Überschätzung der Verdünnungsquote führt und die dem BVA gemeldete Population besser geeignet wäre, um den methodisch bedingten Patientenüberschuss in den Abrechnungsdaten zu quantifizieren. Leider ist es auch nach Anfrage beim BVA nicht möglich, die grundsätzlich öffentlich zugänglichen Informationen zu dieser Längsschnittpopulation zusätzlich unterteilt in Region, Geschlecht, Alter und Kassenart zu erhalten. Diese Aufteilung wäre aber erforderlich, um die Verdünnungsquote ausreichend spezifizieren zu können. Aus diesem Grund konnte als Referenz nur die KM 6-Statistik verwendet werden, ohne die Validität dieser Datenquelle sicher bewerten zu können.

Auf der anderen Seite gibt es aber auch Faktoren, die wiederum zu einer Unterschätzung der Verdünnungsquote führen. So haben nach Angaben der BARMER GEK 8,5 % der Versicherten und nach Ergebnissen einer vom BKK Bundesverband in Auftrag gegebenen Umfrage sogar 20 % der gesetzlich Krankenversicherten innerhalb eines Jahres überhaupt keinen Kontakt zu einem niedergelassenen Arzt (2,3).

Da die Abrechnungsdaten nur Personen mit mindestens einem Arztkontakt enthalten, müsste die Vergleichspopulation entsprechend auf Personen reduziert werden, die ebenfalls

mindestens einen Arztkontakt im Bezugsjahr hatten. Da die beschriebenen Angaben zur Nicht-Inanspruchnahme-Rate aber nicht ausreichen, um diese valide aus den Bezugsdaten herausrechnen zu können, kann an dieser Stelle lediglich aufgezeigt werden, dass die Verdünnungsquote hierdurch eine Unterschätzung erfährt, die vermutlich den mit der Bezugspopulation verbundenen Überschätzungseffekt deutlich übersteigt. Da die Nicht-Inanspruchnahmerate bei Männern höher ist als bei Frauen (2), ist zu vermuten, dass dieser Unterschätzungseffekt bei Männern stärker ausfällt und dadurch ein Teil der Unterschiede in den geschlechtsspezifischen Verdünnungsquoten erklärt werden kann.

Darüber hinaus sei daraufhingewiesen, dass die in diesem Bericht für die Berechnung der Verdünnungsquote verwendete Population aus den Abrechnungsdaten nur eine Teilpopulation darstellt. So werden für Versorgungsforschungsfragen, die sich insbesondere mit regionalen Unterschieden beschäftigen, belastbare Angaben zu Alter, Geschlecht und regionaler Zuordnung der Patienten benötigt. In der Folge mussten alle Patienten, für die diesbezüglich keine validen Angaben vorlagen, aus der ursprünglichen Population ausgeschlossen werden. Demzufolge ist die Verdünnungsquote ohne Ausschluss dieser Patienten deutlich höher als in diesem Bericht beschrieben. Bei der Interpretation der Ergebnisse, die auf Grundlage dieser Untersuchungspopulation dargestellt werden, muss stets beachtet werden, dass mit dem Patientenausschluss möglicherweise auch Leistungen und Diagnosen ausgeschlossen werden, die theoretisch zu einem Patienten der Untersuchungspopulation gehören aber diesem nicht eindeutig zugeordnet werden können, so dass es in diesem Fall zu einer zusätzlichen Verdünnung der Auswertungsergebnisse kommen kann.

Grundsätzlich zeigen die beschriebenen Ergebnisse zur Verdünnungsquote, dass populationsbezogene Auswertungsergebnisse, die auf Grundlage der vertragsärztlichen Abrechnungsdaten erstellt werden, immer in Abhängigkeit von der Fragestellung und unter Beachtung der alters-, geschlechts- und regionsspezifischen Verdünnungsquote interpretiert werden müssen. So ist bei Vergleichen zwischen Männern und Frauen immer darauf zu achten, dass die Ergebnisse für Frauen (insbesondere bei der

Darstellung aufsummierter Ereignisse oder der Bestimmung der Prävalenz seltener Ereignisse) tendenziell zu niedrig („verdünnt“) ausfallen werden. Gleiches gilt bei der Betrachtung von Altersabhängigkeiten, bei der z.B. Ergebnisse für Patienten in der höchsten Altersgruppe in der Tendenz unterschätzt werden. Vor allem ist aber die Varianz der regionsspezifischen Verdünnungsquoten bei der Interpretation der für den Versorgungsatlas zentralen Ergebnisse zu populationsbezogenen und regionalen Unterschieden zu berücksichtigen. So ist im Jahr 2007 insbesondere bei den KVen Baden-Württemberg, Rheinland-Pfalz und Mecklenburg-Vorpommern darauf zu achten, dass bei einer Vielzahl von Fragestellungen die regionsspezifischen und populationsbezogenen Ergebnisse im Vergleich zu anderen KVen zu niedrig ausfallen

werden. Gleiches gilt im Jahr 2008 für die KV Bayerns und erneut für die KVen Baden-Württemberg und Rheinland-Pfalz.

Um das Problem der Verdünnungsquote für zukünftige Analysen zu reduzieren, wird die Methodik der Bildung von Patientenentitäten in der Kassenärztlichen Bundesvereinigung (KBV) zusammen mit dem ZI aktuell weiterentwickelt. Ein valider Patientenbezug in den Abrechnungsdaten, wie er nicht nur in der Versorgungsforschung, sondern auch von den KVen für die Erfüllung ihrer gesetzlichen Aufgaben benötigt wird, lässt sich allerdings erst dann herstellen, wenn die gesetzlichen Krankenkassen den KVen die Patientenstammdaten (v.a. Pseudonym der lebenslangen Versichertennummer und Versicherungszeiten) zur Verfügung stellen.

## Literatur

1. Bundesministerium für Gesundheit. Zahlen und Fakten zur gesetzlichen Krankenversicherung. <http://www.bmg.bund.de/krankenversicherung/zahlen-und-fakten-zur-krankenversicherung/zahlen-und-fakten.html> [Zugriff am 08.07.2011].
2. Maydell von B, Kosack T, Repschläger U, Sievers C, Zeljar R. Achtzehn Arztkontakte im Jahr. Hintergründe und Details. In: Repschläger U., Schulte C., Osterkamp N. (Hrsg): BARMER GEK Gesundheitswesen aktuell 2010. S. 179. BARMER GEK 2010.
3. BKK-Bundesverband. Bevölkerungsumfrage BKK 2011. <http://www.bkk.de/pressepolitik/presse/studien/arztbesuchumfrage-2011> [Zugriff am 07.07.2011]
4. Bundesversicherungsamt. Info-Dateien auf Kassenartenebene. [http://www.bundesversicherungsamt.de/cln\\_108/nn\\_1936018/DE/Risikostrukturausgleich/Info-Dateien\\_\\_und\\_\\_Auswertungen/Info-Dateien\\_\\_Kassenartenebene.html](http://www.bundesversicherungsamt.de/cln_108/nn_1936018/DE/Risikostrukturausgleich/Info-Dateien__und__Auswertungen/Info-Dateien__Kassenartenebene.html). [Zugriff am 08.07.2011].

## Hinweis

Die Langversion des Berichtes ist unter [www.versorgungsatlas.de](http://www.versorgungsatlas.de), im Themenbereich **Methodik der Versorgungsforschung** publiziert.