

**Bericht des Instituts des Bewertungsausschusses zur
Weiterentwicklung des Klassifikationssystems sowie
zur Ermittlung der Veränderungsraten für das Jahr 2021
gemäß § 87a Abs. 5 SGB V**

Freigegeben durch die Trägerorganisationen des Instituts des Bewertungsausschusses am
7. September 2020

Institut des
Bewertungsausschusses



Adresse und Kontakt

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

Zusammenfassung

Der Bewertungsausschuss beschließt jährlich Empfehlungen zur Anpassung der morbiditätsbedingten Gesamtvergütung aufgrund von Veränderungen der Morbiditätsstruktur der Versicherten. Die Empfehlungen des Bewertungsausschusses aus seiner 511. Sitzung vom 11. August 2020 zur Anpassung der morbiditätsbedingten Gesamtvergütungen für das Jahr 2021 beinhalten die demografischen und diagnosebezogenen Veränderungsrate je KV-Bezirk, die in Tabelle 1 dargestellt werden.

KV-Bezirk	Demografische Veränderungsrate	Diagnosebezogene Veränderungsrate
Schleswig-Holstein	0,2795 %	0,5156 %
Hamburg	-0,1076 %	-0,1173 %
Bremen	-0,0989 %	-0,3923 %
Niedersachsen	0,1643 %	0,3625 %
Westfalen-Lippe	0,1604 %	0,7504 %
Nordrhein	0,0766 %	0,6462 %
Hessen	0,0504 %	0,3712 %
Rheinland-Pfalz	0,1243 %	0,7854 %
Baden-Württemberg	0,0013 %	0,2277 %
Bayern	0,0931 %	0,5609 %
Berlin	-0,0367 %	-0,3892 %
Saarland	0,2918 %	0,8948 %
Mecklenburg-Vorpommern	0,3679 %	0,8985 %
Brandenburg	0,1916 %	0,3569 %
Sachsen-Anhalt	0,2331 %	0,8620 %
Thüringen	0,2504 %	0,8355 %
Sachsen	0,0108 %	0,5004 %

Tabelle 1: Demografische und diagnosebezogene Veränderungsrate je KV-Bezirk für das Jahr 2021.

Um die Transparenz über die der Berechnung der Veränderungsrate zugrunde liegenden Daten zu erhöhen, hat der Bewertungsausschuss erstmals auch einen Bericht nach § 87a Abs. 5a SGB V über die Veränderungen der Behandlungsdiagnosen und den Einfluss der jeweiligen Behandlungsdiagnose auf die Veränderungsrate für jeden Bezirk einer Kassenärztlichen Vereinigung erstellt. Dieser Bericht ist unter institut-ba.de veröffentlicht.

Inhaltsverzeichnis

A	Einführung und Grundlagen	6
1	Einführung	8
2	Rechtliche Bezugspunkte	9
3	Datengrundlagen	11
3.1	Geburtsstagsstichprobe der Berichtsjahre 2016 bis 2018	11
3.2	KM6-Daten	16
3.3	SV-Daten	16
3.4	ANZVER-Daten	16
B	Demografische Veränderungsraten	18
1	Vorgehen / Berechnungsschritte zur Berechnung der demografischen Veränderungsraten	20
1.1	Berechnung der Relativgewichte zur Berechnung der demografischen Veränderungsraten	20
1.2	Berechnung der demografischen Veränderungsraten	22
2	Ergebnisse: demografische Veränderungsraten.....	24
3	Empirische Ergebnisse zum Regressionsmodell zur Berechnung der demografischen Veränderungsraten und empirische Ergebnisse zu den demografischen Veränderungsraten	25
3.1	Empirische Ergebnisse zum Regressionsmodell zur Berechnung der demografischen Veränderungsraten	25
3.2	Empirische Ergebnisse zu den demografischen Veränderungsraten	28
C	Diagnosebezogene Veränderungsraten	34
1	Vorgehen / Berechnungsschritte zur Berechnung der diagnosebezogenen Veränderungsraten..	36
1.1	Zuordnung zu Risikokategorien (Klassifikation) zur Berechnung der diagnosebezogenen Veränderungsraten	36
1.2	Berechnung der Relativgewichte zur Berechnung der diagnosebezogenen Veränderungsraten	37
1.3	Berechnung der diagnosebezogenen Veränderungsraten	41
2	Ergebnisse: diagnosebezogene Veränderungsraten.....	48
3	Empirische Ergebnisse zum Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsraten und empirische Ergebnisse zu den diagnosebezogenen Veränderungsraten	50
3.1	Empirische Ergebnisse zum Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsraten	50
3.2	Empirische Ergebnisse zu den diagnosebezogenen Veränderungsraten	59
D	Weiterentwicklungen	68
1	Weiterentwicklung der Datengrundlage	70
2	Weiterentwicklungen zur Berechnung der demografischen Veränderungsraten	70

3 Weiterentwicklungen zur Berechnung der diagnosebezogenen Veränderungsdaten	70
3.1 Übersicht über die Weiterentwicklungen zur Berechnung der diagnosebezogenen Veränderungsdaten	71
3.2 Weiterentwicklung des Klassifikationssystems	75
3.3 Weiterentwicklung der Berechnung der Veränderungsdaten	90
Anhang 1: Abkürzungsverzeichnis	91
Anhang 2: Dokumentation zur Bereitstellung der regionalisierten Geburtstagsstichprobe 2016–2018 für die Gesamtvertragspartner	95

Teil A

Einführung und Grundlagen

1	Einführung	8
2	Rechtliche Bezugspunkte	9
3	Datengrundlagen	11
3.1	Geburtstagsstichprobe der Berichtsjahre 2016 bis 2018	11
3.1.1	Rechtliche Grundlagen der Geburtstagsstichprobe	11
3.1.2	Inhalt des Datenkörpers	12
3.1.3	Qualitätssicherung des Datenkörpers und Ausschlüsse	12
3.1.4	Erweiterung um abgeleitete Merkmale	13
3.2	KM6-Daten	16
3.3	SV-Daten	16
3.4	ANZVER-Daten	16

1 Einführung

Der Bewertungsausschuss beschließt entsprechend den gesetzlichen Vorgaben in § 87a SGB V in einem jährlichen Turnus und separat für jeden Bezirk einer Kassenärztlichen Vereinigung (KV-Bezirk) eine demografische und eine diagnosebezogene Veränderungsrate als Empfehlungen zur Vereinbarung von Veränderungen der Morbiditätsstruktur der Versicherten aller Krankenkassen mit Wohnort im jeweiligen KV-Bezirk. Die Veränderungsrate werden den Kassenärztlichen Vereinigungen und den Landesverbänden der Krankenkassen durch einen Beschluss des Bewertungsausschusses mitgeteilt¹.

Berechnet werden die einzelnen KV-spezifischen Veränderungsrate vom Institut des Bewertungsausschusses auf der Grundlage der Beschlüsse des Bewertungsausschusses oder ggf. des Erweiterten Bewertungsausschusses zu dem „geltenden Klassifikationsmodell“ gemäß § 87a Abs. 5 SGB V. Zur Dokumentation dieser Berechnungen hat der Bewertungsausschuss in seiner 269. Sitzung am 25. Januar 2012 das Institut des Bewertungsausschusses mit einer jährlichen Berichterstattung zum 31. August des betreffenden Jahres beauftragt. Diese Berichte sollen

1. die Dokumentation der Weiterentwicklung des Klassifikationssystems und
2. die Dokumentation der Bestimmung der diagnosebezogenen Veränderungsrate

beinhalten.

Die Berichterstattung dokumentiert die gegenüber dem jeweils zuletzt berichteten Stand vorgenommenen Änderungen

- der Zuordnung von Diagnosen zu Risikokategorien in der unkomprimierten Version,
- der daraus abgeleiteten Risikokategorien der komprimierten Version,
- am Komprimierungsalgorithmus und
- am Hierarchisierungsverfahren.

Ferner soll die Bestimmung der diagnosebezogenen Veränderungsrate mit dem beschlossenen Klassifikationsmodell dargestellt werden. Darüber hinaus wird im vorliegenden Bericht auch die Ermittlung der demografischen Veränderungsrate erläutert.

Mit dem vorliegenden Bericht kommt das Institut des Bewertungsausschusses seiner Dokumentationspflicht gemäß dem Beschluss des Bewertungsausschusses in seiner 269. Sitzung am 25. Januar 2012 nach, einschließlich der auf Arbeitsebene vereinbarten Integration des jährlichen Berichtes zum 31.05. (zur Weiterentwicklung des Klassifikationssystems). Die gemäß Beschluss des Bewertungsausschusses in seiner 332. Sitzung (schriftliche

¹Gemäß § 87a Abs. 5 Satz 1 Nr. 2 SGB V beschließt der Bewertungsausschuss Empfehlungen zur Vereinbarung von Veränderungen der Morbiditätsstruktur nach § 87a Abs. 4 Satz 1 Nr. 2 SGB V. Das Institut des Bewertungsausschusses berechnet dazu für jeden Bezirk einer Kassenärztlichen Vereinigung zwei einheitliche Veränderungsrate (eine diagnosebezogene Veränderungsrate und eine demografische Veränderungsrate). Die Empfehlungen der KV-spezifischen Veränderungsrate für das Jahr 2021 wurden durch den Bewertungsausschuss in seiner 511. Sitzung beschlossen.

Beschlussfassung) unter Nr. 8 vorgenommene Änderung des Beschlusses des Bewertungsausschusses in seiner 269. Sitzung bezogen auf die Berichtspflicht zur Erreichung der vorgegebenen Ziele wurde berücksichtigt und die vorgenommenen Ergänzungen der Anlagen um Relativgewichte und Prävalenzen der Risikokategorien wurden im Zuge des Beschlusses zum aktuellen Klassifikationsmodell (Beschluss des Bewertungsausschusses in seiner 503. Sitzung) in Form der sogenannten Transparenzdaten geliefert.

Der Bericht ist in folgende vier Hauptteile aufgeteilt:

- Teil A: Einführung / Grundlagen,
- Teil B: Demografische Veränderungsdaten,
- Teil C: Diagnosebezogene Veränderungsdaten und
- Teil D: Weiterentwicklungen.

Aus Gründen der leichten Lesbarkeit wird in diesem Bericht auf die gleichzeitige Verwendung männlicher und weiblicher Sprachformen verzichtet. Die ausschließliche Verwendung der männlichen Form dabei ist als geschlechtsunabhängig zu verstehen.

2 Rechtliche Bezugspunkte

Der Auftrag zur regelmäßigen Berichterstattung wurde dem Institut des Bewertungsausschusses mit dem Beschluss des Bewertungsausschusses in seiner 269. Sitzung am 25. Januar 2012 erteilt.

Für die Berichterstattung bzw. für die Weiterentwicklung des Klassifikationssystems und zur Berechnung der Veränderungsdaten müssen weitere rechtliche Grundlagen beachtet werden. An dieser Stelle sollen die wichtigsten rechtlichen Grundlagen, welche für die Ermittlung der demografischen und diagnosebezogenen Veränderungsdaten für das Jahr 2021 Anwendung fanden, mit ihren entsprechenden Regelungen überblicksartig aufgeführt werden:

§ 87a SGB V

In § 87a SGB V finden sich wesentliche Regelungen hinsichtlich der Regionalen Euro-Gebührenordnung, der morbiditätsbedingten Gesamtvergütung und des zu vereinbarenden Behandlungsbedarfs der gesetzlich Versicherten. Unter anderem ist festgelegt, dass der Bewertungsausschuss Empfehlungen zur Vereinbarung von Veränderungen der Morbiditätsstruktur beschließt. Der Gesetzgeber macht in § 87a Abs. 5 Satz 4 SGB V die Vorgabe, dass die Veränderungsdaten auf der Grundlage des Beschlusses des erweiterten Bewertungsausschusses vom 2. September 2009 bestimmt werden mit der Maßgabe, die Datengrundlagen zu aktualisieren.

Beschluss des Erweiterten Bewertungsausschusses in seiner 15. Sitzung am 2.

September 2009

Der Erweiterte Bewertungsausschuss hat in seiner 15. Sitzung am 2. September 2009 beschlossen, dass die durch den Bewertungsausschuss zu beschließende bundesweite Veränderungsrate aus

- einer diagnosebezogenen Komponente und
- einer demografischen Komponente

zu bestehen hat. Ferner wurde in dieser Sitzung beschlossen, dass

- in Zukunft notwendige medizinische und ökonomische Anpassungen vorgenommen werden. Dabei wurde festgelegt, dass das zunächst ausgewählte Modell kein Präjudiz für Entscheidungen in künftigen Jahren ist.

Dieses Klassifikationsmodell wird auch als H15EBA bezeichnet.

Beschluss des Bewertungsausschusses in seiner 503. Sitzung

Der Ausgangspunkt der Weiterentwicklung des Klassifikationsmodells ist das vom Erweiterten Bewertungsausschuss in seiner 29. Sitzung am 19./25. Juni 2012 beschlossene und durch die Beschlüsse des Erweiterten Bewertungsausschusses in seiner 30. Sitzung am 15./30. August 2012 sowie des Bewertungsausschusses in seiner 283. Sitzung am 30. August 2012 und in seiner 288. Sitzung am 22. Oktober 2012 geänderte Klassifikationsmodell (bezeichnet als KM87a_2012). Der Bewertungsausschuss beschloss in den folgenden Jahren Weiterentwicklungen des Klassifikationsmodells

- in seiner 309. Sitzung am 26. Juni 2013 und zusätzlich in seiner 313. Sitzung am 19. August 2013 für die Veränderungsrate für das Jahr 2014 (KM87a_2013),
- in seiner 332. Sitzung (schriftliche Beschlussfassung) für die Veränderungsrate für das Jahr 2015 (KM87a_2014),
- in seiner 355. Sitzung am 23. Juni 2015 für die Veränderungsrate für das Jahr 2016 (KM87a_2015),
- in seiner 382. Sitzung am 31. August 2016 für die Veränderungsrate für das Jahr 2017 (KM87a_2016),
- in seiner 400. Sitzung am 31. August 2017 für die Veränderungsrate für das Jahr 2018 (KM87a_2017) und
- in seiner 430. Sitzung am 12. Dezember 2018 für die Veränderungsrate für das Jahr 2019 (KM87a_2018).
- in seiner 439. Sitzung am 19. Juni 2019 für die Veränderungsrate für das Jahr 2020 (KM87a_2019).

Die Festlegungen für das zur Ermittlung der diagnosebezogenen bzw. demografischen Veränderungsrate für das Jahr 2021 zu verwendende Klassifikationsmodell (KM87a_2020) hat der Bewertungsausschuss in seiner 503. Sitzung (schriftliche Beschlussfassung) be-

schlossen.

3 Datengrundlagen

Gemäß dem Beschluss des Bewertungsausschusses in seiner 503. Sitzung bildet eine bundesweit erhobene Versichertenstichprobe (Geburtstagsstichprobe, kurz GSP) die Datengrundlage.

Im Folgenden werden dementsprechend die Geburtstagsstichprobe der für die Berechnung der demografischen und diagnosebezogenen Veränderungsdaten für das Jahr 2021 relevanten Berichtsjahre 2016 bis 2018 in ihrer verwendeten Form (d.h. inklusive vorgenommener Ausschlüsse und abgeleiteter Merkmale), die amtliche KM6-Statistik, selektivvertragliche Teilnahmedaten (SV-Daten) und die ANZVER-Daten (Daten der Satzart ANZVER87a) vorgestellt.

Weitere Abgrenzungen der Datengrundlage zur Berechnung der Relativgewichte und zur Berechnung der Veränderungsdaten werden in den Teilen B und C beschrieben.

3.1 Geburtstagsstichprobe der Berichtsjahre 2016 bis 2018

Für die Berechnung der Veränderungsdaten für das Jahr 2021 wird die Geburtstagsstichprobe der Berichtsjahre 2016 bis 2018² verwendet. Dieser Datenkörper, einschließlich aller Austausch- und Korrekturlieferungen, wurde mit dem Stand der Qualitätssicherung vom 24. März 2020 verwendet.

3.1.1 Rechtliche Grundlagen der Geburtstagsstichprobe

Gemäß dem Beschluss des Bewertungsausschusses in seiner 348. Sitzung (schriftliche Beschlussfassung) und dem Beschluss des Bewertungsausschusses in seiner 437. Sitzung (schriftliche Beschlussfassung) liegen dem Institut des Bewertungsausschusses bundesweite Versichertenstichproben für die Berichtsjahre 2016 bis 2018 vor. Die Auswahl der Versicherten in der Stichprobe erfolgt anhand ihres Geburtskalendertages (daher die Bezeichnung „Geburtstagsstichprobe“). Es werden jeweils die Versicherten für die Stichprobe ausgewählt, die an einem festgelegten Kalendertag eines beliebigen Monats Geburtstag haben. Tabelle 1 listet diese festgelegten Geburtskalendertage auf. Für das Jahr 2016 sind Versicherte mit Geburtskalendertag an einem 5., 11., 12., 18., 19., 25. oder 26. in der Stichprobe. Für das Jahr 2017 fällt gegenüber dem Jahr 2016 der 11. Kalendertag weg und der 6. kommt hinzu. Für das Jahr 2018 schließlich fällt gegenüber dem Jahr 2017 der 18. Kalendertag weg und der 13. kommt hinzu.

²Im Rahmen der Zusammenführung der kassenseitigen und arztseitigen Datenlieferungen zur Geburtstagsstichprobe (Matching) wurde zusätzlich auch auf Daten der Geburtstagsstichprobe älterer Berichtsjahre zurückgegriffen (vgl. dazu Abschnitt 6.3 der „Dokumentation zur Bereitstellung der regionalisierten Geburtstagsstichprobe 2016-2018 für die Gesamtvertragspartner“ im Anhang dieses Berichtes).

Beschluss	Zu liefernde Berichtsjahre der GSP	Ausgewählte Geburtskalendertage
348. BA	2016	5, 11, 12, 18, 19, 25, 26
348. BA	2017	5, 6, 12, 18, 19, 25, 26
437. BA	2018	5, 6, 12, 13, 19, 25, 26

Tabelle 1: Überblick über die Beschlüsse, welche die Festlegungen zur Lieferung der Daten zur bundesweiten Versichertenstichprobe beinhalten

3.1.2 Inhalt des Datenkörpers

Im Rahmen der Datenlieferungen zur Geburtstagsstichprobe werden für alle gesetzlich Versicherten mit den für das jeweilige Berichtsjahr ausgewählten Geburtskalendertagen pseudonymisierte Versichertenstammdaten und Kostenträgerstammdaten durch die Kassenseite sowie vertragsärztliche Abrechnungsdaten durch die Ärzteseite an die Datenstelle des Instituts des Bewertungsausschusses übermittelt. Die Erhebung ist damit, bezogen auf die ausgewählten Geburtskalendertage, versichertenvollständig konzipiert.

Die Versichertenstammdaten werden durch das sogenannte Matching-Verfahren mit den ärztlichen Abrechnungsdaten zusammengeführt und eindeutig dem Pseudonym der lebenslangen Versichertennummer (der sogenannten „Personen-ID“) zugeordnet. Durch dieses Verfahren entsteht für jeden Versicherten³ der Geburtstagsstichprobe ein Datensatz, der Angaben sowohl über z. B. Alter, Geschlecht, Wohnort und Zahl der Versichertentage als auch über die kodierten Diagnosen und die abgerechneten Leistungen enthält. Eine ausführliche Beschreibung der Geburtstagsstichprobe einschließlich des oben erwähnten Matching-Verfahrens findet sich in Abschnitt 6 der „Dokumentation zur Bereitstellung der regionalisierten Geburtstagsstichprobe 2016-2018 für die Gesamtvertragspartner“ im Anhang dieses Berichtes.

3.1.3 Qualitätssicherung des Datenkörpers und Ausschlüsse

Ebenfalls im Anhang des vorliegenden Berichts ist die Qualitätssicherung der Geburtstagsstichprobe ausführlich dargestellt (siehe Abschnitt 2.3 der „Dokumentation zur Bereitstellung der regionalisierten Geburtstagsstichprobe 2016-2018 für die Gesamtvertragspartner“). Insbesondere werden dort die Gründe aufgeführt, die dazu führen, dass ein Versicherter mit einer Kennzeichnung markiert wird, die im Rahmen der Qualitätssicherung festgestellte Auffälligkeiten anzeigt und zum Ausschluss aus der Datengrundlage führt (Ausschlusskennzeichnung). Personen mit Kostenübernahme nach § 264 SGB V (sogenannte „Betreute“) wurden in der Geburtstagsstichprobe wie im Vorjahr mit einer

³Genauer müsste es heißen „für jedes Versichertenpseudonym [entsteht] ein Datensatz“, da im Einzelfall auch mehrere Pseudonyme für einen Versicherten in den Daten vorhanden sein können. Vereinfachend wird in diesem Bericht immer der Begriff „Versicherter“ verwendet.

Ausschlusskennzeichnung markiert. Darüber hinaus wurde im Rahmen der Qualitätssicherung der Geburtstagsstichprobe die Selektivvertragskennzeichnung einschließlich der Vertragsartkennzeichnung gesetzt (vgl. hierzu Abschnitt 6.5 der „Dokumentation zur Bereitstellung der regionalisierten Geburtstagsstichprobe 2016-2018 für die Gesamtvertragspartner“).

Genauer wurden im Beschluss des Bewertungsausschusses in seiner 503. Sitzung für die Berechnung der diagnosebezogenen Veränderungsraten für das Jahr 2021 folgende Festlegungen getroffen:

- Versicherte, die im Rahmen der Qualitätssicherung zum Stand 24. März 2020 mit einer Ausschlusskennzeichnung markiert sind, werden ausgeschlossen. Ebenso werden alle Abrechnungsdaten (Abrechnungsfälle samt ihren Leistungs- und Diagnoseangaben) dieser Versicherten ausgeschlossen.
- Es werden alle Abrechnungsfälle samt ihren Leistungs- und Diagnoseangaben aus der Datengrundlage ausgeschlossen, die einen Leistungsbedarf von null aufweisen.

3.1.4 Erweiterung um abgeleitete Merkmale

Gemäß Nr. 2.2 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung wird die Geburtstagsstichprobe u. a. um folgende abgeleitete Merkmale erweitert:

- KV-Zuordnung
- Anzahl Versichertentage
- Anzahl Versichertenquartale
- Kennzeichen Versichertenzeitvollständigkeit
- MGV-Leistungsbedarf
- Kennzeichen längsschnittliche Selektivvertragsteilnahme
- Kennzeichen Wohnausländer

Die nachfolgenden Erläuterungen dazu entsprechen den Ausführungen des genannten Beschlusses. Wortwörtlich aus diesem Beschluss übernommene Textpassagen sind kursiv gesetzt.

KV-Zuordnung

Die regionale Zuordnung von Versicherten zu KV-Bezirken erfolgt jahresbezogen für einen Versicherten auf Grundlage des letzten Quartals im jeweiligen Jahr, in dem für diesen Versicherten mindestens ein Datensatz der Satzart 201 in der Datengrundlage vorliegt. Für die Zuordnung wird zunächst allen vorliegenden Datensätzen der Satzart 201 aus diesem Quartal für den Versicherten eine KV-Nummer wie folgt zugeordnet:

- *Enthält das Feld 08 nicht den Wert „AUSLA“, wird die Nummer des im Feld 15 angegebene-*

nen KV-Bezirks zugeordnet.

- Enthält das Feld 08 dagegen den Wert „AUSLA“, so wird die Nummer des KV-Bezirks zugeordnet, in dem die über das Institutionskennzeichen im Feld 02 der Satzart 201 zu ermittelnde Krankenkasse der höchsten Fusionsstufe (Stand: 2. April 2020) ihren Sitz hat.

Bei der jahresbezogenen Zuordnung einer KV-Nummer zu dem Versicherten wird anschließend wie folgt vorgegangen:

1. Wurde durch obige Vorschrift allen im betrachteten Quartal für den Versicherten vorliegenden Datensätzen der Satzart 201 dieselbe KV-Nummer zugeordnet, so wird diese KV-Nummer dem Versicherten zugeordnet.

2. Im Fall, dass durch obige Vorschrift nicht allen vorliegenden Datensätzen dieselbe KV-Nummer zugeordnet wurde, es aber unter den zugeordneten KV-Nummern genau eine gibt, die für den betreffenden Versicherten im Vorquartal gemäß dem obigen Verfahren nicht zugeordnet werden kann, so wird dem Versicherten diese KV-Nummer zugeordnet.

3. In den verbleibenden Fällen, in denen weder durch 1. noch durch 2. eine eindeutige Zuordnung erfolgen kann, werden alle vorliegenden Datensätze der Satzart 201 aus dem Quartal betrachtet, für die im Feld 04 die größte Anzahl an Versichertentagen gemeldet wurde. Ist allen diesen Datensätzen dieselbe KV-Nummer zugeordnet, so wird diese dem Versicherten zugeordnet.

4. Bei immer noch verbliebener Mehrdeutigkeit erfolgt die Zuordnung des Versicherten zu der KV-Nummer des Datensatzes der Satzart 201 aus dem Quartal, für den der mit dem MD5-Hashverfahren aus der Konkatenation von Postleitzahl (Feld 08), PersonenID (Feld 03) und der dem Datensatz zugeordneten KV-Nummer erhaltene Hashwert am größten ist.

Anzahl Versichertentage

Die Zahl der Versichertentage eines Versicherten wird je Quartal aufaddiert und auf die kalendarisch maximal mögliche Anzahl an Tagen dieses Quartals begrenzt.

Anzahl Versichertenquartale

Die Anzahl der Versichertenquartale (AVQ) weist jahresweise aus, wie viele Quartale ein Versicherter bei einer gesetzlichen Krankenkasse versichert war. Ein Quartal wird hierbei gezählt, sobald ein Versicherter in diesem Quartal mindestens einen Tag versichert war.

Kennzeichen Versichertenzeitvollständigkeit

Ein (gesetzlich) Versicherter wird jahresbezogen als versichertenzeitvollständig gekennzeichnet, wenn er in jedem Quartal dieses Jahres mindestens 45 Tage in der gesetzlichen Krankenversicherung (GKV) versichert war. Von dieser Bedingung ausgenommen sind das Quartal des Todes eines Versicherten und ggf. die Folgequartale im Kalenderjahr des Todes.

MGV-Leistungsbedarf

Der Leistungsbedarf, der für die Bestimmung der sogenannten Relativgewichte (siehe die Abschnitte 1.1 in Teil B und 1.2 in Teil C) verwendet wird, ist die für einen Versicherten im Verlauf eines Jahres abgerechnete Leistungsmenge in Punkten. Er beruht auf der Abrechnung vertragsärztlicher Gebührenordnungspositionen nach sachlich-rechnerischer Richtigstellung, ohne Berücksichtigung von Honorarverteilungsmaßstäben. Er ist eingegrenzt auf Leistungen, die der Bewertungsausschuss für seine Empfehlungen der morbiditätsbedingten Gesamtvergütung (MGV) zurechnet.

Die bundeseinheitliche Abgrenzung zwischen den Leistungen der MGV, die bei der Bestimmung von Relativgewichten herangezogen werden, und den Leistungen der sogenannten extrabudgetären Gesamtvergütung (EGV) wird durch die Anlage zum Beschluss des Bewertungsausschusses in seiner 503. Sitzung geregelt. Technisch erfolgt die Ermittlung dieser Leistungen wie folgt:

1. Jede abgerechnete Gebührenordnungsposition wird einem zugehörigen Leistungssegment zugeordnet.
2. Anschließend wird jedes Leistungssegment der MGV oder der EGV zugewiesen.
3. Für die Berechnung der Relativgewichte werden die Leistungssegmente (bzw. die den Leistungssegmenten zugeordneten Leistungen) verwendet, welche der MGV zugeordnet wurden.

Kennzeichen längsschnittliche Selektivvertragsteilnahme

In Nr. 2.2.7 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung wird ein längsschnittlicher Selektivvertragsteilnehmer (SV-Teilnehmer) definiert als Versicherter, der zu irgendeinem Zeitpunkt der Jahre 2017 und 2018 als SV-Teilnehmer in einem bereinigungsrelevanten Selektivvertrag identifiziert wurde. Für die Berechnung der Veränderungsrate für das Jahr 2021 sind dementsprechend als längsschnittliche Nicht-SV-Teilnehmer diejenigen Versicherten definiert, die zu keinem Zeitpunkt der Jahre 2017 und 2018 als SV-Teilnehmer in einem bereinigungsrelevanten Selektivvertrag identifiziert wurden. Die Identifikation der SV-Teilnehmer erfolgt im Rahmen der Qualitätssicherung der Geburtstagsstichprobe (vgl. hierzu Abschnitt 6.5 der „Dokumentation zur Bereitstellung der regionalisierten Geburtstagsstichprobe 2016-2018 für die Gesamtvertragspartner“).

Kennzeichen Wohnausländer

Die Identifikation eines Versicherten als Wohnausländer erfolgt jahresbezogen auf Grundlage des letzten Quartals im jeweiligen Jahr, in dem für diesen Versicherten mindestens ein Datensatz der Satzart 201 in der Datengrundlage vorliegt. Aus diesem Quartal werden für den Versicherten diejenigen Datensätze der Satzart 201 ausgewählt, bei denen die über das Institutionskennzeichen im Feld 02 zu ermittelnde Krankenkasse der höchsten Fusionsstufe (Stand: 2. April 2020) ihren Sitz im ermittelten KV-Bezirk hat (siehe Abschnitt KV-Zuordnung). Enthält einer dieser Datensätze in Feld 08 den Wert "AUSLA", so wird der

Versicherte als Wohnausländer identifiziert.

3.2 KM6-Daten

Die KM6-Statistik ist eine amtliche Stichtagsstatistik, welche, jeweils für den 1. Juli eines Jahres, für 17 Altersgruppen getrennt nach Geschlecht die Zahl der gesetzlich Versicherten aufführt. Sie steht dem Institut auf Ebene der Einzelkasse je Kassenärztlicher Vereinigung zur Verfügung. Versicherte mit Wohnsitz im Ausland werden demjenigen Bezirk einer Kassenärztlichen Vereinigung zugeordnet, in dem ihre Krankenkasse ihren Sitz hat. Bei fusionierten Krankenkassen wird dabei jeweils der letzte bekannte Sitz der Krankenkasse (Stand hier: 2. April 2020) verwendet. Die Übermittlung dieser Datengrundlage erfolgt gemäß Beschluss des Bewertungsausschusses in seiner 348. Sitzung (schriftliche Beschlussfassung). Die KM6-Statistik wird gemäß Nr. 2.7 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung zur Berechnung der sogenannten demografischen Hochrechnungsfaktoren (DHF) verwendet, welche der Hochrechnung der Geburtstagsstichprobe auf die Grundgesamtheit aller gesetzlich Versicherten dienen.

3.3 SV-Daten

Gemäß dem Beschluss des Erweiterten Bewertungsausschusses in seiner 51. Sitzung am 29. März 2017, dem Beschluss des Erweiterten Bewertungsausschusses in seiner 56. Sitzung am 21. August 2018 und gemäß dem Beschluss des Erweiterten Bewertungsausschusses in seiner 61. Sitzung (Teil A) am 29. März 2019 werden selektivvertragliche Teilnahmedaten (SV-Daten) für Verträge, die gemäß § 73b SGB V, §§ 73c und 140a SGB V a.F. und § 140a SGB V i.d.F. des GKV-VSG geschlossen wurden, an das Institut des Bewertungsausschusses übermittelt.

Die SV-Daten werden im Institut des Bewertungsausschusses einer Qualitätssicherung unterzogen. Für die Ermittlung der demografischen und diagnosebezogenen Veränderungsraten für das Jahr 2021 wurden die SV-Daten mit dem Stand der Qualitätssicherung zum 24. März 2020 verwendet.

Ein Teil der SV-Daten liegt als Vollerhebung vor (z. B. Vertragsstammdaten). Daten zu teilnehmenden Versicherten liegen dagegen nur als Versichertenstichprobe vor, wobei für die Auswahl der Versicherten in der Stichprobe für jedes Berichtsjahr dieselben Geburtskalendertage verwendet werden wie für die Geburtstagsstichprobe.

3.4 ANZVER-Daten

Mit den ANZVER87a-Daten liegt dem Institut des Bewertungsausschusses eine weitere Datengrundlage mit Angaben zu Anzahlen von Versicherten vor. In den ANZVER87a-Daten (geliefert beginnend mit dem zweiten Quartal 2012) wurden für jede Krankenkasse und Wohnort-KV die monatlichen Anzahlen von Versicherten, Betreuten und Wohnausländern

gemeldet. Die Datenlieferungen erfolgen gemäß dem Beschluss des Bewertungsausschusses in seiner 348. Sitzung (schriftliche Beschlussfassung), geändert durch den Beschluss des Bewertungsausschusses in seiner 369. Sitzung am 15. Dezember 2015.

Teil B

Demografische Veränderungsraten

1	Vorgehen / Berechnungsschritte zur Berechnung der demografischen Veränderungsdaten .	20
1.1	Berechnung der Relativgewichte zur Berechnung der demografischen Veränderungsdaten.....	20
1.1.1	Abgrenzung der Kalibrierungsmenge zur Berechnung der demografischen Veränderungsdaten	20
1.1.2	Berechnung des gewichteten Mittelwertes des annualisierten MGV-Leistungsbedarfs der Versicherten in der Kalibrierungsmenge.....	20
1.1.3	Ermittlung bundeseinheitlicher Relativgewichte zur Berechnung der demografischen Veränderungsdaten	21
1.2	Berechnung der demografischen Veränderungsdaten.....	22
1.2.1	Abgrenzung der Anwendungsmenge zur Berechnung der demografischen Veränderungsdaten	22
1.2.2	Berechnungsschritte zur Berechnung der demografischen Veränderungsdaten	23
2	Ergebnisse: demografische Veränderungsdaten	24
3	Empirische Ergebnisse zum Regressionsmodell zur Berechnung der demografischen Veränderungsdaten und empirische Ergebnisse zu den demografischen Veränderungsdaten	25
3.1	Empirische Ergebnisse zum Regressionsmodell zur Berechnung der demografischen Veränderungsdaten	25
3.1.1	Relativgewichte zur Berechnung der demografischen Veränderungsdaten	25
3.1.2	Ausgewählte statistische Kennzahlen zum Regressionsmodell zur Berechnung der demografischen Veränderungsdaten	26
3.2	Empirische Ergebnisse zu den demografischen Veränderungsdaten	28
3.2.1	Entwicklung der Altersstruktur von 2017 auf 2018	28
3.2.2	Einzelbeiträge der Risikoklassen zu den regionalen demografischen Veränderungsdaten	33

1 Vorgehen / Berechnungsschritte zur Berechnung der demografischen Veränderungsdaten

Die Berechnung der demografischen Veränderungsdaten erfolgt in mehreren Schritten, die in den folgenden Abschnitten näher dargestellt werden sollen.

1.1 Berechnung der Relativgewichte zur Berechnung der demografischen Veränderungsdaten

1.1.1 Abgrenzung der Kalibrierungsmenge zur Berechnung der demografischen Veränderungsdaten

In einem ersten Schritt wird die sogenannte „Kalibrierungsmenge“ gebildet. Dazu werden gemäß Nr. 2.4 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung aus der in Teil A, Abschnitt 3.1 dieses Berichts beschriebenen Datengrundlage diejenigen Versicherten ausgewählt, die sowohl im Jahr 2016 als auch im Jahr 2018 als versicherungszeitvollständig¹ gekennzeichnet sind (als Konsequenz werden nur Versicherte mit den Geburtskalendertagen 5, 12, 19, 25 und 26 berücksichtigt). Darüber hinaus werden aus der Kalibrierungsmenge diejenigen Versicherten ausgeschlossen, die in einem der Jahre 2016 bis 2018 als SV-Teilnehmer² identifiziert wurden, sowie diejenigen, die zu irgendeinem Zeitpunkt in den Jahren 2016 bis 2018 in Feld 14 der Satzart 201 einen Wert von 1 aufweisen (sogenannte „Betreute“). Weiterhin werden diejenigen Versicherten ausgeschlossen, die in mindestens einem der Jahre 2016 bis 2018 als Wohnausländer³ identifiziert wurden und darüber hinaus im entsprechenden Jahr einem der KV-Bezirke Baden-Württemberg oder Nordrhein zugeordnet sind.

1.1.2 Berechnung des gewichteten Mittelwertes des annualisierten MGW-Leistungsbedarfs der Versicherten in der Kalibrierungsmenge

Für die Versicherten in der Kalibrierungsmenge werden für die weitere Berechnung nur das Alter und das Geschlecht im Jahr 2016 sowie deren MGW-Leistungsbedarf und deren Anzahl an Versichertenquartalen im Jahr 2018 herangezogen. Das Alter und das Geschlecht im Jahr 2016 werden verwendet, um die Versicherten in der Kalibrierungsmenge den Alters- und Geschlechtsgruppen in der Gliederung der KM6-Statistik zuzuordnen. Der MGW-Leistungsbedarf wird für jeden Versicherten mit der Anzahl von Versichertenquartalen im Jahr 2018 annualisiert und im folgenden als annualisierter Leistungsbedarf des Versicherten bezeichnet. Dieser wird gemäß der Formel (1) berechnet.

$$LB_i^{a,2018} = \frac{4 \cdot LB_i^{2018}}{AVQ_i^{2018}} \quad (1)$$

¹Zur Definition von „versicherungszeitvollständig“ siehe Teil A, Abschnitt 3.1.4

²Zur Definition eines „längsschnittlichen Selektivvertragsteilnehmers“ siehe Teil A, Abschnitt 3.1.4

³Zur Definition eines Wohnausländers siehe Teil A, Abschnitt 3.1.4

mit

LB_i^{2018} Leistungsbedarf des i -ten Versicherten im Jahr 2018

AVQ_i^{2018} Anzahl der Versichertenquartale des i -ten Versicherten im Jahr 2018

Für die weitere Berechnung wird zuerst der gewichtete Mittelwert des annualisierten Leistungsbedarfs im Jahr 2018 ($\overline{LB}^{a,2018}$) anhand der Formel (2) ermittelt.

$$\overline{LB}^{a,2018} = \frac{\sum_{i=1}^n \left(\frac{1}{4} \cdot LB_i^{a,2018} \cdot AVQ_i^{2018} \right)}{\sum_{i=1}^n \frac{1}{4} \cdot AVQ_i^{2018}} \quad (2)$$

mit

$LB_i^{a,2018}$ annualisierter Leistungsbedarf des i -ten Versicherten im Jahr 2018

AVQ_i^{2018} Anzahl der Versichertenquartale des i -ten Versicherten im Jahr 2018

n Anzahl der Versicherten in der Kalibrierungsmenge

Der gewichtete Mittelwert des annualisierten Leistungsbedarfs in 2018 in der Kalibrierungsmenge beträgt 3.748 Punkte.

1.1.3 Ermittlung bundeseinheitlicher Relativgewichte zur Berechnung der demografischen Veränderungsdaten

Entsprechend dem vorliegenden Beschluss werden die Relativgewichte wie folgt bestimmt. *Auf Basis des Lebensalters in 2016 und des morbiditätsbedingten Leistungsbedarfs gemäß Nr. 2.2.5 für das Jahr 2018 der Versicherten der Kalibrierungsmenge wird für jede KM6-Alters- und Geschlechtsgruppe das bundeseinheitliche Relativgewicht als Quotient zwischen dem nach Versichertenquartalen gewichteten durchschnittlichen annualisierten morbiditätsbedingten Leistungsbedarf je Versicherten in der jeweiligen KM6-Alters- und Geschlechtsklasse und dem nach Versichertenquartalen gewichteten durchschnittlichen annualisierten morbiditätsbedingten Leistungsbedarf je Versicherten insgesamt berechnet.*

Das entspricht der Berechnung der Relativgewichte zur Bestimmung der demografischen Veränderungsdaten mit einem zweijährig-prospektiven gewichteten Regressionsmodell. Als abhängige Variable (y_i^{2018}) geht der relative Leistungsbedarf eines Versicherten in die multiple lineare Regression ein. Der relative Leistungsbedarf eines Versicherten im Jahr 2018 wird anhand der Formel (3) als Quotient aus dem annualisierten Leistungsbedarf und dem gewichteten Mittelwert des annualisierten Leistungsbedarfs der Versicherten

im Jahr 2018 berechnet.

$$y_i^{2018} = \frac{LB_i^{a,2018}}{\overline{LB}^{a,2018}} \quad (3)$$

mit

$LB_i^{a,2018}$ annualisierter Leistungsbedarf des i -ten Versicherten im Jahr 2018

$\overline{LB}^{a,2018}$ gewichteter Mittelwert des annualisierten Leistungsbedarfs im Jahr 2018

Die Regressionsgleichung ist in Formel (4) dargestellt. Die unabhängigen Variablen dieser Regression ($x_{i,j}^{2016}$) sind sogenannte „Dummy-Variablen“, die die Zuordnung eines Versicherten zu genau einer Alters- und Geschlechtsgruppe gemäß der Gliederung der KM6-Statistik im Jahr 2016 repräsentieren und nur die Werte 0 und 1 annehmen können. Dabei bedeutet der Wert 1, dass der Versicherte der entsprechenden Alters- und Geschlechtsgruppe zugeordnet wurde, und der Wert 0, dass der Versicherte nicht der entsprechenden Alters- und Geschlechtsgruppe zugeordnet wurde.

$$y_i^{2018} = \sum_{j=1}^{34} \beta_j \cdot x_{i,j}^{2016} + \varepsilon_i \quad (4)$$

mit

y_i^{2018} relativer Leistungsbedarf des i -ten Versicherten im Jahr 2018

β_j Modellparameter der j -ten Alters- und Geschlechtsgruppe

$x_{i,j}^{2016}$ Ausprägung der Variablen der j -ten Alters- und Geschlechtsgruppe für den i -ten Versicherten im Jahr 2016

ε_i Fehlerterm der Regression

Es erfolgt eine Gewichtung der Versicherten in der Regression mit der Anzahl ihrer Versichertenquartale im Jahr 2018. Die für die Modellparameter (β_j) berechneten Schätzwerte ($\hat{\beta}_j$) werden als (demografische) Relativgewichte bezeichnet.

1.2 Berechnung der demografischen Veränderungsdaten

1.2.1 Abgrenzung der Anwendungsmenge zur Berechnung der demografischen Veränderungsdaten

Die Anwendungsmenge zur Berechnung der demografischen Veränderungsdaten je KV-Bezirk wird in Nr. 2.6 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung definiert. Die folgende kursiv gesetzte Textpassage ist wortwörtlich aus diesem Beschluss

entnommen.

Datengrundlage ist die amtliche KM6-Statistik in der Version, die vom GKV-Spitzenverband gemäß Beschluss des Bewertungsausschusses in seiner 348. Sitzung (schriftliche Beschlussfassung) in der Satzart ANZVER_KM6 an das Institut des Bewertungsausschusses übermittelt wurde. Zur Ermittlung der demografischen Veränderungsdaten werden die Jahre 2017 und 2018 herangezogen.

Versicherte mit Wohnort im Ausland (Wohnausländer) werden der KV-Region, in der ihre Krankenkasse ihren Sitz hat, zugeordnet. Bei fusionierten Krankenkassen wird dabei jeweils der letzte bekannte Sitz der Krankenkasse (Stand: 2. April 2020) verwendet. Wohnausländer, die hiernach dem KV-Bezirk Nordrhein oder dem KV-Bezirk Baden-Württemberg zugeordnet werden müssten, werden jedoch nicht berücksichtigt.

1.2.2 Berechnungsschritte zur Berechnung der demografischen Veränderungsdaten

Die Berechnung der (demografischen) Relativgewichte erfolgt auf der in Abschnitt 1.1.1 beschriebenen Kalibrierungsmenge. Für die Berechnung der demografischen Veränderungsdaten werden diese Relativgewichte auf die Daten der KM6-Statistik für die Jahre 2017 und 2018 für jeden KV-Bezirk separat angewendet.

Genauer wird je KV-Bezirk ein Demografie-Index bezogen auf das Jahr 2017 und ein Demografie-Index bezogen auf das Jahr 2018 berechnet. In der Formel (5) wird die Berechnung dieses Demografie-Indexes formal dargestellt. Dazu wird die Anzahl der Versicherten je Alters- und Geschlechtsgruppe mit Wohnort in einem KV-Bezirk ($n_j^{t,KV}$) mit dem Relativgewicht der jeweiligen Alters- und Geschlechtsgruppe ($\hat{\beta}_j$) multipliziert und die Werte über alle Alters- und Geschlechtsgruppen summiert. Diese Summe wird dann durch die Versichertenzahl mit Wohnort im Bezirk der Kassenärztlichen Vereinigung dividiert.

$$\bar{y}_{demo}^{t,KV} = \frac{\sum_{j=1}^{34} (\hat{\beta}_j \cdot n_j^{t,KV})}{\sum_{j=1}^{34} n_j^{t,KV}} \quad (5)$$

mit

- $\bar{y}_{demo}^{t,KV}$ Demografie-Index für den KV-Bezirk KV bezogen auf das Jahr t
- $\hat{\beta}_j$ Relativgewicht der Alters- und Geschlechtsgruppe j , mit $j = 1, \dots, 34$
- $n_j^{t,KV}$ Anzahl Versicherte in der Alters- und Geschlechtsgruppe j , im KV-Bezirk KV , im Jahr t
- KV KV-Bezirk, mit $KV = 1, \dots, 17$
- t Jahr, mit $t = 2017, 2018$

Um die demografische Veränderungsrate je KV-Bezirk zu berechnen, wird gemäß Formel (6) der Demografie-Index des Jahres 2018 ($\bar{y}_{demo}^{2018,KV}$) durch den Demografie-Index des Jahres 2017 ($\bar{y}_{demo}^{2017,KV}$) geteilt und davon der Wert eins abgezogen.

$$VR_{2017 \rightarrow 2018}^{KV} = \frac{\bar{y}_{demo}^{2018,KV}}{\bar{y}_{demo}^{2017,KV}} - 1 \quad (6)$$

mit

- $VR_{2017 \rightarrow 2018}^{KV}$ demografische Veränderungsrate für den KV-Bezirk KV
- $\bar{y}_{demo}^{2017,KV}$ Demografie-Index für den KV-Bezirk KV bezogen auf das Jahr 2017
- $\bar{y}_{demo}^{2018,KV}$ Demografie-Index für den KV-Bezirk KV bezogen auf das Jahr 2018
- KV KV-Bezirk, mit $KV = 1, \dots, 17$

2 Ergebnisse: demografische Veränderungsdaten

Unter Verwendung der Vorgaben des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung ergeben sich die in Tabelle 1 dargestellten KV-spezifischen demografischen Veränderungsdaten.

KV-Bezirk	Veränderungsrate
Schleswig-Holstein	0,2795 %
Hamburg	-0,1076 %
Bremen	-0,0989 %
Niedersachsen	0,1643 %
Westfalen-Lippe	0,1604 %
Nordrhein	0,0766 %
Hessen	0,0504 %
Rheinland-Pfalz	0,1243 %
Baden-Württemberg	0,0013 %
Bayern	0,0931 %
Berlin	-0,0367 %
Saarland	0,2918 %
Mecklenburg-Vorpommern	0,3679 %
Brandenburg	0,1916 %
Sachsen-Anhalt	0,2331 %
Thüringen	0,2504 %
Sachsen	0,0108 %

Tabelle 1: Demografische Veränderungsdaten je KV-Bezirk für das Jahr 2021

3 Empirische Ergebnisse zum Regressionsmodell zur Berechnung der demografischen Veränderungsdaten und empirische Ergebnisse zu den demografischen Veränderungsdaten

3.1 Empirische Ergebnisse zum Regressionsmodell zur Berechnung der demografischen Veränderungsdaten

3.1.1 Relativgewichte zur Berechnung der demografischen Veränderungsdaten

In Tabelle 2 werden die Relativgewichte berichtet, welche zur Berechnung der demografischen Veränderungsdaten verwendet wurden.

Altersgruppe gemäß KM6-Statistik	Relativgewicht Frauen	Relativgewicht Männer
bis unter 15 Jahre	56,58 %	59,58 %
15 bis unter 20 Jahre	78,79 %	46,77 %
20 bis unter 25 Jahre	77,75 %	43,70 %
25 bis unter 30 Jahre	83,25 %	46,05 %
30 bis unter 35 Jahre	87,63 %	50,47 %
35 bis unter 40 Jahre	91,78 %	56,66 %
40 bis unter 45 Jahre	98,09 %	64,23 %
45 bis unter 50 Jahre	105,66 %	73,34 %
50 bis unter 55 Jahre	115,45 %	87,32 %
55 bis unter 60 Jahre	125,42 %	106,60 %
60 bis unter 65 Jahre	134,14 %	121,16 %
65 bis unter 70 Jahre	145,83 %	137,51 %
70 bis unter 75 Jahre	168,82 %	165,91 %
75 bis unter 80 Jahre	183,68 %	186,48 %
80 bis unter 85 Jahre	185,12 %	193,70 %
85 bis unter 90 Jahre	183,79 %	193,03 %
90 und mehr Jahre	191,86 %	195,79 %

Tabelle 2: Relativgewichte zur Berechnung der demografischen Veränderungsdaten für das Jahr 2021

3.1.2 Ausgewählte statistische Kennzahlen zum Regressionsmodell zur Berechnung der demografischen Veränderungsdaten

Im Folgenden werden Kennzahlen für das Regressionsmodell zur Berechnung der demografischen Veränderungsdaten für das Jahr 2021 im Vergleich zum Regressionsmodell zur Berechnung der demografischen Veränderungsdaten für das Jahr 2020 dargestellt.

Tabelle 3 enthält die Anzahl der Versicherten in der Kalibrierungsmenge, welche zur Berechnung der Relativgewichte verwendet wurde, sowie die Anzahl der Alters- und Geschlechtsgruppen, welche als unabhängige Variablen im Regressionsmodell verwendet wurden.

Der gewichtete „mean absolute prediction error“ (der gewichtete MAPE) gibt an, wie groß die durchschnittliche absolute Abweichung zwischen dem mit dem Regressionsmodell geschätzten Wert (\hat{y}_i) und dem sich aus den Daten ergebenden Wert (y_i) ist.

Die Berechnung des gewichteten MAPE wird in Formel (7) dargestellt. Er wird berechnet als

Regressionsmodell zur Berechnung der demografischen Veränderungsdaten	Anzahl der Versicherten in der Kalibrierungsmenge	Zahl der Alters- und Geschlechtsgruppen
für das Jahr 2021	10.243.425	34
für das Jahr 2020	10.215.691	34

Tabelle 3: Kennzahlen der Regressionsmodelle zur Berechnung der demografischen Veränderungsdaten für die Jahre 2020 und 2021

Betrag der Differenz des tatsächlichen relativen Leistungsbedarfs je Versicherten im Jahr 2018 ($y_i = y_i^{2018}$) zum mit dem Regressionsmodell geschätzten relativen Leistungsbedarf je Versicherten (\hat{y}_i) multipliziert mit den Versichertenquartalen je Versicherten (AVQ_i). Dieses wird über alle Versicherten $i = 1, \dots, n$ der in Abschnitt 1.1.1 beschriebenen Kalibrierungsmenge summiert und durch die Summe der Versichertenquartale (AVQ_i) der Versicherten in der Kalibrierungsmenge dividiert.

$$MAPE = \frac{\sum_i AVQ_i |y_i - \hat{y}_i|}{\sum_i AVQ_i} \quad (7)$$

mit

- y_i relativer Leistungsbedarf des i -ten Versicherten in der Kalibrierungsmenge im Jahr 2018
- \hat{y}_i geschätzter relativer Leistungsbedarf des i -ten Versicherten in der Kalibrierungsmenge im Jahr 2018
- AVQ_i Anzahl der Versichertenquartale des i -ten Versicherten in der Kalibrierungsmenge im Jahr 2018

Das „**adjustierte** R^2 “ ist das Bestimmtheitsmaß, welches die Anzahl der Variablen im Modell berücksichtigt. Das Bestimmtheitsmaß ist definiert als der Anteil der durch die Regressionsfunktion erklärten Variation in der abhängigen Variable an der Gesamtvariation in der abhängigen Variable. Zur Berechnung wird Formel (8) benutzt, wobei $m = 34$ die Anzahl der Variablen im Modell darstellt und die Summation über die Versicherten in der Kalibrierungsmenge erfolgt ($i = 1, \dots, n$).

$$R_{adj}^2 = 1 - \frac{\frac{1}{n-m-1} \sum_i AVQ_i (y_i - \hat{y}_i)^2}{\frac{1}{n-1} \sum_i AVQ_i (y_i - 1)^2} \quad (8)$$

mit

- y_i relativer Leistungsbedarf des i -ten Versicherten in der Kalibrierungsmenge im Jahr 2018
- \hat{y}_i geschätzter relativer Leistungsbedarf des i -ten Versicherten in der Kalibrierungsmenge im Jahr 2018
- AVQ_i Anzahl der Versicherten quartale des i -ten Versicherten in der Kalibrierungsmenge im Jahr 2018
- n Anzahl der Versicherten in der Kalibrierungsmenge im Jahr 2018
- m Anzahl der Variablen im Modell (= 34)

Der gewichtete MAPE⁴ und das adjustierte R^2 sind für die Regressionsmodelle zur Berechnung der demografischen Veränderungsraten für die Jahre 2020 und 2021 im Vergleich in Tabelle 4 aufgeführt.

Regressionsmodell zur Berechnung der demografischen Veränderungsraten	Gewichteter MAPE	Adjustiertes R^2
für das Jahr 2021	63,64 %	16,48 %
für das Jahr 2020	63,73 %	16,59 %

Tabelle 4: Gewichteter MAPE und adjustiertes R^2 für die Regressionsmodelle zur Berechnung der demografischen Veränderungsraten für die Jahre 2020 und 2021

3.2 Empirische Ergebnisse zu den demografischen Veränderungsraten

3.2.1 Entwicklung der Altersstruktur von 2017 auf 2018

Einen bedeutenden Einfluss auf die demografischen Veränderungsraten hat die Veränderung der Altersstruktur der einzelnen Altersgruppen der KM6-Statistik. Aus diesem Grund wird nachfolgend die Veränderung der entsprechenden Altersstruktur der Versicherten dargestellt.

In Tabelle 5 sind die Anteile der Versichertenzahlen der jeweiligen KM6-Gruppe (aufgeteilt nach Frauen und Männern) an der Gesamtzahl aller gesetzlich Versicherten gemäß KM6-Statistik je KV-Bezirk (Abgrenzung und KV-Zuordnung wie in Nr. 2.6 gemäß Beschluss des Bewertungsausschusses in seiner 503. Sitzung beschrieben) des Jahres 2017 aufgeführt. Analog enthält Tabelle 6 die Anteile der Versichertenzahlen der jeweiligen KM6-Gruppe (aufgeteilt nach Frauen und Männern) an der Gesamtzahl aller GKV-Versicherten gemäß KM6-Statistik je KV-Bezirk des Jahres 2018⁵.

⁴Die Angabe des gewichteten MAPE in Prozent ergibt sich durch die Verwendung relativer Leistungsbedarfe.

⁵Zu beachten ist bei diesen beiden Tabellen, dass gemäß Beschluss des Bewertungsausschusses in seiner

Die Änderungen der Anteile einer KM6-Gruppe des Jahres 2018 gegenüber den entsprechenden Anteilen des Jahres 2017 sind in Tabelle 7 in Prozentpunkten aufgeführt⁶.

503. Sitzung die Versicherten mit Wohnsitz im Ausland gemäß dem Sitz ihrer Krankenkasse einem KV-Bezirk zugeordnet werden. Wohnausländer, die hiernach dem KV-Bezirk Nordrhein oder dem KV-Bezirk Baden-Württemberg zugeordnet werden müssten, werden jedoch nicht berücksichtigt.

⁶Zu beachten ist bei Tabelle 7, dass gemäß Beschluss des Bewertungsausschusses in seiner 503. Sitzung für beide Jahre derselbe, aktuelle Kassensitz verwendet wird, um Wohnausländer regional zuzuordnen. Eine Verlegung des Kassensitzes hat somit allein keine Auswirkungen auf die in Tabelle 7 berichteten Veränderungen.

	KV SH	KV HH	KV HB	KV NI	KV WL	KV NR	KV HE	KV RP	KV BW	KV BY	KV BE	KV SL	KV MV	KV BB	KV ST	KV TH	KV SN
F bis unter 15 Jahre	6.28	6.42	6.33	6.32	6.36	6.28	6.46	6.08	6.39	6.24	6.65	5.56	6.00	6.00	5.64	5.93	6.16
F 15 bis unter 20 Jahre	2.59	2.09	2.28	2.62	2.60	2.41	2.40	2.45	2.54	2.43	1.93	2.27	1.91	1.95	1.83	1.88	1.86
F 20 bis unter 25 Jahre	2.77	2.89	2.99	2.85	2.92	2.85	2.83	2.83	3.00	2.89	2.49	2.70	1.64	1.55	1.64	1.67	1.71
F 25 bis unter 30 Jahre	2.92	4.26	3.62	3.00	3.12	3.38	3.31	3.11	3.33	3.38	4.12	3.01	2.59	2.33	2.50	2.51	2.71
F 30 bis unter 35 Jahre	2.88	4.34	3.36	2.91	2.97	3.28	3.28	2.96	3.21	3.35	4.51	2.90	3.10	3.01	2.91	3.00	3.22
F 35 bis unter 40 Jahre	3.04	4.04	3.13	2.96	2.98	3.26	3.33	3.01	3.20	3.38	3.99	2.84	2.99	3.10	2.78	2.95	3.07
F 40 bis unter 45 Jahre	2.98	3.45	2.95	2.91	2.92	3.05	3.15	2.87	3.04	3.14	3.18	2.68	2.42	2.65	2.33	2.48	2.53
F 45 bis unter 50 Jahre	3.99	3.68	3.47	3.83	3.75	3.79	3.80	3.74	3.74	3.82	3.37	3.57	3.16	3.43	3.30	3.23	3.12
F 50 bis unter 55 Jahre	4.73	4.04	4.03	4.49	4.41	4.45	4.36	4.56	4.34	4.37	3.91	4.57	4.35	4.50	4.19	4.03	3.83
F 55 bis unter 60 Jahre	4.07	3.38	3.74	4.03	4.05	3.98	3.90	4.32	3.97	3.96	3.47	4.56	4.69	4.56	4.39	4.29	3.91
F 60 bis unter 65 Jahre	3.28	2.69	3.23	3.30	3.36	3.28	3.22	3.61	3.24	3.19	2.88	3.86	4.32	3.99	4.17	4.09	3.79
F 65 bis unter 70 Jahre	3.05	2.51	2.92	2.98	2.88	2.90	2.98	3.15	2.85	2.89	2.79	3.38	3.62	3.41	3.76	3.75	3.67
F 70 bis unter 75 Jahre	2.65	2.26	2.51	2.37	2.23	2.38	2.33	2.24	2.23	2.31	2.42	2.27	2.50	2.64	2.97	2.89	3.00
F 75 bis unter 80 Jahre	3.18	2.59	2.94	2.96	2.81	2.85	2.74	2.89	2.81	2.78	2.91	3.11	3.75	3.82	3.98	3.88	3.92
F 80 bis unter 85 Jahre	2.23	1.87	2.19	2.18	2.18	2.16	2.02	2.23	2.09	2.02	2.00	2.52	2.77	2.69	2.83	2.69	2.82
F 85 bis unter 90 Jahre	1.22	1.06	1.26	1.29	1.34	1.25	1.20	1.33	1.25	1.21	1.04	1.44	1.54	1.49	1.64	1.56	1.70
F 90 und mehr Jahre	0.81	0.74	0.82	0.82	0.81	0.75	0.77	0.82	0.77	0.74	0.70	0.84	0.80	0.82	0.92	0.90	1.07
M bis unter 15 Jahre	6.63	6.77	6.69	6.68	6.72	6.64	6.82	6.46	6.75	6.56	7.03	5.92	6.26	6.35	5.97	6.23	6.49
M 15 bis unter 20 Jahre	2.73	2.21	2.41	2.79	2.78	2.56	2.58	2.60	2.72	2.61	2.02	2.45	2.04	2.08	2.00	2.02	1.98
M 20 bis unter 25 Jahre	3.00	2.88	3.14	3.10	3.18	3.04	3.01	3.07	3.29	3.15	2.45	2.98	1.79	1.80	1.83	1.82	1.83
M 25 bis unter 30 Jahre	3.06	4.00	3.86	3.25	3.41	3.49	3.45	3.32	3.63	3.61	3.89	3.33	2.71	2.50	2.76	2.71	2.90
M 30 bis unter 35 Jahre	2.85	4.24	3.60	3.05	3.16	3.29	3.33	3.06	3.38	3.49	4.39	3.10	3.20	3.11	3.18	3.25	3.42
M 35 bis unter 40 Jahre	2.83	3.94	3.33	2.93	2.99	3.12	3.20	2.92	3.14	3.30	3.96	2.85	3.11	3.13	3.07	3.15	3.27
M 40 bis unter 45 Jahre	2.70	3.32	2.98	2.79	2.82	2.83	2.91	2.66	2.83	2.96	3.09	2.57	2.50	2.57	2.53	2.59	2.65
M 45 bis unter 50 Jahre	3.58	3.52	3.46	3.60	3.57	3.48	3.50	3.38	3.38	3.50	3.23	3.28	3.04	3.23	3.36	3.28	3.12
M 50 bis unter 55 Jahre	4.24	3.85	3.96	4.22	4.21	4.10	4.04	4.13	3.97	4.02	3.70	4.16	4.02	4.22	4.06	3.88	3.64
M 55 bis unter 60 Jahre	3.59	3.09	3.39	3.72	3.76	3.61	3.54	3.86	3.55	3.57	3.23	4.16	4.37	4.34	4.14	4.09	3.65
M 60 bis unter 65 Jahre	2.78	2.31	2.73	2.94	3.03	2.84	2.80	3.15	2.79	2.75	2.52	3.46	3.97	3.73	3.79	3.78	3.37
M 65 bis unter 70 Jahre	2.52	2.01	2.38	2.58	2.46	2.37	2.46	2.70	2.34	2.34	2.21	2.99	3.22	3.04	3.26	3.34	3.16
M 70 bis unter 75 Jahre	2.07	1.77	1.98	1.90	1.80	1.84	1.88	1.80	1.77	1.81	1.85	1.80	2.05	2.14	2.41	2.39	2.38
M 75 bis unter 80 Jahre	2.38	1.87	2.14	2.25	2.08	2.13	2.09	2.18	2.09	2.04	2.10	2.23	2.79	2.93	3.02	2.95	2.97
M 80 bis unter 85 Jahre	1.50	1.17	1.36	1.45	1.41	1.46	1.38	1.52	1.43	1.31	1.27	1.61	1.78	1.85	1.82	1.77	1.90
M 85 bis unter 90 Jahre	0.62	0.51	0.57	0.67	0.69	0.67	0.65	0.72	0.69	0.63	0.53	0.75	0.78	0.80	0.79	0.77	0.88
M 90 und mehr Jahre	0.24	0.21	0.24	0.26	0.25	0.23	0.26	0.26	0.25	0.24	0.19	0.27	0.21	0.24	0.26	0.26	0.31

Tabelle 5: Anteile der Versichertenzahlen der jeweiligen KM6-Altersgruppe (aufgeteilt nach Frauen (F) und Männern (M)) an der Gesamtzahl aller gesetzlich Versicherten gemäß KM6-Statistik je KV-Bezirk für das Jahr 2017

	KV SH	KV HH	KV HB	KV NI	KV WL	KV NR	KV HE	KV RP	KV BW	KV BY	KV BE	KV SL	KV MV	KV BB	KV ST	KV TH	KV SN
F bis unter 15 Jahre	6.28	6.49	6.43	6.35	6.38	6.32	6.51	6.12	6.42	6.26	6.72	5.61	6.04	6.06	5.72	5.98	6.23
F 15 bis unter 20 Jahre	2.51	2.04	2.24	2.53	2.53	2.34	2.34	2.36	2.45	2.35	1.92	2.21	1.93	1.97	1.84	1.92	1.89
F 20 bis unter 25 Jahre	2.74	2.85	2.98	2.83	2.88	2.82	2.81	2.80	2.96	2.84	2.48	2.62	1.74	1.63	1.72	1.74	1.79
F 25 bis unter 30 Jahre	2.88	4.14	3.52	2.96	3.09	3.35	3.26	3.07	3.29	3.34	3.90	2.96	2.26	2.05	2.21	2.20	2.42
F 30 bis unter 35 Jahre	2.93	4.39	3.42	2.96	3.01	3.35	3.33	3.03	3.27	3.40	4.60	2.97	3.17	3.06	2.96	3.00	3.28
F 35 bis unter 40 Jahre	3.05	4.07	3.19	2.98	3.00	3.28	3.33	3.04	3.23	3.39	4.09	2.87	3.05	3.15	2.84	2.99	3.12
F 40 bis unter 45 Jahre	2.99	3.49	2.95	2.90	2.91	3.06	3.16	2.88	3.04	3.17	3.27	2.72	2.56	2.80	2.42	2.55	2.66
F 45 bis unter 50 Jahre	3.64	3.48	3.27	3.55	3.48	3.51	3.55	3.45	3.50	3.58	3.18	3.29	2.92	3.16	3.05	3.05	2.93
F 50 bis unter 55 Jahre	4.73	4.00	3.97	4.45	4.36	4.40	4.30	4.46	4.27	4.31	3.78	4.46	4.08	4.26	4.00	3.84	3.67
F 55 bis unter 60 Jahre	4.24	3.49	3.81	4.16	4.15	4.09	4.00	4.43	4.07	4.07	3.60	4.63	4.78	4.70	4.48	4.36	3.98
F 60 bis unter 65 Jahre	3.39	2.75	3.31	3.39	3.47	3.36	3.29	3.69	3.32	3.28	2.87	3.99	4.33	3.99	4.12	4.06	3.71
F 65 bis unter 70 Jahre	3.02	2.46	2.92	3.01	2.94	2.92	2.96	3.25	2.87	2.88	2.78	3.50	3.92	3.65	4.00	3.95	3.86
F 70 bis unter 75 Jahre	2.63	2.20	2.46	2.36	2.23	2.36	2.35	2.26	2.21	2.32	2.25	2.35	2.39	2.42	2.83	2.80	2.77
F 75 bis unter 80 Jahre	3.06	2.48	2.79	2.80	2.64	2.71	2.61	2.71	2.65	2.64	2.86	2.91	3.57	3.68	3.87	3.75	3.83
F 80 bis unter 85 Jahre	2.46	2.00	2.34	2.36	2.33	2.31	2.17	2.37	2.23	2.15	2.17	2.66	3.00	2.94	3.06	2.92	3.04
F 85 bis unter 90 Jahre	1.20	1.02	1.21	1.26	1.31	1.23	1.18	1.31	1.23	1.20	1.03	1.43	1.58	1.49	1.62	1.54	1.67
F 90 und mehr Jahre	0.83	0.74	0.83	0.84	0.84	0.78	0.79	0.85	0.79	0.75	0.71	0.87	0.85	0.86	0.96	0.94	1.10
M bis unter 15 Jahre	6.63	6.87	6.85	6.71	6.74	6.69	6.87	6.50	6.77	6.58	7.13	5.97	6.30	6.41	6.02	6.29	6.57
M 15 bis unter 20 Jahre	2.67	2.18	2.37	2.71	2.71	2.50	2.52	2.53	2.64	2.53	2.01	2.38	2.07	2.10	2.03	2.06	2.02
M 20 bis unter 25 Jahre	2.98	2.89	3.17	3.09	3.15	3.02	3.02	3.03	3.29	3.14	2.46	2.91	1.87	1.89	1.90	1.93	1.91
M 25 bis unter 30 Jahre	3.04	3.95	3.78	3.27	3.41	3.49	3.44	3.32	3.64	3.61	3.71	3.28	2.36	2.20	2.45	2.40	2.61
M 30 bis unter 35 Jahre	2.95	4.33	3.67	3.15	3.25	3.40	3.43	3.18	3.52	3.59	4.50	3.16	3.24	3.15	3.22	3.26	3.47
M 35 bis unter 40 Jahre	2.86	3.99	3.38	2.99	3.05	3.18	3.26	3.00	3.24	3.39	4.07	2.95	3.17	3.23	3.17	3.25	3.36
M 40 bis unter 45 Jahre	2.73	3.43	3.01	2.82	2.85	2.87	2.97	2.71	2.89	3.03	3.23	2.61	2.67	2.76	2.65	2.69	2.82
M 45 bis unter 50 Jahre	3.29	3.34	3.27	3.35	3.33	3.24	3.28	3.14	3.20	3.30	3.07	3.03	2.84	3.03	3.17	3.13	3.01
M 50 bis unter 55 Jahre	4.24	3.86	3.94	4.18	4.16	4.05	3.99	4.04	3.91	3.98	3.60	4.06	3.79	4.00	3.92	3.74	3.51
M 55 bis unter 60 Jahre	3.75	3.23	3.50	3.87	3.88	3.72	3.66	3.98	3.66	3.69	3.38	4.22	4.44	4.46	4.25	4.17	3.75
M 60 bis unter 65 Jahre	2.87	2.39	2.81	3.03	3.14	2.95	2.88	3.23	2.89	2.86	2.54	3.57	3.96	3.73	3.76	3.75	3.31
M 65 bis unter 70 Jahre	2.51	1.95	2.35	2.60	2.51	2.38	2.44	2.77	2.33	2.33	2.20	3.09	3.48	3.25	3.46	3.52	3.30
M 70 bis unter 75 Jahre	2.05	1.71	1.93	1.90	1.80	1.83	1.89	1.83	1.77	1.81	1.73	1.90	1.99	1.98	2.29	2.31	2.22
M 75 bis unter 80 Jahre	2.27	1.80	2.05	2.12	1.96	2.00	1.99	2.03	1.97	1.93	2.06	2.08	2.65	2.84	2.93	2.86	2.88
M 80 bis unter 85 Jahre	1.65	1.26	1.47	1.58	1.52	1.56	1.48	1.62	1.52	1.41	1.38	1.70	1.95	2.02	1.99	1.96	2.06
M 85 bis unter 90 Jahre	0.63	0.50	0.57	0.66	0.68	0.67	0.65	0.72	0.69	0.64	0.54	0.77	0.81	0.82	0.80	0.79	0.88
M 90 und mehr Jahre	0.26	0.23	0.26	0.28	0.28	0.26	0.28	0.29	0.28	0.26	0.21	0.29	0.25	0.28	0.28	0.29	0.35

Tabelle 6: Anteile der Versichertenzahlen der jeweiligen KM6-Altersgruppe (aufgeteilt nach Frauen (F) und Männern (M)) an der Gesamtzahl aller gesetzlich Versicherten gemäß KM6-Statistik je KV-Bezirk für das Jahr 2018

	KV SH	KV HH	KV HB	KV NI	KV WL	KV NR	KV HE	KV RP	KV BW	KV BY	KV BE	KV SL	KV MV	KV BB	KV ST	KV TH	KV SN
F bis unter 15 Jahre	0.00	0.07	0.10	0.02	0.01	0.04	0.05	0.04	0.03	0.02	0.07	0.05	0.04	0.05	0.08	0.05	0.07
F 15 bis unter 20 Jahre	-0.08	-0.04	-0.04	-0.09	-0.07	-0.07	-0.06	-0.09	-0.09	-0.08	-0.01	-0.07	0.02	0.02	0.00	0.04	0.04
F 20 bis unter 25 Jahre	-0.03	-0.03	-0.00	-0.02	-0.04	-0.04	-0.02	-0.04	-0.04	-0.05	-0.00	-0.08	0.10	0.08	0.07	0.07	0.08
F 25 bis unter 30 Jahre	-0.04	-0.12	-0.10	-0.03	-0.02	-0.03	-0.05	-0.03	-0.04	-0.04	-0.22	-0.05	-0.33	-0.28	-0.28	-0.31	-0.29
F 30 bis unter 35 Jahre	0.05	0.05	0.06	0.06	0.05	0.07	0.05	0.07	0.06	0.05	0.08	0.07	0.06	0.05	0.05	0.01	0.05
F 35 bis unter 40 Jahre	0.01	0.03	0.06	0.02	0.02	0.02	0.00	0.03	0.03	0.01	0.10	0.03	0.06	0.05	0.06	0.05	0.05
F 40 bis unter 45 Jahre	0.01	0.05	-0.00	-0.02	-0.01	0.02	0.02	0.01	0.00	0.03	0.10	0.03	0.14	0.15	0.10	0.07	0.13
F 45 bis unter 50 Jahre	-0.35	-0.20	-0.21	-0.28	-0.27	-0.28	-0.25	-0.28	-0.25	-0.25	-0.19	-0.29	-0.25	-0.27	-0.25	-0.18	-0.18
F 50 bis unter 55 Jahre	0.00	-0.03	-0.06	-0.04	-0.05	-0.05	-0.06	-0.11	-0.07	-0.06	-0.14	-0.12	-0.27	-0.24	-0.19	-0.18	-0.16
F 55 bis unter 60 Jahre	0.17	0.11	0.06	0.13	0.10	0.11	0.10	0.11	0.09	0.11	0.13	0.07	0.09	0.14	0.09	0.07	0.08
F 60 bis unter 65 Jahre	0.12	0.06	0.08	0.09	0.11	0.09	0.06	0.08	0.08	0.08	-0.01	0.13	0.01	-0.01	-0.04	-0.03	-0.08
F 65 bis unter 70 Jahre	-0.03	-0.06	-0.01	0.03	0.07	0.02	-0.02	0.10	0.03	-0.00	-0.01	0.13	0.30	0.24	0.24	0.20	0.19
F 70 bis unter 75 Jahre	-0.03	-0.06	-0.04	-0.01	0.01	-0.02	0.02	0.02	-0.02	0.01	-0.17	0.08	-0.11	-0.22	-0.15	-0.09	-0.23
F 75 bis unter 80 Jahre	-0.12	-0.11	-0.15	-0.16	-0.17	-0.14	-0.13	-0.18	-0.16	-0.15	-0.05	-0.19	-0.18	-0.14	-0.12	-0.13	-0.09
F 80 bis unter 85 Jahre	0.23	0.13	0.15	0.18	0.14	0.15	0.14	0.14	0.14	0.14	0.18	0.14	0.24	0.25	0.22	0.23	0.22
F 85 bis unter 90 Jahre	-0.02	-0.04	-0.05	-0.03	-0.03	-0.02	-0.03	-0.02	-0.02	-0.01	-0.01	-0.01	0.03	0.00	-0.02	-0.02	-0.03
F 90 und mehr Jahre	0.02	0.01	0.01	0.02	0.03	0.03	0.02	0.03	0.02	0.01	0.01	0.02	0.05	0.04	0.05	0.03	0.03
M bis unter 15 Jahre	0.00	0.09	0.15	0.03	0.02	0.05	0.06	0.05	0.02	0.02	0.10	0.04	0.04	0.06	0.06	0.06	0.08
M 15 bis unter 20 Jahre	-0.06	-0.03	-0.04	-0.08	-0.07	-0.05	-0.06	-0.08	-0.08	-0.09	-0.00	-0.07	0.03	0.02	0.03	0.04	0.04
M 20 bis unter 25 Jahre	-0.02	0.01	0.02	-0.01	-0.03	-0.02	0.01	-0.04	-0.00	-0.01	0.01	-0.07	0.08	0.09	0.08	0.11	0.09
M 25 bis unter 30 Jahre	-0.02	-0.05	-0.08	0.01	0.00	-0.00	-0.01	-0.00	0.01	-0.00	-0.17	-0.05	-0.35	-0.30	-0.31	-0.30	-0.29
M 30 bis unter 35 Jahre	0.11	0.09	0.07	0.10	0.10	0.11	0.10	0.12	0.14	0.10	0.10	0.07	0.04	0.04	0.04	0.02	0.05
M 35 bis unter 40 Jahre	0.04	0.05	0.05	0.06	0.07	0.07	0.06	0.08	0.10	0.09	0.11	0.09	0.06	0.10	0.11	0.10	0.09
M 40 bis unter 45 Jahre	0.03	0.11	0.04	0.02	0.03	0.04	0.06	0.05	0.06	0.07	0.14	0.04	0.18	0.19	0.11	0.10	0.16
M 45 bis unter 50 Jahre	-0.29	-0.19	-0.19	-0.25	-0.23	-0.24	-0.21	-0.24	-0.18	-0.20	-0.16	-0.24	-0.19	-0.20	-0.19	-0.14	-0.12
M 50 bis unter 55 Jahre	0.00	0.00	-0.02	-0.04	-0.05	-0.05	-0.04	-0.09	-0.06	-0.03	-0.11	-0.10	-0.23	-0.22	-0.14	-0.14	-0.13
M 55 bis unter 60 Jahre	0.16	0.13	0.11	0.14	0.12	0.12	0.13	0.13	0.12	0.12	0.15	0.06	0.07	0.12	0.11	0.08	0.10
M 60 bis unter 65 Jahre	0.09	0.08	0.07	0.09	0.11	0.11	0.08	0.07	0.10	0.11	0.02	0.11	-0.00	0.00	-0.03	-0.03	-0.06
M 65 bis unter 70 Jahre	-0.01	-0.06	-0.03	0.02	0.05	0.01	-0.03	0.07	-0.00	-0.01	-0.01	0.09	0.25	0.21	0.20	0.17	0.15
M 70 bis unter 75 Jahre	-0.02	-0.06	-0.06	-0.00	0.01	-0.01	0.01	0.03	-0.00	0.00	-0.12	0.09	-0.07	-0.16	-0.12	-0.08	-0.17
M 75 bis unter 80 Jahre	-0.11	-0.07	-0.09	-0.12	-0.12	-0.13	-0.10	-0.15	-0.13	-0.11	-0.04	-0.14	-0.14	-0.10	-0.09	-0.09	-0.09
M 80 bis unter 85 Jahre	0.15	0.09	0.11	0.13	0.10	0.10	0.10	0.10	0.09	0.10	0.11	0.09	0.17	0.17	0.17	0.18	0.16
M 85 bis unter 90 Jahre	0.01	-0.01	0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.01	0.00	0.01	0.01	0.04	0.02	0.02	0.02	0.01
M 90 und mehr Jahre	0.02	0.02	0.01	0.02	0.03	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.04	0.04	0.03	0.03	0.04

Tabelle 7: Änderung der Anteile der Versichertenzahlen einer KM6-Altersgruppe an der Gesamtzahl aller gesetzlich Versicherten des Jahres 2018 gegenüber den entsprechenden Anteilen der Versichertenzahlen des Jahres 2017

3.2.2 Einzelbeiträge der Risikoklassen zu den regionalen demografischen Veränderungsraten

Die Einzelbeiträge der Alters- und Geschlechtsgruppen zur regionalen demografischen Veränderungsrate sind in einer Excel-Mappe getrennt nach KV-Bezirken unter institut-ba.de veröffentlicht. Die Berechnung der Einzelbeiträge erfolgt anhand der Formeln (9) und (10), wobei in Formel (9) zunächst die Zerlegung des Demografie-Index in Partialindizes dargestellt ist und Formel (10) die Zerlegung der demografischen Veränderungsrate mithilfe der Partialindizes beinhaltet.

$$\begin{aligned}
 \bar{y}_{demo}^{t,KV} &= \frac{\sum_{j=1}^{34} \hat{\beta}_j \cdot n_j^{t,KV}}{\sum_{j=1}^{34} n_j^{t,KV}} \\
 &= \frac{\hat{\beta}_1 \cdot n_1^{t,KV} + \hat{\beta}_2 \cdot n_2^{t,KV} + \dots + \hat{\beta}_{34} \cdot n_{34}^{t,KV}}{\sum_{j=1}^{34} n_j^{t,KV}} \quad (9) \\
 &= \underbrace{\frac{\hat{\beta}_1 \cdot n_1^{t,KV}}{\sum_{j=1}^{34} n_j^{t,KV}}}_{Partialindex_1^{t,KV}} + \underbrace{\frac{\hat{\beta}_2 \cdot n_2^{t,KV}}{\sum_{j=1}^{34} n_j^{t,KV}}}_{Partialindex_2^{t,KV}} + \dots + \underbrace{\frac{\hat{\beta}_{34} \cdot n_{34}^{t,KV}}{\sum_{j=1}^{34} n_j^{t,KV}}}_{Partialindex_{34}^{t,KV}}
 \end{aligned}$$

mit

- $\bar{y}_{demo}^{t,KV}$ Demografie-Index im KV-Bezirk KV , mit $KV = 1, \dots, 17$ für das Jahr t
- $\hat{\beta}_j$ Relativgewicht der Alters- und Geschlechtsgruppe j , mit $j = 1, \dots, 34$
- $n_j^{t,KV}$ Anzahl der Versicherten in der Alters- und Geschlechtsgruppe j im KV-Bezirk KV für das Jahr t
- t Jahr, mit $t = 2017, 2018$

$$\begin{aligned}
 VR_{2017 \rightarrow 2018}^{KV} &= \frac{\bar{y}_{demo}^{2018,KV}}{\bar{y}_{demo}^{2017,KV}} - 1 = \frac{\bar{y}_{demo}^{2018,KV} - \bar{y}_{demo}^{2017,KV}}{\bar{y}_{demo}^{2017,KV}} \quad (10) \\
 &= \sum_{j=1}^{34} \left(\frac{Partialindex_j^{2018,KV} - Partialindex_j^{2017,KV}}{\bar{y}_{demo}^{2017,KV}} \right)
 \end{aligned}$$

mit

- $VR_{2017 \rightarrow 2018}^{KV}$ demografische Veränderungsrate für den KV-Bezirk KV

Die Zerlegung der jeweiligen Veränderungsrate erfolgt also additiv. Der Einzelbeitrag der Alters- und Geschlechtsgruppe j entspricht der Entwicklung des j -ten Partialindex in der Formel (10) in Relation zum Wert $\bar{y}_{demo}^{2017,KV}$, dem Demografie-Index für den KV-Bezirk KV bezogen auf das Jahr 2017.

Teil C

Diagnosebezogene Veränderungsraten

1	Vorgehen / Berechnungsschritte zur Berechnung der diagnosebezogenen Veränderungsdaten	36
1.1	Zuordnung zu Risikokategorien (Klassifikation) zur Berechnung der diagnosebezogenen Veränderungsdaten.....	36
1.2	Berechnung der Relativgewichte zur Berechnung der diagnosebezogenen Veränderungsdaten	37
1.2.1	Einleitung	37
1.2.2	Komprimierung der Risikokategorien zur Berechnung der diagnosebezogenen Veränderungsdaten	38
1.2.3	Kalibrierung der komprimierten Risikokategorien zur Berechnung der diagnosebezogenen Veränderungsdaten	39
1.2.4	Ermittlung bundeseinheitlicher Relativgewichte zur Berechnung der diagnosebezogenen Veränderungsdaten.....	41
1.3	Berechnung der diagnosebezogenen Veränderungsdaten.....	41
1.3.1	Abgrenzung der Anwendungsmenge zur Berechnung der diagnosebezogenen Veränderungsdaten	41
1.3.2	Demografische Hochrechnungsfaktoren zur Berechnung der diagnosebezogenen Veränderungsdaten	42
1.3.3	Schritte zur Berechnung der diagnosebezogenen Veränderungsdaten	43
2	Ergebnisse: diagnosebezogene Veränderungsdaten.....	48
3	Empirische Ergebnisse zum Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsdaten und empirische Ergebnisse zu den diagnosebezogenen Veränderungsdaten	50
3.1	Empirische Ergebnisse zum Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsdaten.....	50
3.1.1	Komprimierung: Auswahl THCC und Bildung von RHCC zum Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsdaten.....	50
3.1.2	Kalibrierung des Regressionsmodells zur Berechnung der diagnosebezogenen Veränderungsdaten	52
3.1.3	Ausgewählte statistische Kennzahlen zum Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsdaten	58
3.2	Empirische Ergebnisse zu den diagnosebezogenen Veränderungsdaten	59
3.2.1	Entwicklung der Häufigkeiten von Diagnosen, Risikokategorien und Risikoklassen von 2017 auf 2018	59
3.2.2	Einzelbeiträge der Risikoklassen zu den regionalen diagnosebezogenen Veränderungsdaten...	65

1 Vorgehen / Berechnungsschritte zur Berechnung der diagnosebezogenen Veränderungsrate

Die Berechnung der diagnosebezogenen Veränderungsrate erfolgt in mehreren Schritten, die in den folgenden Abschnitten näher erläutert werden sollen.

1.1 Zuordnung zu Risikokategorien (Klassifikation) zur Berechnung der diagnosebezogenen Veränderungsrate

Zunächst werden jedem Versicherten aufgrund seines Alters und Geschlechts eine Alters- und Geschlechtsgruppe (AGG) sowie weiterhin aufgrund seiner Diagnosen (sofern vorhanden) eine oder mehrere sogenannte „hierarchisierte Condition Categories“ (HCC) zugeordnet. Hierfür ergeben sich, wie in Abbildung 1 schematisch dargestellt, aus den Diagnosen des Versicherten zunächst eine oder mehrere Condition Categories (CC) und dann aus der Menge der dem Versicherten so zugeordneten CC die diesem zugeordneten HCC. Dieser gesamte Vorgang wird als Klassifikation bezeichnet (für weitere Informationen siehe auch den Abschnitt zu den medizinischen Weiterentwicklungen in Teil D).

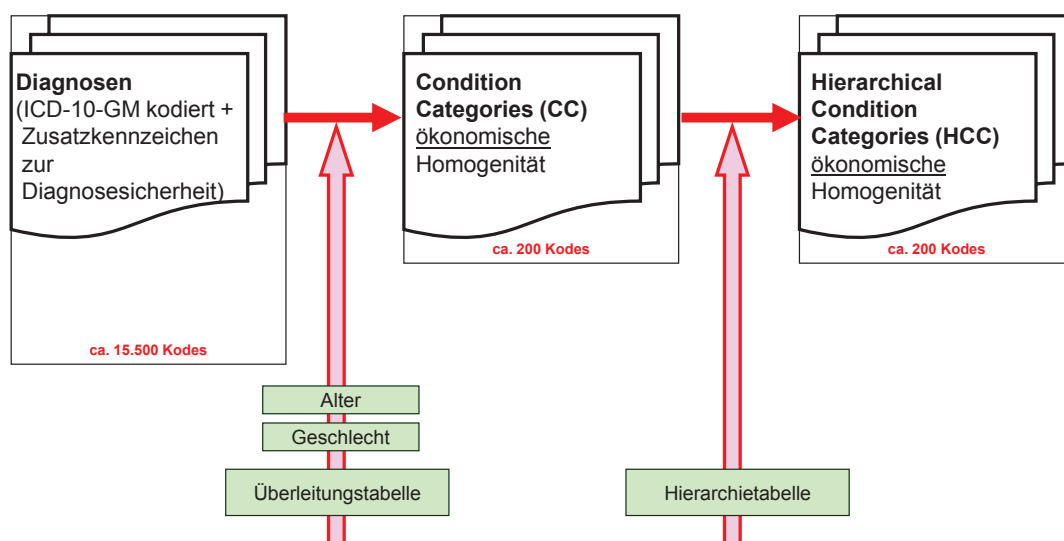


Abbildung 1: Bildung der diagnosebezogenen Risikokategorien

Die in Abbildung 1 dargestellte Zuordnung von CC aufgrund vorliegender Diagnosen erfolgt mit Hilfe sogenannter „Überleitungstabellen“. Dadurch werden die ca. 15.500 existierenden Diagnosen zu ca. 200 CC verdichtet. Die Überleitungstabellen enthalten neben der Zuordnung einer Diagnose zu einer CC¹ teilweise auch Zusatzbedingungen hinsichtlich des Alters und des Geschlechts der Versicherten. Einem Versicherten wird in einem solchen

¹In einzelnen Fällen ist auch die Zuordnung einer Diagnose zu zwei CC möglich.

Fall eine CC aufgrund einer Diagnose nur dann zugewiesen, wenn das Alter und Geschlecht des Versicherten innerhalb der Vorgaben liegt, die in den Überleitungstabellen zu den einzelnen Diagnosen angegeben sind. Zur anschließenden Durchführung der Hierarchisierung liegt eine sogenannte „Hierarchietabelle“ vor, die einzelne CC „über“ andere CC einordnet/hierarchisiert (z. B. CC1 über CC2, CC2 über CC3)². Werden einem Versicherten mehrere CC (z. B. die CC1 und die CC2) zugeordnet, so werden nur die in der Hierarchie am höchsten stehenden CC weiter betrachtet (im Beispiel wird dem Versicherten nur die CC1 zugeordnet). Die nach der Hierarchisierung verbleibenden CC eines Versicherten werden als HCC des Versicherten bezeichnet, und es gilt, dass die Anzahl der CC eines Versicherten immer größer oder gleich der Anzahl der HCC dieses Versicherten ist. Die vollständige formale Beschreibung der Klassifikation einschließlich der oben erwähnten Überleitungstabellen und Hierarchietabellen wird als Klassifikationssystem bezeichnet.

Die Beschreibung, welche HCC einem Versicherten zugeordnet wurden, erfolgt in den nachfolgend beschriebenen Berechnungen durch die Definition einer Dummy-Variablen für jede HCC, welche jeweils nur den Wert 1 (einem Versicherten wurde die HCC zugeordnet) oder 0 (einem Versicherten wurde die HCC nicht zugeordnet) annehmen kann. Ebenso wird durch solche Dummy-Variablen beschrieben, welche AGG einem Versicherten zugeordnet wurde. Das Ergebnis der Klassifikation ist somit eine Tabelle, die je Versicherten eine Zeile mit den Ausprägungen (0 oder 1) der AGG-Variablen und der HCC-Variablen³ enthält.

1.2 Berechnung der Relativgewichte zur Berechnung der diagnosebezogenen Veränderungsraten

1.2.1 Einleitung

Die (diagnosebezogenen) Relativgewichte zur Berechnung der diagnosebezogenen Veränderungsraten werden ebenso wie bei den demografischen Veränderungsraten mit einem zweijährig-prospektiven Regressionsmodell berechnet, welches mit Hilfe einer gewichteten multiplen linearen Regression erstellt wird. Datengrundlage für diese Regression ist die wie in Teil B, Abschnitt 1.1.1 abgegrenzte Kalibrierungsmenge. Die abhängige Variable der Regression ist der in Teil B, Abschnitt 1.1.3 definierte relative Leistungsbedarf eines Versicherten. Die unabhängigen Variablen des Regressionsmodells sind die AGG- und HCC-Variablen (im Folgenden auch als Risikokategorien bezeichnet) bzw. daraus abgeleitete Variablen (im Folgenden als Risikoklassen bezeichnet). Die verwendete

²Die Hierarchisierung ist transitiv, d.h. wenn CC1 über CC2 und CC2 über CC3 gilt, dann gilt auch CC1 über CC3.

³Das gemäß Beschluss des Bewertungsausschusses in seiner 503. Sitzung zu verwendende Klassifikationssystem umfasst 34 AGG und 188 CC bzw. HCC.

Regressionsgleichung ist in Formel (1) schematisch dargestellt.

$$y_i^{2018} = \sum_{j=1}^m \beta_j \cdot x_{i,j}^{2016} + \varepsilon_i \quad (1)$$

mit

- y_i^{2018} relativer Leistungsbedarf des i -ten Versicherten im Jahr 2018
- β_j Modellparameter der j -ten unabhängigen Variablen (Risikokategorie oder Risikoklasse)
- $x_{i,j}^{2016}$ Ausprägung der j -ten unabhängigen Variablen beim i -ten Versicherten im Jahr 2016
- m Anzahl der unabhängigen Variablen der Regression
- ε_i Fehlerterm der Regression

Die Gewichtung der Versicherten in der Regression erfolgt mit der Anzahl ihrer Versichertenquartale im Jahr 2018. Die für die Modellparameter (β_j) berechneten Schätzwerte ($\hat{\beta}_j$) werden als (diagnosebezogene) Relativgewichte bezeichnet.

1.2.2 Komprimierung der Risikokategorien zur Berechnung der diagnosebezogenen Veränderungsdaten

In diesem und dem nächsten Abschnitt wird beschrieben, wie die Risikokategorien ggf. weiter verdichtet und dabei neue Variablen abgeleitet werden, die dann als unabhängige Variablen (Risikoklassen) im Regressionsmodell verwendet werden. Der dazu zuerst durchgeführte und als Komprimierung bezeichnete Vorgang erfolgt gemäß dem in Nr. 3.1 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung beschriebenen Komprimierungsalgorithmus (wortwörtlich aus diesem Beschluss übernommene Textpassagen sind im Folgenden wieder kursiv gesetzt):

1. *Auf Basis der Risikokategorien der Versicherten des Jahres 2016 als unabhängige Variablen und des annualisierten morbiditätsbedingten Leistungsbedarfs dieser Versicherten des Jahres 2018, dividiert durch den mit der Anzahl der Versichertenquartale gewichteten Mittelwert des annualisierten Leistungsbedarfs in der Kalibrierungsmenge, als abhängiger Variable werden zunächst zur Bestimmung der ökonomischen Relevanz der HCCs vorläufige Relativgewichte durch eine gewichtete multiple lineare Regression bestimmt. Die ökonomische Relevanz jeder HCC wird ermittelt als das Produkt aus ihrem vorläufig bestimmten Relativgewicht und der Anzahl der Versicherten im Jahre 2016 mit dieser HCC in der Kalibrierungsmenge gemäß Nr. 2.3. Die HCCs werden dann entsprechend ihrer ökonomischen Relevanz*

in eine absteigende Rangfolge sortiert. In das Klassifikationsmodell werden jene HCCs entsprechend des Rangplatzes ihrer ökonomischen Relevanz unkomprimiert aufgenommen, bis die Summe ihrer ökonomischen Relevanzen erstmals mindestens 70 % der Summe der ökonomischen Relevanzen aller HCCs erreicht hat. Die auf diese Weise in das Klassifikationsmodell aufgenommenen HCCs werden als THCCs bezeichnet.

- 2. Zusätzlich werden jene HCCs als THCCs in das Klassifikationsmodell aufgenommen, deren vorläufiges Relativgewicht mindestens 1,5 beträgt.*
- 3. Darüber hinaus werden die erklärenden Variablen zusätzlich zu Schritt 1 und Schritt 2 ggf. um weitere HCCs aufgrund ihres maximalen Beitrags zur erklärten Varianz des Klassifikationsmodells ergänzt. Dazu wird ein Regressionsmodell mit den 34 Alters- und Geschlechtsgruppen und einer Anzahl von HCCs gebildet, wobei die Anzahl der in dieses Regressionsmodell aufzunehmenden HCCs auf die gleiche Anzahl beschränkt wird, wie THCCs nach Schritt 1 ausgewählt wurden. Die so ermittelten HCCs werden zusätzlich als THCCs in das Klassifikationsmodell aufgenommen, sofern diese nicht bereits zuvor als THCC in das Klassifikationsmodell aufgenommen worden sind.*
- 4. Die nicht nach Schritt 1 bis 3 ausgewählten HCCs werden entsprechend ihrer jeweiligen organsystembezogenen Zuordnung zusammengefasst und als RHCCs in das Klassifikationsmodell aufgenommen.*

THCCs und RHCCs bilden, mit Ausnahme der RHCC024 (Krankheit und Zustände der Schwangerschaft, Geburt und Wochenbett), zusammen mit den 34 Alters- und Geschlechtsgruppen (AGG) die unabhängigen Variablen des diagnosebezogenen Klassifikationsmodells und werden im Folgenden als komprimierte Risikokategorien bezeichnet.

1.2.3 Kalibrierung der komprimierten Risikokategorien zur Berechnung der diagnosebezogenen Veränderungsraten

Im Anschluss an den Komprimierungsalgorithmus werden die komprimierten Risikokategorien gemäß Nr. 3.2 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung kalibriert, d.h. das Klassifikationsmodell wird mittels deterministischer Regeln um negative und insignifikante Relativgewichte bereinigt.

Zur Vorbereitung des Kalibrierungsalgorithmus wird das Regressionsmodell mit den komprimierten Risikokategorien und dem annualisierten morbiditätsbedingten Leistungsbedarf, dividiert durch den mit der Anzahl der Versicherten quartale gewichteten Mittelwert des annualisierten Leistungsbedarfs in der Kalibrierungsmenge, berechnet. Eventuell auftretende negative oder insignifikante (p -Wert $\geq 0,05$) Relativgewichte komprimierter Risikokategorien werden innerhalb des Kalibrierungsalgorithmus in der nachfolgenden Schrittfolge behandelt:

- 1. Bei negativen Relativgewichten von RHCCs:*

- a) Die HCC mit dem größten negativen Gewicht (kleinsten Gewicht) innerhalb der RHCC gleich null setzen (dazu vorläufige Relativgewichte aus Nr. 3.1 Schritt 1 verwenden); die betroffene RHCC setzt sich aus den verbliebenen HCCs zusammen; erneute Regression
- b) Schritt a wird solange wiederholt, bis die RHCC ein positives Gewicht hat (bzw. die RHCC gleich null gesetzt ist)

2. Bei insignifikanten Relativgewichten (p -Wert $\geq 0,05$) von RHCCs:

- a) Die HCC mit dem größten p -Wert innerhalb der RHCC gleich null setzen (dazu vorläufige p -Werte aus Nr. 3.1 Schritt 1 verwenden); die betroffene RHCC setzt sich aus den verbliebenen HCCs zusammen; erneute Regression
- b) Schritt a wird solange wiederholt, bis die RHCC ein signifikantes Gewicht hat (bzw. die RHCC gleich null gesetzt ist)

Ist mehr als eine RHCC negativ (bzw. insignifikant), wird Schritt 1 (bzw. Schritt 2) zunächst für die RHCC mit dem größten negativen Relativgewicht (bzw. dem größten p -Wert) durchgeführt. Wenn diese RHCC nach Durchlaufen des Algorithmus ein positives Relativgewicht hat (bzw. signifikant ist) oder null gesetzt ist, wird die RHCC mit dem dann größten negativen Relativgewicht (größten p -Wert) gemäß Schritt 1 (bzw. Schritt 2) bearbeitet.

Dieser Prozess wird solange wiederholt, bis keine RHCC mehr negativ (bzw. insignifikant) ist.

3. AGGs mit negativen bzw. insignifikanten Relativgewichten (p -Wert $\geq 0,05$) werden symmetrisch für beide Geschlechtsgruppen, beginnend mit den betroffenen AGG mit dem höchsten Alter, schrittweise mit der nächstjüngeren AGG (nächstältere AGG, falls die jüngste AGG betroffen ist) zusammengefasst und nach jedem Ausschluss eine erneute Regression durchgeführt.

Ist nach Abschluss von Schritt 3 eine oder mehrere RHCCs negativ (bzw. insignifikant), so wird der Algorithmus ab Schritt 1 erneut durchlaufen.

4. Sollten nach dem Algorithmus gemäß Schritt 1 bis 3 THCCs mit negativen oder insignifikanten Relativgewichten resultieren, sind diese in die ihnen nach organ-systembezogener Zuordnung entsprechenden RHCCs einzugliedern und erneut auf negative oder insignifikante verbleibende Risikokategorien zu prüfen und ggf. der Algorithmus ab Schritt 1 erneut anzuwenden.

5. Abschließende Regression mit den sich so ergebenden AGGs, THCCs und RHCCs.

Nach der Kalibrierung verbleiben stets THCCs, RHCCs und AGGs mit positiven und signifikanten Relativgewichten im Klassifikationsmodell. Diese werden als Risikoklassen des Klassifikationsmodells bezeichnet.

1.2.4 Ermittlung bundeseinheitlicher Relativgewichte zur Berechnung der diagnosebezogenen Veränderungsdaten

Die bundeseinheitlichen Relativgewichte zur Berechnung der Veränderungsdaten werden anhand einer gewichteten linearen Regression mit dem relativen Leistungsbedarf als abhängige Variable und den Risikoklassen (AGGs, THCCs, RHCCs) als unabhängige Variablen auf der Kalibrierungsmenge bestimmt. Die sich so ergebenden Schätzwerte für die Modellparameter werden als bundeseinheitliche Relativgewichte bezeichnet.

1.3 Berechnung der diagnosebezogenen Veränderungsdaten

1.3.1 Abgrenzung der Anwendungsmenge zur Berechnung der diagnosebezogenen Veränderungsdaten

Die Berechnung der diagnosebezogenen Veränderungsdaten wird auf der sogenannten „Anwendungsmenge“ durchgeführt. Diese wird gemäß Nr. 2.5 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung wie folgt gebildet:

Die Datengrundlage zur Ermittlung der Morbiditätsindizes und der diagnosebezogenen Veränderungsdaten wird hinsichtlich der Versicherten der Geburtstagsstichprobe nach Nr. 2.1 wie folgt abgegrenzt:

- 1. In den Jahren 2017 und 2018 werden die Versicherten mit Versichertenzeit im jeweiligen Jahr ausgewählt, die jahresspezifisch mindestens 120 Tage Versichertenzeit aufweisen oder im jeweiligen Jahr gestorben sind oder für die in mindestens einem Quartal des betreffenden Kalenderjahres ohne Versichertenzeit Abrechnungsfälle vorliegen.*
- 2. Längsschnittliche SV-Teilnehmer werden nur dann in die Anwendungsmenge nach Ziffer 1 aufgenommen, wenn sie einem in Nr. 2.2.8 genannten KV-Bezirk zugeordnet wurden und gemäß der Kennzeichnung zur Selektivvertragsteilnahme nach Nr. 2.1 für sämtliche Quartale der Jahre 2017 und 2018 mit Teilnahme als Teilnehmer ausschließlich an Selektivverträgen nach § 73b SGB V identifizierbar sind.*
- 3. Von der Versichertenmenge nach Ziffern 1 und 2 werden weiterhin diejenigen Versicherten ausgenommen, die zu irgendeinem Zeitpunkt in den Jahren 2017 bis 2018 in Feld 14 der Satzart 201 einen Wert von 1 aufweisen.*
- 4. Von der Versichertenmenge nach Ziffern 1 und 2 werden weiterhin diejenigen Versicherten ausgenommen, die in mindestens einem der Jahre 2017 bis 2018 Wohnausländer gemäß Nr. 2.2.9 und darüber hinaus im entsprechenden Jahr gemäß der Zuordnungsvorschrift in Nr. 2.2.1 einem der KV-Bezirke Baden-Württemberg oder Nordrhein zugeordnet sind.*

Die erwähnten KV-Bezirke nach Nr. 2.2.8 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung sind Bayern und Baden-Württemberg. Die erwähnte Wohnausländer-

Definition nach Nr. 2.2.9 und die erwähnte Zuordnungsvorschrift nach Nr. 2.2.1 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung werden in Teil A, Abschnitt 3.1.4 beschrieben.

1.3.2 Demografische Hochrechnungsfaktoren zur Berechnung der diagnosebezogenen Veränderungsdaten

Mit den demografischen Hochrechnungsfaktoren (DHF) wird die Versichertenstichprobe, im vorliegenden Fall die Anwendungsmenge zur Berechnung der diagnosebezogenen Veränderungsdaten, mit ihrer Alters-Geschlechts-Verteilung auf alle GKV-Versicherten im jeweiligen KV-Bezirk in der Alters-Geschlechts-Struktur der KM6-Statistik hochgerechnet. Die DHF, welche bei der Ermittlung der diagnosebezogenen Veränderungsdaten unter Verwendung des Klassifikationsmodells für das Jahr 2021 genutzt werden, sind in Nr. 2.7 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung beschrieben.

Die Ermittlung und Anwendung der DHF erfolgt jahres- und KV-spezifisch. Zusätzlich erfolgt eine Differenzierung der DHF nach Wohnausländern bzw. Nicht-Wohnausländern. Die Nicht-Wohnausländer werden auf die Versicherten der ANZVER-Daten⁴ unter Nutzung der Alters-Geschlechtsverteilung der KM6-Statistik hochgerechnet. Die Wohnausländer werden unter Beachtung der Alters-Geschlechtsverteilung auf die entsprechenden Versicherten der KM6-Statistik hochgerechnet. Dabei werden Wohnausländer sowohl in der Stichprobe als auch in der KM6-Statistik der KV-Region zugeordnet, in der ihre Krankenkasse ihren Sitz hat. Bei fusionierten Krankenkassen wird dabei jeweils der letzte bekannte Sitz der Krankenkasse (Stand: 2. April 2020) verwendet.

Für die KV-Bezirke Baden-Württemberg und Nordrhein werden die DHF nur für die Nicht-Wohnausländer berechnet.

Die DHF für die Jahre $t = 2017, 2018$ berechnen sich wie folgt:

$$DHF_{AG,WA}^{t,KV} = \frac{N_{AG,WA}^{t,KV}}{n_{AG,WA}^{t,KV}} \quad (2)$$

wobei

$$N_{AG,WA}^{t,KV} = \text{Anzahl_Vers_KM6}_{AG,WA}^{t,KV} \cdot A_{WA}^{t,KV} \quad (3)$$

mit

$$A_{WA}^{t,KV} = \begin{cases} \frac{\text{Anzahl_Vers_ANZVER}_{AG,WA}^{t,KV}}{\sum_{AG} \text{Anzahl_Vers_KM6}_{AG,WA}^{t,KV}} & \text{für } WA = \text{Nicht_Wohnausländer} \\ 1 & \text{für } WA = \text{Wohnausländer} \end{cases} \quad (4)$$

⁴Für die ANZVER-Daten werden die Felder „Durchschnittliche quartalsbezogene Anzahl Versicherte“ und „Durchschnittliche quartalsbezogene Anzahl betreuter Personen“ der Satzart ANZVER87a_IK gemäß 369. BA herangezogen.

und

$$n_{AG,WA}^{t,KV} = \frac{1}{4} ANZAHL_VERS_QUARTALE_{AG,WA}^{t,KV} \quad (5)$$

$n_{AG,WA}^{t,KV}$ wird für die Jahre $t = 2017, 2018$ auf der Menge der längsschnittlichen Nicht-SV-Teilnehmer sowie der längsschnittlichen SV-Teilnehmer mit Wohnort in den KV-Bezirken Bayern und Baden-Württemberg, abgegrenzt nach Nr. 2.5 Ziffern 1, 3 und 4 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung berechnet. Falls $n_{AG,WA}^{t,KV} = 0$ ist, dann wird $DHF_{AG,WA}^{t,KV}$ auf null gesetzt.

Je KV-Bezirk, Alters- und Geschlechtsgruppe und Wohnausländerstatus ergibt sich jahresspezifisch ein DHF, der den entsprechenden Versicherten in der Anwendungsmenge zugeordnet wird.

Für die Ermittlung der diagnosebezogenen Veränderungsdaten werden von den SV-Teilnehmern bei der Berechnung von relativen Risikowerten gemäß der Formel (7) in Abschnitt 1.3.3 nur diejenigen Versicherten berücksichtigt, die ausschließlich an Verträgen gemäß § 73b SGB V teilgenommen haben. Um alle weiteren SV-Teilnehmer ebenfalls zu berücksichtigen, wird ein Korrekturfaktor für nur-§ 73b-Versicherte aufgenommen. Der Beschluss des Bewertungsausschusses in seiner 503. Sitzung legt in Nr. 2.8 die Bestimmung dieses Korrekturfaktors für die KV-Bezirke⁵ Bayern und Baden-Württemberg für jede Alters- und Geschlechtsgruppe und jedes Jahr $t = 2017, 2018$ fest:

$$K_{AG}^{t,KV} = \frac{n_{AG,nicht_nur_73b}^{t,KV} + n_{AG,nur_73b}^{t,KV}}{n_{AG,nur_73b}^{t,KV}} \quad (6)$$

Dabei ist $n_{AG,i}^{t,KV}$ die Anzahl der Versichertenjahre, nach Quartalen gebildet wie bei der Berechnung der DHF. Die Größe wird berechnet für die Kollektive i mit $i = nur_73b$ bzw. $i = nicht_nur_73b$ und die Jahre $t = 2017, 2018$. Falls $n_{AG,nur_73b}^{t,KV} = 0$, dann wird $K_{AG}^{t,KV}$ auf null gesetzt.

1.3.3 Schritte zur Berechnung der diagnosebezogenen Veränderungsdaten

Die Berechnung der diagnosebezogenen Veränderungsdaten erfolgt KV-spezifisch. Die Berechnung kann dabei in drei Schritte unterteilt werden:

1. Berechnung des relativen Risikowertes für jeden Versicherten der Anwendungsmenge
2. Berechnung von KV-spezifischen Morbiditätsindizes
3. Berechnung von KV-spezifischen Veränderungsdaten

⁵Dieser Korrekturfaktor wird ausschließlich bei der Ermittlung der diagnosebezogenen Veränderungsdaten für die KV-Bezirke Bayern und Baden-Württemberg genutzt, da nur in diesen beiden KV-Bezirken SV-Teilnehmer, die ausschließlich an Verträgen gemäß § 73b SGB V teilgenommen haben, bei der Berechnung von relativen Risikowerten berücksichtigt werden.

Im ersten Schritt werden für jeden Versicherten der Anwendungsmenge jahresspezifisch für die Jahre 2017 und 2018 relative Risikowerte ($\hat{y}_{i,VB}^{t,KV}$) gemäß der Formel (7) berechnet.

$$\hat{y}_{i,VB}^{t,KV} = \sum_{j=1}^r \hat{\beta}_j \cdot x_{i,j,VB}^{t,KV} \quad (7)$$

mit

- $\hat{y}_{i,VB}^{t,KV}$ relativer Risikowert des i -ten Versicherten der entsprechenden KV bezogen auf das Jahr t
 VB (Versorgungsbereich) = „gesamt“ (für alle Diagnosen von Nicht-SV-Teilnehmern) und
 VB (Versorgungsbereich) = „Nicht-HA“ (für alle Diagnosen aus dem nicht hausärztlichen Versorgungsbereich von SV-Teilnehmern mit Teilnahme an Verträgen ausschließlich nach § 73b SGB V in den KV-Bezirken Bayern und Baden-Württemberg)
- $x_{i,j,VB}^{t,KV}$ Ausprägung der Risikoklasse j beim i -ten Versicherten der entsprechenden KV im Jahr t für den Versorgungsbereich VB
- $\hat{\beta}_j$ bundeseinheitliches Relativgewicht der Risikoklasse j
- KV Kennzeichnung des KV-Bezirks, für den der Risikowert gebildet wird, entsprechend der Wohnortzuordnung des Versicherten i im Jahr t
- r Anzahl der Risikoklassen
- t Jahr, mit $t = 2017, 2018$

Die relativen Risikowerte $\hat{y}_{i,VB}^{t,KV}$ werden entsprechend der Modellbildung (siehe Abschnitt 1.2) bestimmt. In Formel (7) werden die für die Modellparameter berechneten Schätzwerte $\hat{\beta}_j$ sowohl für das Jahr 2017 als auch für das Jahr 2018 verwendet.

- $\hat{\beta}_j$ ist das Ergebnis der Modellierung gemäß Abschnitt 1.2 auf Grundlage der Jahre 2016 (Alter, Geschlecht, Diagnosen) bzw. 2018 (individueller relativer Leistungsbedarf zwei Jahre später).
- Der relative Risikowert für den Versicherten i auf Grundlage des Jahres t für den Versorgungsbereich VB ergibt sich aus seinen Diagnosen des Jahres t im Versorgungsbereich VB , seinem Alter in t , seinem Geschlecht und den normativ verwendeten Relativgewichten $\hat{\beta}_j$.

Aus den relativen Risikowerten aller Versicherten mit Wohnort im Bereich einer Kassenärztlichen Vereinigung werden im zweiten Schritt mit der Formel (8) Morbiditätsindizes für längsschnittliche Nicht-SV-Teilnehmer in allen KV-Bezirken und mit der Formel (9) Morbiditätsindizes für längsschnittliche SV-Teilnehmer in den KV-Bezirken Bayern und

Baden-Württemberg berechnet⁶.

$$\bar{y}_{gesamt}^{t,KV} = \frac{\sum_{i \in \text{NichtSVTN}} (\hat{y}_{i,gesamt}^{t,KV} \cdot AVQ_i^{t,KV} \cdot DHF_i^{t,KV})}{\sum_{i \in \text{NichtSVTN}} (AVQ_i^{t,KV} \cdot DHF_i^{t,KV})} \quad (8)$$

$$\bar{y}_{\text{Nicht-HA}}^{t,KV} = \frac{\sum_{i \in \text{SVTN}_{\text{nur}_{73b}}} (\hat{y}_{i,\text{Nicht-HA}}^{t,KV} \cdot AVQ_i^{t,KV} \cdot DHF_i^{t,KV} \cdot K_i^{t,KV})}{\sum_{i \in \text{SVTN}_{\text{nur}_{73b}}} (AVQ_i^{t,KV} \cdot DHF_i^{t,KV} \cdot K_i^{t,KV})} \quad (9)$$

mit

NichtSVTN längsschnittliche Nicht-SV-Teilnehmer des Jahres t in der Anwendungsmenge

SVTN_{nur_73b} längsschnittliche SV-Teilnehmer des Jahres t in der Anwendungsmenge

$\bar{y}_{gesamt}^{t,KV}$ Morbiditätsindex je KV-Bezirk ($KV = 1, \dots, 17$) des Jahres t mit VB (Versorgungsbereich) = „gesamt“ (für alle Diagnosen von Nicht-SV-Teilnehmern)

$\hat{y}_{i,gesamt}^{t,KV}$ relativer Risikowert des i -ten Versicherten der entsprechenden KV bezogen auf das Jahr t für den Versorgungsbereich „gesamt“

$\bar{y}_{\text{Nicht-HA}}^{t,KV}$ Morbiditätsindex je KV-Bezirk ($KV = 1, \dots, 17$) des Jahres t mit VB (Versorgungsbereich) = „Nicht-HA“ (für alle Diagnosen aus dem nicht hausärztlichen Versorgungsbereich von SV-Teilnehmern mit Teilnahme an Verträgen ausschließlich nach § 73b SGB V in den KV-Bezirken Bayern und Baden-Württemberg)

$\hat{y}_{i,\text{Nicht-HA}}^{t,KV}$ relativer Risikowert des i -ten Versicherten der entsprechenden KV bezogen auf das Jahr t für den Versorgungsbereich „Nicht-HA“

$AVQ_i^{t,KV}$ Anzahl der Versichertenquartale des i -ten Versicherten der entsprechenden KV im Jahr t

$DHF_i^{t,KV}$ Demografischer Hochrechnungsfaktor des i -ten Versicherten der entsprechenden KV für das Jahr t

$K_i^{t,KV}$ Korrekturfaktor für längsschnittliche SV-Teilnehmer des i -ten Versicherten der entsprechenden KV für das Jahr t

t Jahr, mit $t = 2017, 2018$

Hierzu werden auf der Anwendungsmenge, bezogen auf die beiden Versichertengruppen Nicht-SV-Teilnehmer und SV-Teilnehmer mit Teilnahme an Verträgen ausschließlich nach § 73b SGB V, KV-spezifische Morbiditätsindizes berechnet. Der Morbiditätsindex eines KV-Bezirks ist der mit der Anzahl der Versichertenquartale und den demografischen Hochrechnungsfaktoren gewichtete Mittelwert der relativen Risikowerte der Versicherten

⁶Für die KV-Bezirke Bayern und Baden-Württemberg werden je zwei Morbiditätsindizes mit den in der Formel (7) angegebenen Abgrenzungen für den Versorgungsbereich berechnet.

mit Wohnort im Bezirk einer Kassenärztlichen Vereinigung.

Im dritten Schritt erfolgt die Berechnung der Veränderungsdaten gemäß den im Beschluss des Bewertungsausschusses in seiner 503. Sitzung festgelegten Formeln. Danach bestimmen sich die diagnosebezogenen Veränderungsdaten aller KV-Bezirke, ausgenommen die KV-Bezirke Baden-Württemberg und Bayern, gemäß der Formel (10) als Quotient aus dem Morbiditätsindex des Folgejahres 2018 und dem Morbiditätsindex des Basisjahres 2017 abzüglich eins.

$$VR_{2017 \rightarrow 2018}^{KV} = \frac{\bar{y}_{gesamt}^{2018, KV}}{\bar{y}_{gesamt}^{2017, KV}} - 1 \quad (10)$$

mit

$VR_{2017 \rightarrow 2018}^{KV}$	diagnosebezogene Veränderungsrate für KV-Bezirk KV
$\bar{y}_{gesamt}^{2017, KV}$	Morbiditätsindex des Jahres 2017 für KV-Bezirk KV unter Verwendung aller Diagnosen des Jahres 2017 (wird ausschließlich für Nicht-SV-Teilnehmer verwendet)
$\bar{y}_{gesamt}^{2018, KV}$	Morbiditätsindex des Jahres 2018 für KV-Bezirk KV unter Verwendung aller Diagnosen des Jahres 2018 (wird ausschließlich für Nicht-SV-Teilnehmer verwendet)
KV	KV-Bezirk (außer Baden-Württemberg und Bayern)

Die diagnosebezogenen Veränderungsdaten für die KV-Bezirke Baden-Württemberg und Bayern ergeben sich als gewichtete Summe

- des Quotienten aus dem Morbiditätsindex des Jahres 2018 und dem Morbiditätsindex des Jahres 2017 der längsschnittlichen SV-Teilnehmer abzüglich eins und
- des Quotienten aus dem Morbiditätsindex des Jahres 2018 und dem Morbiditätsindex des Jahres 2017 der längsschnittlichen Nicht-SV-Teilnehmer abzüglich eins

gemäß der Formel (11):

$$VR_{2017 \rightarrow 2018}^{KV} = G_{KV} \cdot \left(\frac{\bar{y}_{Nicht-HA}^{2018, KV}}{\bar{y}_{Nicht-HA}^{2017, KV}} - 1 \right) + (1 - G_{KV}) \cdot \left(\frac{\bar{y}_{gesamt}^{2018, KV}}{\bar{y}_{gesamt}^{2017, KV}} - 1 \right) \quad (11)$$

mit

$VR_{2017 \rightarrow 2018}^{KV}$	diagnosebezogene Veränderungsrate für KV-Bezirk KV
$\bar{y}_{gesamt}^{2017, KV}$	Morbiditätsindex des Jahres 2017 für KV-Bezirk KV unter Verwendung aller Diagnosen des Jahres 2017 (wird ausschließlich für Nicht-SV-Teilnehmer verwendet)

$\bar{y}_{gesamt}^{2018,KV}$	Morbiditätsindex des Jahres 2018 für KV-Bezirk <i>KV</i> unter Verwendung aller Diagnosen des Jahres 2018 (wird ausschließlich für Nicht-SV-Teilnehmer verwendet)
$\bar{y}_{Nicht-HA}^{2017,KV}$	Morbiditätsindex des Jahres 2017 für KV-Bezirk <i>KV</i> unter Verwendung von ausschließlich nicht hausärztlichen Diagnosen des Jahres 2017 (wird ausschließlich für SV-Teilnehmer in den KV-Bezirken Baden-Württemberg und Bayern verwendet)
$\bar{y}_{Nicht-HA}^{2018,KV}$	Morbiditätsindex des Jahres 2018 für KV-Bezirk <i>KV</i> unter Verwendung von ausschließlich nicht hausärztlichen Diagnosen des Jahres 2018 (wird ausschließlich für SV-Teilnehmer in den KV-Bezirken Baden-Württemberg und Bayern verwendet)
G_{KV}	Gewichtungsfaktor (MGV-Leistungsbedarfsanteil der SV-Teilnehmer des KV-Bezirks Baden-Württemberg bzw. Bayern am für die Berechnung der Veränderungsdaten verwendeten Gesamt-MGV-Leistungsbedarf des jeweiligen KV-Bezirks)
<i>KV</i>	KV-Bezirk (ausschließlich Baden-Württemberg und Bayern)

Die Gewichtung der Veränderung des Morbiditätsindex für die längsschnittlichen SV-Teilnehmer erfolgt mit dem Faktor G_{KV} , dem Anteil des mit den DHF hochgerechneten tatsächlichen Leistungsbedarfs der morbiditätsbedingten Gesamtvergütung der SV-Teilnehmer des Jahres 2018 an dem mit den DHF hochgerechneten tatsächlichen Leistungsbedarf der morbiditätsbedingten Gesamtvergütung aller Versicherten der Anwendungsmenge mit Wohnort im jeweiligen KV-Bezirk für das Jahr 2018 gemäß der Formel (12).

$$G_{KV} = \frac{\sum_{i \in SVTN18 \cap KV} LB_Punkte_i^{2018,MGV} \cdot DHF_i^{2018,KV}}{\sum_{i \in KV} LB_Punkte_i^{2018,MGV} \cdot DHF_i^{2018,KV}} \quad (12)$$

mit

<i>KV</i>	KV-Bezirk Baden-Württemberg, KV-Bezirk Bayern
$LB_Punkte_i^{2018,MGV}$	morbiditätsbedingter Leistungsbedarf des Versicherten <i>i</i> in Punkten gemäß Nr. 2.2.5 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung.
<i>SVTN18</i>	SV-Teilnehmer im Jahr 2018 gemäß der Definition in Nr. 2.2.6 des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung

Der komplementäre Anteil, der als Differenz zu 100 % verbleibt, bestimmt die Gewichtung der Veränderung des Morbiditätsindex der Nicht-SV-Teilnehmer.

2 Ergebnisse: diagnosebezogene Veränderungsdaten

In Abschnitt 1 wird die Ermittlung der Veränderungsdaten dargestellt. Für die KV-Bezirke Bayern und Baden-Württemberg werden die diagnosebezogenen Veränderungsdaten jeweils aus zwei Komponenten berechnet, welche anschließend zu einer Veränderungsrate des jeweiligen KV-Bezirks zusammengefasst werden. Die separaten Komponenten der diagnosebezogenen Veränderungsdaten für die KV-Bezirke Bayern und Baden-Württemberg werden

- für Nicht-SV-Teilnehmer unter Verwendung aller Diagnosen und
- für SV-Teilnehmer einzig unter Verwendung der Diagnosen des nicht hausärztlichen Versorgungsbereichs

ermittelt. Diese für die Berechnung der diagnosebezogenen Veränderungsdaten der KV-Bezirke Bayern und Baden-Württemberg ermittelten Zwischenergebnisse sind in Tabelle 1 aufgeführt.

KV-Bezirk	Komponenten der Veränderungsrate	
	SV-TN	Nicht-SV-TN
Baden-Württemberg	0,6049 %	0,1411 %
Bayern	1,3237 %	0,4660 %

Tabelle 1: Komponenten der Veränderungsdaten für die Berechnung der diagnosebezogenen Veränderungsdaten der KV-Bezirke Baden-Württemberg und Bayern

Um die zwei separaten Komponenten der diagnosebezogenen Veränderungsdaten der KV-Bezirke Bayern und Baden-Württemberg zu jeweils einer finalen diagnosebezogenen Veränderungsrate zusammenzuführen, wird als Gewichtungsfaktor der Anteil des Leistungsbedarfs der morbiditätsbedingten Gesamtvergütung der SV-Teilnehmer im jeweiligen KV-Bezirk am gesamten Leistungsbedarf der morbiditätsbedingten Gesamtvergütung des KV-Bezirks im Jahr 2018 herangezogen. Hierfür wurde die Formel (12) in Abschnitt 1.3.3 verwendet. Die berechneten Gewichtungsfaktoren sind in Tabelle 2 aufgeführt.

KV-Bezirk	Gewichtungsfaktor
Baden-Württemberg	0,186806
Bayern	0,110663

Tabelle 2: Gewichtungsfaktoren bei der Berechnung der diagnosebezogenen Veränderungsdaten für die KV-Bezirke Baden-Württemberg und Bayern

Unter Verwendung der Vorgaben des Beschlusses des Bewertungsausschusses in seiner

503. Sitzung ergeben sich die in Tabelle 3 dargestellten KV-spezifischen diagnosebezogenen Veränderungsdaten.

KV-Bezirk	Veränderungsrate
Schleswig-Holstein	0,5156 %
Hamburg	-0,1173 %
Bremen	-0,3923 %
Niedersachsen	0,3625 %
Westfalen-Lippe	0,7504 %
Nordrhein	0,6462 %
Hessen	0,3712 %
Rheinland-Pfalz	0,7854 %
Baden-Württemberg	0,2277 %
Bayern	0,5609 %
Berlin	-0,3892 %
Saarland	0,8948 %
Mecklenburg-Vorpommern	0,8985 %
Brandenburg	0,3569 %
Sachsen-Anhalt	0,8620 %
Thüringen	0,8355 %
Sachsen	0,5004 %

Tabelle 3: Diagnosebezogene Veränderungsdaten je KV-Bezirk für das Jahr 2021. Für die Berechnung der diagnosebezogenen Veränderungsdaten der KV-Bezirke Bayern und Baden-Württemberg werden die Gewichtungsfaktoren aus Tabelle 2 und die Komponenten der Veränderungsdaten aus Tabelle 1 verwendet.

3 Empirische Ergebnisse zum Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsdaten und empirische Ergebnisse zu den diagnosebezogenen Veränderungsdaten

3.1 Empirische Ergebnisse zum Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsdaten

3.1.1 Komprimierung: Auswahl THCC und Bildung von RHCC zum Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsdaten

Auf Grundlage des im Beschluss des Bewertungsausschusses in seiner 503. Sitzung angegebenen und in Abschnitt 1.2.2 beschriebenen Komprimierungsalgorithmus wurden 41 Risikokategorien als THCC ausgewählt. Die Auswahl der THCC erfolgte

- in Schritt 1 aufgrund ihrer ökonomischen Relevanz (Ökonom. Relevanz),
- in Schritt 2 aufgrund ihres Relativgewichts ($RG \geq 1,5$) oder
- in Schritt 3 aufgrund ihres Beitrags zur erklärten Varianz (R^2).

In Tabelle 4 ist dargestellt, bei welchem Schritt des Komprimierungsalgorithmus die THCC jeweils ins Modell aufgenommen wurden. Sofern eine THCC über den ersten Aufnahmegrund hinaus auch in einem folgenden Schritt ins Modell gelangt wäre, ist dieser Aufnahmegrund ebenfalls angegeben.

Tabelle 4: Auswahl der THCC im Modell für die Berechnung der diagnosebezogenen Veränderungsdaten für das Jahr 2021

THCC im Modell		Aufnahmegrund
THCC001	HIV/AIDS	$RG \geq 1, 5, R^2$
THCC006	Andere Infektionskrankheiten	Ökonom. Relevanz, R^2
THCC007	Sekundäre bösartige Neubildungen	R^2
THCC010	Bösartige Neubildungen Dickdarm, Rektum, Anus, Mamma, Geschlechtsorgane, Niere und ableitende Harnwege, Auge, Schilddrüse, maligne Melanome	Ökonom. Relevanz, R^2
THCC013	Andere Neubildungen	Ökonom. Relevanz
THCC022	Andere schwerwiegende endokrine und Stoffwechselerkrankungen	R^2
THCC024	Andere endokrine, metabolische und Ernährungsstörungen	Ökonom. Relevanz, R^2
THCC036	Andere gastrointestinale Erkrankungen	Ökonom. Relevanz, R^2
THCC038	Entzündliche Systemerkrankungen des rheumatischen Formenkreises	Ökonom. Relevanz, R^2

Fortsetzung auf der nächsten Seite...

Tabelle 4 – Fortsetzung

	THCC im Modell	Aufnahmegrund
THCC040	Cox- oder Gonarthrose	Ökonom. Relevanz
THCC041	Erkrankungen des Knochens und des Knorpels	Ökonom. Relevanz, R^2
THCC043	Andere Erkrankungen der Muskeln, Sehnen, Gelenke und des Bindegewebes	Ökonom. Relevanz
THCC050	Alterserkrankungen, nicht-psychotische hirnorganische Erkrankungen/Zustände	Ökonom. Relevanz, R^2
THCC055	Bipolare Störungen, depressive Episoden	Ökonom. Relevanz, R^2
THCC057	Persönlichkeitsstörungen	Ökonom. Relevanz, R^2
THCC058	Depressive Episode, Angst, Dysthymia	Ökonom. Relevanz, R^2
THCC060	Angststörungen, sexuelle Störungen, Belastungs- und Anpassungsstörungen	Ökonom. Relevanz, R^2
THCC071	Polyneuropathie, Myasthenie, Myopathie	Ökonom. Relevanz, R^2
THCC072	Multiple Sklerose	R^2
THCC074	Epilepsie und nichtepileptische Anfälle	R^2
THCC076	Radikulopathien, andere neurologische Erkrankungen und Folgen von Verletzungen des Nervensystems	Ökonom. Relevanz, R^2
THCC080	Kardiomyopathien, Herzinsuffizienz, PAH	R^2
THCC086	Erkrankungen der Herzklappen und rheumatische Herzerkrankungen	R^2
THCC090	Hypertensive Herzerkrankung	R^2
THCC091	Essentielle Hypertonie	Ökonom. Relevanz, R^2
THCC106	Erkrankungen der Arterien, Venen, Lymphgefäße	Ökonom. Relevanz
THCC110	Asthma bronchiale	Ökonom. Relevanz, R^2
THCC115	Akute Bronchitis, Grippe	Ökonom. Relevanz
THCC123	Katarakt	Ökonom. Relevanz, R^2
THCC124	Andere Erkrankungen des Auges	Ökonom. Relevanz, R^2
THCC126	Hörverlust	Ökonom. Relevanz
THCC127	Andere Erkrankungen HNO-Bereich	Ökonom. Relevanz
THCC134	Inkontinenz	R^2
THCC135	Entzündungen der ableitenden Harnwege	Ökonom. Relevanz
THCC139	Störungen der Menstruation, klimakterische Störungen, andere Erkrankungen der weiblichen Geschlechtsorgane	Ökonom. Relevanz, R^2
THCC140	Erkrankungen des männlichen Genitalapparates	Ökonom. Relevanz

Fortsetzung auf der nächsten Seite...

Tabelle 4 – Fortsetzung

	THCC im Modell	Aufnahmegrund
THCC153	Verbrennungen 1. und 2. Grades, andere Erkrankungen der Haut und Hautanhangsgebilde	Ökonom. Relevanz, R^2
THCC162	Extremitätenfrakturen, andere Verletzungen	Ökonom. Relevanz, R^2
THCC186	Multiples Myelom/Plasmozytom und andere chronische Leukämien	R^2
THCC190	Knöcherne Erkrankungen der Wirbelsäule, Erkrankungen der Bandscheiben, Stenosen des Spinalkanals und der Foramina intervertebralia	Ökonom. Relevanz, R^2
THCC194	Migräne und sonstige chronische Kopfschmerzen	Ökonom. Relevanz

In Tabelle 5 sind diejenigen THCC aufgelistet, die im Modell zur Berechnung der Veränderungsdaten für das Jahr 2021 gegenüber dem Modell zur Berechnung der Veränderungsdaten für das Jahr 2020 weggefallen oder hinzugekommen sind.

THCC ist nur im Modell zur Berechnung der Veränderungsdaten für das Jahr 2020 (weggefallene THCC)	THCC ist nur im Modell zur Berechnung der Veränderungsdaten für das Jahr 2021 (hinzugekommene THCC)	Beschriftung
THCC084		Koronare Herzkrankheit / andere chronisch-ischämische Erkrankungen des Herzens
THCC128		Funktionsstörung, Versagen und Abstoßung eines Nierentransplantates

Tabelle 5: Weggefallene und hinzugekommene THCC im Modell zur Berechnung der diagnosebezogenen Veränderungsdaten für das Jahr 2021 im Vergleich zum Modell zur Berechnung der diagnosebezogenen Veränderungsdaten für das Jahr 2020

3.1.2 Kalibrierung des Regressionsmodells zur Berechnung der diagnosebezogenen Veränderungsdaten

Es erfolgte keine Kalibrierung der komprimierten Risikokategorien anhand des in Abschnitt 1.2.3 beschriebenen Kalibrierungsalgorithmus, da keine der komprimierten Risikokategorien ein negatives oder insignifikantes Relativgewicht aufwies. Die komprimierten Risikokategorien sind damit gleichzeitig die Risikoklassen des Modells zur Berechnung der Veränderungsdaten für das Jahr 2021. Tabelle 6 enthält neben den Relativgewichten

(Schätzwerte für die Modellparameter $\hat{\beta}_j$ aus Formel (1)) der Risikoklassen auch die Anzahl der Stichprobenversicherten (N) in der Kalibrierungsmenge, denen die entsprechende Risikoklasse zugeordnet wurde.

Tabelle 6: *Relativgewichte im Modell zur Berechnung der diagnosebezogenen Veränderungsdaten für das Jahr 2021. N(2016) ist die Anzahl der Versicherten in der Kalibrierungsmenge, denen die entsprechende Risikoklasse im Jahr 2016 zugeordnet wurde.*

Risikoklassen		Relativgewicht	N(2016) in Kalibrierungsmenge
AGG001	Weiblich, 0 Jahre	22,83 %	6.146
AGG002	Weiblich, 1 bis 5 Jahre	21,02 %	233.651
AGG003	Weiblich, 6 bis 12 Jahre	19,02 %	324.049
AGG004	Weiblich, 13 bis 17 Jahre	36,32 %	249.730
AGG005	Weiblich, 18 bis 24 Jahre	26,20 %	378.011
AGG006	Weiblich, 25 bis 34 Jahre	30,47 %	671.017
AGG007	Weiblich, 35 bis 44 Jahre	31,76 %	644.235
AGG008	Weiblich, 45 bis 54 Jahre	34,67 %	868.100
AGG009	Weiblich, 55 bis 59 Jahre	36,96 %	411.781
AGG010	Weiblich, 60 bis 64 Jahre	35,15 %	345.579
AGG011	Weiblich, 65 bis 69 Jahre	37,77 %	301.039
AGG012	Weiblich, 70 bis 74 Jahre	45,79 %	246.916
AGG013	Weiblich, 75 bis 79 Jahre	47,20 %	310.212
AGG014	Weiblich, 80 bis 84 Jahre	39,71 %	202.326
AGG015	Weiblich, 85 bis 89 Jahre	38,42 %	117.106
AGG016	Weiblich, 90 bis 94 Jahre	49,29 %	49.756
AGG017	Weiblich, 95 bis 124 Jahre	64,93 %	10.899
AGG018	Männlich, 0 Jahre	29,03 %	6.453
AGG019	Männlich, 1 bis 5 Jahre	25,13 %	245.343
AGG020	Männlich, 6 bis 12 Jahre	20,57 %	341.159
AGG021	Männlich, 13 bis 17 Jahre	17,82 %	263.175
AGG022	Männlich, 18 bis 24 Jahre	13,16 %	393.193
AGG023	Männlich, 25 bis 34 Jahre	16,72 %	677.729
AGG024	Männlich, 35 bis 44 Jahre	20,73 %	611.102
AGG025	Männlich, 45 bis 54 Jahre	26,81 %	797.786
AGG026	Männlich, 55 bis 59 Jahre	35,90 %	370.676
AGG027	Männlich, 60 bis 64 Jahre	35,65 %	299.113

Fortsetzung auf der nächsten Seite...

Tabelle 6 – Fortsetzung

	Risikoklassen	Relativgewicht	N(2016) in Kalibrierungs- menge
AGG028	Männlich, 65 bis 69 Jahre	39,19 %	248.727
AGG029	Männlich, 70 bis 74 Jahre	49,74 %	193.785
AGG030	Männlich, 75 bis 79 Jahre	53,60 %	224.893
AGG031	Männlich, 80 bis 84 Jahre	47,49 %	126.587
AGG032	Männlich, 85 bis 89 Jahre	42,41 %	56.606
AGG033	Männlich, 90 bis 94 Jahre	46,48 %	14.502
AGG034	Männlich, 95 bis 124 Jahre	57,15 %	2.043
THCC001	HIV/AIDS	171,48 %	11.395
THCC006	Andere Infektionskrankheiten	6,87 %	2.677.228
THCC007	Sekundäre bösartige Neubildungen	121,69 %	38.055
THCC010	Bösartige Neubildungen Dickdarm, Rektum, Anus, Mamma, Geschlechtsorgane, Niere und ableitende Harnwege, Auge, Schilddrüse, maligne Melanome	32,54 %	352.796
THCC013	Andere Neubildungen	9,88 %	697.495
THCC022	Andere schwerwiegende endokrine und Stoffwechselerkrankungen	36,06 %	112.058
THCC024	Andere endokrine, metabolische und Ernährungsstörungen	12,40 %	3.514.973
THCC036	Andere gastrointestinale Erkrankungen	9,71 %	2.360.763
THCC038	Entzündliche Systemerkrankungen des rheumatischen Formenkreises	39,09 %	373.441
THCC040	Cox- oder Gonarthrose	13,09 %	946.540
THCC041	Erkrankungen des Knochens und des Knorpels	13,92 %	978.858
THCC043	Andere Erkrankungen der Muskeln, Sehnen, Gelenke und des Bindegewebes	8,44 %	2.367.081
THCC050	Alterserkrankungen, nicht-psychotische hirnorganische Erkrankungen/Zustände	20,26 %	334.055
THCC055	Bipolare Störungen, depressive Episoden	43,47 %	500.312
THCC057	Persönlichkeitsstörungen	27,07 %	402.253
THCC058	Depressive Episode, Angst, Dysthymia	21,44 %	734.247

Fortsetzung auf der nächsten Seite...

Tabelle 6 – Fortsetzung

	Risikoklassen	Relativgewicht	N(2016) in Kalibrierungsmenge
THCC060	Angststörungen, sexuelle Störungen, Belastungs- und Anpassungsstörungen	10,61 %	1.007.023
THCC071	Polyneuropathie, Myasthenie, Myopathie	25,07 %	364.898
THCC072	Multiple Sklerose	65,37 %	37.176
THCC074	Epilepsie und nichtepileptische Anfälle	33,97 %	127.461
THCC076	Radikulopathien, andere neurologische Erkrankungen und Folgen von Verletzungen des Nervensystems	12,27 %	1.686.387
THCC080	Kardiomyopathien, Herzinsuffizienz, PAH	17,53 %	370.111
THCC086	Erkrankungen der Herzklappen und rheumatische Herzerkrankungen	14,72 %	416.477
THCC090	Hypertensive Herzerkrankung	18,52 %	209.839
THCC091	Essentielle Hypertonie	14,23 %	2.327.899
THCC106	Erkrankungen der Arterien, Venen, Lymphgefäße	5,85 %	1.144.283
THCC110	Asthma bronchiale	19,29 %	764.634
THCC115	Akute Bronchitis, Grippe	5,73 %	1.155.928
THCC123	Katarakt	8,80 %	810.177
THCC124	Andere Erkrankungen des Auges	9,73 %	1.634.942
THCC126	Hörverlust	17,20 %	624.747
THCC127	Andere Erkrankungen HNO-Bereich	8,28 %	4.034.857
THCC134	Inkontinenz	11,22 %	420.918
THCC135	Entzündungen der ableitenden Harnwege	11,13 %	646.862
THCC139	Störungen der Menstruation, klimakterische Störungen, andere Erkrankungen der weiblichen Geschlechtsorgane	13,41 %	2.012.154
THCC140	Erkrankungen des männlichen Genitalapparates	10,80 %	668.573
THCC153	Verbrennungen 1. und 2. Grades, andere Erkrankungen der Haut und Hautanhangsgebilde	7,92 %	2.428.549
THCC162	Extremitätenfrakturen, andere Verletzungen	8,39 %	1.813.332
THCC186	Multiples Myelom/Plasmozytom und andere chronische Leukämien	149,39 %	9.426

Fortsetzung auf der nächsten Seite...

Tabelle 6 – Fortsetzung

	Risikoklassen	Relativgewicht	N(2016) in Kalibrierungsmenge
THCC190	Knöchernen Erkrankungen der Wirbelsäule, Erkrankungen der Bandscheiben, Stenosen des Spinalkanals und der Foramina intervertebralia	18,65 %	1.173.701
THCC194	Migräne und sonstige chronische Kopfschmerzen	11,31 %	591.022
RHCC001	Infektiöse und parasitäre Krankheiten (nach Komprimierung)	9,30 %	43.687
RHCC002	Bösartige Neubildungen (nach Komprimierung)	52,54 %	98.829
RHCC003	Sonstige Neubildungen und Krankheiten der Mamma [Brustdrüse] (nach Komprimierung)	8,47 %	1.039.934
RHCC004	Diabetes mellitus	15,08 %	987.799
RHCC005	Sonstige endokrine, Ernährungs- und Stoffwechselkrankheiten (nach Komprimierung)	12,44 %	27.985
RHCC006	Krankheiten der Leber	2,47 %	698.359
RHCC007	Sonstige Krankheiten des Verdauungssystems (nach Komprimierung)	17,33 %	432.050
RHCC008	Krankheiten des Muskel-Skelett-Systems und des Bindegewebes (nach Komprimierung)	7,74 %	427.269
RHCC009	Krankheiten des Blutes und der blutbildenden Organe sowie bestimmte Störungen mit Beteiligung des Immunsystems	12,34 %	715.721
RHCC010	Störungen des Bewusstseins und neuropsychologische Symptome (nach Komprimierung)	17,14 %	217.338
RHCC011	Psychische und Verhaltensstörungen durch Alkohol-, Drogen und sonstige Substanzen	5,14 %	686.752
RHCC012	Sonstige psychische Krankheiten (nach Komprimierung)	24,83 %	194.588
RHCC013	Entwicklungs- und Lernstörungen, Intelligenzminderung, Ticstörungen und komplexe Fehlbildungs-Syndrome	14,09 %	561.396

Fortsetzung auf der nächsten Seite...

Tabelle 6 – Fortsetzung

	Risikoklassen	Relativgewicht	N(2016) in Kalibrierungs- menge
RHCC014	Sonstige neurologische Krankheiten (nach Komprimierung)	31,76 %	146.893
RHCC015	Herzstillstand, Ateminsuffizienz, Tod und Vorhandensein eines Tracheostomas	12,21 %	84.788
RHCC016	Krankheiten des Herz-Kreislaufsystems (nach Komprimierung)	13,30 %	1.231.750
RHCC017	Zerebrovaskuläre Krankheiten	11,34 %	552.348
RHCC018	Sonstige Blutgefäßkrankheiten (nach Komprimierung)	12,07 %	608.574
RHCC019	Krankheiten der unteren Atemwege (nach Komprimierung)	10,38 %	628.733
RHCC020	Augenkrankheiten (nach Komprimierung)	17,61 %	954.055
RHCC021	HNO- und stomatologische Krankheiten (nach Komprimierung)	10,76 %	319.245
RHCC022	Krankheiten der Nieren und ableitenden Harnwege, inkl. Zustände nach Nieren-Transplantation (nach Komprimierung)	9,61 %	922.290
RHCC023	Krankheiten der Genitalorgane (nach Komprimierung)	16,87 %	387.367
RHCC025	Krankheiten der Haut und der Unterhaut (nach Komprimierung)	11,58 %	489.604
RHCC026	Sonstige Verletzungen, Vergiftungen und Zustände nach sonstigen medizinischen Maßnahmen, sowie motorische und kognitive Funktionseinschränkungen (nach Komprimierung)	7,81 %	1.663.073
RHCC027	Symptome, Krankheitszeichen, Störungen und Kontakttänfälle	6,88 %	448.161
RHCC028	Krankheiten und Zustände des Neugeborenen	9,41 %	14.108
RHCC029	Krankheiten und Zustände nach sonstigen Transplantationen und nach besonderen medizinischen Maßnahmen	31,38 %	35.698

Fortsetzung auf der nächsten Seite...

Tabelle 6 – Fortsetzung

Risikoklassen	Relativgewicht	N(2016) in Kalibrierungsmenge
RHCC030 Untersuchungen, prophylaktische Maßnahmen, anamnestiche Angaben und Zustände nach weiteren medizinischen Maßnahmen	13,41 %	719.312
RHCC031 Sonstige Symptome, Krankheitszeichen, Störungen und Kontakianlässe ohne Vorliegen von Diagnosen außerhalb der ACC031	9,39 %	45.318

3.1.3 Ausgewählte statistische Kennzahlen zum Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsdaten

Im Folgenden werden Kennzahlen für das Modell zur Berechnung der Veränderungsdaten für das Jahr 2021 im Vergleich zum Modell zur Berechnung der Veränderungsdaten für das Jahr 2020 dargestellt.

Tabelle 7 weist die Anzahlen der Risikoklassen nach Komprimierung und Kalibrierung für die beiden Klassifikationsmodelle zur Berechnung der Veränderungsdaten für das Jahr 2020 und für das Jahr 2021 aus.

Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsdaten für	Anzahl der Versicherten in der Kalibrierungsmenge	Zahl der Risikoklassen				Gesamt (m)	Anteil der Alters- und Geschlechtsklassen an der ökon. Relevanz ⁷
		Alters- und Geschlechtsklassen	THCC	RHCC			
das Jahr 2021	10.243.425	34	41	30	105	30,59 %	
das Jahr 2020	10.215.691	34	43	30	107	30,70 %	

Tabelle 7: Kennzahlen der Regressionsmodelle zur Berechnung der diagnosebezogenen Veränderungsdaten für die Jahre 2020 und 2021

⁷Die ökonomische Relevanz einer Risikoklasse ist das Produkt aus ihrer Häufigkeit und ihrem Relativgewicht. Der Anteil der Alters- und Geschlechtsgruppen an der ökonomischen Relevanz ergibt sich aus der Summe der ökonomischen Relevanzen der Alters- und Geschlechtsgruppen geteilt durch die Summe der ökonomischen Relevanzen aller Risikoklassen (inkl. der Alters- und Geschlechtsgruppen).

In Tabelle 8 werden der gewichtete MAPE⁸ und das adjustierte R^2 zur Beurteilung der Regressionsmodelle angegeben. Diese beiden Kennzahlen werden, wie schon im Kontext der demografischen Veränderungsrate in Abschnitt 3.1.2 in Teil B beschrieben, auf der Kalibrierungsmenge berechnet.

Regressionsmodell zur Berechnung der diagnosebezogenen Veränderungsdaten	Gewichteter MAPE	Adjustiertes R^2
für das Jahr 2021	52,24 %	37,01 %
für das Jahr 2020	52,35 %	37,44 %

Tabelle 8: Gewichteter MAPE und adjustiertes R^2 für die Regressionsmodelle zur Berechnung der diagnosebezogenen Veränderungsdaten für die Jahre 2020 und 2021

3.2 Empirische Ergebnisse zu den diagnosebezogenen Veränderungsdaten

3.2.1 Entwicklung der Häufigkeiten von Diagnosen, Risikokategorien und Risikoklassen von 2017 auf 2018

Die Risikoklassen des Klassifikationsmodells zur Berechnung der Veränderungsdaten für das Jahr 2021 sind Ergebnis der Anwendung des Komprimierungsalgorithmus auf hierarchisierte Risikokategorien (HCC), die sich wiederum aus den kodierten Diagnosen ergeben (vgl. Abschnitt 1.1).

Da die Entwicklung der Häufigkeiten von Risikoklassen von 2017 auf 2018 auf der Entwicklung der Häufigkeiten der Risikokategorien und letztlich der Diagnosen beruht, werden die mittlere Anzahl distinkter Diagnosen⁹, Risikokategorien und Risikoklassen je KV-Bezirk und je Versicherten für die Jahre 2017 und 2018 berichtet. Für die KV-Bezirke Baden-Württemberg und Bayern werden separate Auswertungen für SV-Teilnehmer (basierend auf Diagnosen aus dem nicht hausärztlichen Versorgungsbereich) und für Nicht-SV-Teilnehmer (basierend auf allen Diagnosen) durchgeführt¹⁰.

Als klassifizierbare Diagnosen werden diejenigen Diagnosen bezeichnet, die (im Zusammenspiel mit dem Zusatzkennzeichen für die Diagnosesicherheit (ZK-DS), Alter und/oder

⁸Die Angabe des gewichteten MAPE in Prozent ergibt sich durch die Verwendung relativer Leistungsbedarfe.

⁹Als Diagnose wird hier die Kombination eines ICD-Kodes mit einem Zusatzkennzeichen für die Diagnosesicherheit verstanden.

¹⁰Sowohl für die Bestimmung diagnosebezogener Veränderungsdaten als auch in diesem Abschnitt zur Entwicklung der Häufigkeiten von Diagnosen usw. sind SV-Teilnehmer „längsschnittlich“ definiert, d.h. dass die Teilnahme in mindestens einem Quartal der Jahre 2017 bzw. 2018 zur Zuordnung als Teilnehmer führt. Da die Teilnahme an Selektivverträgen freiwillig ist, gibt es Fluktuationen zwischen Teilnehmern und Nicht-SV-Teilnehmern. In früheren oder späteren Berichten sind die Teilkollektive der sogenannten „längsschnittlichen“ SV-Teilnehmer nicht unbedingt gleich, da sich die „Längsschnittlichkeit“ auf unterschiedliche Zweijahreszeiträume bezieht. Bei größeren Unterschieden zwischen den Teilkollektiven kann das dazu führen, dass die hier für das Jahr 2017 oder 2018 berichteten Werte nicht übereinstimmen mit Werten, die für das gleiche Jahr in einem anderen Bericht angegeben werden.

Geschlecht) zur Zuordnung einer CC führen. Dabei ist zu beachten, dass Diagnosen mit dem Zusatzkennzeichen „Z“ überhaupt nur in Zusammenhang mit den Diagnosen der CC200 („Zustand nach Neubildung“) berücksichtigt bzw. klassifiziert werden.

Die mittleren Anzahlen an distinkten klassifizierbaren Diagnosen werden zusätzlich differenziert nach den Zusatzkennzeichen für die Diagnosesicherheit (ZK-DS) berichtet (Tabelle 9).

Tabelle 9: Durchschnittliche Anzahl der distinkten und klassifizierbaren Diagnosen je KV-Bezirk und je Versicherten für die Jahre 2017 und 2018. Die Diagnosenzählung basiert auf den Versicherten der Anwendungsmenge des jeweiligen Jahres, gewichtet mit der Versichertenzeit in Quartalen und den demografischen Hochrechnungsfaktoren.

KV-Bezirk	ZK-DS	ØAnzahl Diagnosen		Veränderung in Prozent
		2017	2018	
Schleswig-Holstein	G	12,94	13,05	0,88 %
	Z	0,07	0,07	2,34 %
	zusammen	13,00	13,12	0,89 %
Hamburg	G	12,15	12,16	0,11 %
	Z	0,06	0,06	1,90 %
	zusammen	12,21	12,22	0,11 %
Bremen	G	13,25	13,23	-0,20 %
	Z	0,06	0,06	-0,16 %
	zusammen	13,31	13,29	-0,20 %
Niedersachsen	G	14,18	14,30	0,84 %
	Z	0,06	0,07	0,63 %
	zusammen	14,25	14,37	0,84 %
Westfalen-Lippe	G	13,78	13,98	1,47 %
	Z	0,06	0,06	1,67 %
	zusammen	13,84	14,04	1,48 %
Nordrhein	G	13,57	13,73	1,18 %
	Z	0,06	0,06	2,41 %
	zusammen	13,63	13,79	1,19 %
Hessen	G	13,34	13,46	0,86 %
	Z	0,06	0,06	1,75 %
	zusammen	13,40	13,51	0,87 %
Rheinland-Pfalz	G	14,39	14,63	1,67 %
	Z	0,06	0,06	1,86 %

Fortsetzung auf der nächsten Seite...

Tabelle 9 – Fortsetzung

KV-Bezirk	ZK-DS	ØAnzahl Diagnosen		Veränderung in Prozent
		2017	2018	
	zusammen	14,45	14,69	1,67 %
Baden-Württemberg (Nicht-SV-TN)	G	11,90	12,01	0,89 %
	Z	0,05	0,05	-0,64 %
	zusammen	11,96	12,06	0,88 %
Baden-Württemberg (SV-TN)	G	6,48	6,56	1,23 %
	Z	0,04	0,04	0,20 %
	zusammen	6,52	6,60	1,22 %
Bayern (Nicht-SV-TN)	G	13,18	13,34	1,24 %
	Z	0,06	0,06	1,40 %
	zusammen	13,24	13,40	1,24 %
Bayern (SV-TN)	G	8,37	8,53	1,95 %
	Z	0,07	0,07	2,02 %
	zusammen	8,44	8,60	1,95 %
Berlin	G	14,15	14,15	-0,01 %
	Z	0,06	0,06	0,62 %
	zusammen	14,21	14,21	-0,01 %
Saarland	G	14,46	14,72	1,79 %
	Z	0,06	0,06	1,48 %
	zusammen	14,52	14,78	1,79 %
Mecklenburg-Vorpommern	G	15,30	15,55	1,62 %
	Z	0,08	0,08	2,52 %
	zusammen	15,38	15,63	1,62 %
Brandenburg	G	15,00	15,13	0,84 %
	Z	0,06	0,06	1,74 %
	zusammen	15,06	15,19	0,85 %
Sachsen-Anhalt	G	15,82	16,07	1,61 %
	Z	0,06	0,06	3,13 %
	zusammen	15,88	16,14	1,62 %
Thüringen	G	16,01	16,27	1,68 %
	Z	0,07	0,07	3,16 %
	zusammen	16,07	16,34	1,69 %

Fortsetzung auf der nächsten Seite...

Tabelle 9 – Fortsetzung

KV-Bezirk	ZK-DS	ØAnzahl Diagnosen		Veränderung in Prozent
		2017	2018	
Sachsen	G	14,30	14,49	1,35 %
	Z	0,07	0,07	2,20 %
	zusammen	14,37	14,56	1,35 %

Die Veränderung (in Prozent) der mittleren Anzahl der distinkten und klassifizierbaren Diagnosen pro Versicherten ist in Abbildung 2 dargestellt.

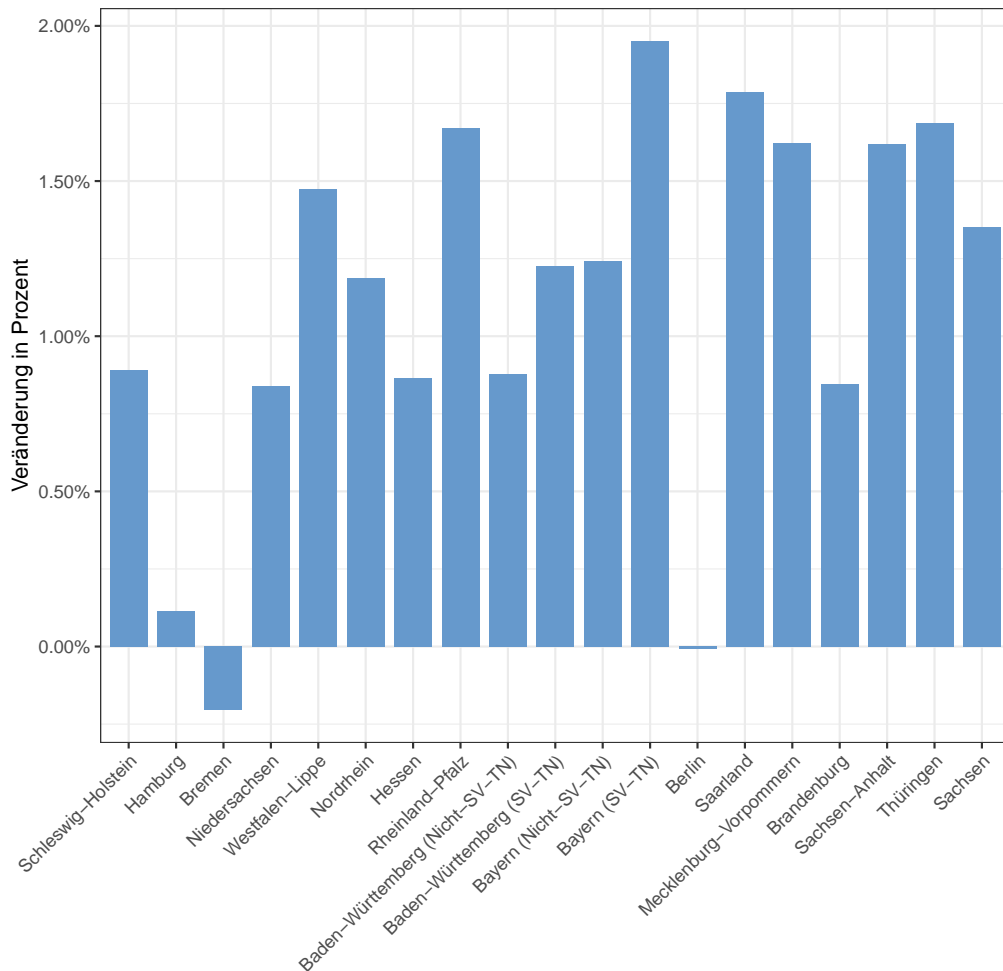


Abbildung 2: Veränderung (in Prozent) der mittleren Anzahl der distinkten und klassifizierbaren Diagnosen pro Versicherten von 2017 auf 2018 je KV-Bezirk

In der Abbildung 3 werden die Anteile der mittleren Anzahlen an distinkten klassifizierbaren Diagnosen nach den Zusatzkennzeichen für die Diagnosesicherheit „G“ und „Z“ je

KV-Bezirk für die Jahre 2017 und 2018 abgebildet.

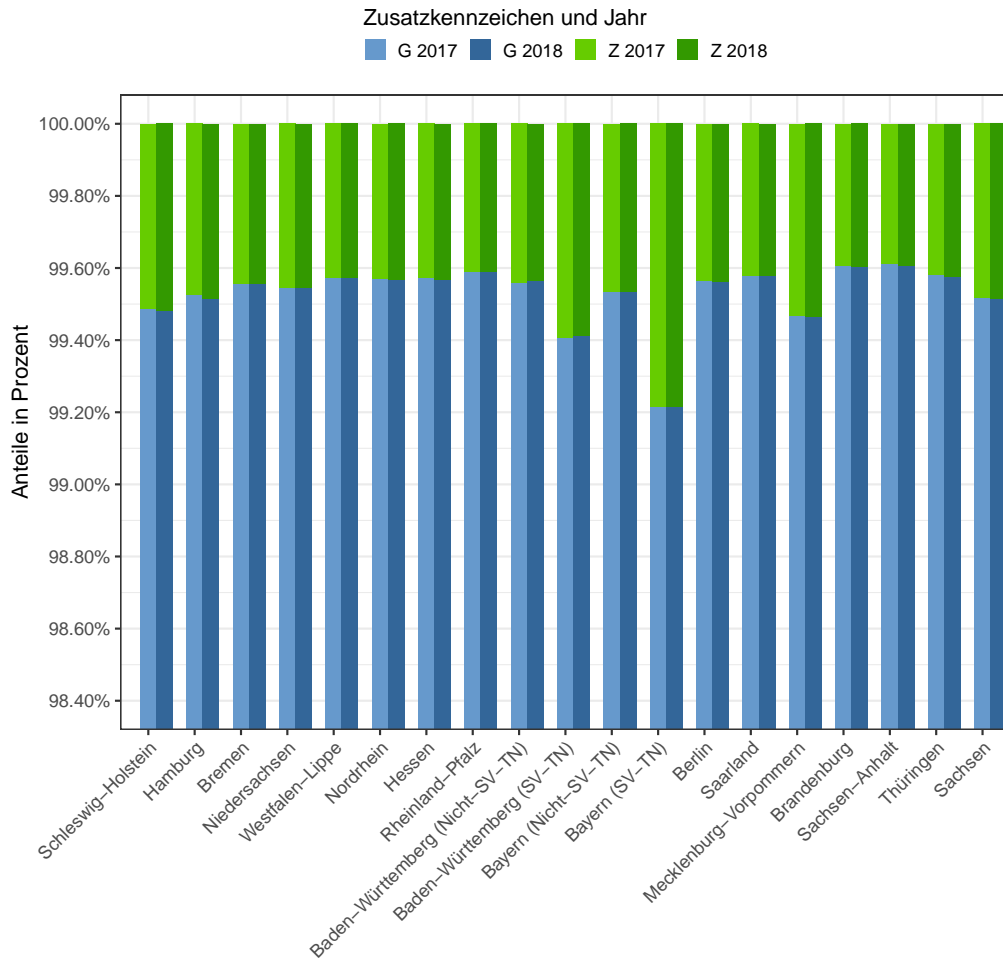


Abbildung 3: Verteilung der Zusatzkennzeichen für die Diagnosesicherheit „G“ und „Z“ unter den mittleren Anzahlen distinkter und klassifizierbarer Diagnosen pro Versicherten je KV-Bezirk für die Jahre 2017 und 2018

In Tabelle 10 ist die durchschnittliche Anzahl der Risikokategorien je KV-Bezirk und je Versicherten für die Jahre 2017 und 2018 dargestellt. Die sogenannte „NonCW-HCC“ (HCC201) sowie die Alters- und Geschlechtsgruppen werden dabei nicht mitgezählt.

KV-Bezirk	ØAnzahl HCC		Veränderung in Prozent
	2017	2018	
Schleswig-Holstein	5,70	5,75	0,78 %
Hamburg	5,25	5,24	-0,07 %
Bremen	5,68	5,66	-0,40 %
Niedersachsen	6,04	6,08	0,56 %
Westfalen-Lippe	5,89	5,96	1,19 %
Nordrhein	5,83	5,88	0,93 %
Hessen	5,75	5,79	0,63 %
Rheinland-Pfalz	6,16	6,24	1,24 %
Baden-Württemberg (N-TN)	5,21	5,23	0,36 %
Baden-Württemberg (SV-TN)	2,84	2,86	0,69 %
Bayern (N-TN)	5,56	5,60	0,74 %
Bayern (SV-TN)	3,61	3,68	1,82 %
Berlin	5,77	5,76	-0,24 %
Saarland	6,17	6,24	1,13 %
Mecklenburg-Vorpommern	6,59	6,67	1,21 %
Brandenburg	6,42	6,47	0,72 %
Sachsen-Anhalt	6,66	6,73	1,16 %
Thüringen	6,73	6,82	1,27 %
Sachsen	6,22	6,27	0,85 %

Tabelle 10: Durchschnittliche Anzahl der Risikokategorien (HCC) je KV-Bezirk und je Versicherten für die Jahre 2017 und 2018

Tabelle 11 führt die durchschnittliche Anzahl der Risikoklassen je KV-Bezirk und je Versicherten für die Jahre 2017 und 2018 auf. Die „RHCC024 Krankheit und Zustände bei Schwangerschaft, Geburt und Wochenbett“ (vgl. Abschnitt 1.2.2) sowie die Alters- und Geschlechtsgruppen werden dabei nicht mitgezählt.

KV-Bezirk	ØAnzahl THCC/RHCC		Veränderung in Prozent
	2017	2018	
Schleswig-Holstein	5,54	5,59	0,76 %
Hamburg	5,10	5,10	-0,03 %
Bremen	5,52	5,50	-0,41 %
Niedersachsen	5,87	5,90	0,54 %
Westfalen-Lippe	5,72	5,79	1,20 %
Nordrhein	5,67	5,72	0,91 %
Hessen	5,59	5,63	0,62 %
Rheinland-Pfalz	5,99	6,06	1,20 %
Baden-Württemberg (N-TN)	5,08	5,10	0,35 %
Baden-Württemberg (SV-TN)	2,76	2,78	0,67 %
Bayern (N-TN)	5,41	5,45	0,71 %
Bayern (SV-TN)	3,50	3,57	1,82 %
Berlin	5,58	5,57	-0,22 %
Saarland	5,99	6,06	1,14 %
Mecklenburg-Vorpommern	6,38	6,46	1,17 %
Brandenburg	6,21	6,26	0,76 %
Sachsen-Anhalt	6,42	6,49	1,14 %
Thüringen	6,50	6,58	1,28 %
Sachsen	6,02	6,07	0,85 %

Tabelle 11: Durchschnittliche Anzahl der Risikoklassen (THCC/RHCC) je KV-Bezirk und je Versicherten für die Jahre 2017 und 2018

3.2.2 Einzelbeiträge der Risikoklassen zu den regionalen diagnosebezogenen Veränderungsdaten

Die Einzelbeiträge sämtlicher Risikoklassen zur regionalen diagnosebezogenen Veränderungsrate sind, getrennt nach KV-Bezirken, in einer Excelmappe enthalten und unter institut-ba.de veröffentlicht. Die Berechnung der Einzelbeiträge erfolgt anhand der For-

meln (13) und (14).

$$\begin{aligned}
 \bar{y}_{VB}^{t,KV} &= \frac{\sum_{i=1}^{n_{VB}^{t,KV}} \hat{y}_{i,VB}^{t,KV} \cdot w_i^{t,KV}}{\sum_{i=1}^{n_{VB}^{t,KV}} w_i^{t,KV}} \\
 &= \frac{\sum_{i=1}^{n_{VB}^{t,KV}} (\hat{\beta}_1 \cdot x_{i,1,VB}^{t,KV} + \hat{\beta}_2 \cdot x_{i,2,VB}^{t,KV} + \dots + \hat{\beta}_r \cdot x_{i,r,VB}^{t,KV}) \cdot w_i^{t,KV}}{\sum_{i=1}^{n_{VB}^{t,KV}} w_i^{t,KV}} \\
 &= \underbrace{\frac{\sum_{i=1}^{n_{VB}^{t,KV}} w_i^{t,KV} \cdot \hat{\beta}_1 \cdot x_{i,1,VB}^{t,KV}}{\sum_{i=1}^{n_{VB}^{t,KV}} w_i^{t,KV}}}_{\text{Partialindex}_{1,VB}^{t,KV}} + \underbrace{\frac{\sum_{i=1}^{n_{VB}^{t,KV}} w_i^{t,KV} \cdot \hat{\beta}_2 \cdot x_{i,2,VB}^{t,KV}}{\sum_{i=1}^{n_{VB}^{t,KV}} w_i^{t,KV}}}_{\text{Partialindex}_{2,VB}^{t,KV}} + \dots \\
 &\quad + \underbrace{\frac{\sum_{i=1}^{n_{VB}^{t,KV}} w_i^{t,KV} \cdot \hat{\beta}_r \cdot x_{i,r,VB}^{t,KV}}{\sum_{i=1}^{n_{VB}^{t,KV}} w_i^{t,KV}}}_{\text{Partialindex}_{r,VB}^{t,KV}}
 \end{aligned} \tag{13}$$

mit

VB Versorgungsbereich ($VB = \text{„gesamt“}$ für alle Diagnosen von Nicht-SV-Teilnehmern, $VB = \text{„Nicht-HA“}$ für alle nicht hausärztlichen Diagnosen von SV-Teilnehmern in den KV-Bezirken Bayern und Baden-Württemberg)

KV KV-Bezirk ($KV = 1, \dots, 17$)

t Jahr ($t = 2017, 2018$)

$n_{VB}^{t,KV}$ Anzahl der Versicherten in der Anwendungsmenge für das Jahr t , die dem Versorgungsbereich VB und dem KV-Bezirk KV zugeordnet wurden.

$\bar{y}_{VB}^{t,KV}$ Morbiditätsindex für den Versorgungsbereich VB im KV-Bezirk KV für das Jahr t

$\hat{y}_{i,VB}^{t,KV}$ relativer Risikowert des Versicherten i der entsprechenden KV im Jahr t für den Versorgungsbereich VB

$\hat{\beta}_j$ bundeseinheitliches Relativgewicht der Risikoklasse j ($j = 1, \dots, r$)

$x_{i,j,VB}^{t,KV}$ Ausprägung der Risikoklasse j des Versicherten i der entsprechenden KV im Jahr t für den Versorgungsbereich VB

$w_i^{t,KV}$ Gewichtungsvariable für den Versicherten i der entsprechenden KV im Jahr t ($w_i^{t,KV} = AVQ_i^{t,KV} \cdot DHF_i^{t,KV} \cdot K_i^{t,KV}$)

$AVQ_i^{t,KV}$ Anzahl der Versichertenquartale des Versicherten i der entsprechenden KV im Jahr t

$DHF_i^{t,KV}$ Demografischer Hochrechnungsfaktor des Versicherten i der entsprechenden KV für das Jahr t bezogen auf die Anzahl aller Versicherten im KV-Bezirk

$K_i^{t,KV}$ Korrekturfaktor für längsschnittliche SV-Teilnehmer des Versicherten i der entsprechenden KV für das Jahr t (der Korrekturfaktor nimmt für $VB = \text{„gesamt“}$ [für alle Nicht-SV-Teilnehmer] den Wert 1 an)

$$\begin{aligned}
 VR_{VB,2017 \rightarrow 2018}^{KV} &= \frac{\bar{y}_{VB}^{2018,KV}}{\bar{y}_{VB}^{2017,KV}} - 1 = \frac{\bar{y}_{VB}^{2018,KV} - \bar{y}_{VB}^{2017,KV}}{\bar{y}_{VB}^{2017,KV}} \\
 &= \sum_{j=1}^r \left(\frac{Partialindex_{j,VB}^{2018,KV} - Partialindex_{j,VB}^{2017,KV}}{\bar{y}_{VB}^{2017,KV}} \right) \quad (14)
 \end{aligned}$$

mit

$VR_{VB,2017 \rightarrow 2018}^{KV}$ diagnosebezogene Veränderungsrate für den Versorgungsbereich VB im KV-Bezirk KV

Die Zerlegung der jeweiligen Veränderungsrate erfolgt also additiv, der Einzelbeitrag der Risikoklasse j entspricht der Entwicklung des j -ten Partialindex in Relation zum Morbiditätsindex für den Versorgungsbereich VB im KV-Bezirk KV für das Jahr 2017 ($\bar{y}_{VB}^{2017,KV}$).

Da in den KV-Bezirken Bayern und Baden-Württemberg separate Komponenten für SV-Teilnehmer und Nicht-SV-Teilnehmer berechnet werden, die dann zur diagnosebezogenen Veränderungsrate für diese KV-Bezirke zusammengeführt werden, werden in diesen beiden Regionen auch die Einzelbeiträge der Risikoklassen für SV-Teilnehmer und Nicht-SV-Teilnehmer separat ermittelt.

Teil D

Weiterentwicklungen

1	Weiterentwicklung der Datengrundlage	70
2	Weiterentwicklungen zur Berechnung der demografischen Veränderungsraten	70
3	Weiterentwicklungen zur Berechnung der diagnosebezogenen Veränderungsraten	70
3.1	Übersicht über die Weiterentwicklungen zur Berechnung der diagnosebezogenen Veränderungsraten	71
3.2	Weiterentwicklung des Klassifikationssystems.....	75
3.2.1	Einleitung	75
3.2.2	Grundlagen zum Klassifikationssystem gemäß § 87a Abs. 5 SGB V (KS87a).....	76
3.2.3	Grundsätze der Bearbeitungen	79
3.2.4	Umsetzung der ICD-Änderungen bei der Schlüsselnummer U07.0/U07.0! im Klassifikationssystem	88
3.2.5	Anpassung des Klassifikationssystems an die Version 2020 der ICD-10-GM	89
3.3	Weiterentwicklung der Berechnung der Veränderungsraten.....	90

1 Weiterentwicklung der Datengrundlage

Mit Beschluss des Bewertungsausschusses in seiner 437. Sitzung (schriftliche Beschlussfassung) haben sich die zulässigen Ausprägungen für das Feld 05 (Geschlecht) der Satzart 201 (Versicherten-Stammdaten) der Geburtstagsstichprobe geändert. Ab dem Berichtsjahr 2018 ist neben den bisher möglichen Ausprägungen „weiblich“ und „männlich“ auch die Ausprägung „unbestimmt“ möglich. Die Ausprägung „unbestimmt“ trat in der verwendeten Geburtstagsstichprobe im Berichtsjahr 2018 bei sechs Versicherten auf.

2 Weiterentwicklungen zur Berechnung der demografischen Veränderungsraten

Im Beschluss des Bewertungsausschusses in seiner 503. Sitzung wurde gegenüber dem Vorjahresbeschluss der Umgang mit der neuen Geschlechtsausprägung „unbestimmt“ geregelt: Versicherten mit der Geschlechtsausprägung „unbestimmt“ in Feld 05 der Satzart 201 wird die Geschlechtsausprägung „weiblich“ zugewiesen. Bei der Berechnung der demografischen Veränderungsdaten wird die Geschlechtsausprägung bei der Zuweisung von Versicherten zu einer Alters- und Geschlechtsgruppe in der Kalibrierungsmenge zur Berechnung der Relativgewichte verwendet.

3 Weiterentwicklungen zur Berechnung der diagnosebezogenen Veränderungsdaten

Der Erweiterte Bewertungsausschuss hat in seiner 15. Sitzung am 2. September 2009 die Weiterentwicklung des Klassifikationssystems beschlossen. Der Bewertungsausschuss hat anschließend in seiner 269. Sitzung am 25. Januar 2012 Vorgaben zur Weiterentwicklung beschlossen. Als Ausgangspunkt für die Weiterentwicklung wurde dort das vom Erweiterten Bewertungsausschuss in seiner 15. Sitzung am 2. September 2009 in Beschlussteil B, Nr. 2.3 beschlossene Modell unter Verwendung der Version p06 des Klassifikationssystems festgelegt. Dieses Klassifikationsmodell¹ wurde für die Bestimmung der diagnosebezogenen Veränderungsdaten für das Jahr 2010 verwendet. Es folgte eine Weiterentwicklung zum Klassifikationsmodell für die Ermittlung der diagnosebezogenen und demografischen Veränderungsdaten des Jahres 2013². Darauf aufbauend wurden in den Folgejahren die Klassifikationsmodelle für die Ermittlung der Veränderungsdaten für die Jahre 2014, 2015,

¹Das Klassifikationsmodell wird ausführlich beschrieben im „Bericht des Instituts des Bewertungsausschusses zu Vollständigkeits- und Repräsentativitätsprüfungen gemäß Beschluss des Bewertungsausschusses in seiner 184. Sitzung vom 20. Mai 2009 und den Ermittlungen der Strukturveränderungen der Morbidität zwischen den Jahren 2006 und 2008“.

²Vgl. „Bericht des Instituts des Bewertungsausschusses zur Weiterentwicklung des Klassifikationssystems sowie zur Ermittlung der Veränderungsdaten für das Jahr 2013 gemäß § 87a Abs. 5 SGB V“ sowie analoge Berichte für die Veränderungsdaten der Jahre 2015-2020.

2016, 2017, 2018, 2019 und 2020 entwickelt und angewendet. In den folgenden Abschnitten wird die Weiterentwicklung des Klassifikationsmodells für die Berechnung der Veränderungsdaten für das Jahr 2020 hin zu dem Klassifikationsmodell für die Berechnung der Veränderungsdaten für das Jahr 2021 dargestellt.

3.1 Übersicht über die Weiterentwicklungen zur Berechnung der diagnosebezogenen Veränderungsdaten

Tabelle 1 gibt einen Überblick über die Weiterentwicklungen vom Klassifikationsmodell für die Berechnung der Veränderungsdaten für das Jahr 2020 im Jahr 2019 (KM87a_2019) hin zum Klassifikationsmodell für die Berechnung der Veränderungsdaten für das Jahr 2021 im Jahr 2020 (KM87a_2020). Der Vollständigkeit halber sind in dieser Tabelle auch Merkmale aufgelistet, deren Ausprägung sich zwischen den beiden Modellen nicht unterscheidet.

Die in der Tabelle aufgelisteten Merkmale der Modelle sind untergliedert in die Bereiche:

- Datengrundlage,
- Klassifikation,
- Berechnung der Relativgewichte und
- Berechnung der Veränderungsdaten.

Tabelle 1: Vergleich des Klassifikationsmodells für die Berechnung der Veränderungsdaten für das Jahr 2021 mit dem Vorjahresmodell

Merkmal	Klassifikationsmodell für die Berechnung der Veränderungsdaten	
	für das Jahr 2020 (KM87a_2019)	für das Jahr 2021 (KM87a_2020)
Datengrundlage		
Berichtsjahre	2015 - 2017	2016 - 2018
GSP-Daten	Stand vom 26.04.2019	Stand vom 24.03. 2020
SV-Daten	Stand vom 26.04.2019	Stand vom 24.03. 2020
Qualitätssicherung	Ausschluss, wenn Markierung erfolgt ist	
Fälle ohne Leistungsbedarf	Ausschluss	
Identifikation Selektivvertragsteilnehmer	mindestens ein Datensatz mit einer Kennzeichnung zur Selektivvertragsteilnahme	
Identifikation Wohnausländer	Auf Grundlage des letzten Versichertenquartals je Jahr werden für den Versicherten diejenigen Datensätze der Satzart 201 ausgewählt, bei denen die über das Institutionskennzeichen zu ermittelnde Krankenkasse ihren Sitz im ermittelten KV-Bezirk hat. Enthält einer dieser Datensätze in Feld 08 den Wert "AUSLA", so wird der Versicherte als Wohnausländer identifiziert.	
Diagnosen von nicht zur Diagnose verpflichteten Ärzten	kein Ausschluss	
Version der KM6-Statistik	2016 und 2017	2017 und 2018
Geschlechtsausprägung „unbestimmt“		Umwandlung in „weiblich“
Klassifikation		
Version des Klassifikationssystems	p13a	p14a
Berechnung der Relativgewichte		
verwendete Geburtskalendertage	5, 12, 18, 19, 25	5, 12, 19, 25, 26
Jahresbezüge des prospektiven Modells	2015 und 2017	2016 und 2018
MGV-Leistungsbedarf	in der Einheit Punkte	

Fortsetzung auf der nächsten Seite...

Tabelle 1 – Fortsetzung

Merkmal	Klassifikationsmodell für die Berechnung der Veränderungsraten	
	für das Jahr 2020	für das Jahr 2021
Annualisierung des Leistungsbedarfs	ja, mit Anzahl an Versichertenquartalen im jeweiligen Jahr	
Abhängige Variable im Regressionsmodell	relativer MGV-Leistungsbedarf	
Regionale Abgrenzung	bundeseinheitlich	
versicherenzzeitvollständig	nur Berücksichtigung von Versicherten, die mindestens 45 Tage in jedem Quartal eines Jahres versichert waren (Ausnahme: Quartal des Todes und die Folgequartale im Jahr des Todes)	
SV-Teilnehmer in der Stichprobe	Ausschluss	
Personen mit Kostenübernahme nach § 264 SGB V (sog. Be treute)	Ausschluss von Personen mit Wert 1 in Feld 14 der Satzart 201 in mindestens einem der Berichtsjahre	
Wohnausländer	Ausschluss von Versicherten, die in mindestens einem der Berichtsjahre als Wohnausländer identifiziert wurden und die im entsprechenden Jahr einem der KV-Bezirke Baden-Württemberg oder Nordrhein zugeordnet sind	
Kompression	Auswahl von HCCs auf Basis der ökonomischen Relevanz, der Größe des Relativgewichts und des maximalen Beitrags zur erklärten Varianz als THCCs. Nicht ausgewählte HCCs gehen in RHCCs ein.	
Kalibrierung	Sicherstellung von ausschließlich positiven und signifikanten Relativgewichten durch mögliches Nullsetzen von HCCs in RHCCs, Zusammenlegen von AGGs, Eingliedern von THCCs in RHCCs	
Berechnung der Veränderungsraten		
Anwendungsjahre	2016 und 2017	2017 und 2018
verwendete Geburtskalendertage	2016: 5, 11, 12, 18, 19, 25, 26 2017: 5, 6, 12, 18, 19, 25, 26	2017: 5, 6, 12, 18, 19, 25, 26 2018: 5, 6, 12, 13, 19, 25, 26
Regionale Abgrenzung	KV-spezifisch	

Fortsetzung auf der nächsten Seite...

Tabelle 1 – Fortsetzung

Merkmal	Klassifikationsmodell für die Berechnung der Veränderungsdaten	
	für das Jahr 2020	für das Jahr 2021
versicherungszeitlang	keine Berücksichtigung von Versicherten, die weniger als 120 Tage im Jahr versichert waren, bei der Berechnung von relativen Risikowerten. Ausnahme: Verstorbene und Versicherte, für die in einem Quartal des Jahres ohne Versicherungszeit Abrechnungsfälle vorliegen.	
SV-Teilnehmer in der Stichprobe	Keine Berücksichtigung von Versicherten, die in einem der Anwendungsjahre als SV-Teilnehmer identifiziert wurden, bei der Berechnung von relativen Risikowerten. Ausnahme: Für die KV-Bezirke Bayern und Baden-Württemberg werden auch für SV-Teilnehmer ausschließlich nach § 73b SGB V relative Risikowerte berechnet und damit getrennte Raten für SV-Teilnehmer und Nicht-SV-Teilnehmer berechnet.	
Personen mit Kostenübernahme nach § 264 SGB V (sog. Be treute)	Ausschluss von Personen mit Wert 1 in Feld 14 der Satzart 201 in mindestens einem der Anwendungsjahre	
Wohnausländer	Ausschluss von Versicherten, die in mindestens einem der Anwendungsjahre als Wohnausländer identifiziert wurden und die im entsprechenden Jahr einem der KV-Bezirke Baden-Württemberg oder Nordrhein zugeordnet sind	
Diagnosen (Nicht-SV-Teilnehmer)	Verwendung aller Diagnosen	
Diagnosen (SV-Teilnehmer)	Verwendung von ausschließlich nicht hausärztlichen Diagnosen	
SV-Diagnosen	keine Verwendung	
Demografische Hochrechnungsfaktoren (DHF)	Berechnung mit Hochrechnungsfaktoren je KV, Alters- und Geschlechtsgruppe, Wohnausländerstatus und Jahr. Beachtung der Verteilung der Alters- und Geschlechtsgruppen der KM6-Daten. Wohninländer: Hochrechnung auf Anzahl Versicherte gemäß Anzver-Daten. Wohnausländer: Hochrechnung auf Anzahl Versicherte gemäß KM6-Daten.	
Korrekturfaktor der DHF für „nur-§ 73b-Versicherte“	Berechnung mit Korrekturfaktoren in den KV-Bezirken Baden-Württemberg und Bayern je Alters- und Geschlechtsgruppe und Jahr	

Fortsetzung auf der nächsten Seite...

Tabelle 1 – Fortsetzung

Merkmal	Klassifikationsmodell für die Berechnung der Veränderungsraten	
	für das Jahr 2020	für das Jahr 2021
Berechnung der Veränderungsrate ³	$VR_{2016 \rightarrow 2017}^{KV} = \frac{\bar{y}_{gesamt}^{2017, KV}}{\bar{y}_{gesamt}^{2016, KV}} - 1$ <p>Ausnahme: Berechnung der Veränderungsraten für die KV Bezirke Bayern und Baden Württemberg. In diesen beiden KV-Bezirken wird ein gewichtetes Mittel aus getrennten Raten für SV-Teilnehmer (mit Teilnahme an ausschließlich Verträgen nach § 73b SGB V) und Nicht-SV-Teilnehmern berechnet:</p> $VR_{2016 \rightarrow 2017}^{KV} = G_{KV} \cdot \left(\frac{\bar{y}_{Nicht-HA}^{2017, KV}}{\bar{y}_{Nicht-HA}^{2016, KV}} - 1 \right) + (1 - G_{KV}) \cdot \left(\frac{\bar{y}_{gesamt}^{2017, KV}}{\bar{y}_{gesamt}^{2016, KV}} - 1 \right)$	$VR_{2017 \rightarrow 2018}^{KV} = \frac{\bar{y}_{gesamt}^{2018, KV}}{\bar{y}_{gesamt}^{2017, KV}} - 1$ <p>Ausnahme: Berechnung der Veränderungsraten für die KV Bezirke Bayern und Baden Württemberg. In diesen beiden KV-Bezirken wird ein gewichtetes Mittel aus getrennten Raten für SV-Teilnehmer (mit Teilnahme an ausschließlich Verträgen nach § 73b SGB V) und Nicht-SV-Teilnehmern berechnet:</p> $VR_{2017 \rightarrow 2018}^{KV} = G_{KV} \cdot \left(\frac{\bar{y}_{Nicht-HA}^{2018, KV}}{\bar{y}_{Nicht-HA}^{2017, KV}} - 1 \right) + (1 - G_{KV}) \cdot \left(\frac{\bar{y}_{gesamt}^{2018, KV}}{\bar{y}_{gesamt}^{2017, KV}} - 1 \right)$

3.2 Weiterentwicklung des Klassifikationssystems

3.2.1 Einleitung

Im Jahr 2019 erfolgte die Ermittlung der diagnosebezogenen Veränderungsraten gemäß § 87a SGB V für das Jahr 2020 mit dem Klassifikationsmodell „KM87a_2019“. Das zu diesem Klassifikationsmodell gehörende Klassifikationssystem (KS87a) war die Version p13a.

Gemäß Beschluss des Bewertungsausschusses in seiner 269. Sitzung am 25. Januar 2012 zur Weiterentwicklung des Klassifikationssystems gemäß § 87a Abs. 5 SGB V (siehe Abschnitt 3.2.3.1) erfolgte die jährliche Weiterentwicklung von der Version p13a zur Version p14a.

Die Version p14a wurde am 3. Juni 2020 vom Arbeitsausschuss des Bewertungsausschusses freigegeben. Am 19. Juni 2020 hat der Bewertungsausschuss in seiner 503. Sitzung (schriftliche Beschlussfassung) das zur Ermittlung der diagnosebezogenen bzw. demografischen Veränderungsraten zu verwendende Klassifikationsmodell für das Jahr 2021 beschlossen („KM87a_2020“). Das zu diesem Klassifikationsmodell gehörende Klassifikationssystem ist die Version p14a.

Die Weiterentwicklungen des Klassifikationssystems von der Version p13a zur Version p14a umfassen insgesamt:

³Die Bezeichnung der verwendeten Variablen ist analog zu den Bezeichnungen in den Formeln in Teil C.

- die Umsetzung der ICD-Änderungen bei der Schlüsselnummer U07.0/U07.0! im Klassifikationssystem (siehe Abschnitt 3.2.4) und,
- die Anpassung des Klassifikationssystems an die Version 2020 der ICD-10-GM (siehe Abschnitt 3.2.5).

In der folgenden Tabelle 2 sind die Parameter der KS87a-Versionen p13a und p14a gegenübergestellt.

Parameter	Version p13a	Version p14a
Alters- und Geschlechtsgruppen (AGG)	34	34
Diagnosegruppen (DxG)	846	847
Risikokategorien (CC/HCC)	188	188
Hierarchien (CC→HCC)	268	268
Aggregierte Risikokategorien (ACC)	32	32
Klassifizierbare ICD-Jahresversionen	2013 bis 2019	2014 bis 2020

Tabelle 2: Parameter der Versionen p13a und p14a des Klassifikationssystems

3.2.2 Grundlagen zum Klassifikationssystem gemäß § 87a Abs. 5 SGB V (KS87a)

3.2.2.1 Überleitung der Schlüsselnummern der ICD-10-GM in die Risikokategorien des Klassifikationssystems

Die Zuordnung von Diagnosen zu diagnosebezogenen Risikokategorien erfolgt bis zur Ebene der unkomprimierten Risikokategorien in zwei Gruppierungsschritten. Im ersten Schritt werden alle gültigen ICD-Kodes (ca. 14.700) in aktuell 847 Diagnosegruppen (Dx-Gruppen, DxG) zusammengefasst. Im zweiten Schritt werden die Dx-Gruppen den aktuell 188 diagnosebezogenen Risikokategorien (Condition Categories, CC bzw. Hierarchical Condition Categories, HCC) zugeordnet.

Zum Zweck einer zusammenfassenden Übersicht ist jede CC bzw. HCC genau einer von 32 aggregierten Risikokategorien (Aggregated Condition Categories, ACC) zugeordnet.

In den folgenden Ausführungen werden die Gruppierungsschritte der Schlüsselnummern der ICD-10-GM in die Diagnosegruppen und Risikokategorien des Klassifikationssystems meist verkürzt als ICD-DxG-Überleitungen, DxG-(H)CC-Überleitungen, ICD-DxG-(H)CC-Überleitungen oder ICD-(H)CC-Überleitungen bezeichnet.

3.2.2.1.1 Zusatzkennzeichen für die Diagnosensicherheit und Seitenlokalisierung

Im Klassifikationssystem KS87a werden die Schlüsselnummern der ICD-10-GM immer nur im Zusammenhang mit ihrem Zusatzkennzeichen für die Diagnosensicherheit (ZK-DS) in Risikokategorien übergeleitet.

In Verbindung mit dem ZK-DS „G“ (gesichert) erfolgt für alle ICD-Kodes eine Überleitung in Risikokategorien. Von den ICD-Kodes mit dem ZK-DS „Z“ (Zustand nach) wird nur eine festgelegte Auswahl (Schlüsselnummern für Tumoren) übergeleitet. Für ICD-Kodes mit den ZK-DS „A“ (ausgeschlossen) oder „V“ (Verdacht auf) erfolgt keine Überleitung in Risikokategorien.

Wenn in den folgenden Ausführungen ICD-(H)CC-Überleitungen ohne Nennung des ZK-DS aufgeführt werden, handelt es sich immer um Überleitungen mit dem Zusatzkennzeichen „G“ (gesichert).

Die Zusatzkennzeichen für die Seitenlokalisierung werden im Klassifikationssystem derzeit nicht benutzt.

3.2.2.1.2 Alter und Geschlecht als Kriterien für ICD-DxG-(H)CC-Überleitungen

Weitere Kriterien für ICD-DxG-(H)CC-Überleitungen sind das Alter und das Geschlecht der Versicherten. Die meisten Überleitungen erfolgen unabhängig von Alter und Geschlecht, einige jedoch nur, wenn definierte Bedingungen für Alter und Geschlecht erfüllt sind. Dabei ist es auch möglich, dass eine Schlüsselnummer in Abhängigkeit von Alter und Geschlecht in unterschiedliche Risikokategorien übergeleitet wird.

3.2.2.1.3 Einfache und doppelte ICD-(H)CC-Überleitungen

Anhand der in den Abschnitten 3.2.2.1.1 und 3.2.2.1.2 genannten Kriterien (Zusatzkennzeichen für die Diagnosensicherheit, Alter, Geschlecht) ergibt sich für eine ICD-Schlüsselnummer in der Regel eine eindeutige Zuordnung zu einer einzelnen Risikokategorie („einfache ICD-(H)CC-Überleitung“).

In wenigen Einzelfällen wird eine ICD-Schlüsselnummer unabhängig von den genannten Kriterien immer gleichzeitig in zwei Risikokategorien übergeleitet. Zum Beispiel (KS87a-Version p14a) wird „I13.20 Hypertensive Herz- und Nierenkrankheit mit (kongestiver) Herzinsuffizienz und Niereninsuffizienz: Ohne Angabe einer hypertensiven Krise“ gleichzeitig der „(H)CC080 Kardiomyopathien, Herzinsuffizienz, PAH“ und auch der „(H)CC207 Progrediente Niereninsuffizienz“ zugeordnet. In den folgenden Ausführungen wird in diesen Fällen die Begrifflichkeit „doppelte ICD-(H)CC-Überleitung“ verwendet.

3.2.2.2 Hierarchisierung der Risikokategorien des Klassifikationssystems

Die Hierarchisierung der Risikokategorien gehört zu den Grundsätzen von diagnosebasierten Versichertenklassifikationssystemen, bei denen Versicherte nicht in disjunkte Risikokategorien klassifiziert werden. Die Hierarchien dienen dazu, dass die schwerwiegenderen Erkrankungen oder Zustände in der Regel die weniger schwerwiegenden dominieren.

Im Klassifikationssystem p14a stehen die meisten diagnosebezogenen Risikokategorien (140 von 188) untereinander in einer oder mehreren hierarchischen Beziehungen. Das heißt, sie sind einander über- oder untergeordnet. Die folgende Abbildung 1 soll dies beispielhaft verdeutlichen.

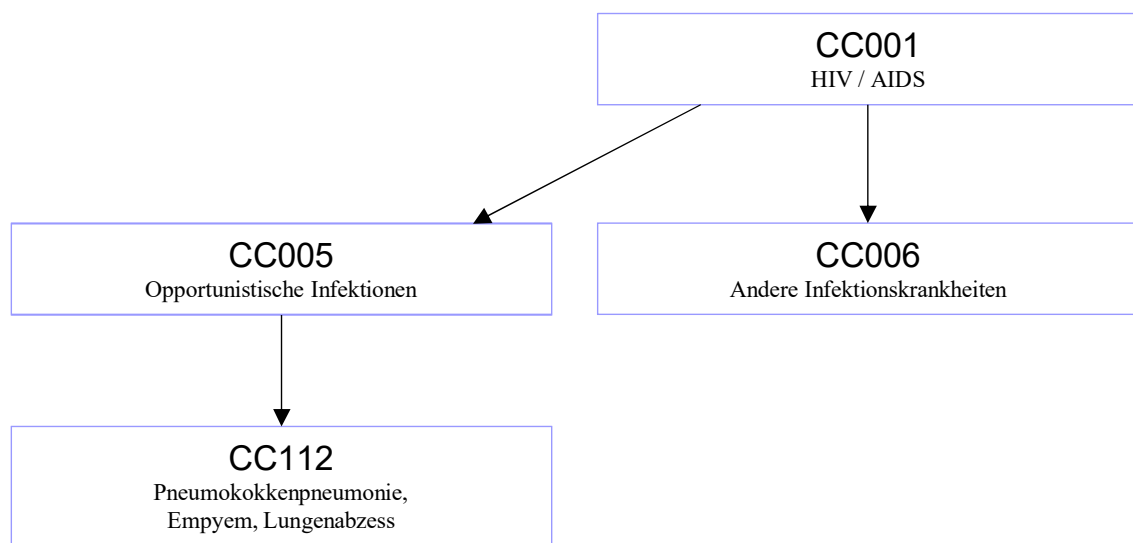


Abbildung 1: Beispielhafte Darstellung von Hierarchiebeziehungen zwischen Risikokategorien

Im Beispiel ist die CC005 der CC001 untergeordnet und der CC112 übergeordnet. Gleichzeitig ist auch die CC001 der CC112 übergeordnet, was als transitive Hierarchiebeziehung bezeichnet wird.

Wird einem Versicherten aufgrund seiner kodierten Diagnosen sowohl die CC001 als auch die CC005 zugewiesen, bewirkt die Hierarchisierung der Risikokategorien des Klassifikationssystems, dass die CC005 bei diesem Versicherten unterdrückt (gelöscht) und die CC001 (Condition Category 001) zur HCC001 (Hierarchical Condition Category 001) wird.

3.2.2.3 Alters-und Geschlechtsgruppen (AGG) des Klassifikationssystems

Neben den diagnosebezogenen Risikokategorien enthält das Klassifikationssystem KS87a auch Risikokategorien, die den Versicherten aufgrund ihres Alters und Geschlechts zugeordnet werden. Jedem Versicherten wird dabei genau eine Alters- und Geschlechtsgruppe zugeordnet. Diese Klassifizierung ist unabhängig davon, ob für den Versicherten auch Diagnosen im Datensatz vorliegen oder nicht. Sie hat keinerlei Bezug zur im Abschnitt

3.2.2.1.2 beschriebenen Funktion von Alter und Geschlecht als Kriterien für ICD-DxG-CC-Überleitungen.

3.2.3 Grundsätze der Bearbeitungen

Basis für die Grundsätze der Bearbeitungen des Klassifikationssystems sind die Vorgaben aus dem Beschluss des Bewertungsausschusses in seiner 269. Sitzung am 25. Januar 2012 zur Weiterentwicklung des Klassifikationssystems gemäß § 87a Abs. 5 SGB V.

3.2.3.1 *Beschluss des Bewertungsausschusses in seiner 269. Sitzung am 25. Januar 2012 zur Weiterentwicklung des Klassifikationssystems gemäß § 87a Abs. 5 SGB V*

Mit dem Beschluss des Bewertungsausschusses in seiner 332. Sitzung über das zur Ermittlung der diagnosebezogenen bzw. demografischen Veränderungsdaten zu verwendende Klassifikationsmodell für das Jahr 2015 gemäß § 87a Abs. 5 SGB V wurde dieser Beschluss mit Wirkung zum 31. August 2013 geändert. Diese Änderung betraf jedoch nur die Vorgaben zur Berichtslegung und nicht die hier thematisierten Festlegungen zur Weiterentwicklung des Klassifikationssystems.

3.2.3.1.1 *Anwendungsbedingungen*

Gemäß Beschluss des Bewertungsausschusses in seiner 269. Sitzung ist bei der Weiterentwicklung des Klassifikationssystems als Anwendungsbedingung insbesondere zu berücksichtigen, dass

- keine ambulanten Kodierrichtlinien zur Sicherstellung einheitlicher Diagnosequalität vorhanden sind,
- weitere regional unterschiedliche Gegebenheiten, insbesondere unterschiedliche Abgrenzungen der morbiditätsbedingten Gesamtvergütung, vorliegen können und
- die Ermittlung der diagnosebezogenen Veränderungen der Morbiditätsstruktur gemäß § 87a Abs. 4 SGB V auf der Ebene aller Versicherten aller Krankenkassen mit Wohnort im Bezirk einer Kassenärztlichen Vereinigung zu erfolgen hat.

Darüber hinaus sind Veränderungen der Struktur der vertragsärztlichen Versorgung durch Gesetzesänderungen und deren Umsetzung zu berücksichtigen.

3.2.3.1.2 *Kriterien und Ziele*

Gemäß Beschluss des Bewertungsausschusses in seiner 269. Sitzung sind bei der Weiterentwicklung des Klassifikationssystems die folgenden Kriterien und Ziele zu berücksichtigen.

- Das Klassifikationssystem soll alle Diagnosen berücksichtigen, die für die MGV-relevante Veränderung der Morbiditätsstruktur maßgeblich sind.
- Auf der untersten Ebene sollen die Diagnosen klinisch homogenen Diagnosegruppen zugeordnet werden, die eine Differenzierung sowohl akuter als auch chronischer, schwerwiegender und/oder besonders häufiger Krankheiten gewährleisten. Die Zuordnung von Diagnosen zu Risikokategorien erfolgt unter Verwendung dieser Diagnosegruppen. Die Änderung von Zuordnungen von Diagnosen zu Diagnosegruppen und Risikokategorien sowie die Anpassung der Risikokategorien, insbesondere ihre Zusammenlegung, Aufteilung und hierarchische Anordnung, haben regelbasiert zu erfolgen.
- Um den Einfluss von Variationen in der Kodierung multimorbider Patienten oder solcher mit einem bestimmten Schweregrad einer Erkrankung einzugrenzen, sind die Risikokategorien des weiterentwickelten Klassifikationssystems erstens in Hierarchien anzuordnen und zweitens zu komprimieren. Hierbei kann die Anordnung in Hierarchien nicht nur nach Schweregraden, sondern auch in organsystemübergreifenden Hierarchien erfolgen, in denen (eine) dominierende Risikokategorie(n) zur Beschreibung MGV-relevanter Veränderungen der Morbiditätsstruktur geeignet ist (sind).
- Die Risikokategorien des Klassifikationssystems sollen möglichst ökonomisch homogen sein.
- Das Klassifikationssystem und die damit ermittelten diagnosebezogenen Veränderungsdaten sollen hinsichtlich der statistischen Eigenschaften den Anforderungen in Bezug auf Reliabilität, Validität, Signifikanz, Stabilität und Schätzgenauigkeit genügen und insbesondere robust sein gegenüber Ausreißerwerten und verbliebenen Mängeln in den zugrunde liegenden Daten.
- Das Klassifikationssystem soll resistent sein gegenüber möglichen Anreizwirkungen zur Manipulation der Dokumentationsgrundlagen.
- Die jährlichen Anpassungen des Klassifikationssystems aufgrund von Änderungen in der ICD-10-GM sollen regelbasiert berücksichtigt werden.

3.2.3.1.3 *Rahmenvorgaben*

Der Beschluss des Bewertungsausschusses in seiner 269. Sitzung legt den Rahmen für die Anpassung des Klassifikationssystems fest. Die konkrete Umsetzung erfolgt durch das Institut des Bewertungsausschusses nach den jeweiligen Vorgaben der Trägerorganisationen in der Arbeitsgruppe Grouperanpassung bzw. medizinische Grouperanpassung des Bewertungsausschusses.

Die Ausgestaltung eines Regelwerks sowie ggf. notwendige Konkretisierungen der Methoden für die jährliche regelmäßige Anpassung des Klassifikationssystems sind im Verlauf

der Arbeiten zur Weiterentwicklung des Klassifikationssystems unter Berücksichtigung der Anwendungsbedingungen (siehe Abschnitt 3.2.3.1.1) und der Kriterien und Ziele (siehe Abschnitt 3.2.3.1.2) vorzunehmen.

Die Weiterentwicklung der Diagnosezuordnungen des Klassifikationssystems erfolgt auf der Ebene des jeweils aktualisierten Standes der für den definierten Verwendungszweck weiterentwickelten Version des Klassifikationssystems vor dem Hintergrund der Anwendungsbedingungen (siehe Abschnitt 3.2.3.1.1) und der Kriterien und Ziele (siehe Abschnitt 3.2.3.1.2). Dabei können neben den jährlichen Änderungen am ICD-Katalog auch weitere Aspekte, wie eine selektive Behebung von festgestellten Mängeln in der medizinisch und/oder ökonomisch begründeten Zuordnung, berücksichtigt werden.

3.2.3.2 Ausgestaltung des Regelwerks für die Weiterentwicklung des Klassifikationssystems

Wie im Beschluss des Bewertungsausschusses in seiner 269. Sitzung festgelegt, soll die Ausgestaltung des Regelwerks für die Weiterentwicklung des Klassifikationssystems im Verlauf der entsprechenden Bearbeitungen durch das Institut des Bewertungsausschusses nach den jeweiligen Vorgaben der Trägerorganisationen in der Arbeitsgruppe Grupperanpassung bzw. medizinische Grupperanpassung des Bewertungsausschusses erfolgen. Dementsprechend wurden im Rahmen der bisherigen Weiterentwicklungen des Klassifikationssystems ein Regelwerk für die Anpassungen an neue Jahresversionen der ICD-10-GM (siehe Abschnitt 3.2.3.2.1), Kriterien für die Beurteilung der klinischen Homogenität von Diagnosegruppen (siehe Abschnitt 3.2.3.2.2), eine Begriffsdefinition und ein Regelwerk für die Bearbeitung von Hierachieverletzungen (siehe Abschnitt 3.2.3.2.3) sowie ein noch vorläufiges Verfahren zur Operationalisierung der ökonomischen Homogenität der diagnosebezogenen Risikokategorien, insbesondere im Zusammenhang mit Änderungen von ICD-(H)CC-Überleitungen (siehe Abschnitt 3.2.3.2.4) erstellt.

3.2.3.2.1 Grundlagen und Regelwerk für die Anpassung des Klassifikationssystems an neue Jahresversionen der ICD-10-GM

In der vertragsärztlichen Versorgung sind die Diagnosen nach der Internationalen Klassifikation der Krankheiten in der jeweiligen vom Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM)⁴ herausgegebenen deutschen Fassung (ICD-10-GM) zu verschlüsseln (§ 295 Absatz 1 SGB V). Die ICD-10-GM wird vom BfArM in jedem Kalenderjahr in einer neuen geänderten Fassung herausgegeben. Dadurch ist es notwendig, auch das Klassifikationssystem regelmäßig an die jeweils neue Jahresversion der ICD-10-GM anzupassen.

Im Folgenden werden die drei wichtigsten Anpassungsschritte des Klassifikationssystems

⁴bis zum 26.05.2020: Deutsches Institut für Medizinische Dokumentation und Information (DIMDI)

an neue Jahresversionen der ICD-10-GM dargestellt. Diese erfolgen für alle endständigen (terminalen) und die nichtendständigen (nichtterminalen) vierstelligen Codes. Nichtendständige dreistellige ICD-Schlüsselnummern werden nicht mit einbezogen, da sie laut § 295 SGB V (Abrechnung ärztlicher Leistungen) zur Kodierung im vertragsärztlichen Bereich nicht zugelassen sind.

Bei der Beschreibung der Anpassungen des Klassifikationssystems an neue Jahresversionen der ICD-10-GM werden die Überleitungen der ICD-Kodes in die diagnosebezogenen Risikokategorien (ICD-(H)CC-Überleitungen) ohne die erste Gruppierungsstufe, die Zusammenfassung in Diagnosegruppen (siehe Abschnitt 3.2.2.1), dargestellt.

Neue ICD-(H)CC-Überleitungen für neue Schlüsselnummern der ICD-10-GM

Die jeweilige Jahresversion der ICD-10-GM enthält im Regelfall neue Schlüsselnummern im Vergleich zur Vorversion. Diese neuen Diagnosecodes werden in das Klassifikationssystem integriert, indem für sie jeweils Überleitungen in die bereits vorhandenen diagnosebezogenen Risikokategorien erstellt werden. Neben der Auswahl der zuzuordnenden Risikokategorie werden für diese neuen ICD-(H)CC-Überleitungen auch die weiteren Parameter – Zusatzkennzeichen für die Diagnosesicherheit, untere Altersgrenze, obere Altersgrenze und Geschlecht – festgelegt. Außerdem erhalten sie eine Markierung für das Gültigkeitsjahr der neuen Version der ICD-10-GM. Für alle Vorjahre werden sie als ungültig markiert.

Bei der Überleitung der neuen Codes gilt der Grundsatz: Eine neue Schlüsselnummer wird genauso übergeleitet wie die Schlüsselnummer, die ein Arzt voraussichtlich ausgewählt hätte, wenn er die gleiche Diagnose mit der ICD-Vorjahresversion kodiert hätte. Am Beispiel des Codes „M75.6 Läsion des Labrums bei degenerativer Veränderung des Schultergelenkes“, der 2019 neu in die ICD-10-GM aufgenommen wurde, soll dieser Grundsatz veranschaulicht werden.

Die ICD-Version 2018 ermöglicht die Kodierung degenerativer Schulterläsionen nur mit den folgenden Schlüsselnummern:

- M75.0 Adhäsive Entzündung der Schultergelenkkapsel
- M75.1 Läsionen der Rotatorenmanschette
- M75.2 Tendinitis des M. biceps brachii
- M75.3 Tendinitis calcarea im Schulterbereich
- M75.4 Impingement-Syndrom der Schulter
- M75.5 Bursitis im Schulterbereich
- M75.8 Sonstige Schulterläsionen
- M75.9 Schulterläsion, nicht näher bezeichnet

Wollte ein Arzt in 2018 die degenerative Läsion des Schultergelenk-Labrums verschlüsseln, hat er korrekterweise den ICD-Code „M75.8 Sonstige Schulterläsionen“ angegeben. Die Schlüsselnummer M75.8 wird im Klassifikationssystem in die Risikokategorie „(H)CC043

Andere Erkrankungen der Muskeln, Sehnen, Gelenke und des Bindegewebes“ übergeleitet. Ab 2019 gibt es für die degenerative Läsion des Schultergelenk-Labrums den spezifischen ICD-Kode „M75.6 Läsion des Labrums bei degenerativer Veränderung des Schultergelenkes“. Will ein Arzt jetzt diese Krankheit verschlüsseln, gibt er korrekterweise den neuen, genaueren Kode an. Die Integration der neuen Schlüsselnummer M75.6 in das Klassifikationssystem erfolgt entsprechend dem oben genannten Grundsatz so, dass sie – genau wie der Kode M75.8 – in die Risikokategorie „(H)CC043 Andere Erkrankungen der Muskeln, Sehnen, Gelenke und des Bindegewebes“ übergeleitet wird.

Wichtigste Orientierungshilfe für die Integration neuer Schlüsselnummern in das Klassifikationssystem ist die mit jeder Jahresversion der ICD-10-GM vom BfArM herausgegebene Überleitungstabelle zwischen der neuen und der Vorjahresversion. Diese wird durch das BfArM parallel zur Überarbeitung der ICD-10-GM erstellt, um die einheitliche Anpassung der Systeme, die die ICD-10-GM nutzen, zu gewährleisten.

Ungültige ICD-(H)CC-Überleitungen durch Streichung von Schlüsselnummern in neuen Versionen der ICD-10-GM

Mit neuen Jahresversionen der ICD-10-GM werden im Regelfall auch Schlüsselnummern der jeweiligen Vorversion aus dem Diagnosen-Katalog gestrichen. Bei der Anpassung des Klassifikationssystems werden alle ICD-(H)CC-Überleitungen dieser Schlüsselnummern ab dem Jahr des Inkrafttretens der neuen Version als ungültig markiert.

Prüfung der ICD-DxG-(H)CC-Überleitungen von Schlüsselnummern mit inhaltlichen Textänderungen in neuen Versionen der ICD-10-GM

Neben der Neuaufnahme und Streichung von Schlüsselnummern erfolgen mit neuen Jahresversionen der ICD-10-GM auch Änderungen der Textbezeichnungen bestehender Diagnose-Kodes. Diese Textänderungen sind überwiegend nur redaktionell, gelegentlich können sie aber auch die medizinisch inhaltliche Aussage einer ICD-Schlüsselnummer verändern. Bei den Anpassungen des Klassifikationssystems an neue Jahresversionen der ICD-10-GM werden daher auch alle ICD-(H)CC-Überleitungen der Schlüsselnummern mit inhaltlichen Textänderungen im Sinne ihrer weiteren medizinischen Korrektheit überprüft und gegebenenfalls geändert.

3.2.3.2.2 Kriterien für „klinische Homogenität von Diagnosegruppen“

Als erster Schritt für die Bearbeitung dieser Zielstellung wurden die folgenden grundsätzlichen Kriterien für die „klinische Homogenität von Diagnosegruppen“ formuliert.

Die „klinische Homogenität“ wurde definiert als Zusammengehörigkeit (Ähnlichkeit) von Diagnosen bezüglich

- Lokalisation (Organ, Organsystem, Körperteil, -abschnitt, -seite),
- Ätiologie,

- Komplikation(en),
- Komorbiditäten,
- Schweregrad/Stadium,
- Verlauf,
- akute oder chronische Erkrankung und
- Behandlung stationär oder ambulant.

Die Kriterien für die „klinische Homogenität von Diagnosegruppen“ sollen grundsätzlich herangezogen werden, wenn Weiterentwicklungen des Klassifikationssystems mit der Zielsetzung der Verbesserung der Zuordnung von Diagnosen zu klinisch homogenen Diagnosegruppen erfolgen. Aber in Bezug auf die jeweiligen spezifischen Bearbeitungsbereiche werden einerseits nicht immer alle Kriterien zutreffen können, andererseits wird es eventuell erforderlich sein, die vorliegenden Kriterien zu erweitern. Die aufgeführte Liste ist somit nicht als abschließend zu betrachten. Beispielsweise wurde bei einzelnen Anpassungen des Klassifikationssystems als weiteres Kriterium auch „Krankheit/Symptom“ berücksichtigt.

Prinzipiell nicht für den jeweiligen Bearbeitungsbereich nutzbar sind einzelne Kriterien, wenn die entsprechende Information nicht aus dem Diagnoseeintrag in der ICD-10-GM ableitbar ist.

Zum Beispiel kann aus den ICD-Schlüsselnummern

- I50.00 Primäre Rechtsherzinsuffizienz
- I50.01 Sekundäre Rechtsherzinsuffizienz

kein Schweregrad abgeleitet werden, wie bei den Codes

- I50.02! Rechtsherzinsuffizienz ohne Beschwerden (NYHA-Stadium I)
- I50.03! Rechtsherzinsuffizienz mit Beschwerden bei stärkerer Belastung (NYHA-Stadium II)
- I50.04! Rechtsherzinsuffizienz mit Beschwerden bei leichter Belastung (NYHA-Stadium III)
- I50.05! Rechtsherzinsuffizienz mit Beschwerden in Ruhe (NYHA-Stadium IV).

Letztere enthalten wiederum nicht die Information, ob die Krankheit primär oder sekundär ist, die gegebenenfalls für weitere Kriterien notwendig sein kann.

Ableitbarkeit entsprechender Informationen aus dem Diagnoseeintrag in der ICD-10-GM bedeutet nicht, dass die Information immer direkt im Text der Schlüsselnummer stehen muss. Zum Beispiel kann aufgrund des aktuellen medizinischen Wissensstandes mit hoher Wahrscheinlichkeit abgeleitet werden, dass bei der ICD-Diagnose „S00.31 Schürfwunde der Nase“ eine ambulante und bei „S02.7 Multiple Frakturen der Schädel- und Gesichtsschädelknochen“ eine stationäre Behandlung durchgeführt wird.

Neben den grundsätzlichen Kriterien für die „klinische Homogenität von Diagnosegruppen“

sind gegebenenfalls bei besonderen Fragestellungen auch die Häufigkeiten des Auftretens der einzelnen Diagnosen (Prävalenzen) mit zu berücksichtigen.

Ergänzend wurde festgelegt, dass die für die einzelnen Bearbeitungsbereiche zutreffenden Kriterien jeweils zu priorisieren sind. Die Auswahl und Priorisierung der Kriterien für „klinische Homogenität“ erfolgt für die einzelnen Bearbeitungsbereiche erstmals vor Beginn der Bearbeitung und wird im Verlauf kontinuierlich überprüft und gegebenenfalls angepasst.

3.2.3.2.3 *Begriffsdefinition und Regelwerk für die Bearbeitung von Hierarchieverletzungen*

Grundsätzliches zur Hierarchisierung der Risikokategorien des Klassifikationssystems wird bereits im Abschnitt 3.2.2.2 beschrieben. Die Hierarchien dienen dazu, dass die schwerwiegenderen Erkrankungen oder Zustände die weniger schwerwiegenden dominieren. Liegt eine Hierarchieverletzung vor, kann dieses Ziel möglicherweise nicht erreicht werden.

Begriffsdefinition „Hierarchieverletzung“

Eine Hierarchieverletzung zwischen zwei untereinander hierarchisierten Risikokategorien liegt vor,

- wenn die Kostengewichte beider Risikokategorien positiv und signifikant von Null verschieden sind (Signifikanzniveau 5%) und
- wenn das Kostengewicht der untergeordneten Risikokategorie größer ist als das Kostengewicht der übergeordneten Risikokategorie und
- wenn die Differenz zwischen den Kostengewichten der beiden Risikokategorien statistisch signifikant von Null abweicht. Hierfür muss ein Signifikanzniveau von 5% (p-Wert: $\leq 0,05$) vorliegen.

Die Signifikanzprüfung für die Differenz der Kostengewichte erfolgt mit dem F-Test zum Vergleich zweier Modelle: ein Modell mit Restriktionen (Nullhypothesenmodell) und ein Modell ohne Restriktionen (unrestringiertes Modell). Das unrestringierte Modell entspricht hier dem Regressionsmodell aus Formel (1) in Teil C. Das Nullhypothesenmodell enthält hingegen die Restriktion, dass die Kostengewichte der beiden CCs (CC1 und CC2) der zu untersuchenden Hierarchie gleich sein sollen ($H_0: \hat{\beta}_{CC1} = \hat{\beta}_{CC2}$). Die Teststatistik zur Prüfung der Nullhypothese berechnet sich nach der folgenden Formel (Formel für die Berechnung des F-Wertes):

$$F = \frac{(SQR_{Null} - SQR_{unrestr}) \cdot (T - K)}{SQR_{unrestr} \cdot L}$$

- SQR_{Null} Summe der Residuenquadrate für das Nullhypothesenmodell
(mit Restriktionen)
- $SQR_{unrestr}$ Summe der Residuenquadrate für das unrestringierte Modell
- L Anzahl der Restriktionen
- T Zahl der Beobachtungen (Zahl der Versicherten)
- K Zahl der unabhängigen Variablen im unrestringierten Modell
- F F-Wert (unter der Nullhypothese ist dieser F-verteilt mit L
Freiheitsgraden im Zähler und $T - K$ Freiheitsgraden im Nenner)

In der folgenden Abbildung 2 ist die Feststellung einer Hierarchieverletzung graphisch dargestellt.

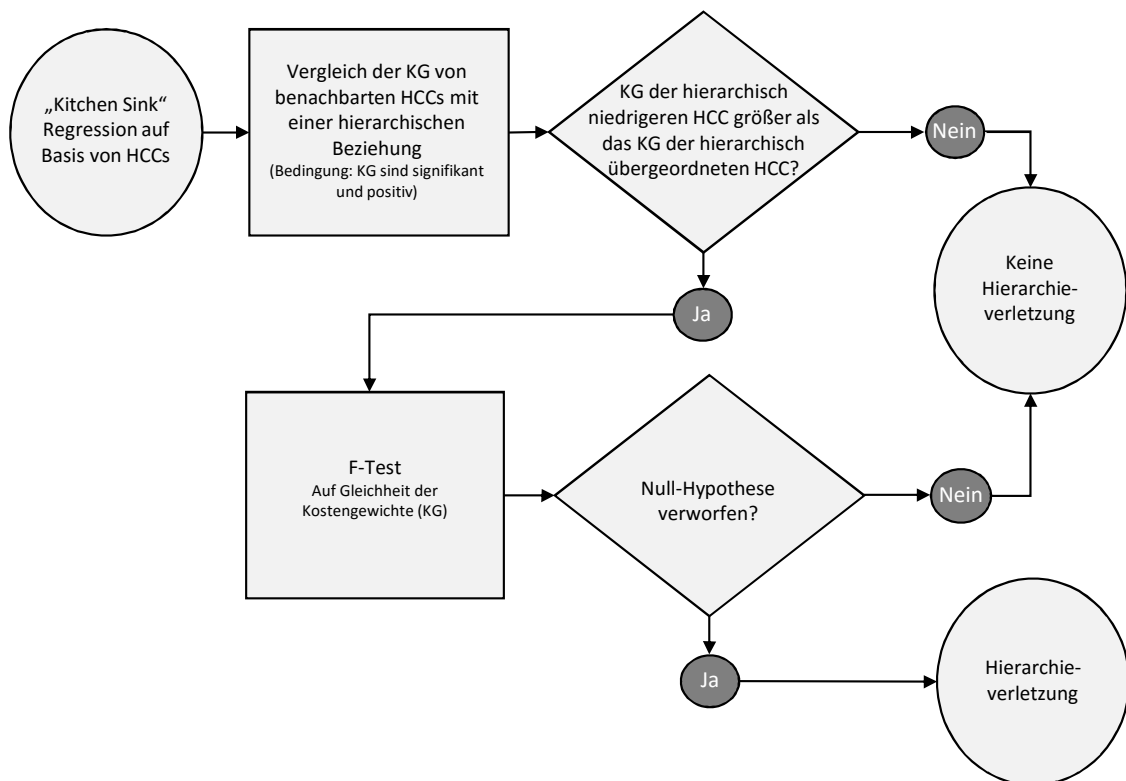


Abbildung 2: Flussdiagramm zur Feststellung einer Hierarchieverletzung

Regelwerk für die Bearbeitung von Hierarchieverletzungen

Gemeinsam mit der Definition des Begriffs „Hierarchieverletzung“ wurde das folgende Regelwerk für die Bearbeitung von Hierarchieverletzungen erarbeitet.

- Die Feststellung einer Hierarchieverletzung erfolgt grundsätzlich definitionsgemäß, also insbesondere nach Validierung mit dem F-Test (Signifikanzniveau: 5%).

- Zusätzlich wird untersucht, ob diese Hierarchieverletzung eine gewisse Konstanz hat, d.h. ob sie bereits bei früheren Versionen des Klassifikationssystems und/oder bei anderen Datenabgrenzungen vorhanden war.
- Bei jeder festgestellten und mit dem F-Test validierten Hierarchieverletzung ist zu untersuchen, ob sie gegebenenfalls unter medizinischen Gesichtspunkten tolerabel ist.
- Erst wenn dies nicht zutrifft, wird die Bearbeitung dieser Hierarchieverletzung eine Aufgabe im Rahmen der medizinischen Weiterentwicklung des Klassifikationssystems.

Für die Bearbeitung einer Hierarchieverletzung, bzw. allgemein für Überarbeitungen der Hierarchisierung der diagnosebezogenen Risikokategorien des Klassifikationssystems, stehen die folgenden Maßnahmen zur Verfügung:

- Entfernung der Hierarchie
- Änderung (Umkehrung) der Hierarchie
- Änderung der ICD-Zuordnung zu den untereinander hierarchisierten Risikokategorien
- Neuaufnahme von Hierarchien
- Akzeptanz der Hierarchieverletzung (im Ausnahmefall),
 - wenn die vorliegende Hierarchie medizinisch sinnvoll ist und/oder
 - wenn die Bewertungen der beiden untereinander hierarchisierten Risikokategorien sehr nah beieinander liegen, wobei dabei zunächst unbestimmt bleibt, was als „sehr nah“ anzusehen ist.

3.2.3.2.4 *Operationalisierung der Zielsetzung der ökonomischen Homogenität der diagnosebezogenen Risikokategorien*

Im Beschluss des Bewertungsausschusses in seiner 269. Sitzung am 25. Januar 2012 zur Weiterentwicklung des Klassifikationssystems gemäß § 87a Abs. 5 SGB V wurde unter „Kriterien und Ziele“ (siehe Abschnitt 3.2.3.1.2) formuliert: „Die Risikokategorien des Klassifikationssystems sollen möglichst ökonomisch homogen sein.“

Die Operationalisierung dieser Zielsetzung hat – auch abhängig von der präzisen Definition, worauf sich die ökonomische Homogenität der Risikokategorien genau beziehen sollte – zu unterschiedlichen Ansätzen geführt. Angestrebt wird ein quantitatives Konzept für die Beurteilung der ökonomischen Homogenität der diagnosebezogenen Risikokategorien im Klassifikationssystem. Dafür wurden mehrere Methoden zur Ermittlung geeigneter Messgrößen diskutiert. Bisher konnte keine Methode eine vollständig zufriedenstellende Beurteilung der ökonomischen Homogenität von diagnosebezogenen Risikokategorien liefern.

Im Rahmen der aktuellen Weiterentwicklung des Klassifikationssystems wird derzeit ein neu entwickeltes quantitatives Konzept getestet.

Im Vordergrund der Beurteilung der ökonomischen Homogenität von Risikokategorien steht nicht die Beurteilung der im Klassifikationssystem bereits existierenden diagnosebezogenen Risikokategorien. Zunächst soll die Betrachtung der ökonomischen Homogenität einen Beitrag zur Beurteilung von Weiterentwicklungen des Klassifikationssystems leisten, z.B. wenn neue diagnosebezogene Risikokategorien gebildet oder die ICD-Zuordnungen zu den vorhandenen geändert werden, bzw. wenn vorhandene diagnosebezogene Risikokategorien gesplittet oder fusioniert werden.

Solche Weiterentwicklungen des Klassifikationssystems sind grundsätzlich durch medizinisch fachliche Gründe veranlasst. Als vorläufige Lösungsmöglichkeit bei der Operationalisierung der Zielsetzung der ökonomischen Homogenität der diagnosebezogenen Risikokategorien wurde deshalb eine Kombination aus den Methoden zur Ermittlung geeigneter Messgrößen und dem Aspekt der jeweiligen medizinisch fachlichen Begründung wie folgt in das Regelwerk für die Weiterentwicklung des Klassifikationssystems aufgenommen:

- Bei Änderungen im Klassifikationssystem mit Umstrukturierung bzw. Neubildung von diagnosebezogenen Risikokategorien, bei denen auch die Auswirkungen auf die ökonomische Homogenität der Risikokategorien zu prüfen sind, sollen nach Möglichkeit entsprechende Kennzahlen gemäß den von Ärzte- bzw. Kassenseite entwickelten Berechnungsverfahren bzw. deren Weiterentwicklungen ermittelt und berücksichtigt werden.
- Die Umsetzung der vorgesehenen Änderungen ist jedoch nicht an die Prüfung(en) zur ökonomischen Homogenität gebunden, sofern Einvernehmen darin besteht, dass die aus medizinisch fachlicher Sicht erforderlichen Änderungen auch ohne diese Prüfung(en) vorgenommen werden können.
- Dieses Vorgehen stellt einen Konsens der Trägerorganisationen aufgrund des gegenwärtigen Diskussions- und Wissensstandes dar. Es ist nicht ausgeschlossen, dass zukünftig im Einvernehmen von dem oben genannten Vorgehen abgewichen werden kann.

3.2.4 Umsetzung der ICD-Änderungen bei der Schlüsselnummer U07.0/U07.0! im Klassifikationssystem

Nach Vorgabe durch die WHO hat das DIMDI⁵ für die ICD-10-GM mit Wirkung zum 01.11.2019 bei der Schlüsselnummer

U07.0 Nicht belegte Schlüsselnummer U07.0

⁵seit 26.05.2020 Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM)

die Textbezeichnung und den Primär-/Sekundärkode-Status wie folgt geändert
U07.0! Gesundheitsstörung im Zusammenhang mit dem Gebrauch von E-Zigaretten
[Vaporizer].

Diese Änderungen wurden mit Wirkung ab 2020 in das Klassifikationssystem übernommen.
Die ICD-DxG-(H)CC-ACC-Zuordnung von U07.0/U07.0! wurde nicht geändert.

3.2.5 Anpassung des Klassifikationssystems an die Version 2020 der ICD-10-GM

3.2.5.1 Neue ICD-CC-Überleitungen für neue Schlüsselnummern

Gegenüber der 2019er Version hat die Version 2020 der ICD-10-GM die folgenden sechs neuen Schlüsselnummern:

- D76.4 Zytokinfreisetzungssyndrom [cytokine release syndrome]
- H35.3- Degeneration der Makula und des hinteren Poles
- H35.30 Altersbedingte feuchte Makuladegeneration
- H35.31 Altersbedingte trockene Makuladegeneration
- H35.38 Sonstige Degeneration der Makula und des hinteren Poles
- Z29.22 HIV-Präexpositionsprophylaxe

Für jede der neuen ICD-Schlüsselnummern wurde eine neue Überleitung in eine diagnosebezogene Risikokategorie in das Klassifikationssystem aufgenommen. Alle neuen Überleitungen erfolgen als einfache Überleitungen (siehe Abschnitt 3.2.2.1.3), nur in Verbindung mit dem Zusatzkennzeichen für die Diagnosesicherheit „G“ und immer unabhängig von Alter und Geschlecht (siehe Abschnitt 3.2.2.1.2). Alle ICD-(H)CC-Überleitungen der neuen Schlüsselnummern der ICD-10-GM Version 2020 werden im Folgenden detailliert dargestellt.

D76.4 Zytokinfreisetzungssyndrom [cytokine release syndrome]
wird übergeleitet in die Risikokategorie
(H)CC045 Angeborene und erworbene Erkrankungen des Immunsystems.

- H35.3- Degeneration der Makula und des hinteren Poles
- H35.30 Altersbedingte feuchte Makuladegeneration
- H35.31 Altersbedingte trockene Makuladegeneration
- H35.38 Sonstige Degeneration der Makula und des hinteren Poles

werden übergeleitet in die Risikokategorie
(H)CC121 Makuladegeneration sowie diabetische Retinopathie und andere Netzhautaffektionen.

Z29.22 HIV-Präexpositionsprophylaxe
wird übergeleitet in die Risikokategorie
(H)CC202 Hierarchisch niedrigste Symptom-(H)CC.

3.2.5.2 Ab 2020 nicht mehr gültige ICD-CC-Überleitungen von Schlüsselnummern der ICD-10-GM 2019

Die Schlüsselnummer „H35.3 Degeneration der Makula und des hinteren Poles“ der Version 2019 der ICD-10-GM wurden ab 2020 aus dem Diagnosen-Katalog gestrichen. Ihre Überleitung in die diagnosebezogene Risikokategorie „(H)CC121 Makuladegeneration sowie diabetische Retinopathie und andere Netzhautaffektionen“ wurden daher ab dem Diagnosejahr 2020 als ungültig markiert.

3.2.5.3 Neue Diagnosegruppen im Klassifikationssystem

Im Zusammenhang mit der Aufnahme einer neuen Überleitung in eine diagnosebezogene Risikokategorie in das Klassifikationssystem für die ab 2020 neue ICD-Schlüsselnummer „Z29.22 HIV-Präexpositionsprophylaxe“ wurde die folgende Diagnosegruppe neu gebildet und der Risikokategorie „(H)CC202 Hierarchisch niedrigste Symptom-(H)CC“ zugeordnet: DXG0022 HIV-Präexpositionsprophylaxe.

3.3 Weiterentwicklung der Berechnung der Veränderungsdaten

Im Beschluss des Bewertungsausschusses in seiner 503. Sitzung wurde gegenüber dem Vorjahresbeschluss der Umgang mit der neuen Geschlechtsausprägung „unbestimmt“ geregelt: Versicherten mit der Geschlechtsausprägung „unbestimmt“ in Feld 05 der Satzart 201 wird die Geschlechtsausprägung „weiblich“ zugewiesen. Hintergrund sind die Regelungen im Klassifikationssystem und bei der demografischen Hochrechnung. Im Klassifikationssystem erfolgt die Zuordnung von Diagnosen zu Diagnosegruppen und Risikokategorien teilweise abhängig von den Geschlechtsausprägungen „weiblich“ oder „männlich“. Bei der demografischen Hochrechnung wird auf die amtliche KM6-Statistik, die die Versichertenzahlen nach „weiblich“ und „männlich“ ausweisen, hochgerechnet.

Anhang 1: Abkürzungsverzeichnis

ACC	Aggregated Condition Category(ies)
AGG	AG-Gruppe(n), Alters-Geschlechts-Gruppe(n), Alters- und Geschlechtsgruppe(n)
BA	Bewertungsausschuss nach § 87 Abs. 1 Satz 1 SGB V
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte
CC	Condition Category(ies)
DHF	Demografische Hochrechnungsfaktoren
DIMDI	Deutsches Institut für Medizinische Dokumentation und Information
DxG	Dx-Gruppe(n), Diagnosegruppe(n)
EBA	Erweiterter Bewertungsausschuss nach § 87 Abs. 4 SGB V
EGV	extrabudgetäre Gesamtvergütung
F-Test	statistischer Test, bei dem die Testgröße bei Gültigkeit der Nullhypothese einer F-Verteilung folgt
GKV	Gesetzliche Krankenversicherung
GSP	Geburtstagsstichprobe
H15EBA	Kürzel für das Klassifikationsmodell gemäß Beschluss des 15. EBA
HCC	Hierarchical Condition Category(ies)
ICD-10-GM	Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme - 10. Revision - German Modification
ICD-Kode	Schlüsselnummer der ICD-10-GM
KG	Kostengewicht
KM87a	Klassifikationsmodell gemäß § 87a Abs. 5 SGB V
KM87a_2012	Kürzel für das Klassifikationsmodell gemäß Beschluss des 29. EBA mit Anpassungen der Datengrundlage gemäß Beschluss des 30. EBA und des 288. BA
KM87a_2013	Kürzel für das Klassifikationsmodell gemäß Beschluss des 309. BA ergänzt durch den Beschluss des 313. BA
KM87a_2014	Kürzel für das Klassifikationsmodell gemäß Beschluss des 332. BA
KM87a_2015	Kürzel für das Klassifikationsmodell gemäß Beschluss des 355. BA in Verbindung mit dem Beschluss des 332. BA
KM87a_2016	Kürzel für das Klassifikationsmodell gemäß Beschluss des

	382. BA
KM87a_2017	Kürzel für das Klassifikationsmodell gemäß Beschluss des 400. BA
KM87a_2018	Kürzel für das Klassifikationsmodell gemäß Beschluss des 430. BA
KM87a_2019	Kürzel für das Klassifikationsmodell gemäß Beschluss des 439. BA
KM87a_2020	Kürzel für das Klassifikationsmodell gemäß Beschluss des 503. BA
KS87a	Klassifikationssystem gemäß § 87a Abs. 5 SGB V
KV	Kassenärztliche Vereinigung
KV SH	Kassenärztliche Vereinigung Schleswig-Holstein
KV HH	Kassenärztliche Vereinigung (Hansestadt) Hamburg
KV HB	Kassenärztliche Vereinigung (Hansestadt) Bremen
KV NI	Kassenärztliche Vereinigung Niedersachsen
KV WL	Kassenärztliche Vereinigung Westfalen-Lippe
KV NR	Kassenärztliche Vereinigung Nordrhein
KV HE	Kassenärztliche Vereinigung Hessen
KV RP	Kassenärztliche Vereinigung Rheinland-Pfalz
KV BW	Kassenärztliche Vereinigung Baden-Württemberg
KV BY	Kassenärztliche Vereinigung Bayerns
KV BE	Kassenärztliche Vereinigung Berlin
KV SL	Kassenärztliche Vereinigung Saarland
KV MV	Kassenärztliche Vereinigung Mecklenburg-Vorpommern
KV BB	Kassenärztliche Vereinigung Brandenburg
KV ST	Kassenärztliche Vereinigung Sachsen-Anhalt
KV TH	Kassenärztliche Vereinigung Thüringen
KV SN	Kassenärztliche Vereinigung Sachsen
MAPE	Mean Absolute Prediction Error
MGV	morbiditybedingte Gesamtvergütung
Nicht-SV-TN	Nichtteilnehmer an Selektivverträgen
NonCW-HCC	NonCostWeight Hierarchical Condition Category (HCC mit einem normativen Relativgewicht von null)

p-Wert	Wahrscheinlichkeit, bei Gültigkeit der Nullhypothese den errechneten oder einen extremeren Wert der Teststatistik zu erhalten
RHCC	„Rest“ Hierarchical Condition Category, zusammengefasste hierarchisierte Risikokategorie
SGB V	Sozialgesetzbuch V, Das Fünfte Buch Sozialgesetzbuch Gesetzliche Krankenversicherung (Artikel 1 des Gesetzes vom 20. Dezember 1988, BGBl. I S. 2477, 2482)
SV-Teilnehmer / SV-TN	Selektivvertragsteilnehmer
THCC	„Top“ Hierarchical Condition Category, ausgewählte hierarchisierte Risikokategorie
ZK-DS	Zusatzkennzeichen für die Diagnosesicherheit

Anhang 2: Dokumentation zur Bereitstellung der regionalisierten Geburtstagsstichprobe 2016–2018 für die Gesamtvertragspartner

Version 1.1a

Stand 01.09.2020

Inhaltsverzeichnis

1	Vorbemerkungen	1
1.1	Beschlusslage des Bewertungsausschusses	1
1.2	Aufbau dieses Dokuments	2
1.3	Änderungen im Vergleich zur GSPR 2017	3
1.4	Ansprechpartner	3
2	Beschreibung des Datenkörpers	4
2.1	Datenmodell	4
2.2	Felder der einzelnen Satzarten	5
2.3	Markierung von Auffälligkeiten in einzelnen Satzarten – Flagging	12
2.3.1	Einführung	12
2.3.2	Flag der Satzart DS201 – Versichertenstammdaten	13
2.3.3	Flag der Satzart DS202 – Ambulante Abrechnungen (KV-Fall)	18
2.3.4	Flag der Satzart DS219 – Kostenträgerverzeichnis (KT-Stamm)	19
2.3.5	Flag der Satzart DS220 – Kostenträgerhistorie	19
2.4	Ausgewählte Ergebnisse der Datenprüfung	20
3	Einschränkungen der Nutzung	33
3.1	Ursachen für eingeschränkte Nutzungsmöglichkeiten	33
3.2	Intra- und interregionale Vergleiche	35
3.3	Längsschnittbetrachtung	35
3.4	Demografische Hochrechnungsfaktoren	36
4	Hinweise zur Nutzung	36
4.1	Allgemeine Nutzungshinweise	36
4.2	Hinweise zur Verknüpfung mit anderen Datenkörpern	39
5	Kommentierte Datensatzbeschreibung	40
5.1	Satzart DS201 – Versichertenstammdaten	41
5.2	Satzart DS202 – Ambulante Abrechnungen (KV-Fall)	51
5.3	Satzart DS203 – Diagnosen der ambulanten Behandlung (KV-Fall-Diagnosen)	54
5.4	Satzart DS210 – Gebührenordnungspositionen der ambulanten Behandlung	56
5.5	Satzart DS211 – Betriebsstättenverzeichnis	57
5.6	Satzart DS213 – ICD-10-Stammdatei	59
5.7	Satzart DS215 – Gebührenordnungspositions-Stammdatei	60
5.8	Satzart DS219 – Kostenträgerverzeichnis (KT-Stamm)	61
5.9	Satzart DS220 – Kostenträgerhistorie	62
5.10	Festlegungen zur Datenübermittlung an die Gesamtvertragspartner	66
5.10.1	Arztseitig	66
5.10.2	Kassenseitig	66
6	Geburtsstichprobe (GSP)	67
6.1	Historie und Beschlusslage des (Erweiterten) Bewertungsausschusses	67
6.2	Übermittlung der Daten an die Datenstelle des BA	70
6.2.1	Datenlieferung	70
6.2.2	Pseudonymisierungsverfahren	70
6.2.3	Satzarten der Geburtsstichprobe	71
6.3	Zusammenführung der Datenlieferungen zur GSP – Matching	71
6.3.1	Definition	71
6.3.2	Aufgabenstellung	72
6.3.3	Ziel des Matchingverfahrens	73
6.3.4	Stufen des Matchingverfahrens	73
6.3.5	Matching-Quoten	77

6.4	Verbleibende Mängel der Datenqualität der zusammengeführten Daten.....	81
6.5	Identifizierung von Teilnehmern an Selektivverträgen	81
A	Anhang	82
A.1	Glossar	82
A.2	Geschlechtsspezifische Abrechnungsfälle	84

Abbildungsverzeichnis

1	Überblick über die Satzarten der GSPR 2018	6
---	--	---

Tabellenverzeichnis

1	Anteil auffälliger Datensätze in der Satzart DS201	14
2	Fusionskassen mit falsch pseudonymisierten Personen-IDs	15
3	Lieferquoten der Versichertenstammdaten nach KV-Bezirk	22
4	Anzahl der Fusionskassen nach Lieferquote	23
5	Anzahl der Versicherten nach Lieferquote	24
6	Anzahl der Fusionskassen nach Lieferquote	26
7	Anzahl der Versicherten nach Lieferquote	27
8	Inanspruchnahmequote nach KV-Bezirk.....	29
9	Ausschöpfungsgrad des Leistungsbedarfs nach KV-Bezirk	31
10	Anzahl der Fusionskassen nach Ausschöpfungsgrad des Leistungsbedarfs	32
11	Anzahl der Versicherten nach Ausschöpfungsgrad des Leistungsbedarfs	33
12	Matching-Quote nach Matchingstufen	78
13	Matching-Quote nach KV	79
14	Matching-Quote nach Kassenart	80
A1	Geschlechtsspezifische Diagnosen	85
A2	Geschlechtsspezifische GOPen.....	87

1 Vorbemerkungen

1.1 Beschlusslage des Bewertungsausschusses

Der Bewertungsausschuss hat in seiner 274. Sitzung, Teil A (schriftliche Beschlussfassung), geändert durch die Beschlüsse des Bewertungsausschusses in seiner 288. Sitzung und in seiner 308. Sitzung (schriftliche Beschlussfassung), zuletzt geändert durch den Beschluss des Bewertungsausschusses in seiner 438. Sitzung (schriftliche Beschlussfassung), beschlossen, den Trägerorganisationen des Bewertungsausschusses Daten der sogenannten regionalisierten Geburtstagsstichprobe (GSPR) zu übermitteln, die diese den Vertragspartnern nach § 87a Abs. 2 Satz 1 SGB V zur Verfügung stellen. Die Vertragspartner sollen dadurch Kenntnis von der Datengrundlage zur Ermittlung der demografischen sowie diagnosenbezogenen Veränderungsrate nehmen können. Die Vertragspartner sollen die Datengrundlage insbesondere hinsichtlich der Kodierqualität und der Plausibilität prüfen können. Eine abweichende Berechnung demografischer und diagnosenbezogener Veränderungsrate auf regionaler Ebene ist nicht vorgesehen. Jedoch kann aufgrund der Einsicht in die Datengrundlagen eine regionalspezifische Gewichtung der beiden Raten gemäß § 87a Absatz 4 Satz 3 SGB V erfolgen.

Gegenstand der Datenübermittlung sind die an die Datenstelle des Bewertungsausschusses gelieferten Daten der Geburtstagsstichprobe (GSP), die in der Datenstelle gemäß den Vorgaben der Arbeitsebene des Bewertungsausschusses und unter Mitwirkung des Instituts des Bewertungsausschusses zusammengeführt und qualitätsgesichert wurden. Die daraus abgeleitete GSPR der Berichtsjahre 2016–2018 wird durch die Datenstelle des Bewertungsausschusses an die Trägerorganisationen des Bewertungsausschusses übermittelt. Die Trägerorganisationen leiten diese Daten unter Einsatz der von ihnen eigenständig abzustimmenden sicheren Verfahren innerhalb von vier Wochen, regional nach Wohnortprinzip abgegrenzt, an die jeweiligen Gesamtvertragspartner weiter. Dabei ist sicherzustellen, dass den jeweiligen Vertragspartnern jeweils nur die entsprechend zusammengeführten versichertenbezogenen Stamm- bzw. Abrechnungsdaten der Versicherten mit Wohnort im jeweiligen KV-Bezirk übermittelt werden. In den Daten, die der GKV-Spitzenverband an die regionalen Verbände der Krankenkassen übermittelt, wurde der jeweilige Kassenbezug für die nicht der Empfängerkassenart angehörigen Krankenkassen in einzelnen Satzarten entfernt und das entsprechende Feld leer oder mit einem Ersatzwert übermittelt.

Die Daten der Geburtstagsstichprobe, die den Gesamtvertragspartnern zur Verfügung gestellt werden, wurden entsprechend der vom Bewertungsausschuss beschlossenen Datensatzbeschreibung und entsprechenden Festlegungen – auch bezüglich der Qualitätssicherung – auf der Arbeitsebene des Bewertungsausschusses im Umfang reduziert und angepasst (für die Datensatzbeschreibung siehe Kapitel 5). Die Datengrundlage der vorliegenden GSPR 2018 wurde im Rahmen der AG IT am 24.03.2020 von den Trägerorganisationen abgenommen.

Der so entstandene Datenkörper wird im Weiteren als regionalisierte Geburtstagsstichprobe 2018 (GSPR 2018) bezeichnet. Wie oben erwähnt, umfasst die GSPR 2018 den Berichtszeitraum 2016–2018 (genauer die Quartale 2016/1 bis 2018/4) und wurde zuletzt am 15.04.2020 von der Datenstelle des Bewertungsausschusses an die Trägerorganisationen des Bewertungsausschusses übermittelt.

1.2 Aufbau dieses Dokuments

Die vorliegende Dokumentation stellt für die Empfänger der GSPR 2018 Informationen zusammen, die die Nutzung des Datenkörpers erleichtern sollen. Es ist nicht Ziel dieser Dokumentation, konkrete Nutzungsmöglichkeiten aufzuzählen oder zu beschreiben.

Kapitel 2 richtet sich in erster Linie an Leser ohne vertiefte Kenntnisse der Geburtstagsstichprobe und beschreibt den Inhalt sowie den Umfang der GSPR 2018. Dabei werden zunächst das Datenmodell skizziert und anschließend die einzelnen Satzarten genauer beschrieben. Zudem wird auf einige bei der Datenprüfung der GSPR 2018 beobachtete Auffälligkeiten eingegangen und beschrieben, welche Datensätze aufgrund dieser Auffälligkeiten markiert wurden. Kapitel 2 schließt mit einem Überblick über ausgewählte Kennzahlen zur Beschreibung der Datenqualität.

In Kapitel 3 werden Einschränkungen der Nutzungsmöglichkeiten der GSPR 2018 und ihre Gründe ausgeführt.

Kapitel 4 gibt allgemeine Hinweise zur Nutzung des Datenkörpers. Wie bereits erwähnt, wird dabei nicht auf konkrete Nutzungsmöglichkeiten eingegangen.

Kapitel 5 dokumentiert die vom Bewertungsausschuss in seiner 438. Sitzung (schriftliche Beschlussfassung) als Anlage beschlossene technische Datensatzbeschreibung. Zum besseren Verständnis für den Nutzer enthält diese zusätzliche Erläuterungen.

Kapitel 6 richtet sich an Leser, die mehr Informationen zu der der GSPR 2018 zugrunde liegenden Versichertenstichprobe des Bewertungsausschusses erhalten möchten. Dabei handelt es sich um die eingangs bereits erwähnte Geburtstagsstichprobe, deren Aufbereitung kurz skizziert wird. Insbesondere wird der Zusammenführungsprozess („Matching“) der kassenseitig und arztseitig gelieferten Daten beschrieben und kurz auf die Identifikation von an einem Selektivvertrag teilnehmenden Versicherten eingegangen.

Im Anhang dieses Dokuments finden sich neben einigen Begriffsdefinitionen (Anhang A.1) eine Übersicht über Diagnosen und Gebührenordnungspositionen, die auf das Geschlecht von Versicherten schließen lassen und bei der Qualitätsprüfung von Bedeutung sind (Anhang A.2).

1.3 Änderungen im Vergleich zur GSPR 2017

Für Leser, die mit der GSPR der Vorjahres (GSPR 2017) vertraut sind, seien an dieser Stelle die wesentlichen Änderungen der GSPR 2018 genannt:

- **Berichtszeitraum:** Der Datenkörper GSPR 2018 umfasst die Berichtsjahre 2016–2018 (GSPR 2017: 2015–2017).
- **Austausch- bzw. Korrekturdatenlieferungen:** Aufgrund von Austausch- bzw. Korrekturdatenlieferungen für die Jahre 2016 und 2017 können sich die Daten für diese beiden Jahre zwischen der GSPR 2017 und der GSPR 2018 unterscheiden.
- **Fusionsstand:** Die Daten der GSPR 2018 berücksichtigen den Fusionsstand der gesetzlichen Krankenkassen am 02.04.2020 (GSPR 2017: 23.01.2019).
- **Markierung von Datensätzen der Satzart DS201:** Bei der Qualitätsprüfung der Geburtstagsstichprobe werden Kriterien zur Markierung auffälliger Datensätze festgelegt. Diese Markierungen wirken sich auch auf die Markierung von Datensätzen in der Satzart DS201 aus (siehe Abschnitt 2.3).
- **Matchingverfahren:** Dem Matchingverfahren liegen Daten der Berichtsjahre 2015–2018 zugrunde. Vor diesem Hintergrund können die in der Dokumentation zur GSPR 2017 berichteten Matching-Quoten von den hier berichteten Matching-Quoten der Jahre 2016 und 2017 abweichen. Eine weitere Veränderung beim Matchingverfahren ist der aktualisierte Fusionsstand der gesetzlichen Krankenkassen.
- **Hinzukommende und wegfallende Geburtskalendertage:** Für die Berichtsjahre 2016 und 2017 enthalten die GSPR 2017 und die GSPR 2018 jeweils Daten zu Versicherten mit den gleichen Geburtskalendertagen. Das Berichtsjahr 2018 ist nur in der GSPR 2018 enthalten und umfasst Daten zu Versicherten mit den Geburtskalendertagen 5, 6, 12, 13, 19, 25 und 26.

1.4 Ansprechpartner

Als Ansprechpartner stehen die jeweiligen Trägerorganisationen – GKV-Spitzenverband bzw. Kassenärztliche Bundesvereinigung – zur Verfügung.

2 Beschreibung des Datenkörpers

2.1 Datenmodell

Mit der GSPR 2018 liegt ein Datenkörper vor, in dem kassen- und kassenartübergreifend arzt- und kassenseitig gelieferte Daten miteinander nach Wohnortprinzip und aktuellem Kassenfusionsstand (02.04.2020) verknüpft vorliegen und nach KV-Bezirken sowie nach Kassenarten regionalisiert abgrenzbar sind bzw. abgegrenzt werden. Die GSPR 2018 enthält somit nicht nur relevante Angaben zu Versicherten in der gesetzlichen Krankenversicherung, sondern auch differenzierte Angaben zur Inanspruchnahme vertragsärztlicher Leistungen dieser Versicherten über einen Berichtszeitraum von drei Jahren (Quartale 2016/1 bis 2018/4). Für die GSPR 2018 werden im genannten Zeitraum für das Berichtsjahr 2016 diejenigen Versicherten ausgewählt (Stichprobenversicherte), die am 5., 11., 12., 18., 19., 25. oder 26. Kalendertag eines beliebigen Monats Geburtstag haben; für das Berichtsjahr 2017 diejenigen Versicherten, die am 5., 6., 12., 18., 19., 25. oder 26. Kalendertag eines beliebigen Monats Geburtstag haben und für das Berichtsjahr 2018 diejenigen Versicherten, die am 5., 6., 12., 13., 19., 25. oder 26. Kalendertag eines beliebigen Monats Geburtstag haben.

Die arztseitig gelieferten Abrechnungsdaten stellen differenzierte Informationen über die Inanspruchnahme kollektivvertragsärztlicher Leistungen der Stichprobenversicherten zur Verfügung. Im Zentrum der arztseitig gelieferten Daten steht der Abrechnungsfall. Für den einzelnen Abrechnungsfall sind nicht nur die zugehörige Betriebsstätte und der abgerechnete Leistungsbedarf bekannt, sondern auch die in diesem Abrechnungsfall dokumentierten Diagnosen und abgerechneten Gebührenordnungspositionen (GOPen). Durch das mitgelieferte Betriebsstättenverzeichnis werden der Betriebsstätte (Arztpraxis) Informationen wie z. B. der Praxistyp und die Abrechnungsgruppe zugeordnet. Die ebenfalls beigefügte ICD-10-Stammdatei und die Gebührenordnungspositions-Stammdatei erlauben die Zuordnung von Schlüsselnummern zur Bezeichnung bzw. geben Aufschluss über die Bewertung der abgerechneten GOPen. Jedem der in der GSPR 2018 ausgelieferten Abrechnungsfälle ist eine Personen-ID eindeutig zugeordnet. Über diese Personen-ID können die arztseitig gelieferten Daten mit den kassenseitig gelieferten Daten verknüpft werden.

Die kassenseitig gelieferten Daten zu gesetzlich versicherten Personen in der GSPR 2018 umfassen unter anderem Geburtsjahr, Geschlecht, Versichertenzeiten, Wohnort und Informationen über die Teilnahme an Selektivverträgen. Jede Person ist dabei zeitraumunabhängig eindeutig über ihre Personen-ID zu identifizieren.

Durch das in der GSPR 2018 enthaltene Kostenträgerverzeichnis und die Kostenträgerhistorie ist außerdem die historische Kassenzugehörigkeit jeder Person dokumentiert, d. h. es ist nachzuvollziehen, wie zurückliegende Versichertenverhältnisse einer Person (also z. B. im ersten Quartal 2016) zu den zwischenzeitlich existierenden Fusionskassen in Beziehung stehen. Es könnten also zum Beispiel kassenbezogene Auswertungen durchge-

führt werden, die dem Fusionsstand am 02.04.2020 entsprechen, auch wenn die Versichertenstammdaten sich auf weiter zurückliegende Quartale beziehen, in denen Kostenträger angegeben sind, die nicht mehr autonom existierten.

Einschränkend ist hinzuzufügen, dass den Gesamtvertragspartnern jeweils nicht der gesamte Datenkörper der GSPR 2018 zur Verfügung gestellt wird. Den arztseitigen Empfängern wird eine regional aufgeteilte GSPR 2018 zur Verfügung gestellt, d. h. den Empfängern werden entsprechend nur die versichertenbezogenen Stamm- und Abrechnungsdaten der Versicherten mit Wohnort im jeweiligen KV-Bezirk übermittelt. Den kassenseitigen Empfängern werden die versichertenbezogenen Stamm- und Abrechnungsdaten der Versicherten derjenigen Kassen, die nicht zur eigenen Kassenart gehören, zwar geliefert, allerdings ist der Kassenbezug in diesen Fällen maskiert. Oben beschriebene kassenbezogene Auswertungen können damit nur für Fusionskassen der eigenen Kassenart durchgeführt werden.

Dennoch stellt dieser Datenbestand, der für jede Region Informationen aus allen Kassen aller Kassenarten umfasst und hinsichtlich der regionalen Versicherten gegebenenfalls auch zurückliegende Inanspruchnahme von Ärzten aus anderen KV-Bezirken umfasst, eine umfangreiche Datengrundlage dar.

Mit der GSPR 2018 liegt somit ein Datenkörper vor, mit dem die Gesamtvertragspartner ihre Vorbereitungen zur Anpassung des morbiditätsbedingten Behandlungsbedarfs gemäß § 87a Abs. 4 SGB V durchführen können. Die in Kapitel 3 aufgeführten Einschränkungen der Nutzungsmöglichkeiten sind dabei zu beachten.

2.2 Felder der einzelnen Satzarten

Die GSPR 2018 umfasst die folgenden Satzarten:

- DS201: Versichertenstammdaten
- DS202: Ambulante Abrechnungen (KV-Fall)
- DS203: Diagnosen der ambulanten Behandlung (KV-Fall-Diagnosen)
- DS210: Gebührenordnungspositionen der ambulanten Behandlung
- DS211: Betriebsstättenverzeichnis
- DS213: ICD-10-Stammdatei
- DS215: Gebührenordnungspositions-Stammdatei
- DS219: Kostenträgerverzeichnis (KT-Stamm)
- DS220: Kostenträgerhistorie

Für die technische Datensatzbeschreibung siehe Kapitel 5.

Vier der neun gelieferten Satzarten enthalten versicherten- bzw. abrechnungsfallspezifische Felder (DS201, DS202, DS203, DS210), während die fünf verbleibenden Satzarten

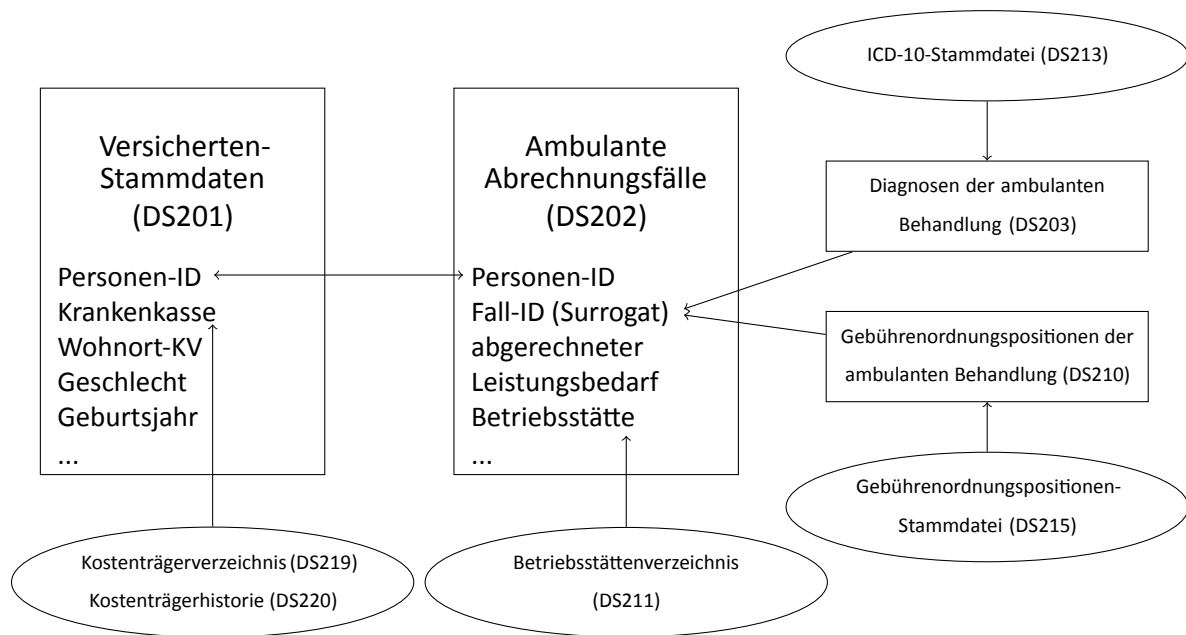


Abbildung 1: Überblick über die Satzarten der GSPR 2018

als Stammdatenverzeichnisse zu bezeichnen sind. Abbildung 1 gibt einen Überblick über die Beziehung der Satzarten zueinander.

Die kassenseitigen Empfänger erhalten gemäß Beschluss des Bewertungsausschusses die Datensätze der Satzarten DS201, DS202, DS203 und DS210 nur dann mit Kassenbezug, wenn die entsprechende Kasse bzw. deren Rechtsnachfolger¹ zur Kassenart gehört, für die der kassenseitige Empfänger benannt ist (Stand 02.04.2020). Die für eine Kasse zu verwendende Kassenart wird dabei der Kostenträgerhistorie der Satzart DS220 (Feld 07 – Kassenart des direkten Rechtsnachfolgers) entnommen.

Den arztseitigen Empfängern werden die zusammengeführten versichertenbezogenen Stamm- bzw. Abrechnungsdaten für Versicherte mit Wohnort im jeweiligen KV-Bezirk übermittelt (bezieht sich auf die Datensätze der Satzarten DS201, DS202, DS203, DS210, DS211 und DS215).

Im Folgenden werden einige der in den genannten Satzarten erfassten Felder aufgezählt und ggf. erläutert.

DS201 – Versichertenstammdaten

Die Satzart DS201 enthält die mit der bundesweiten Versichertenstichprobe erhobenen Versichertenstammdaten. Jeder Stichprobenversicherte ist für den gesamten Berichtszeitraum eindeutig über eine Personen-ID identifizierbar. Zu einer Personen-ID können in einem Quartal auch mehrere Datensätze vorliegen. Für jede Personen-ID liegt je ge-

¹Wie bereits ausgeführt, ist es mit dem vorliegenden Datenkörper möglich, die Fusionshistorie der am 02.04.2020 existierenden Krankenkassen (hier „Fusionskassen“ genannt) nachzuvollziehen. Als Rechtsnachfolger kommt damit nur eine dieser Fusionskassen in Frage.

liefertem Quartal und Institutionskennzeichen (gemäß Feld 02) genau ein Datensatz vor. Versicherte mit Wohnsitz im Ausland und Versicherte mit Übernahme der Krankenbehandlung gegen Kostenerstattung nach § 264 Abs. 2 SGB V² (sogenannte „Betreute“) sind für den gesamten Berichtszeitraum 2016/1–2018/4 in den vorliegenden Datenkörper einbezogen, wobei alle Datensätze der letztgenannte Gruppe als auffällig markiert sind (siehe Abschnitt 2.3.2).

Pro Datensatz liegen neben den zuvor genannten Informationen die Angabe zur Anzahl der Versichertentage, das Geschlecht des Versicherten, das Geburtsjahr, der Versicherungsstatus, die Angabe, ob der Versicherte in diesem Quartal verstorben ist, sowie die Wohnort-KV vor. Jeder Versicherte bekommt für alle Quartale eines Kalenderjahres dieselbe Wohnort-KV zugeordnet. Als Wohnort-KV des Versicherten ist dabei die Kassenärztliche Vereinigung am Wohnort³ im letzten Versicherungsquartal des betreffenden Kalenderjahres zu ermitteln. Das Merkmal „Kennzeichen Wohnort“ gibt an, ob sich der Wohnsitz des Versicherten im Inland oder im Ausland befindet. Haben Versicherte innerhalb eines Quartals die Krankenkasse gewechselt, so liegen für diese Versicherten in diesen Quartalen mehrere Datensätze (für verschiedene Krankenkassen) vor.

Neben den genannten Attributen wird außerdem das Institutionskennzeichen (IK) des sogenannten Rechtsnachfolgers angegeben. Jedem Datensatz sind damit zwei Angaben zum Kostenträger zugeordnet (Feld 02 und Feld 12, siehe Datensatzbeschreibung in Kapitel 5). Die Versichertenstammdaten enthalten pro Datensatz die Angabe eines Institutionskennzeichens (Feld 02) der Krankenkasse, bei der der Versicherte im Berichtsquartal versichert war. Es kann sich um ein Krankenversicherungskarten-IK (KVK-IK) oder ein anderes IK der entsprechenden Krankenkasse handeln (im Folgenden als „Teilkasse“ bezeichnet). Bei dieser Krankenkasse muss es sich nicht zwangsläufig um eine Fusionskasse handeln. Außerdem ist zu beachten, dass es im Berichtszeitraum der GSPR 2018 (d. h. zwischen 2016/1 und 2018/4) zu Fusionen von Krankenkassen gekommen ist. Die ursprünglich gemeldete Krankenkasse (ggf. Teilkasse) muss also am Stichtag nicht mehr eigenständig existieren. Aus diesem Grund ist der Satzart DS201 das Rechtsnachfolger-Institutionskennzeichen angefügt worden (Feld 12). Anhand dieses Kennzeichens kann über die Satzart DS220 dem jeweiligen durch die Krankenkasse gemeldeten Kostenträger in Feld 02 der Name einer der am Stichtag existierenden Fusionskassen als Rechtsnachfolger zugeordnet werden.

Die Satzart DS201 umfasst weiterhin das Feld 13, welches für Gewichtungsfaktoren vor-

²Da für den Personenkreis nach § 264 Abs. 1 SGB V, geändert durch das Asylverfahrensbeschleunigungsgesetz, wegen der dazu nicht rechtzeitig angepassten Implementierung der Datensatzbeschreibung aufgrund der fehlenden Vereinbarung zu einem bundeseinheitlichen Verfahren für die Abrechnung und Nachweisführung der Versorgung dieses Personenkreises keine Abrechnungsdaten übermittelt werden konnten, konnte nicht hinreichend ausgeschlossen werden, dass bei einigen Kassen sowohl Personen der Personengruppe nach § 264 Abs. 2 und/oder Abs. 1 in den Versichertenstammdaten geliefert wurden. Insbesondere aus diesen Gründen wurden die „Betreuten“ im Rahmen der Qualitätssicherung der GSP mit einem Ausschlusskennzeichen markiert. Zum Umgang der „Betreuten“ siehe Abschnitt 2.3.2.

³Für Versicherte mit Wohnsitz im Ausland wird die KV des Kassensitzes herangezogen.

gesehen ist, die immer dann anzuwenden sind, wenn Auswertungen eine Datengrundlage erfordern, die hinsichtlich Alter und Geschlecht repräsentativ sein sollen. Dieses Feld wird in der GSPR 2018, wie schon in den vergangenen Jahren, leer übermittelt. Im Feld 14 wird zudem der Kalendertag des Geburtstags des Versicherten übermittelt und das Feld 15 enthält ein Kennzeichen, ob für den Versicherten eine Kostenübernahme nach § 264 Abs. 2 SGB V⁴ vorliegt.

Aus der Ausprägung des Feldes 16 der Satzart DS201 geht hervor, ob der Versicherte an einem bereinigungsrelevanten Selektivvertrag gemäß §§ 63, 73b, 140a SGB V oder 73c SGB V a. F. teilgenommen hat. Im Feld 17 erfolgt eine Angabe zum selektivvertraglichen Abrechnungsweg und, unabhängig davon, im Feld 18 eine Angabe zur Vollständigkeit des kollektivvertraglichen Versorgungsauftrags⁵.

Eine besondere Bedeutung kommt dem Feld „Flag“ zu. Hier werden Auffälligkeiten der entsprechenden Personen-ID bei der Qualitätsprüfung angezeigt. **Es wird empfohlen, Datensätze mit der Flagausprägung „1“ nicht in möglichen Auswertungen zu verwenden.** Gründe, die zur Markierung von Datensätzen führen, sind in Abschnitt 2.3.2 ausgeführt.

DS202 – Ambulante Abrechnungen (KV-Fall)

Die vertragsärztlichen Abrechnungsfälle sind in der Satzart DS202 enthalten. Bestandteil der GSPR 2018 sind alle Abrechnungsfälle von Versicherten der Geburtstagsstichprobe, denen erfolgreich eine Personen-ID der Versichertenstammdaten zugeordnet werden konnte (zum Verfahren der Zuordnung siehe Abschnitt 6.3). Diejenigen Abrechnungsfälle, denen keine Personen-ID zugeordnet werden kann, sind nicht Bestandteil der GSPR 2018. Jedem Abrechnungsfall der Satzart DS202 ist damit über die Personen-ID die in der Satzart DS201 enthaltene Information über den Versicherten zugeordnet. Die Zuordnung zur Personen-ID erfolgt ggf. quartalsübergreifend. Nicht für jeden Abrechnungsfall existiert daher zwangsläufig ein *quartalsgleicher* Datensatz in der Satzart DS201. Allerdings sind in der GSPR 2018 nur solche Abrechnungsfälle enthalten, denen ein *jahresgleicher* Datensatz in der Satzart DS201 zugeordnet werden kann.

Für jeden Datensatz, also für jeden Abrechnungsfall, liegen in der Satzart DS202 das Abrechnungsquartal, das Rechtsnachfolger-IK (d. h. das Institutionskennzeichen der Fusionskasse des Abrechnungsfalles zum Stand 02.04.2020) sowie die Personen-ID vor. Zusätzlich wird ein künstlicher Schlüssel zur Identifikation des Abrechnungsfalles, die sogenannte Fall-ID, angegeben. Dieser künstliche Schlüssel dient dazu, die Diagnosen der Satzart DS203 und die Gebührenordnungspositionen der Satzart DS210 einem Abrechnungsfall zuzuordnen. Jedem Abrechnungsfall ist darüber hinaus die Wohnort-KV zugeordnet.

⁴Siehe Fußnote 2 auf Seite 7.

⁵In der GSPR 2018 kann das Feld auch eine Ausprägung annehmen, die nicht im Beschlusstext des Bewertungsausschusses in seiner 438. Sitzung definiert ist. Einzelheiten sind in der Anmerkung zum Feld 18 in Abschnitt 5.1 zu finden.

Weiterhin ist zu jedem Abrechnungsfall der KV-Bezirk der Betriebsstätte (Arztpraxis) enthalten. Nimmt der Versicherte Leistungen außerhalb seines Wohnort-KV-Bezirks in Anspruch („Fremdfall“), so wird die KV-Nummer der Betriebsstätte durch „00“ gekennzeichnet und das Betriebsstättenpseudonym ist in diesen Fällen durch den Wert „9999xxxx“ maskiert, wobei „xxxx“ die vierstellige Abrechnungsgruppenzuordnung der Betriebsstätte bezeichnet.

In zwei Feldern der Satzart DS202 wird der pro Abrechnungsfall abgerechnete Leistungsbedarf laut EBM, nach sachlich-rechnerischer Richtigstellung und vor Anwendung honorarwirksamer Begrenzungsregelungen berichtet. Es wird dabei die Summe des abgerechneten Leistungsbedarfs der in Punkten bewerteten Gebührenordnungspositionen des EBM in Zehntelpunkten sowie die Summe des abgerechneten Leistungsbedarfs der in Euro bewerteten Gebührenordnungspositionen des EBM in Cent angegeben. In einem weiteren Feld wird die Summe des abgerechneten Leistungsbedarfs laut regionaler Euro-Gebührenordnung nach sachlich-rechnerischer Richtigstellung und vor Anwendung honorarwirksamer Begrenzungsregelungen in Cent berichtet. Weiterhin ist die Anzahl der Leistungstage mit Abrechnung von Gebührenordnungspositionen angegeben.

Die Satzart DS202 enthält außerdem ein Feld mit der Bezeichnung „Flag“, welches zur Markierung auffälliger Datensätze dient. Gründe für eine Markierung sind in Abschnitt 2.3.3 beschrieben. In der vorliegenden GSPR 2018 wurde kein Datensatz mit der Ausprägung „1“ als auffällig markiert.

DS203 – Diagnosen der ambulanten Behandlung (KV-Fall-Diagnosen)

Die Diagnosen der Abrechnungsfälle sind in der Satzart DS203 enthalten. Für jeden Abrechnungsfall der Satzart DS202 werden in der Satzart DS203 neben dem Abrechnungsquartal und der Fall-ID alle gemeldeten Diagnosen⁶ durch die Nennung des ICD-10-GM-Codes aufgezählt. Zur weiteren Beschreibung der Diagnosen werden außerdem die Diagnosesicherheit und die Seitenlokalisierung angegeben. Anhand des Feldes „Diagnosezähler“ kann die Anzahl der insgesamt gemeldeten Diagnosen pro Abrechnungsfall ermittelt werden. Das Feld „Rechtsnachfolger-IK“ wird leer übermittelt. Im Falle der Übermittlung des Ersatzwertes „UUU“ als Diagnose in Feld 05 wird als Diagnosesicherheit (Feld 06) kein Eintrag übermittelt.

Die abrechnungsfallspezifischen Diagnosen der Satzart DS203 sind über den künstlichen Schlüssel „Fall-ID“ den Abrechnungsfällen der Satzart DS202 zugeordnet.

DS210 – Gebührenordnungspositionen der ambulanten Behandlung

Die abgerechneten Gebührenordnungspositionen (GOPen) der Abrechnungsfälle sind in der Satzart DS210 enthalten. Für jeden Abrechnungsfall der Satzart DS202 werden in

⁶Für eine vollständige Transparenz werden die Diagnosedatensätze der in der GSPR 2018 enthaltenen Abrechnungsfälle den Gesamtvertragspartnern so zur Verfügung gestellt, wie die Datenstelle sie erhalten hat.

der Satzart DS210 neben dem Abrechnungsquartal und der Fall-ID alle abgerechneten GOPen aufgezählt. Zur Beschreibung der GOPen wird neben der EBM-Ziffer⁷ die pro Abrechnungsfall abgerechnete Anzahl der einzelnen GOPen, der für diese GOPen abgerechnete Leistungsbedarf nach EBM (inklusive Angabe des Kennzeichens der Einheit: Zehntelpunkte, Cent oder sonstige) sowie die Art der Inanspruchnahme des Leistungsscheins angegeben. Zusätzlich wird der Leistungsbedarf der abgerechneten GOPen nach Euro-Gebührenordnung angeführt. Das Feld „Rechtsnachfolger-IK“ wird leer übermittelt.

Die abrechnungsfallspezifischen Gebührenordnungspositionen der Satzart DS210 sind über den künstlichen Schlüssel „Fall-ID“ den Abrechnungsfällen der Satzart DS202 zugeordnet.

DS211 – Betriebsstättenverzeichnis

In der Satzart DS211 werden quartalsweise alle Betriebsstätten aufgeführt, die im Berichtszeitraum (2016/1 bis 2018/4) Leistungen für Stichprobenversicherte erbracht haben. Daraus ergibt sich, dass die Satzart DS211 pro Quartal nicht zwangsläufig praxisvollständig ist. Es kann nicht ausgeschlossen werden, dass einzelne Betriebsstätten in den jeweiligen Quartalen keine Leistungen für Stichprobenversicherte abgerechnet haben. Sie sind dann im Betriebsstättenverzeichnis für dieses Quartal nicht enthalten.

Neben dem Abrechnungsquartal und dem Betriebsstättenpseudonym sind die KV-Nummer, die Abrechnungsgruppe sowie der Praxistyp (z. B. Einzelpraxis, MVZ usw.) der Betriebsstätte angegeben. Weiterhin ist das Kennzeichen des Versorgungsbereichs (Hausarzt, Facharzt oder übergreifend) ausgewiesen.

Über das Betriebsstättenpseudonym sind die aufgezählten Attribute des Betriebsstättenverzeichnisses den Betriebsstätten der einzelnen Abrechnungsfälle der Satzart DS202 zugeordnet, dies allerdings nur dann, wenn das Betriebsstättenpseudonym der Satzart DS202 nicht aufgrund von Fremdfällen maskiert wurde (siehe Ausführungen zur Satzart DS202).

DS213 – ICD-10-Stammdatei

In der ICD-10-Stammdatei werden alle ICD-10-GM-Diagnosecodes aufgeführt, die im Rahmen der Abrechnung verwendet werden können und nach SGB V zulässig sind. Neben dem ICD-10-Code werden der Gültigkeitszeitraum, der Klartext des ICD-10-Codes sowie das Notationskennzeichen angegeben. Mit dem Notationskennzeichen wird der Diagnosticsyp (Kreuzcode, Sterncode oder Ausrufezeichencode) beschrieben.

Dieses Schlüsselverzeichnis dient zur Beschreibung der in der Satzart DS203 verwendeten ICD-10-Codes.

⁷Die angegebene Gebührenordnungsposition kann auch eine regional vergebene Ziffer sein (siehe Ausführungen zur Satzart DS215).

DS215 – Gebührenordnungspositions-Stammdatei

In der Gebührenordnungspositions-Stammdatei werden alle GOPen, die in den einzelnen Quartalen in den jeweiligen KV-Bezirken gültig sind (inklusive regional vergebener GOPen), aufgeführt. Neben dem Leistungsquartal und der KV-Nummer werden für die einzelnen GOPen der Wert der GOP nach EBM in Zehntelpunkten oder Cent, unterschieden in ambulant und stationär, angegeben. Außerdem wird der Wert der GOP nach regionaler Euro-Gebührenordnung in Cent, wiederum unterschieden nach ambulant und stationär, aufgeführt.

DS219 – Kostenträgerverzeichnis (KT-Stamm)

Jeder Abrechnungsfall in der Satzart DS202 wird mit einem Institutionskennzeichen geliefert (Feld 02). Dabei kann es sich um ein Geschäftsstellen-IK, um ein Praxisnetz-IK, Abrechnungs-IK oder anderes handeln. Im Kostenträgerverzeichnis der Satzart DS219 erfolgt die Zuordnung aller dem GKV-Spitzenverband bekannten Institutionskennzeichen (im Feld KVK-IK) zu einem übergeordneten Kassensitz-IK, das die Krankenkasse in der GSPR 2018 eindeutig kennzeichnet. Soweit nicht vom Kostenträger selbst eindeutig bestimmt, wird eines dieser Institutionskennzeichen vom GKV-Spitzenverband ausgewählt und fortan als Kassensitz-IK in den Satzarten DS219 und DS220 angegeben. Das Kassensitz-IK dient der Zusammenfassung einzelner KVK-IKs zu einer übergeordneten Kasse, die zu einem nicht näher definierten Zeitpunkt als eigenständige Kasse existiert hat.

Auch in der Satzart DS219 ist ein Feld mit der Bezeichnung „Flag“ zu finden. Diese Markierung dient Dokumentationszwecken: Hier werden diejenigen Kostenträger durch die Feldausprägung „1“ markiert, bei denen Versicherte aufgrund bestimmter Auffälligkeiten in der Satzart DS201 markiert wurden. Für eine genauere Beschreibung des Flags siehe Abschnitt 2.3.4.

DS220 – Kostenträgerhistorie

Die Kostenträgerhistorie der Satzart DS220 macht Angaben zu den einzelnen Kostenträgern/Krankenkassen. Im Zentrum steht dabei das Kassensitz-IK aus der Satzart DS219. Für jedes Kassensitz-IK kann in der Satzart DS220 der Name des Kostenträgers und sämtliche Rechtsnachfolger nachvollzogen werden. Hat ein Kostenträger nicht nur einen, sondern (durch mehrfache, nacheinander erfolgte Fusionen) mehrere Rechtsnachfolger, so lässt sich die gesamte Fusionshistorie in den einzelnen Fusionsstufen nachvollziehen. Für alle Kassensitz-IKs sind der Name und die Kassenart sowohl des direkten Rechtsnachfolgers als auch des aktuellen Rechtsnachfolgers entsprechend dem Fusionsstand zum 02.04.2020 angegeben. Zusätzlich wird angegeben, in welchem Zeitraum der direkte Rechtsnachfolger gleich dem Rechtsnachfolger der höchsten Fusionsstufe gemäß dem in diesem Zeitraum vorliegenden Fusionsstand war bzw. ist. In der Fusionsstufe 1 ist der Kostenträger nicht fusioniert, d. h. der Name des Kostenträgers, der Name des direkten

Rechtsnachfolgers und der Name des aktuellen Rechtsnachfolgers sind identisch. Bei Fusionsstufen größer als 1 ist dies nicht mehr der Fall. Die Anzahl der Fusionen, die der Kostenträger bis zum aktuellen Rechtsnachfolger durchlaufen hat, ist gleich der Fusionsstufe minus eins.

Die Satzart DS220 enthält neben den genannten Informationen außerdem die Information über den KV-Bezirk des Kassensitz-IKs. Diese Angabe bezieht sich ausschließlich auf die Kassensitz-IKs der aktuellen Rechtsnachfolger, d. h. das entsprechende Feld ist nur für diejenigen Kassensitz-IKs ausgefüllt, die einem der aktuellen Rechtsnachfolger-IK entsprechen, sonst bleibt das Feld zum KV-Bezirk leer.

Das Feld „Flag“ der Satzart DS220 dient wiederum Dokumentationszwecken: Hier werden diejenigen Kostenträger durch die Felddarstellung „1“ markiert, bei denen Versicherte aufgrund bestimmter Auffälligkeiten in der Satzart DS201 markiert wurden. Für eine genauere Beschreibung des Flags siehe Abschnitt 2.3.5.

2.3 Markierung von Auffälligkeiten in einzelnen Satzarten – Flagging

2.3.1 Einführung

Wie in Abschnitt 2.2 ausgeführt, ist in vier Satzarten (DS201, DS202, DS219 und DS220) ein Feld mit der Bezeichnung „Flag“ enthalten. Hiermit werden üblicherweise Auffälligkeiten in den Daten markiert. Diese Auffälligkeiten wurden im Rahmen einer Datenprüfung der Geburtstagsstichprobe identifiziert und haben verschiedene Ursachen. Diese Ursachen werden in den folgenden Abschnitten beschrieben.

Das Flag der Satzart DS201 unterscheidet sich dabei von den Flags der Satzarten DS219 und DS220. Das erstgenannte Flag markiert bestehende Auffälligkeiten eines Datensatzes. Immer wenn das Flag in der Satzart DS201 die Ausprägung „1“ aufweist, ist der Datensatz als auffällig markiert. **Es wird empfohlen, diese Datensätze aus einer möglichen Auswertung auszuschließen, da sie Auswertungsergebnisse deutlich verzerren können.**

Auch die Satzart DS202 enthält ein Feld mit der Bezeichnung „Flag“. Wie bei der Satzart DS201 wird auch hier bei der Ausprägung „1“ der Ausschluss des jeweiligen Datensatzes bei Auswertungen empfohlen. In der vorliegenden GSPR 2018 werden allerdings keine Abrechnungsfälle als auffällig markiert und das Flag-Feld besitzt ausschließlich die Ausprägung „0“.

Die Flags der Satzarten DS219 und DS220 dienen hingegen Dokumentationszwecken. Hier werden Teilkassen bzw. Fusionskassen markiert, bei denen Versicherte in der Satzart DS201 markiert wurden.

Im Folgenden werden die Flags der einzelnen Satzarten und ihre Ursachen genauer beschrieben und die Kassen genannt, in denen die beschriebenen Probleme auftreten.

2.3.2 Flag der Satzart DS201 – Versichertenstammdaten

Die Markierung eines Versichertenstammdatensatzes (Satzart DS201) wird über das Feld 09 („Flag“) angegeben, welches entweder die Ausprägung „0“ oder die Ausprägung „1“ hat. Die Ausprägung „0“ entspricht dabei keiner Markierung im eigentlichen Sinne, sondern dient der Kennzeichnung eines Datensatzes ohne Auffälligkeiten. Die Ausprägung „1“ markiert hingegen einen Datensatz der Satzart DS201 mit einer oder mehreren Auffälligkeiten. **Es wird empfohlen, Datensätze mit Auffälligkeiten (Flagausprägung „1“) für Auswertungen nicht zu verwenden.** Die in der Datensatzbeschreibung der Satzart DS201 außerdem aufgeführte Ausprägung „2“ (Datensatz aus Fehlwertersatz) ist für die GSPR 2018 nicht relevant, da kein Fehlwertersatzverfahren angewendet wurde.

Zu den Auffälligkeiten, die zu einer Markierung mit „1“ führen, gehören:

- Nicht eindeutige Personen-ID
- Falsch pseudonymisierte Personen-ID
- Dummy-Personen-ID bzw. -Versichertennummer
- Kollidierende Personen-ID
- Datensätze nach dem Sterbequartal
- Fehlende Datensätze nach Abgleich mit Daten zu Selektivverträgen
- Versicherte mit Kostenübernahme nach §264 Abs. 2 SGB V⁸
- Falschmatcher
- Auffälliger Ausschöpfungsgrad

Für Leser, die mit der GSPR 2017 vertraut sind, sei an dieser Stelle darauf hingewiesen, dass sich bei der GSPR 2018 die Gründe, die zu einer Auffälligkeitsmarkierung in der Satzart DS201 geführt haben, teilweise geändert haben. Mit Hilfe des Auffälligkeitsflags lässt sich ermitteln, inwiefern sich die Anzahl der in der GSPR 2018 als auffällig identifizierten Versicherten für die Jahre 2016 und 2017 gegenüber den gleichen Jahren in der GSPR 2017 geändert hat. Aufgrund der unterschiedlichen Pseudonymisierung der Personen-IDs in den beiden Versionen der GSPR ist jedoch nicht erkennbar, für welche Personen-IDs eine Änderung eingetreten ist.

Tabelle 1 liefert einen Überblick über den Anteil der markierten Datensätze für die Berichtsjahre 2016 bis 2018⁹. In der letzten Spalte werden außerdem für jedes Jahr die beiden häufigsten Markierungsgründe genannt. Die aufgeführten Werte basieren auf Auswertungen der gesamten GSPR 2018.

Im Folgenden werden die oben genannten Gründe näher erläutert, die für die GSPR 2018

⁸Siehe Fußnote 2 auf Seite 7.

⁹Die berichteten Anteile berücksichtigen Datensätze bzw. Personen-IDs gegebenenfalls mehrfach. Die ausgewiesenen Zahlen addieren sich demnach nicht notwendigerweise zur Gesamtzahl markierter Datensätze.

Jahr	Anteil markierter Datensätze	Relevanteste Markierungsgründe und Anteil an markierten Datensätzen
2016	0,42 %	Personen mit Kostenübernahme-Kennzeichen: 90,54 %
2017	0,56 %	Personen mit Kostenübernahme-Kennzeichen: 92,18 %
2018	0,48 %	Personen mit Kostenübernahme-Kennzeichen: 92,48 %

Tabelle 1: Anteil von als auffällig markierten Datensätzen in der Satzart DS201. In der letzten Spalte sind die beiden häufigsten Markierungsgründe und ihre relative Häufigkeit an allen markierten Datensätzen genannt.

zu einer Markierung geführt haben.

Nicht eindeutige Personen-ID

In den von den Krankenkassen gelieferten Daten kann es vorkommen, dass einem Versichertenverhältnis (eindeutig identifiziert anhand des gemeldeten Institutionskennzeichens und der pseudonymisierten Versichertennummer) mehrere Personen-IDs zugeordnet sind. Alle Personen-IDs, die solchen Versichertenverhältnissen zugeordnet sind, werden in der Satzart DS201 *über den gesamten Berichtszeitraum* markiert. Diese Markierung ist unabhängig vom jeweiligen Versichertenverhältnis. Wechselt der Versicherte seine Krankenkasse, so ist die Personen-ID entsprechend auch in diesem Versichertenverhältnis markiert. Hat ein Versicherter mehrfache Versicherungswechsel, so wird dieses Vorgehen auch auf die folgenden Versichertenverhältnisse angewendet.

Insgesamt sind aus dem genannten Grund 81.392 Datensätze der Satzart DS201 in der GSPR 2018 markiert (0,04 % aller Datensätze). Dies betrifft 10.351 distinkte Personen-IDs (0,05 % aller Personen-IDs).

Falsch pseudonymisierte Personen-ID

Es kann vorkommen, dass Krankenkassen falsch pseudonymisierte Versichertennummern und/oder Personen-IDs liefern¹⁰. Dies kann alle Versicherten einer Fusions- oder Teilkasse betreffen, oder nur für einzelne Versicherte (z. B. diejenigen mit einem bestimmten Geburtstagskalendertag) der Fall sein. Alle Personen-IDs der mit Fehlern behaf-

¹⁰Sind bei einem Versicherten die Personen-ID und die Versichertennummer falsch pseudonymisiert, so kann diesem Versicherten kein Abrechnungsfall zugeordnet werden. Dieser Versicherte wird damit gegebenenfalls fälschlicherweise als ein Nie-Inanspruchnehmer definiert, da er gemäß den Daten keine ärztlichen Leistungen in Anspruch nimmt. Findet man nun für Teilkassen bzw. Fusionskassen mit gleichbleibenden und nicht zu kleinen Versichertenanzahlen in einem oder mehreren Quartalen eine sehr geringe Inanspruchnahmequote, so kann dies auf einen Pseudonymisierungsfehler in einzelnen Quartalen hindeuten. Die Inanspruchnahmequote gibt dabei den Anteil der Versicherten mit mindestens einem Abrechnungsfall an allen Stichprobenversicherten an (siehe Abschnitt 2.4).

teten Geburtskalendertage der betroffenen Fusionskassen (bzw. Teilkassen) werden in allen ihren Versichertenverhältnissen *für das entsprechende Jahr* markiert. Wechselt der Versicherte in diesem Jahr seine Krankenkasse, so ist die Personen-ID entsprechend auch in diesem Versichertenverhältnis markiert. Hat ein Versicherter in diesem Jahr mehrfache Versicherungswechsel, so wird dieses Vorgehen auch auf die folgenden Versichertenverhältnisse angewendet.

Für die betroffenen Fusionskassen sind in Tabelle reftab:2 die Größenklasse und das Jahr, in dem die Personen-IDs von Versicherten markiert werden, aufgeführt. Insgesamt sind aufgrund falsch pseudonymisierter Personen-IDs 187 Datensätze in der Satzart DS201 markiert (0,00 % aller Datensätze). Diese betreffen 24 distinkte Personen-IDs (0,00 % aller Personen-IDs).

Größenklasse der Fusionskasse	Jahr	Geburtskalendertage
Groß	2016	25
Groß	2017	25

Tabelle 2: *Kombinationen aus Fusionskasse und Geburtskalendertag, die für ein Berichtsjahr aufgrund falsch pseudonymisierter Personen-IDs markierte Datensätze in der Satzart DS201 enthalten. Einteilung der Fusionskassen in Größenklassen anhand der Versichertenanzahl im Quartal 2018/4: bis 99.999 Versicherte (klein), 100.000 bis 499.999 Versicherte (mittel), ab 500.000 Versicherte (groß)*

Die Markierung wird an die Satzart DS219 weitergegeben (Abschnitt 2.3.4).

Dummy-Personen-ID bzw. -Versichertennummern (kassenspezifisch und/oder lebenslang)

Aufgrund von Informationslücken werden von den Krankenkassen für fehlende Versichertennummern (kassenspezifisch und/oder lebenslang) möglicherweise sogenannte Dummy-Werte verwendet, für die dann Pseudonyme an die Datenstelle des Bewertungsausschusses übermittelt werden. Diese (Dummy-)Personen-IDs weisen auffällig viele Versichertenverhältnisse oder Abrechnungsfälle auf. Wird eine solche Personen-ID identifiziert, so wird diese *für den gesamten Berichtszeitraum* und in allen ihren Versichertenverhältnissen markiert. In der GSPR 2018 sind aus dem genannten Grund 135 Datensätze der Satzart DS201 markiert (0,00 % aller Datensätze). Diese betreffen 10 distinkte Personen-IDs (0,00 % aller Personen-IDs).

Auch in den arztseitig gelieferten Abrechnungsfällen sind möglicherweise Pseudonyme für Dummy-Versichertennummern enthalten, die nur dann mit den kassenseitigen Pseudonymen für Dummy-Versichertennummern matchen¹¹, wenn identische Dummy-Werte verwendet wurden. Die meisten dieser Dummy-Personen-IDs wurden wegen der

¹¹Das Matchingverfahren der Geburtstagsstichprobe wird in Abschnitt 6.3 ausführlich erläutert. Zum Verständnis der folgenden Ausführungen zu den Auffälligkeiten soll hier nur eine kurze Definition gegeben werden: Im vorliegenden Dokument beschreibt der Begriff Matching die Zuordnung von arztseitig gelieferten Abrechnungsfällen zu den kassenseitig gelieferten Versichertenstammdaten anhand des Pseu-

nicht-eindeutigen Angaben zu Geschlecht und/oder Geburtsquartal aus der GSPR 2018 ausgeschlossen. Alle übrigen Pseudonyme für Dummy-Werte in den arztseitig übermittelten Daten haben nicht mit den kassenseitigen Versichertenpseudonymen gematcht und sind daher ebenfalls nicht in der GSPR 2018 enthalten.

Kollidierende Personen-ID

Bei der Zuordnung der arztseitig gelieferten Abrechnungsfälle zu den kassenseitig gelieferten Versichertenstammdaten kann es vorkommen, dass Abrechnungsfälle mehreren verschiedenen Personen-IDs aus den Versichertenstammdaten zugeordnet werden können. Diese Fälle gelten dann als nicht matchbar und werden nicht mit der GSPR 2018 ausgeliefert. Alle diesen Fällen zuordenbaren Personen-IDs werden *in dem betreffenden Jahr* und für alle ihre Versichertenverhältnisse markiert. Wechselt der Versicherte in diesem Jahr seine Krankenkasse, so ist die Personen-ID entsprechend auch in diesem Versichertenverhältnis markiert. Hat ein Versicherter mehrfache Versicherungswechsel, so wird dieses Vorgehen auch auf die folgenden Versichertenverhältnisse angewendet.

Insgesamt sind aus dem genannten Grund 319 Datensätze in der Satzart DS201 der GSPR 2018 markiert (0,00 % aller Datensätze). Diese betreffen 82 distinkte Personen-IDs (0,00 % aller Personen-IDs).

Datensätze nach dem Sterbequartal

Es kann vorkommen, dass auch nach dem Quartal, in dem ein Versicherter verstirbt, noch Versichertenstammdaten zu diesem Versicherten von den Krankenkassen für einzelne Quartale geliefert werden. In der GSPR 2018 sind alle Datensätze zu einem Versicherten in der Satzart DS201 markiert, die sich auf ein Quartal nach dem Sterbequartal dieses Versicherten beziehen.

Insgesamt sind aus diesem Grund 1.169 Datensätze in der Satzart DS201 der GSPR 2018 markiert (0,00 % aller Datensätze). Dies betrifft 241 distinkte Personen-IDs (0,00 % aller Personen-IDs).

Fehlende Datensätze nach Abgleich mit Daten zu Selektivverträgen

Die mit der GSP gelieferten Daten zu Selektivverträgen (gemäß Beschluss des Erweiterten Bewertungsausschuss in seiner 51., 56. und 61. Sitzung) können Versicherte beinhalten, die zwar in den Versichertenstammdaten der Geburtstagsstichprobe enthalten sind, allerdings nicht in allen Quartalen, in denen die Teilnahme an einem Selektivvertrag berichtet wird. Trifft dies für alle Quartale eines Jahres zu, in denen ein Versicherter in den Selektivvertragsdaten enthalten ist, so werden *in dem betreffenden Jahr* alle Versicher-

donyms der Versichertennummer und eines der Institutionskennzeichen der Kasse. Ziel des Matchingverfahrens ist es dabei, allen Abrechnungsfällen eine eindeutige Personen-ID der Versichertenstammdaten zuzuordnen. Diejenigen Abrechnungsfälle, denen keine Personen-ID zugeordnet werden kann, sind nicht Bestandteil der GSPR 2018.

tenverhältnisse dieses Versicherten markiert. Wechselt der Versicherte in diesem Jahr seine Krankenkasse, so ist die Personen-ID entsprechend auch in diesem Versichertenverhältnis markiert. Hat ein Versicherter mehrfache Versicherungswechsel, so wird dieses Vorgehen auch auf die folgenden Versichertenverhältnisse angewendet.

Insgesamt sind aus dem genannten Grund 1.098 Datensätze in der Satzart DS201 der GSPR 2018 markiert (0,00 % aller Datensätze). Diese betreffen 502 distinkte Personen-IDs (0,00 % aller Personen-IDs).

Versicherte mit Kostenübernahme nach §264 Abs. 2 SGB V

Versicherte, die in einem Quartal mit einer Markierung zur Kostenübernahme nach §264 Abs. 2 SGB V¹² versehen sind (sogenannte „Betreute“), werden jeweils für das gesamte Kalenderjahr und für alle Versichertenverhältnisse markiert.

Insgesamt sind aufgrund der Kennzeichnung für Versicherte mit Kostenübernahme 908.980 Datensätze in der Satzart DS201 markiert (0,45 % aller Datensätze). Diese betreffen 155.460 distinkte Personen-IDs (0,68 % aller Personen-IDs). Die genannten Zahlen sind möglicherweise höher als die Anzahl der Datensätze mit Kennzeichen zur Kostenübernahme in der Satzart DS201. Dies lässt sich damit erklären, dass für einen Versicherten alle Datensätze eines Berichtsjahres markiert wurden, wenn mindestens ein Datensatz mit Kennzeichen zur Kostenübernahme vorliegt.

Falschmatcher

Im Rahmen der Datenzusammenführung (Matching, siehe Abschnitt 6.3) kann es zu sogenannten Falschmatchern kommen. Bei diesen Versicherten besteht der Verdacht einer irrtümlichen Zuordnung des Abrechnungsfalles. Anhand von geschlechtsspezifischen GOPen und/oder Diagnosen kann in Einzelfällen auf das vermeintliche Geschlecht des Versicherten geschlossen werden. Stimmt dieses nicht mit dem gelieferten Geschlecht überein, wird angenommen, dass die Zuordnung des Abrechnungsfalles zu einem Versicherten fehlerhaft war und es sich um einen Falschmatcher handelt. Markiert werden jeweils alle Versicherten einer Geburtskalendertag-Quartal-IK-Kombination mit einem auffällig hohen Anteil an Falschmatchern.

Im Rahmen der Datenprüfung waren die Anteile der Falschmatcher stets unterhalb der festgelegten Schwellenwerte, die zu einer Markierung von Datensätzen in der Satzart DS201 führen. In der GSPR 2018 liegen daher keine aus diesem Grund markierten Datensätze vor. Daraus folgt allerdings nicht, dass für alle Versicherten das aus den dokumentierten Angaben für Diagnosen und Leistungen abgeleitete Geschlecht mit dem Geschlecht der Personen-ID übereinstimmt. Möchte man tatsächlich alle betroffenen Personen bzw. Fälle von einer Auswertung ausschließen, kann dies mit Hilfe der in Abschnitt A.2 aufgeführten GOPen und/oder Diagnosen geschehen.

¹²Siehe Fußnote 2 auf Seite 7.

Auffälliger Ausschöpfungsgrad des Leistungsbedarfs

Eine Kennzahl, die im Rahmen der Datenprüfung berechnet wird, ist der sogenannte „Ausschöpfungsgrad des Leistungsbedarfs“. Der Ausschöpfungsgrad setzt den durchschnittlichen Leistungsbedarf je Versicherten der bundesweiten Versichertenstichprobe ins Verhältnis zum durchschnittlichen Leistungsbedarf aller Versicherten (siehe dazu auch Abschnitt 2.4). Beträgt der Ausschöpfungsgrad genau 100 %, entspricht der durchschnittliche Leistungsbedarf je Versicherten in der Stichprobe genau dem durchschnittlichen Leistungsbedarf je Versicherten in der Grundgesamtheit.

Es ist nicht zu erwarten, dass der Ausschöpfungsgrad für alle Fusionskassen und in allen Quartalen 100 % beträgt. Sehr starke Abweichungen lassen allerdings vermuten, dass Datenprobleme dafür ursächlich sind. Beispielsweise kann eine fehlerhafte Zuordnung von Abrechnungsfällen zu Personen-IDs zu einem verringerten Ausschöpfungsgrad führen.

Um Datenprobleme zu identifizieren, wird der Ausschöpfungsgrad für jede KV-Quartals-Fusionskassen-Geburtskalendertags-Kombination berechnet. Die Markierung erfolgt dann auf Grundlage einer der folgenden Bedingungen:

1. Die KV-Quartals-Fusionskassen-Geburtskalendertags-Kombination umfasst mehr als 2.000 Personen-IDs und weist einen Ausschöpfungsgrad kleiner als 0,85·99 % oder größer als $(1/0,85) \cdot 99$ % auf.
2. Die KV-Quartals-Fusionskassen-Geburtskalendertags-Kombination umfasst mehr als 1.000 Personen-IDs und weist einen Ausschöpfungsgrad kleiner als 0,7·99 % oder größer als $(1/0,7) \cdot 99$ % auf.
3. Die KV-Quartals-Fusionskassen-Geburtskalendertags-Kombination umfasst mehr als 800 Personen-IDs und weist einen Ausschöpfungsgrad kleiner als 0,6·99 % oder größer als $(1/0,6) \cdot 99$ % auf.

Wird also eine KV-Quartals-Fusionskassen-Geburtskalendertags-Kombination identifiziert, für die mindestens eine dieser drei Bedingungen erfüllt ist, werden *alle* Personen-IDs, die dieser Kombination angehören, im *betreffenden Jahr* in allen ihren Versicherungsverhältnissen markiert. Wechselt ein markierter Versicherter seine Krankenkasse in diesem Jahr, so ist die Personen-ID entsprechend auch in diesem Versicherungsverhältnis markiert. Hat ein Versicherter in dem Jahr mehrfache Versicherungswechsel, so wird dieses Vorgehen auch auf die folgenden Versicherungsverhältnisse angewendet.

In der GSPR 2018 wurden keine Datensätze in der Satzart DS201 aus dem genannten Grund markiert.

2.3.3 Flag der Satzart DS202 – Ambulante Abrechnungen (KV-Fall)

Wie bei den Versichertenstammdaten (Abschnitt 2.3.2) sind auch in der Satzart DS202 auffällige Datensätze durch die Ausprägung „1“ im Feld „Flag“ gekennzeichnet. In der

GSPR 2018 werden solche Markierungen ausschließlich aus folgendem Grund vorgenommen.

Fälle mit Dummy-Versichertennummern

Abrechnungsfälle, bei denen die Vermutung nahe liegt, dass die gelieferten Versichertennummern Dummy-Nummern sind, werden markiert. Genauer gesagt, werden alle Fälle geflaggt, die zu einem Versichertennummernpseudonym gehören, das im entsprechenden Quartal entweder mehr als 200 Fälle oder mehr als 100 Fälle in mehr als 10 KVen besitzt.

In der GSPR 2018 nimmt dieses Feld für keinen Datensatz die Ausprägung „1“ an, d. h. es wurden keine Abrechnungsfälle als auffällig markiert. Nichtsdestotrotz kann der Ausschluss von gewissen Abrechnungsfällen bei Auswertungen sinnvoll sein – nämlich dann, wenn die zugehörige Person in den Versichertenstammdaten (Satzart DS201) als auffällig markiert wurde.

2.3.4 Flag der Satzart DS219 – Kostenträgerverzeichnis (KT-Stamm)

In der Satzart DS219 werden sämtliche dem GKV-Spitzenverband bekannte Institutionskennzeichen von Teilkassen aufgeführt. Für die Markierung einer Teilkasse in der Satzart DS219 werden neben den bisher beschriebenen keine zusätzlichen QS-Maßnahmen durchgeführt. Die Markierung einer Teilkasse, die über die Ausprägung des Feldes 03 realisiert ist, dient ausschließlich Dokumentationszwecken.

Einer Teilkasse wird im Feld 03 die Ausprägung „1“ zugewiesen, wenn Versicherte dieser Teilkasse aufgrund der Auffälligkeit „falsch pseudonymisierte Personen-ID“ markiert sind.

Die Feldausprägung „0“ entspricht keiner Markierung im eigentlichen Sinn, sondern zeigt an, dass für die Versicherten dieser Teilkasse das Problem „falsch pseudonymisierte Personen-ID“ nicht vorliegt. Es ist keinesfalls der Rückschluss zu ziehen, dass die Versichertenstammdatensätze derjenigen Teilkassen, für die die Feldausprägung gleich „0“ ist, in der Satzart DS201 nicht markiert wären, da die Möglichkeit besteht, dass Versichertenstammdatensätze in der Satzart DS201 aus anderen Gründen markiert sind.

In der GSPR 2018 wurden insgesamt vier Teilkassen mit der Flagausprägung „1“ markiert. Diese gehören zu einer Fusionskasse

2.3.5 Flag der Satzart DS220 – Kostenträgerhistorie

Die Satzart DS220 gibt Auskunft über die Fusion von Kostenträgern. Hier kann nachvollzogen werden, welche Rechtsvorgänger zu den aktuell gültigen Fusionskassen vorliegen. Für die Markierung einer Fusionskasse und all ihrer Rechtsvorgänger in der Satzart DS220 werden, neben den bisher beschriebenen, keine zusätzlichen QS-Maßnahmen durchge-

führt. Die Markierung einer Fusionskasse (bzw. Teilkasse) dient Dokumentationszwecken. Das Institutionskennzeichen einer Fusionskasse wird im Feld 13 mit der Ausprägung „1“ markiert, wenn für Versicherte in einem beliebigen Jahr das Problem „auffälliger Ausschöpfungsgrad des Leistungsbedarfs“ vorliegt (siehe Abschnitt 2.3.5).

Die Feldausprägung „0“ im Feld 13 der Satzart DS220 entspricht keiner Markierung im eigentlichen Sinn, sondern zeigt an, dass für Versicherte dieser Fusionskasse und ihrer Rechtsvorgänger das Problem „auffälliger Ausschöpfungsgrad des Leistungsbedarfs“ nicht besteht. Es ist keinesfalls der Rückschluss zu ziehen, dass die Versicherten der Fusionskassen, für die die Feldausprägung gleich „0“ ist, in der Satzart DS201 nicht markiert wären. Es lassen sich Fusionskassen finden, deren Versicherten teilweise aufgrund anderer Auffälligkeiten in der Satzart DS201 markiert sind, ohne dass die zugehörige Fusionskasse in der Satzart DS220 die Flagausprägung „1“ aufweist.

In der GSPR 2018 sind keine Fusionskassen in der Satzart DS220 markiert.

2.4 Ausgewählte Ergebnisse der Datenprüfung

Im Folgenden werden ausgewählte Ergebnisse der Datenprüfung der GSPR 2018 zugrunde liegenden bundesweiten Versichertenstichprobe dargestellt. Weitere Informationen zu dieser Stichprobe sind in Kapitel 6 zu finden. Hier werden lediglich Ergebnisse zu den Kennzahlen Lieferquote der Versichertenstammdaten, Inanspruchnahmequote sowie Ausschöpfungsgrad des Leistungsbedarfs berichtet. Dabei ist, wie bei der ausgelieferten GSPR 2018, der Fusionsstand der Krankenkassen zum 02.04.2020 berücksichtigt. Zu diesem Zeitpunkt gab es 105 Fusionskassen.

Lieferquote der Versichertenstammdaten: Um die Lieferquote zu berechnen, werden pro Quartal die in der Stichprobe enthaltenen Versicherten der jeweiligen Wohnort-KV gezählt, die in diesem Quartal nicht wegen Auffälligkeiten markiert wurden. Die so ermittelte Anzahl von Versicherten wird durch die von den Krankenkassen in den ANZVER-Daten¹³ berichtete Anzahl der Versicherten geteilt. Wie in Abschnitt 2.1 beschrieben, umfasst die GSPR in jedem Berichtsjahr Daten von Versicherten mit sieben unterschiedlichen Geburtskalendertagen und damit aus Versicherten mit insgesamt 84 Geburtstagen (7 Tage pro Monat). Unter der Annahme, dass jeder dieser Geburtskalendertage die gleiche relative Häufigkeit besitzt, liegt der Zielwert für die Lieferquoten demnach bei ca. $84/365,25 \approx 23,00\%$. In der Realität liegt jedoch eine Verzerrung der Häufigkeiten von Geburtskalendertagen vor (insbesondere der Geburtskalendertag 1 sowie die durch 5 teilbaren Geburtskalendertage sind in der Grundgesamtheit überrepräsentiert). Im Rahmen der Qualitätssicherung wurden auf Grundlage der GSP deshalb sogenannte *Hochrechnungsfaktoren* für Geburtskalendertage ermittelt, die für jeden Geburtskalendertag den Anteil an der Grundgesamtheit schätzen und hier als Zielquoten dienen. Die berechneten Zielwerte sind sowohl KV- als auch fusionskassenabhängig und liegen zwi-

¹³Satzart ANZVER87a gemäß Beschluss des Bewertungsausschusses in seiner 277. und 369. Sitzung.

schen 22,38 und 23,08 %. Abgesehen von der Unsicherheit in den geschätzten Zielquoten selbst, können Lieferlücken einzelner Krankenkassen oder eine häufige Markierung wegen Auffälligkeiten in den Versichertenstammdaten zu Abweichungen zwischen den Liefer- und Zielquoten führen. Nicht auszuschließen ist allerdings auch, dass die Werte durch den Unterschied zwischen tagesgenauer Zählung in der Versichertenstichprobe und Stichtagszählung in den als Referenz herangezogenen ANZVER-Daten oder durch den Einfluss bestimmter Versichertengruppen (wie zum Beispiel Wohnausländern) verzerrt werden. Die Lieferquoten geben damit einen Eindruck vom Stichprobenumfang bezogen auf die Grundgesamtheit, wenn – wie empfohlen – auf die Verwendung der mit „1“ (Auffälligkeiten) markierten Datensätze der Satzart DS201 in der GSPR 2018 verzichtet wird.

Tabelle 3 gibt die Lieferquoten der Versichertenstammdaten für die einzelnen KV-Bezirke in den Quartalen 2016/1 bis 2018/4 an. Die Tabellen 4 und 5 liefern Informationen zu den Lieferquoten der Fusionskassen für den Berichtszeitraum. Aufgrund der Vielzahl der Fusionskassen wurden diese in fünf Gruppen eingeteilt. Die Gruppeneinteilung orientiert sich dabei an der Höhe der Lieferquoten. Für jede Gruppe sind für jedes Quartal die Anzahl der Fusionskassen, die eine solche Lieferquote aufweisen bzw. die Anzahl der Versicherten, die bei diesen Fusionskassen versichert sind, angegeben.

Wohnort-KV	2016/1	2016/2	2016/3	2016/4	2017/1	2017/2	2017/3	2017/4	2018/1	2018/2	2018/3	2018/4
Baden-Württt.	22,99 %	22,97 %	22,99 %	22,96 %	22,98 %	22,92 %	22,94 %	22,96 %	22,88 %	22,85 %	22,88 %	22,83 %
Bayerns	23,09 %	23,07 %	23,07 %	23,06 %	23,03 %	22,97 %	22,99 %	22,99 %	22,95 %	22,92 %	22,86 %	22,94 %
Berlin	22,89 %	22,86 %	22,88 %	22,85 %	22,85 %	22,75 %	22,74 %	22,75 %	22,75 %	22,74 %	22,82 %	22,79 %
Brandenburg	23,12 %	23,08 %	23,07 %	23,05 %	23,05 %	22,99 %	22,94 %	22,99 %	22,99 %	22,97 %	22,86 %	22,90 %
Bremen	22,58 %	22,45 %	22,48 %	22,47 %	22,61 %	22,52 %	22,54 %	22,55 %	22,64 %	22,60 %	22,66 %	22,66 %
Hamburg	22,60 %	22,52 %	22,55 %	22,57 %	22,70 %	22,67 %	22,69 %	22,71 %	22,74 %	22,72 %	22,82 %	22,78 %
Hessen	22,88 %	22,86 %	22,87 %	22,85 %	22,87 %	22,80 %	22,78 %	22,78 %	22,76 %	22,72 %	22,74 %	22,73 %
Mecklenburg-Vorp.	23,17 %	23,13 %	23,12 %	23,12 %	23,14 %	23,09 %	23,10 %	23,11 %	23,09 %	23,06 %	23,11 %	23,11 %
Niedersachsen	22,96 %	22,94 %	22,96 %	22,94 %	22,96 %	22,88 %	22,88 %	22,89 %	22,88 %	22,84 %	22,87 %	22,85 %
Nordrhein	22,92 %	22,88 %	22,87 %	22,84 %	22,90 %	22,81 %	22,80 %	22,79 %	22,77 %	22,76 %	22,78 %	22,76 %
Rheinland-Pfalz	23,07 %	23,03 %	23,05 %	23,02 %	23,04 %	22,98 %	22,97 %	22,97 %	22,93 %	22,89 %	22,90 %	22,92 %
Saarland	22,99 %	22,96 %	22,93 %	22,89 %	22,92 %	22,88 %	22,93 %	22,94 %	22,89 %	22,83 %	22,87 %	22,92 %
Sachsen	23,12 %	23,06 %	23,11 %	23,10 %	23,15 %	23,09 %	23,09 %	23,09 %	23,07 %	23,01 %	23,04 %	23,03 %
Sachsen-Anhalt	23,08 %	23,04 %	23,06 %	23,05 %	23,11 %	23,06 %	23,06 %	23,07 %	23,07 %	23,09 %	23,05 %	23,04 %
Schleswig-Holstein	23,07 %	23,01 %	23,01 %	22,96 %	23,03 %	22,93 %	22,91 %	22,91 %	22,98 %	22,96 %	22,99 %	22,95 %
Thüringen	23,02 %	22,96 %	23,03 %	23,01 %	23,00 %	22,98 %	22,96 %	22,97 %	22,99 %	22,94 %	22,99 %	22,98 %
Westfalen-Lippe	22,88 %	22,85 %	22,86 %	22,83 %	22,87 %	22,80 %	22,80 %	22,80 %	22,78 %	22,77 %	22,79 %	22,76 %

Tabelle 3: Lieferquoten der Versichertenstammdaten nach KV-Bezirk (ohne als auffällig markierte Versicherte). Wohnausländer sind in die Auswertung nicht mit einbezogen. Die Zellenfarbe deutet den Grad der Unter- (rot) bzw. Überschreitung (blau) der Zielquote an.

Quartal	Zielquotenbereich	Zielquote um mehr als 1,0 Prozentpunkte unterschritten	Zielquote zwischen 0,2 und 1,0 Prozentpunkte unterschritten	Abweichung von Zielquote kleiner als 0,2 Prozentpunkte	Zielquote zwischen 0,2 und 1,0 Prozentpunkte überschritten	Zielquote um mehr als 1,0 Prozentpunkte überschritten	Anzahl Fusionskassen im Quartal insgesamt
2016/1	22,62 %–23,07 %	1	22	55	27	0	105
2016/2	22,62 %–23,07 %	1	20	59	25	0	105
2016/3	22,62 %–23,07 %	1	14	63	27	0	105
2016/4	22,62 %–23,07 %	2	14	69	20	0	105
2017/1	22,70 %–23,08 %	1	18	62	24	0	105
2017/2	22,70 %–23,08 %	1	24	58	22	0	105
2017/3	22,70 %–23,08 %	1	22	59	23	0	105
2017/4	22,70 %–23,08 %	1	22	61	21	0	105
2018/1	22,62 %–23,02 %	1	20	59	24	1	105
2018/2	22,62 %–23,02 %	1	23	59	22	0	105
2018/3	22,62 %–23,02 %	0	21	61	23	0	105
2018/4	22,62 %–23,02 %	0	22	60	23	0	105

Tabelle 4: Anzahl der Fusionskassen nach Lieferquote (ohne als auffällig markierte Versicherte). Wohnausländer sind in die Auswertung einbezogen. Berücksichtigt ist der Fusionsstand der Krankenkassen am 02.04.2020.

Quartal	Zielquotenbereich	Zielquote um mehr als 1,0 Prozentpunkte unterschritten	Zielquote zwischen 0,2 und 1,0 Prozentpunkte unterschritten	Abweichung von Zielquote kleiner als 0,2 Prozentpunkte	Zielquote zwischen 0,2 und 1,0 Prozentpunkte überschritten	Zielquote um mehr als 1,0 Prozentpunkte überschritten	Anzahl Versicherte im Quartal insgesamt
2016/1	22,62 %–23,07 %	53.786	825.490	12.032.061	3.473.259	0	16.384.597
2016/2	22,62 %–23,07 %	53.852	1.378.674	14.000.787	997.263	0	16.430.576
2016/3	22,62 %–23,07 %	53.985	1.277.639	14.205.472	934.687	0	16.471.785
2016/4	22,62 %–23,07 %	56.170	1.914.754	13.733.162	822.788	0	16.526.873
2017/1	22,70 %–23,08 %	54.853	1.700.134	13.863.810	969.361	0	16.588.158
2017/2	22,70 %–23,08 %	55.008	4.243.528	11.480.420	835.271	0	16.614.227
2017/3	22,70 %–23,08 %	55.298	3.941.059	11.694.035	950.457	0	16.640.849
2017/4	22,70 %–23,08 %	56.179	2.984.211	12.895.918	754.784	0	16.691.092
2018/1	22,62 %–23,02 %	442	2.673.415	13.236.445	775.706	1.680	16.687.688
2018/2	22,62 %–23,02 %	442	3.533.841	12.536.558	637.735	0	16.708.576
2018/3	22,62 %–23,02 %	0	3.730.748	12.179.990	809.155	0	16.719.893
2018/4	22,62 %–23,02 %	0	4.090.378	11.910.670	752.855	0	16.753.903

Tabella 5: Anzahl der Versicherten nach Lieferquote der Fusionskasse, bei der sie versichert sind (ohne als auffällig markierte Versicherte). Wohnausländer sind in die Auswertung einbezogen. Berücksichtigt ist der Fusionsstand der Krankenkassen am 02.04.2020.

Die angegebenen Lieferquoten, die die Anzahl der gelieferten Versicherten ohne Auffälligkeitsmarkierung zur Anzahl der Versicherten der jeweiligen Fusionskasse ins Verhältnis setzen, können auch als Netto-Lieferquote interpretiert werden. In Ergänzung dazu werden in Tabelle 6 und in Tabelle 7 auch diejenigen Lieferquoten der Versichertenstammdaten berichtet, die die Anzahl der ursprünglich von den jeweiligen Fusionskassen gelieferten Versicherten (also alle Versicherten unabhängig von markierten Auffälligkeiten) ins Verhältnis zur Anzahl der Versicherten der jeweiligen Fusionskasse setzen. Diese wird als „ursprüngliche Lieferquote“ bezeichnet.

Quartal	Zielquotenbereich	Zielquote um mehr als 1,0 Prozentpunkte unterschritten	Zielquote zwischen 0,2 und 1,0 Prozentpunkte unterschritten	Abweichung von Zielquote kleiner als 0,2 Prozentpunkte	Zielquote zwischen 0,2 und 1,0 Prozentpunkte überschritten	Zielquote um mehr als 1,0 Prozentpunkte überschritten	Anzahl Fusionskassen im Quartal insgesamt
2016/1	22,62 %–23,07 %	0	20	52	33	0	105
2016/2	22,62 %–23,07 %	0	18	59	28	0	105
2016/3	22,62 %–23,07 %	0	12	60	33	0	105
2016/4	22,62 %–23,07 %	1	11	67	26	0	105
2017/1	22,70 %–23,08 %	0	15	63	27	0	105
2017/2	22,70 %–23,08 %	0	17	64	24	0	105
2017/3	22,70 %–23,08 %	0	16	63	26	0	105
2017/4	22,70 %–23,08 %	0	16	66	23	0	105
2018/1	22,62 %–23,02 %	1	13	64	26	1	105
2018/2	22,62 %–23,02 %	1	15	67	22	0	105
2018/3	22,62 %–23,02 %	0	13	64	28	0	105
2018/4	22,62 %–23,02 %	0	13	68	24	0	105

Tabelle 6: Anzahl der Fusionskassen nach ursprünglicher Lieferquote (inklusive der als auffällig markierten Versicherten). Wohnausländer sind in die Auswertung einbezogen. Berücksichtigt ist der Fusionsstand der Krankenkassen am 02.04.2020.

Quartal	Zielquotenbereich	Zielquote um mehr als 1,0 Prozentpunkte unterschritten	Zielquote zwischen 0,2 und 1,0 Prozentpunkte unterschritten	Abweichung von Zielquote kleiner als 0,2 Prozentpunkte	Zielquote zwischen 0,2 und 1,0 Prozentpunkte überschritten	Zielquote um mehr als 1,0 Prozentpunkte überschritten	Anzahl Versicherte im Quartal insgesamt
2016/1	22,62 %–23,07 %	0	272.463	11.522.529	4.641.747	0	16.436.738
2016/2	22,62 %–23,07 %	0	214.690	14.169.693	2.108.073	0	16.492.456
2016/3	22,62 %–23,07 %	0	108.464	12.550.265	3.883.488	0	16.542.217
2016/4	22,62 %–23,07 %	1.965	106.578	14.392.328	2.102.910	0	16.603.780
2017/1	22,70 %–23,08 %	0	146.721	15.237.193	1.282.521	0	16.666.435
2017/2	22,70 %–23,08 %	0	369.722	15.238.874	1.092.613	0	16.701.209
2017/3	22,70 %–23,08 %	0	151.260	15.107.706	1.473.074	0	16.732.040
2017/4	22,70 %–23,08 %	0	99.806	15.714.272	972.209	0	16.786.288
2018/1	22,62 %–23,02 %	442	91.284	15.582.594	1.087.078	1.681	16.763.079
2018/2	22,62 %–23,02 %	442	109.495	16.037.825	638.104	0	16.785.866
2018/3	22,62 %–23,02 %	0	90.704	15.466.496	1.240.637	0	16.797.838
2018/4	22,62 %–23,02 %	0	90.421	15.929.896	813.562	0	16.833.880

Tabelle 7: Anzahl der Versicherten nach ursprünglicher Lieferquote der Fusionskasse, bei der sie versichert sind (inklusive der als auffällig markierten Versicherten). Wohnausländer sind in die Auswertung einbezogen.

Inanspruchnahmequote: Weiterhin wird die sogenannte Inanspruchnahmequote berichtet. Zwar wird allen mit der GSPR 2018 ausgelieferten Fällen eine Personen-ID zugeordnet, aber nicht jedem in den Versichertenstammdaten enthaltenen Versicherten ist in jedem Quartal auch ein Abrechnungsfall zugeordnet.

Die Inanspruchnahmequoten in Tabelle 8 geben pro KV-Bezirk und Quartal den Anteil der Versicherten mit mindestens einem Abrechnungsfall an allen Stichprobenversicherten an. Wiederum werden zur Berechnung der Quote nur Versicherte und Abrechnungsfälle einbezogen, die keine Markierung wegen Auffälligkeiten tragen.

Wohnort-KV	2016/1	2016/2	2016/3	2016/4	2017/1	2017/2	2017/3	2017/4	2018/1	2018/2	2018/3	2018/4
Baden-Württt.	70,15 %	69,37 %	67,41 %	69,33 %	70,57 %	67,47 %	66,98 %	68,30 %	70,76 %	67,76 %	66,46 %	68,50 %
Bayerns	74,35 %	73,48 %	71,79 %	74,08 %	75,24 %	72,08 %	71,65 %	73,23 %	75,73 %	72,55 %	71,75 %	73,67 %
Berlin	75,36 %	73,73 %	71,79 %	73,93 %	75,17 %	72,08 %	71,62 %	72,90 %	75,48 %	72,21 %	70,90 %	73,18 %
Brandenburg	79,13 %	77,56 %	75,90 %	78,31 %	79,58 %	76,52 %	76,01 %	77,74 %	80,36 %	76,81 %	75,75 %	78,31 %
Bremen	76,34 %	75,06 %	73,30 %	75,31 %	76,76 %	73,53 %	73,19 %	74,36 %	76,84 %	73,79 %	72,13 %	74,23 %
Hamburg	75,28 %	73,96 %	72,06 %	74,02 %	75,21 %	72,79 %	71,71 %	72,73 %	75,53 %	72,71 %	71,27 %	73,03 %
Hessen	77,14 %	75,53 %	73,96 %	75,70 %	77,28 %	73,90 %	73,61 %	74,58 %	77,38 %	74,08 %	73,11 %	74,85 %
Mecklenburg-Vorp.	79,97 %	78,56 %	76,81 %	79,26 %	80,32 %	77,81 %	76,95 %	78,71 %	80,94 %	77,93 %	76,76 %	79,36 %
Niedersachsen	78,06 %	76,71 %	75,18 %	77,10 %	78,88 %	75,41 %	75,25 %	76,45 %	79,22 %	75,98 %	74,62 %	76,60 %
Nordrhein	76,87 %	75,67 %	74,07 %	75,75 %	77,26 %	74,26 %	73,84 %	75,02 %	77,96 %	74,57 %	73,58 %	75,16 %
Rheinland-Pfalz	78,02 %	76,99 %	75,38 %	76,68 %	78,68 %	75,35 %	75,04 %	75,89 %	79,22 %	75,70 %	74,77 %	76,20 %
Saarland	78,01 %	77,27 %	76,13 %	77,18 %	79,18 %	76,11 %	75,88 %	76,45 %	79,95 %	76,59 %	75,86 %	77,41 %
Sachsen	79,28 %	78,32 %	76,60 %	78,62 %	79,75 %	77,08 %	76,61 %	77,94 %	80,22 %	77,21 %	75,78 %	78,28 %
Sachsen-Anhalt	79,88 %	78,76 %	77,54 %	79,66 %	80,69 %	78,09 %	77,91 %	79,44 %	81,53 %	78,34 %	77,35 %	79,86 %
Schleswig-Holstein	77,54 %	75,93 %	74,56 %	76,45 %	77,98 %	74,85 %	74,44 %	75,46 %	78,73 %	74,67 %	74,00 %	75,58 %
Thüringen	80,00 %	78,57 %	77,10 %	79,22 %	80,70 %	77,53 %	77,23 %	78,49 %	81,18 %	78,19 %	76,96 %	79,32 %
Westfalen-Lippe	77,52 %	76,11 %	74,48 %	76,30 %	77,97 %	74,63 %	74,28 %	75,31 %	78,50 %	74,84 %	73,91 %	75,57 %

Tabelle 8: Inanspruchnahmequote nach KV-Bezirk (ohne als auffällig markierte Versicherte). Wohnausländer sind in der Auswertung nicht berücksichtigt.

Ausschöpfungsgrad des Leistungsbedarfs: Der Ausschöpfungsgrad setzt den durchschnittlichen Leistungsbedarf je Versicherten in der Stichprobe ins Verhältnis zum durchschnittlichen Leistungsbedarf aller Versicherten. Datengrundlage für den kollektivvertraglichen Leistungsbedarf aller Versicherten des Berichtszeitraums 2016–2018 sind die von der Kassenärztlichen Bundesvereinigung gelieferten ARZTRG-Daten (Satzart ARZTRG87aKA). Zur Ermittlung der Anzahl aller Versicherten werden wiederum die ANZVER-Daten verwendet. Ein Ausschöpfungsgrad von kleiner (größer) als 100 % kann daher ein Anzeichen dafür sein, dass der durchschnittliche Leistungsbedarf je Versicherten in der Stichprobe kleiner (größer) ist als der durchschnittliche Leistungsbedarf aller Versicherten. Der Leistungsbedarf eines Versicherten berücksichtigt dabei die regionale Euro-Gebührenordnung. In die Berechnung des Ausschöpfungsgrades werden nur Versicherte einbezogen, die keine Markierung wegen Auffälligkeiten in den Versichertenstammdaten tragen.

Tabelle 9 gibt die Ausschöpfungsgrade der einzelnen KV-Bezirke in den Quartalen 2016/1 bis 2018/4 an. In den Tabellen 10 und 11 sind Informationen zu den Ausschöpfungsgraden der Fusionskassen für den Berichtszeitraum gesammelt. Aufgrund der Vielzahl der Fusionskassen wurden diese in vier Gruppen eingeteilt. Die Gruppeneinteilung orientiert sich dabei an der Höhe der Ausschöpfungsgrade. Für jede Gruppe sind für jedes Quartal die Anzahl der Fusionskassen, die einen solchen Ausschöpfungsgrad aufweisen, bzw. die Anzahl der Versicherten, die bei diesen Fusionskassen versichert sind, angegeben.

Wohnort-KV	2016/1	2016/2	2016/3	2016/4	2017/1	2017/2	2017/3	2017/4	2018/1	2018/2	2018/3	2018/4
Baden-Württt.	98,02 %	98,25 %	98,47 %	98,16 %	98,25 %	98,43 %	98,43 %	98,21 %	98,71 %	98,66 %	98,64 %	98,68 %
Bayerns	97,98 %	98,00 %	98,14 %	98,24 %	98,27 %	98,44 %	98,43 %	98,47 %	98,65 %	98,74 %	99,10 %	98,71 %
Berlin	97,73 %	97,91 %	97,96 %	97,89 %	98,01 %	98,35 %	98,64 %	98,54 %	98,13 %	98,09 %	97,86 %	97,98 %
Brandenburg	98,82 %	99,12 %	99,13 %	99,10 %	98,96 %	99,07 %	98,87 %	99,21 %	98,60 %	98,85 %	99,24 %	99,42 %
Bremen	97,85 %	99,32 %	98,54 %	98,63 %	97,87 %	98,49 %	98,52 %	98,30 %	97,82 %	98,32 %	98,91 %	98,58 %
Hamburg	98,02 %	98,26 %	97,72 %	97,96 %	97,87 %	97,56 %	97,60 %	97,82 %	97,70 %	97,44 %	97,27 %	96,91 %
Hessen	98,10 %	98,28 %	98,28 %	98,32 %	98,26 %	98,74 %	98,53 %	98,68 %	98,74 %	98,82 %	98,90 %	98,78 %
Mecklenburg-Vorp.	98,73 %	98,51 %	98,83 %	98,47 %	98,61 %	98,88 %	98,89 %	98,95 %	98,86 %	99,25 %	99,06 %	98,94 %
Niedersachsen	98,62 %	98,59 %	98,75 %	98,84 %	98,90 %	98,94 %	98,84 %	98,91 %	98,57 %	98,75 %	98,70 %	98,79 %
Nordrhein	97,70 %	98,04 %	98,33 %	98,32 %	97,47 %	98,46 %	97,87 %	98,35 %	98,81 %	98,83 %	98,75 %	98,42 %
Rheinland-Pfalz	98,05 %	98,46 %	98,45 %	98,32 %	98,65 %	98,93 %	98,98 %	99,20 %	99,23 %	99,47 %	99,50 %	99,55 %
Saarland	97,48 %	98,23 %	98,37 %	98,16 %	97,84 %	97,98 %	98,41 %	98,53 %	98,32 %	98,58 %	98,55 %	98,33 %
Sachsen	98,75 %	99,28 %	98,81 %	98,87 %	99,08 %	99,27 %	99,12 %	99,06 %	99,14 %	99,29 %	99,16 %	99,35 %
Sachsen-Anhalt	99,13 %	99,42 %	99,35 %	99,25 %	99,04 %	99,70 %	99,16 %	99,41 %	99,11 %	99,29 %	99,08 %	99,28 %
Schleswig-Holstein	98,66 %	98,83 %	98,82 %	98,78 %	98,42 %	99,02 %	99,32 %	99,08 %	98,90 %	98,85 %	98,45 %	98,65 %
Thüringen	99,45 %	99,78 %	99,51 %	99,90 %	99,54 %	99,93 %	99,52 %	99,52 %	99,56 %	99,69 %	99,41 %	99,34 %
Westfalen-Lippe	98,30 %	98,40 %	98,52 %	98,60 %	98,45 %	98,59 %	98,81 %	98,73 %	98,72 %	98,83 %	98,94 %	98,97 %

Tabelle 9: Ausschöpfungsgrad des Leistungsbedarfs nach KV-Bezirk (ohne als auffällig markierte Versicherte). Wohnausländer sind in der Auswertung nicht berücksichtigt.

Quartal	Ausschöpfungsgrad kleiner als 94 %	Ausschöpfungsgrad zwischen 94 % und 98 %	Ausschöpfungsgrad zwischen 98 % und 102 %	Ausschöpfungsgrad größer als 102 %	Gesamtanzahl Fusionskassen
2016/1	2	30	62	11	105
2016/2	3	30	65	7	105
2016/3	2	22	73	8	105
2016/4	3	31	63	8	105
2017/1	3	31	65	6	105
2017/2	1	26	69	9	105
2017/3	0	27	65	13	105
2017/4	2	21	76	6	105
2018/1	0	20	74	11	105
2018/2	0	22	75	8	105
2018/3	2	14	78	11	105
2018/4	0	15	79	11	105

Tabelle 10: Anzahl der Fusionskassen nach Ausschöpfungsgrad des Leistungsbedarfs (ohne als auffällig markierte Versicherte). Wohnausländer sind in die Auswertung einbezogen. Berücksichtigt ist der Fusionsstand der Krankenkassen am 02.04.2020.

Quartal	Ausschöpfungsgrad kleiner als 94 %	Ausschöpfungsgrad zwischen 94 % und 98 %	Ausschöpfungsgrad zwischen 98 % und 102 %	Ausschöpfungsgrad größer als 102 %	Gesamtanzahl Versicherten im Quartal
2016/1	3.095	5.112.715	11.176.849	91.940	16.384.599
2016/2	2.682	3.532.230	12.823.967	71.699	16.430.578
2016/3	1.155	3.410.115	12.985.204	75.314	16.471.788
2016/4	2.695	3.000.919	13.448.919	74.340	16.526.873
2017/1	7.345	3.816.673	12.751.268	12.873	16.588.159
2017/2	1.013	1.770.851	14.817.033	25.323	16.614.220
2017/3	0	2.505.431	14.100.243	35.174	16.640.848
2017/4	2.397	1.431.413	15.246.607	10.673	16.691.090
2018/1	0	967.464	15.683.866	36.355	16.687.685
2018/2	0	352.847	16.303.768	51.958	16.708.573
2018/3	1.427	1.468.026	15.188.594	61.843	16.719.890
2018/4	0	1.204.423	15.502.583	46.890	16.753.896

Tabelle 11: Anzahl der Versicherten nach Ausschöpfungsgrad des Leistungsbedarfs der Fusionskasse, bei der sie versichert sind (ohne als auffällig markierte Versicherte). Wohnausländer sind in die Auswertung einbezogen.

3 Einschränkungen der Nutzung

3.1 Ursachen für eingeschränkte Nutzungsmöglichkeiten

Die Nutzung des vorliegenden Datenkörpers zur Auswertung bestimmter Fragestellungen unterliegt einigen Einschränkungen. Ursächlich hierfür sind Probleme in den ursprünglich gelieferten Daten, Probleme bei der Zusammenführung kassen- und arztseitiger Datenlieferungen, die Erfassungsgrenzen der Abrechnungsdaten sowie die Eigenschaften der Stichprobe.

Soweit möglich, sind Problemlösungen gefunden oder – wenn die Probleme zwar identifizierbar, aber nicht lösbar waren – die problembehafteten Datensätze markiert worden (siehe Abschnitt 2.3). **Es wird empfohlen, die markierten Datensätze (Ausprägung „1“ des Feldes 11 in der Satzart DS201) nicht mit in Auswertungen einzubeziehen.**

Im Folgenden werden die vier genannten Probleme ausgeführt. In den Abschnitten 3.2 und 3.3 werden dann Hinweise auf Einschränkungen bei Auswertungen zu ausgewählten Fragestellungen gegeben.

Probleme in den ursprünglich gelieferten Daten

Fehlerhafte Datenlieferungen von Teilkassen können sich auf alle KV-Bezirke auswirken. Beispielsweise kann die fehlende Lieferung von Versicherten einer Teilkasse zu Lieferquoten führen, die deutlich unter den in Abschnitt 2.4 angegebenen Zielquoten der KV-Bezirke liegen. Dieser Effekt ist besonders stark, wenn es sich um Teilkassen von Fusionskassen mit einem hohen Anteil von Versicherten innerhalb einzelner KV-Bezirke handelt.

Ein weiteres – wenn auch nicht so bedeutsames – Problem ist die unter Umständen falsche Pseudonymisierung der arzt- und kassenseitig gelieferten Versichertennummern.

Probleme bei der Datenzusammenführung

Mit dem zur Datenzusammenführung verwendeten Matchingverfahren können sehr viele, aber nicht alle Abrechnungsfälle sicher den Stichprobenversicherten zugeordnet werden (siehe Abschnitt 6.3). Einschränkungen in den Nutzungsmöglichkeiten ergeben sich z. B. dann, wenn Abrechnungsfälle von Versicherten oder die Stammdaten von Versicherten bestimmter Teilkassen mit hohem Versicherungsanteil in einzelnen KV-Bezirken als problembehaftet markiert werden müssen und in der Konsequenz aus der Auswertung ausgeschlossen werden sollten. Gründe für die Markierung von Personen sind nicht-eindeutige Personen-IDs, falsch pseudonymisierte Personen-IDs, die Verwendung einer Dummy-Versichertennummer, kollidierende Personen-IDs, Datensätze nach dem Sterbequartal, auffälliger Ausschöpfungsgrad des Leistungsbedarfs sowie fehlende Datensätze nach Abgleich mit Daten zu Selektivverträgen (siehe Abschnitt 2.3.2).

Erfassungsgrenzen der Abrechnungsdaten

Insgesamt stehen Abrechnungsdaten nur aus vertragsärztlicher Behandlung von Versicherten in der gesetzlichen Krankenversicherung zur Verfügung. Die Abrechnungsdaten aus vertragsärztlicher Behandlung entstammen zudem ausschließlich Abrechnungen, die über die zuständigen Kassenärztlichen Vereinigungen erfolgten. In den bereitgestellten Daten können somit keine Angaben aus knappschaftsärztlicher Versorgung nach § 167 SGB V enthalten sein. Auch Abrechnungsfälle mit Abrechnung gegenüber dem Patienten und anschließender Kostenerstattung durch seine Krankenkasse nach §§ 13, 14 SGB V sind nicht enthalten. Für die selektivvertragliche Versorgung nach §§ 63, 73b, 140a SGB V oder 73c SGB V a. F. sind nur Angaben aus KV-Abrechnungen eingeschlossen. Dadurch können bei bestimmten Versicherten größere Daten- bzw. Informationslücken vorhanden sein.

Stichprobeneigenschaften

Die Geburtstagsstichprobe ist eine Pseudozufallsstichprobe und sollte damit keine großen Verzerrungen aufweisen. Allerdings können Lücken in der Stichprobe dazu führen, dass die Stichprobe unter Umständen für einzelne KV-Bezirke nicht repräsentativ ist. Lü-

cken entstehen z. B. aufgrund nicht-liefernder Krankenkassen bzw. durch Daten, die zwar geliefert wurden, sachlich aber nicht korrekt sind. Diese wurden entsprechend markiert, d. h. bei Nicht-Verwendung dieser Datensätze wird weiteres Stichprobenmaterial von einer möglichen Auswertung ausgeschlossen.

Da die Stichprobenversicherten in der ursprünglich erhobenen Geburtstagsstichprobe nur einen Anteil an der Grundgesamtheit ausmachen und dieser Anteil in der GSPR 2018 durch Ausschlüsse nochmals weiter reduziert wird, ist zu beachten, dass nicht alle Praxen in der Stichprobe enthalten sind und für die in der Stichprobe enthaltenen Betriebsstätten (Praxen) die große Mehrzahl der Abrechnungsfälle in der GSPR 2018 nicht enthalten ist. Betriebsstättenbezogene Auswertungen weisen somit nennenswerte Stichprobenfehler auf; Praxen mit nennenswerten Anteilen von Fremdfällen (Patienten aus anderen KV-Bezirken) können in der nach Wohnort von Versicherten regional aufgeteilten GSPR 2018 grob unvollständig repräsentiert und eventuell sogar in ihrer Patientenstruktur verzerrt sein.

3.2 Intra- und interregionale Vergleiche

Wenn in KV-Bezirken Teilkassen mit ihren entsprechenden Versicherten fehlen bzw. diese als problembehaftet markiert sind, wird das Versorgungsgeschehen in der Regel unvollständig abgebildet.

Nicht alle Kassen bzw. Kassenarten verursachen diese Lückenhaftigkeit gleichermaßen. Auch kann nicht ausgeschlossen werden, dass die Lücken in einigen Teilkassen systematischer Natur sind, dass also das Fehlen von Versicherten bzw. die zusätzliche (notwendige) Herausnahme von markierten Versicherten die Repräsentativität der Stichprobe beeinflussen. Jegliche Schlussfolgerungen aus kassen- oder kassenartübergreifenden Vergleichen sind daher vorsichtig zu interpretieren. Insbesondere ist die Bemessung von Anteilen der einzelnen Kassen oder auch von Kassenarten weder anhand der Anzahl der Versicherten noch anhand des Leistungsbedarfs fehlerfrei möglich.

Auch der Vergleich des Versorgungsgeschehens zwischen Regionen (KV-Bezirken) bedarf einer vorsichtigen Interpretation. Wenn gerade die Versicherten einer Fusionskasse, die in einem KV-Bezirk einen höheren Anteil an Versicherten als in anderen KV-Bezirken aufweist, aus der Auswertung ausgeschlossen werden müssen, kann der Vergleich von Kennzahlen zwischen KV-Bezirken zu falschen Schlussfolgerungen führen. Auch ist zu beachten, dass Selektivvertragsteilnehmer, deren Abrechnungsdaten nicht vollständig vorliegen, sich nicht über alle KV-Bezirke gleichermaßen verteilen.

3.3 Längsschnittbetrachtung

Der vorliegende Datenkörper umfasst Daten aus den Jahren 2016 bis 2018. Es ist damit also theoretisch möglich, Veränderungsdaten von Kennzahlen für diesen Zeitraum zu be-

rechnen. Wie bereits ausgeführt, sind aufgrund von Problemen diverse Datensätze markiert worden, die in Auswertungen nicht verwendet werden sollten. Die Markierung von Datensätzen bezieht sich in den meisten Fällen jeweils auf das Jahr, in dem das entsprechende Datenproblem im Rahmen der Datenprüfung aufgetreten ist. Es ist also möglich, dass einzelne Versicherte oder auch alle Versicherten einer Fusionskasse in einem Berichtsjahr markiert sind und in einem anderen nicht. Jahresvergleiche von Kennzahlen können damit problematisch sein. Dies sollte bei der Interpretation der Kennzahlen einer Längsschnittbetrachtung berücksichtigt werden.

3.4 Demografische Hochrechnungsfaktoren

Aufgrund der durch Qualitätssicherungsmaßnahmen erforderlichen Ausschlussmarkierungen von Versicherten kann es im Vergleich zur amtlichen KM6-Statistik zu Verzerrungen der demografischen Struktur der Stichprobe kommen. Um die demografische Repräsentativität der Geburtstagsstichprobe zu gewährleisten, hat der Bewertungsausschuss in der Vergangenheit sogenannte demografische Hochrechnungsfaktoren verwendet (vgl. etwa Beschluss des Bewertungsausschusses in seiner 503. Sitzung zur Berechnung der diagnosenbezogenen Veränderungsdaten für das Jahr 2021).

Zur Bereitstellung der demografischen Hochrechnungsfaktoren ist in der Satzart DS201 ein eigenes Feld vorgesehen, welches in der GSPR 2018 allerdings leer übermittelt wird. Wie in den vergangenen Jahren erfolgt stattdessen eine Übermittlung der Hochrechnungsfaktoren sowie der Korrekturfaktoren für Teilnehmer ausschließlich an Selektivverträgen nach § 73b SGB V an die Trägerorganisationen des Instituts des Bewertungsausschusses nach Festlegung des Klassifikationsmodell zur Berechnung der diagnosenbezogenen bzw. demografischen Veränderungsdaten für das Jahr 2021 separat.

4 Hinweise zur Nutzung

4.1 Allgemeine Nutzungshinweise

Im vorliegenden Abschnitt werden Nutzungshinweise für eine versichertenorientierte Auswertung unter Einbeziehung von Kassenwechslern gegeben. Dies schließt Nutzungshinweise zu den Diagnose-Datensätzen der Satzart DS203, zu den GOP-Datensätzen der Satzart DS210 und zu den Abrechnungsfall-Datensätzen der Satzart DS202 ein.

Weitere Nutzungshinweise werden in Kapitel 5 entlang einzelner Felder der Satzarten gegeben.

Versichertenorientierte Auswertungen unter Einbezug von Kassenwechslern

Das Datenmodell der GSPR 2018 ist so gestaltet, dass versichertenorientierte Auswertungen insbesondere auf Jahresebene mit hoher Vollständigkeit ermöglicht werden. So sind

für die berichteten Versicherten die wesentlichen Inanspruchnahmedaten, insbesondere die vertragsärztlichen Diagnosen, mit hoher Vollständigkeit zusammengestellt (mit Ausnahme der Diagnosen aus selektivvertraglicher Abrechnung). Beispielsweise könnte bei versichertenbezogenen Auswertungen die Auswahl der Versicherten wie folgt getroffen werden:

- Auswahl derjenigen Versicherten, die in die Auswertung einbezogen werden sollen, anhand der Personen-IDs der Satzart DS201. Die Auswahl geschieht beispielsweise durch Festlegung des Auswertungsjahres (ermittelt anhand der ersten vier Stellen des Versichertenquartals) und Auswahl derjenigen Personen-IDs, die in diesem Zeitraum eine vom Nutzer festgelegte Anzahl von Versichertentagen aufweisen.
- Diesen Versicherten sind je Jahr eindeutige Attribute aus der Satzart DS201 zuzuordnen, um eine Vervielfachung von Datensätzen zu vermeiden. Es wird empfohlen, zunächst die für einen Versicherten in einem Quartal angegebenen Versichertenzeiten aufzuaddieren und dabei auf die maximale kalendarisch mögliche Anzahl Versichertentage dieses Quartals zu begrenzen und anschließend diese Versichertenzeiten für die Quartale des Jahres aufzuaddieren. Hinsichtlich der anderen Attribute wird empfohlen, bei unterschiedlichen Angaben möglichst die zeitlich letzte Angabe des Auswertungsjahres zu verwenden.
- Auswahl der Abrechnungsfälle zu diesen ausgewählten Versicherten anhand der Satzart DS202 über die Personen-ID und das Auswertungsjahr (ermittelt anhand der ersten vier Stellen des Abrechnungsquartals).
- Auswahl der Diagnosen (Satzart DS203) und Gebührenordnungspositionen (Satzart DS210) zu den ausgewählten Abrechnungsfällen der Satzart DS202 (und damit zu den ausgewählten Versicherten) verknüpft durch die Fall-ID (künstlicher Schlüssel zur Identifikation des Abrechnungsfalles) in den Satzarten DS202, DS203 und DS210. Ein Bezug zur ICD-10-Stammdatei (Satzart DS213) und/oder GOP-Stammdatei (Satzart DS215) ist, je nach Auswertung, zu berücksichtigen.

Wie in Abschnitt 6.3 ausgeführt, erfolgte die Zuordnung einer Personen-ID zum Abrechnungsfall möglicherweise quartalsübergreifend. Es ist deshalb nicht ausgeschlossen, dass einem Abrechnungsfall in einem bestimmten Quartal eine Personen-ID zugeordnet ist, deren Versichertenstammdaten sich für das betreffende Quartal in der Satzart DS201 nicht finden lassen. Die eben beschriebene Auswahl von Abrechnungsfällen über Personen-IDs der Satzarten DS201 und DS202 unter Berücksichtigung eines ganzen Auswertungsjahres führt in der Konsequenz üblicherweise nicht zum gleichen Auswahlresultat wie eine quartalsgenaue Auswahl.

Weitere Hinweise:

- **Versicherte, die in einem Zeitraum eine Auffälligkeitsmarkierung aufweisen (Ausprägung „1“ des Feldes „Flag“ in der Satzart DS201) sollten für Auswertungen**

nicht herangezogen werden (vgl. Abschnitt 2.3.2).

- Flags in den Satzarten DS219 oder DS220 sollten nicht automatisch zum Ausschluss führen (vgl. Abschnitte 2.3.4 und 2.3.5).

Versicherte, die innerhalb eines Quartals ihre Kasse wechseln, weisen für dieses Quartal in der Satzart DS201 üblicherweise mehr als einen Datensatz auf. Versicherte mit Beschäftigung (z. B. geringfügig oder in Teilzeit) und gleichzeitigem Leistungsbezug von der Bundesagentur für Arbeit, der Deutschen Rentenversicherung oder anderen Leistungsträgern, können unter Umständen für eine gewisse Zeit bei mehreren Kassen pflichtversichert sein, solange sich die Versichertenverhältnisse nicht geklärt haben. Auch für diese Versicherten können für die entsprechenden Quartale in der Satzart DS201 mehrere Datensätze gefunden werden. In derartigen Konstellationen sollte das weitere Vorgehen anhand des beabsichtigten Auswertungskonzepts festgelegt werden.

- Bei Auswertungen mit Kassenbezug sollten dem Versicherten (ermittelt anhand der Personen-ID der Satzarten DS201 bzw. DS202) die entsprechenden Abrechnungsfälle, Diagnosen, Gebührenordnungspositionen *je Kassenzugehörigkeit* (ermittelt anhand des Rechtsnachfolger-IKs der Satzarten DS201 und DS202¹⁴) zugeordnet werden.
- Bei versichertenbezogenen Auswertungen sollte die Personen-ID der Satzart DS201 ohne Berücksichtigung der Kassenzugehörigkeit herangezogen werden. Sichergestellt werden muss in diesem Fall, dass die Zuordnung der Abrechnungsfälle, Diagnosen und Gebührenordnungspositionen zum Versicherten durch die gewählte Abfragetechnik keine Vervielfachung erzeugt.

Wie bereits erläutert, muss bei einer Zählung der Versichertenquartale oder einer Summation der Versichertenzeiten bei Kassenwechslern und anderen oben genannten Konstellationen durch die gewählte Abfragetechnik sichergestellt werden, dass keine unplausiblen Quartals- und Jahreswerte erzeugt werden.

Nutzungshinweise zu Diagnosen-Datensätzen der Satzart DS203:

Bei der Nutzung der Satzart DS203 sind die folgenden Punkte zu beachten:

- Diagnose-Datensätze werden aus der ursprünglichen Lieferung der Ärzteseite ohne Änderungen für die GSPR 2018 übernommen, ungültige Diagnosen wurden nicht entfernt.
- Ein Versicherter kann im Laufe eines Jahres die gleiche Diagnose auf mehreren Quartalsdatensätzen tragen. Auch in einem Quartal können Diagnosen mehrfach auftreten, auch wenn der Patient nur in einer Arztpraxis behandelt wurde.
- Für Auswertungen ist zu beachten, dass die ICD-Kataloge jahresspezifisch sind (Jah-

¹⁴Dies ist allerdings je nach Auswertungszeitraum aufgrund des quartalsübergreifenden Matchings nicht immer vollständig möglich.

resversionen der ICD-10-GM, Satzart DS213).

- Da ungültige Diagnosen nicht ausgeschlossen wurden, sind nicht alle Diagnoseeinträge der Satzart DS203 in der entsprechenden ICD-Stammtabelle (DS213) auffindbar, insbesondere bei sogenannten Auftragsfällen, in denen keine Pflicht zur Diagnosenverschlüsselung besteht, ist in der Satzart DS203 der Ersatzwert „UUU“ eingetragen.

Nutzungshinweise zu GOP-Datensätzen der Satzart DS210:

Bei der Nutzung der Satzart DS210 sind die folgenden Punkte zu beachten:

- Die Satzart DS210 enthält sowohl GOPen des EBM als auch regionale GOPen.
- Bei Verknüpfungen der Satzart DS210 mit der GOP-Stammdatei der Satzart DS215 ist zu beachten, dass die GOP-Stammdatei KV- und quartalspezifisch ist und dass die Verknüpfung in Einzelfällen mehrdeutig ist (bis zum Berichtsjahr 2017 können Unterschiede zwischen BMÄ und E-GO bei derselben GOP existieren).
- Abrechnungsfälle von Versicherten aus dem betreffenden KV-Bezirk, die bei Ärzten anderer KV-Bezirke anfielen („Fremdfälle“) können regionale GOPen anderer KV-Bezirke beinhalten; diese GOPen sind in der GOP-Stammdatei DS215 des KV-Bezirks des Wohnorts des Versicherten ggf. nicht enthalten.

4.2 Hinweise zur Verknüpfung mit anderen Datenkörpern

Verknüpfung über Versicherte, Betriebsstätten oder Abrechnungsfälle

Die Personen-ID der Versicherten, die Nummer der Betriebsstätten (Praxen) und die Fall-ID der Abrechnungsfälle durchlaufen mehrere Pseudonymisierungsschritte. Insbesondere pseudonymisiert die Datenstelle des Bewertungsausschusses vor der Auslieferung der GSPR 2018 diese Identifikationsmerkmale nochmals unter Verwendung eines Schlüssels, der spezifisch für die GSPR 2018 ist. Verknüpfungen von Versicherten, Betriebsstätten oder Abrechnungsfällen zu anderen Datenkörpern sind somit nicht möglich, auch nicht zu den regionalisierten Geburtstagsstichproben vorangegangener Jahre.

Verknüpfung über Institutionskennzeichen der Krankenkassen

Verknüpfungen der Institutionskennzeichen der Krankenkassen (z. B. Rechtsnachfolger-IK in den Satzarten DS201 und DS202) sind grundsätzlich auch zu anderen Datenkörpern möglich (z. B. mit den in den ANZVER-Daten verwendeten IKs). Voraussetzung ist dafür in der Regel, dass die IKs in externen Datenkörpern auf denselben Fusionsstand gebracht werden wie die Institutionskennzeichen in der GSPR 2018 (Stichtag: 02.04.2020); inwieweit die bereitgestellten Satzarten DS219 und DS220 hierzu ausreichend sind, wäre für jeden externen Datenkörper vorab zu prüfen.

Nutzungshinweis für Vergleiche mit den ANZVER-Daten: Sowohl die Zählungen der Versicherten als auch die Abgrenzung der Versicherten sind in der GSPR 2018 und in den ANZVER-Daten nicht identisch.

Leistungssegmentierung (MGV/EGV)

Für die Einteilung der Leistungsbedarfe aus der Satzart DS210 nach morbiditätsbedingter Gesamtvergütung (MGV) und extrabudgetärer Vergütung (EGV) sind externe Listen heranzuziehen. Zum Beispiel ist die für die Berechnung der demografischen und diagnosebezogenen Veränderungsdaten für das Jahr 2021 maßgebliche Liste als Anlage des Beschlusses des Bewertungsausschusses in seiner 503. Sitzung (schriftliche Beschlussfassung) verfügbar.

5 Kommentierte Datensatzbeschreibung

Im Folgenden sind die Datensatzbeschreibungen aller Satzarten der GSPR 2018 gemäß Beschluss des Bewertungsausschusses in seiner 438. Sitzung (Anlage, schriftliche Beschlussfassung) aufgeführt und kommentiert.

Die möglichen Festlegungen in der Spalte „Feldart“ sind für alle Satzarten:

- M (Muss-Feld)
- m (bedingtes Muss-Feld)
- K (Kann-Feld).

Wenn im Rahmen der Datensatzbeschreibungen auf Schlüsselverzeichnisse verwiesen wird, so sind die im Beschluss des Bewertungsausschusses in seiner 386. Sitzung (schriftliche Beschlussfassung) beschriebenen Schlüsselverzeichnisse gemeint. Die jeweils gültige Fassung dieser Schlüsselverzeichnisse ist auf der Internetseite des Instituts des Bewertungsausschusses (<http://institut-ba.de/service/schlüsselverzeichnisse.html>) veröffentlicht.

5.1 Satzart DS201 – Versichertenstammdaten

Dateiinhalt:

Abgrenzung: Für jede versicherte Person wird je Quartal, in dem diese versichert war, und je Krankenkasse ein Datensatz geliefert. Versicherte mit Wohnsitz im Ausland und Personen mit Übernahme der Krankenbehandlung gegen Kostenerstattung nach § 264 SGB V („betreute Personen“) sind entsprechend dem Beschluss des Bewertungsausschusses in seiner 348. Sitzung (schriftliche Beschlussfassung, zuletzt geändert durch Teil C des Beschlusses des Bewertungsausschusses in seiner 389. Sitzung am 21. Februar 2017, sowie entsprechend dem Beschluss des Bewertungsausschusses in seiner 437. Sitzung (schriftliche Beschlussfassung) einbezogen.

Versichertennummern sind nach dem „Pseudonymisierungsverfahren für Datenlieferungen an den Bewertungsausschuss“ gemäß der Anlage des Beschlusses des Bewertungsausschusses in seiner 414. Sitzung am 14. März 2018 bzw. entsprechenden Folgebeschlüssen in der aktuell gültigen Fassung in Verbindung mit den Vorgaben gemäß den Abschnitten 1.1 und 1.5 pseudonymisiert und anonymisiert.

Primärschlüssel: Die Kombination der Felder 01 bis 03 identifiziert einen Datensatz eindeutig.

Anmerkung: Die einzelnen Datenfelder eines Datensatzes sind durch das Trennzeichen „#“ getrennt.

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
00	Satzart	M	5	alphanum.	konstant „DS201“	
01	Versicherungsquartal	M	5	numerisch	Quartal im Format JJJQ	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigen-schaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
02	IK	M	9	alphanum.	Institutionskennzeichen aus der Satzart 201 der GSP. Die kassenseitigen Empfänger erhalten die Datensätze nur dann mit Kassenbezug, wenn die jeweilige Kasse bzw. deren Rechtsnachfolger zur Kassenart des Empfängers gehört. Ist dies nicht der Fall, so wird dieses Feld durch einen Buchstaben (A-Z) in der 1. Stelle und eine laufende Zeilennummer in den weiteren acht Stellen (8-stellig mit führenden Nullen) ersetzt, sodass die Primärschlüsseigen-schaft nicht verletzt wird.	
03	Personen-ID	M	≤9	numerisch	Ordnungsnummer eines Versicherten-nummernpseudonyms, welches nach den Vorgaben aus den Abschnitten 1.1 und 1.5 erzeugt wurde.	
04	Anzahl Versichertentage	M	≤ 2	numerisch	Anzahl der Versichertentage für das Quartal	
05	Geschlecht	M	1	alphanum.	f = female (weiblich) m = male (männlich) x = unbestimmt (ab dem Versicherungsquartal 20181 zulässig) d = divers (ab dem Versicherungsquartal 20194 zulässig)	Personen mit nicht eindeutigen Angaben zum Geschlecht sind nicht Bestandteil der GSPR 2018.

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
06	Geburtsjahr	M	4	numerisch	Geburtsjahr im Format JJJJ; Versicherte, die im Kalenderjahr 2016 95 Jahre und älter sind, erhalten das Geburtsjahr 1921. Versicherte, die im Kalenderjahr 2017 95 Jahre und älter sind, erhalten das Geburtsjahr 1922 und so fort.	Durch die Gleichsetzung des Geburtsjahres von Versicherten über 95 Jahren, soll eine Identifizierung dieser Personen, insbesondere in kleinen KV-Bezirken, verhindert werden.
07	Versichertenstatus	M	1	numerisch	1 = Mitglied 3 = Familienversicherter 5 = Rentner	Personen, die im Berichtszeitraum 2015–2018 mindestens einmal mit Versichertenstatus NULL geliefert wurden, sind nicht Bestandteil der GSPR 2018.
08	Verstorben	M	1	numerisch	Angabe, ob der Versicherte im betreffenden Quartal verstorben ist: 1 = verstorben 0 = sonst	Das Kennzeichen „Verstorben“ wird in der GSPR 2018 für das Quartal gesetzt, für das dieses Kennzeichen in der Geburtstagsstichprobe bei der betreffenden Personen-ID zeitlich zum letzten Mal gesetzt wurde. Dabei werden nur Quartale berücksichtigt, für die die Anzahl der Versichertentage größer als 0 ist. Ist das Kennzeichen in der Geburtstagsstichprobe dagegen nur für Quartale gesetzt, für die die Anzahl der Versichertentage gleich 0 ist, so wird in der GSPR 2018 unter diesen das zeitlich erste mit dem Kennzeichen „Verstorben“ versehen.

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
09	Flag	M	1	numerisch	<p>falls im Rahmen der QS für diesen Versicherten mind. ein Kennzeichen gesetzt wurde</p> <p>2 = Datensatz aus Fehlwertersatz</p> <p>1 = Kennzeichnung vorhanden (Auffälligkeiten)</p> <p>0 = keine Kennzeichnung</p>	<p>In der GSPR 2018 wird kein Fehlwertersatzverfahren vorgenommen. Die Ausprägung „2“ tritt damit nicht auf. Die Ausprägung „1“ entsteht aufgrund einer oder mehrerer der folgenden Auffälligkeiten (siehe dazu Abschnitt 2.3.2):</p> <ul style="list-style-type: none"> • nicht eindeutige Personen-ID • falsch pseudonymisierte Personen-ID • Dummy-Personen-ID bzw. -Versichertennummer • kollidierende Personen-ID • Datensätze nach dem Sterbequartal • fehlende Datensätze nach Abgleich mit Daten zu Selektivverträgen • Versicherte mit Kennzeichen zur Kostenübernahme nach §264 Abs. 2 SGB V • Falschmatcher • auffälliger Ausschöpfungsgrad des Leistungsbedarfs <p>Es wird empfohlen, Datensätze mit der Ausprägung „1“ nicht zu verwenden.</p>

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
10	Gesamtvertragszuständige KV	M	2	alphanum.	<p>Nummer der für den Gesamtvertrag zuständigen Kassenärztlichen Vereinigung gemäß Schlüsselverzeichnis 2 im letzten Versicherungsquartal eines Kalenderjahres.</p> <p>Zu Details, siehe Erläuterungen.</p>	<p>Um jedem Versicherten pro Berichtsjahr genau eine gesamtvertragszuständige KV zuzuordnen, wird folgender Algorithmus angewandt:</p> <p>(i) Zunächst wird jedem mit der Geburtstagsstichprobe gelieferten Datensatz eine gesamtvertragszuständige KV zugeordnet: Für Versicherte, die nicht mit der Postleitzahl „AUSLA“ geliefert wurden, wird die gelieferte gesamtvertragszuständige KV (Feld 15, gemäß Beschluss des 437. BA) genutzt. Hat die gelieferte Postleitzahl hingegen die Ausprägung „AUSLA“, so wird die Nummer des KV-Bezirks zugeordnet, in dem die über das Institutionskennzeichen im Feld 02 der Satzart 201 zu ermittelnde Krankenkasse der höchsten Fusionsstufe ihren Sitz hat.</p> <p>(ii) Auf Grundlage der mit (i) ermittelten gesamtvertragszuständigen KVen wird das folgende Zuordnungsverfahren durchgeführt:</p> <p>Die Zuordnung erfolgt für einen Versicherten auf Grundlage des letzten Quartals im Jahr, in dem für diesen Versicherten mindestens ein Datensatz der Satzart 201 in der Geburtstagsstichprobe vorliegt. Zunächst wird dazu allen vorliegenden Datensätzen der Satzart 201 aus diesem Quartal für den Versicherten eine KV-Nummer wie folgt zugeordnet:</p> <ol style="list-style-type: none"> 1. Wurde durch obige Vorschrift allen im betrachteten Quartal für den Versicherten vorliegenden Datensätzen der Satzart 201 dieselbe KV-Nummer zugeordnet, so wird diese KV-Nummer dem Versicherten zugeordnet.

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
						<p>2. Im Fall, dass durch obige Vorschrift nicht allen vorliegenden Datensätzen dieselbe KV-Nummer zugeordnet wurde, es aber unter den zugeordneten KV-Nummern genau eine gibt, die für den betreffenden Versicherten im Vorquartal gemäß dem obigen Verfahren nicht zugeordnet werden kann, so wird dem Versicherten diese KV-Nummer zugeordnet.</p> <p>3. In den verbleibenden Fällen, in denen weder durch 1. noch durch 2. eine eindeutige Zuordnung erfolgen kann, werden alle vorliegenden Datensätze der Satzart 201 aus dem Quartal betrachtet, für die im Feld 04 die größte Anzahl an Versichertentagen gemeldet wurde. Ist allen diesen Datensätzen dieselbe KV-Nummer zugeordnet, so wird diese dem Versicherten zugeordnet.</p> <p>4. Bei immer noch verbliebener Mehrdeutigkeit erfolgt die Zuordnung des Versicherten zu der KV-Nummer des Datensatzes der Satzart 201 aus dem Quartal, für den der mit dem MD5-Hashverfahren aus der Konkatenation von Postleitzahl (Feld 08), Personen-ID (Feld 03) und der dem Datensatz zugeordneten KV-Nummer erhaltene Hashwert am größten ist.</p>

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
11	Kennzeichen Wohnort	M	1	numerisch	Kennzeichnung zum Wohnsitz des Versicherten: 0 = Wohnsitz im Inland (PLZ in der GSP zuordenbar) 1 = Wohnsitz im Ausland (PLZ in der GSP ist „AUSLA“) 2 = unbekannter Wohnsitz (PLZ in der GSP nicht zuordenbar)	
12	Rechtsnachfolger- IK	m	9	alphanum.	IK des aktuellen Rechtsnachfolgers gemäß Satzart DS220. Die kassenseitigen Empfänger erhalten die Datensätze nur dann mit Kassenbezug, wenn die jeweilige Kasse bzw. deren Rechtsnachfolger zur Kassenart des Empfängers gehört. Ist dies nicht der Fall, so wird dieses Feld geleert	
13	Gewichtungsfaktor	K	7,4	dezimal	Der Gewichtungsfaktor ist immer dann anzuwenden, wenn Auswertungen eine Datengrundlage erfordern, die hinsichtlich Alter und Geschlecht repräsentativ sein muss. Übermittlung ggf. im Rahmen einer Ergänzungsdatenlieferung.	Wie in Abschnitt 2.2 beschrieben, wird dieses Feld leer übermittelt.
14	Kalendertag des Geburtstags	M	≤ 2	numerisch	Kalendertag des Geburtstags des Versicherten, welcher zur Stichprobenziehung geführt hat.	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
15	Kennzeichen Betreute	M	1	numerisch	Angabe gemäß Feld 14 der Satzart 201 der Geburtstagsstichprobe (Kostenübernahme für betreute Personen): 0 = nein 1 = ja	
16	Kennzeichen Bereinigungsrelevanz	M	≤ 5	alphanum.	Angabe, ob der Versicherte im Versicherungsquartal an mindestens einem in Bezug auf den Behandlungsbedarf nach § 87a Abs. 3 Satz 2 SGB V bereinigungsrelevanten Selektivvertrag gemäß § 63 SGB V, § 73b SGB V, § 73c SGB V a. F. oder § 140a SGB V teilgenommen hat. Das Kennzeichen setzt sich aus folgenden Buchstaben zusammen: k = keine Teilnahme an einem Selektivvertrag a = Teilnahme an einem oder mehreren Selektivverträgen nach § 140a SGB V b = Teilnahme an einem oder mehreren Selektivverträgen nach § 73b SGB V c = Teilnahme an einem oder mehreren Selektivverträgen nach § 73c SGB V a. F. d = Teilnahme an einem oder mehreren Selektivverträgen nach § 63 SGB V u = Teilnahme an einem oder mehreren Selektivverträgen, deren Vertragsart nicht feststellbar ist	Alle Datensätze zu einer Personen-ID in einem Quartal enthalten dieselbe Ausprägung dieses Feldes. Die Befüllung dieses Feldes erfolgt basierend auf dem im Rahmen der Qualitätssicherung der Daten zu Selektivverträgen festgestellten Teilnahmestatus (siehe dazu auch Abschnitt 6.5).

Feld Nr	Feld	Feldart	Anzahl Stellen	Feld-eigen-schaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
17	Kennzeichen selektivvertraglicher Abrechnungsweg	K	1	numerisch	<p>Datensätze zu Versicherten mit dem Wert „k“ im Feld 16 erhalten in diesem Feld den Wert „1“, wenn sie in den Daten der Geburtstagsstichprobe als Selektivvertragsteilnehmer mit Abrechnungsweg vollständig über die KV gekennzeichnet sind, sonst wird dieses Feld leer („NULL“) übermittelt.</p> <p>Die Datensätze zu Versicherten mit einem Wert ungleich „k“ im Feld 16 erhalten in diesem Feld den Wert „2“, wenn für sie in dem entsprechenden Datensatz der Geburtstagsstichprobe ein sonstiger Abrechnungsweg dokumentiert wurde, sonst erhalten diese Datensätze in diesem Feld den Wert „1“.</p> <p>Werden für die GSPR mehrere Datensätze der Geburtstagsstichprobe mit unterschiedlichen Abrechnungswegen zu einem Versicherten, einem Institutionskennzeichen und einem Quartal zusammengefasst, so wird der erste auftretende Wert entsprechend der Reihenfolge „2“, „1“, „NULL“ übernommen.</p>	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
18	Kennzeichen Vollständigkeit des Versorgungsauftrags	M	1	numerisch	<p>Unabhängig von der Angabe in Feld 16 ist anzugeben, in welchem Ausmaß ein vollständiger oder unvollständiger kollektivvertraglicher Versorgungsauftrag – vor dem Hintergrund der Teilnahme bzw. Nichtteilnahme des Versicherten an einem oder mehreren Selektivverträgen gemäß § 63 SGB V, § 73b SGB V, § 73c SGB V a. F. und § 140a SGB V – im Versicherungsquartal vorliegt.</p> <p>1 = Kollektivvertraglicher Versorgungsauftrag ist vollständig 2 = Kollektivvertraglicher Versorgungsauftrag ist in Bezug auf hausärztlichen Versorgungsbereich unvollständig und in Bezug auf fachärztlichen Versorgungsbereich vollständig 3 = Kollektivvertraglicher Versorgungsauftrag ist in Bezug auf fachärztlichen Versorgungsbereich unvollständig und in Bezug auf hausärztlichen Versorgungsbereich vollständig 4 = Kollektivvertraglicher Versorgungsauftrag ist sowohl in Bezug auf hausärztlichen als auch fachärztlichen Versorgungsbereich unvollständig</p> <p><u>Hinweis:</u> Reine Add-on-Verträge sind nicht als unvollständiger kollektivvertraglicher Versorgungsauftrag zu werten.</p>	<p>Die Befüllung dieses Feldes basiert auf dem in der Satzart 201 der Geburtstagsstichprobe gelieferten Kennzeichen. Dieses wird abgeglichen mit der im Rahmen der Qualitätssicherung der Daten zu Selektivverträgen festgestellten Vollständigkeit des Versorgungsauftrags und, falls sich daraus eine größere Unvollständigkeit ergibt, als aus dem gelieferten Kennzeichen ersichtlich ist, so erfolgt die Befüllung mit einer Ausprägung, die dieser größeren Unvollständigkeit entspricht.</p> <p>Im Rahmen dieses Korrekturverfahrens kann es passieren, dass für ein Versichertenverhältnis aus der Satzart 201 ein unvollständiger Versorgungsauftrag ermittelt wird, aber keine Aussage über die Art der Unvollständigkeit getroffen werden kann. In diesem Fall, wird das Feld mit dem Wert „5“ befüllt. Zusätzlich zu den genannten Werten kann also die folgende Ausprägung auftreten:</p> <p>5 = Kollektivvertraglicher Versorgungsauftrag ist in Bezug auf hausärztlichen <i>oder</i> fachärztlichen Versorgungsbereich unvollständig</p>

5.2 Satzart DS202 – Ambulante Abrechnungen (KV-Fall)

Dateiinhalt:
<p>Abgrenzung: Zu berücksichtigen sind alle Behandlungsfälle gemäß § 21 BMV-Ä von Versicherten bzw. Fälle von betreuten Personen (§ 264 Abs. 2 SGB V) der Geburtstagsstichprobe, denen in der Datenstelle des Bewertungsausschusses mindestens ein Datensatz aus Satzart DS201 zugeordnet werden konnte.</p> <p>Primärschlüssel: Das Feld 03 identifiziert einen Datensatz eindeutig und verknüpft die Satzart DS202 (KV-Fall) mit der Satzart DS203 (Diagnosen) und der Satzart DS210 (Gebührenordnungspositionen). Für bereichseigene Behandlungsfälle gilt: Die Felder 01, 04 und 05 dienen als Schlüssel zur Verknüpfung mit der Satzart DS211 (Betriebsstättenverzeichnis).</p> <p>Anmerkung: Die einzelnen Datenfelder eines Datensatzes sind durch das Trennzeichen „#“ getrennt. Die Wertangaben in den Feldern 06 bis 09 sind ganzzahlig zu übermitteln.</p>

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
00	Satzart	M	5	alphanum.	konstant „DS202“	
01	Abrechnungsquartal	M	5	numerisch	Quartal im Format JJJQ	Nachtragsfälle aus dem Vorquartal können hier enthalten sein.
02	Rechtsnachfolger- IK	m	9	alphanum.	Institutionskennzeichen der Krankenkasse des gematchten Falls unter Berücksichtigung des aktuellen Fusionsstandes. Die kassenseitigen Empfänger erhalten die Datensätze nur dann mit Kassenbezug, wenn die jeweilige Kasse bzw. deren Rechtsnachfolger zur Kassenart des Empfängers gehört. Ist dies nicht der Fall, so wird dieses Feld geleert.	Es ist zu beachten, dass es Personen-IDs geben kann, für die sich eine Kombination aus Personen-ID und Kassenzugehörigkeit in der Satzart DS202 finden lässt, die sich nicht in der Satzart DS201 (Versichertenstammdaten) widerspiegelt. Bei diesen können in sehr geringem Umfang sogar Institutionskennzeichen vorliegen, die überhaupt nicht in den kassenseitig gelieferten Daten vorkommen. Das Matching dieser Fälle erfolgte dann direkt über die Personen-ID.

¹⁴Siehe Fußnote 2 auf Seite 7.

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
03	Fall_ID	M	≤10	numerisch	Künstlicher Schlüssel zur Datensatzidentifikation. Ordnungsnummer des Pseudonyms einer Fall_ID welches nach den Vorgaben aus den Abschnitten 1.3 und 1.5 erzeugt wurde	
04	Betriebsstättenpseudonym	M	≤8	numerisch	Ordnungsnummer eines Betriebsstättenpseudonyms welches nach den Vorgaben aus den Abschnitten 1.2 und 1.5 erzeugt wurde. Bei Fällen, denen im Feld 05 die KV-Nummer „00“ zugewiesen wurde, wird die Ordnungsnummer des Betriebsstättenpseudonyms durch den Zahlenwert „9999xxxx“ ersetzt, dabei bezeichnet „xxxx“ die vierstellige Abrechnungsgruppenzuordnung der Betriebsstätte.	
05	KV-Nummer	M	2	alphanum.	Nummer der Kassenärztlichen Vereinigung am Ort der Betriebsstätte gemäß Schlüsselverzeichnis 2. Davon abweichende Kennzeichnung für Fremdfälle: 00 = Kennzeichnung für einen Fall (Fremdfall), bei dem die Angaben Wohnort-KV und KV-Nummer verschieden sind	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
06	LB_Punkte	M	≤ 13	numerisch	Summe des abgerechneten Leistungsbedarfs der in Punkten bewerteten Gebührenordnungspositionen des EBM nach sachlich-rechnerischer Richtigstellung und vor Anwendung honorarwirksamer Begrenzungsregelungen in Zehntelpunkten	Hierbei ist zu beachten, dass es rundungsbedingt zu Abweichungen im Vergleich zu den relevanten Summen der Satzart DS210 (Gebührenordnungspositionen der ambulanten Behandlung) kommen kann.
07	LB_Euro	M	≤ 14	numerisch	Summe des abgerechneten Leistungsbedarfs der in Euro bewerteten Gebührenordnungspositionen des EBM nach sachlich-rechnerischer Richtigstellung und vor Anwendung honorarwirksamer Begrenzungsregelungen in Cent	Hierbei ist zu beachten, dass es rundungsbedingt zu Abweichungen im Vergleich zu den relevanten Summen der Satzart DS210 (Gebührenordnungspositionen der ambulanten Behandlung) kommen kann.
08	Leistungstage	M	≤ 3	numerisch	Tage mit Abrechnungen von Gebührenordnungspositionen	
09	LB_EURO_GO	M	≤ 14	numerisch	Summe des abgerechneten Leistungsbedarfs laut regionaler Euro-Gebührenordnung nach sachlich-rechnerischer Richtigstellung und vor Anwendung honorarwirksamer Begrenzungsregelungen in Cent	Hierbei ist zu beachten, dass es rundungsbedingt zu Abweichungen im Vergleich zu den relevanten Summen der Satzart DS210 (Gebührenordnungspositionen der ambulanten Behandlung) kommen kann.
10	PersonenID	M	≤9	numerisch	Ordnungsnummer des Pseudonyms, aus Satzart DS201 Feld 03 übernommen.	
11	Gesamtvertragszuständige KV	M	2	alphanum.	Gesamtvertragszuständige KV des Versicherten, aus Satzart DS201 Feld 10 übernommen.	Der Verweis auf die Satzart DS201 in der Erläuterung bezieht sich auf die Datensätze des gematchten Versicherten, wobei das Versicherungsquartal nicht notwendig mit dem Abrechnungsquartal übereinstimmen muss.

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
12	Flag	M	1	numerisch	falls im Rahmen der QS für diesen Fall-Datensatz mindestens ein Kennzeichen gesetzt wurde: 1 = Kennzeichnung vorhanden (Auffälligkeiten) 0 = keine Kennzeichnung	Wie in Abschnitt 2.3.3 genauer erläutert, werden in der GSPR 2018 keine Abrechnungsfälle als auffällig markiert, das Feld „Flag“ enthält damit keine Informationen.

5.3 Satzart DS203 – Diagnosen der ambulanten Behandlung (KV-Fall-Diagnosen)

Dateiinhalt:						
<p>Abgrenzung: Für jede, ggf. auch mehrfach angegebene Diagnose (Diagnose in Verbindung mit Seitenlokalisierung und Diagnosesicherheit) des Behandlungsfalls aus Satzart DS202 (KV-Fall) wird ein Datensatz übermittelt. Die Diagnosen werden gemäß § 295 Abs. 1 SGB V übermittelt.</p> <p>Primärschlüssel: Die Kombination der Felder 03 und 04 identifiziert einen Datensatz eindeutig. Hierbei verknüpft der künstliche Schlüssel in Feld 03 die Satzart DS203 (KV-Fall-Diagnosen) mit der Satzart DS202 (KV-Fall).</p> <p>Anmerkung: Die einzelnen Datenfelder eines Datensatzes sind durch das Trennzeichen „#“ getrennt.</p>						

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
00	Satzart	M	5	alphanum.	konstant „DS203“	
01	Abrechnungsquartal	M	5	numerisch	Quartal im Format JJJJQ	Diagnosen von Nachtragsfällen aus dem Vorquartal können hier enthalten sein.
02	Rechtsnachfolger- IK	K	9	alphanum.	Dieses Feld wird leer übermittelt, d. h. es folgen zwei Spaltentrennzeichen unmittelbar aufeinander.	
03	Fall_ID	M	≤10	numerisch	aus Feld „Fall_ID“ in KV-Fall	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
04	Diagnosenzähler	M	≤ 4	numerisch	Zähler für die Diagnosen, beginnend mit „1“	
05	Diagnose, kodiert	M	≥ 3, ≤ 10	alphanum.	Angabe des Diagnose-Kodes	Ist eine durch den Vertragsarzt mit Hilfe der ICD-10-GM kodierte Diagnose. Die Ersatzschlüsselnummer „UUU“ ist „kein Bestandteil der vom BfArM veröffentlichten ICD-10-GM, sondern nur ein „Ersatzwert“ für besondere Zwecke in der vertragsärztlichen Versorgung.“ (vgl. ICD-10-GM 2009: Systematisches Verzeichnis, S. XXIV und S. 791) und findet sich somit auch nicht in der ICD-Stammtabelle (Satzart DS213).
06	Diagnosesicherheit	m	1	alphanum.	A = ausgeschlossene Diagnose G = gesicherte Diagnose V = Verdachtsdiagnose Z = symptomloser Zustand nach der betreffenden Diagnose Im Falle der Übermittlung des Ersatzwertes UUU für die Diagnose in Feld 05 wird in Feld 06 kein Eintrag übermittelt.	
07	Seitenlokalisierung	K	1	alphanum.	B = Beidseitig L = Links R = Rechts	Für die vertragsärztliche Versorgung ist die Angabe der Seitenlokalisierung eines ICD-Codes optional.

5.4 Satzart DS210 – Gebührenordnungspositionen der ambulanten Behandlung

Dateiinhalt:
<p>Abgrenzung: Für jede verschiedene Gebührenordnungsposition der Behandlungsfälle aus Satzart DS202 (KV-Fall) wird mindestens ein Datensatz geliefert.</p> <p>Primärschlüssel: Die Kombination der Felder 03 und 04 identifiziert einen Datensatz eindeutig. Hierbei verknüpft der künstliche Schlüssel in Feld 03 die Satzart DS210 (Gebührenordnungspositionen) mit der Satzart DS202 (KV-Fall).</p> <p>Anmerkung: Die einzelnen Datenfelder eines Datensatzes sind durch das Trennzeichen „#“ getrennt. Die Wertangaben in den Feldern 07 und 10 sind ganzzahlig zu übermitteln.</p> <p>Es können abrechnungsbedingte Abweichungen zwischen den Einträgen in Feld 07 bzw. Feld 10 zu den jeweiligen Stammdateneinträgen (Satzart DS215, Felder 05, 06 bzw. 08, 09) bestehen.</p>

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
00	Satzart	M	5	alphanum.	konstant „DS210“	
01	Abrechnungsquartal	M	5	numerisch	Quartal im Format JJJJQ	Gebührenordnungspositionen von Nachtragsfällen aus dem Vorquartal können hier enthalten sein.
02	Rechtsnachfolger- IK	K	9	alphanum.	Dieses Feld wird leer übermittelt, d. h. es folgen zwei Spaltentrennzeichen unmittelbar aufeinander.	
03	Fall_ID	M	≤10	numerisch	aus Feld „Fall_ID“ in KV-Fall	
04	GOP-Zähler	M	≤ 4	numerisch	Zähler für die einzelnen Leistungen des Falles beginnend mit „1“	
05	GOP	M	≤ 6	alphanum.	Gebührenordnungsposition, linksbündig	
06	Anzahl	M	≤ 8	numerisch	Anzahl, so oft wurde obige Gebührenordnungsposition in dem Behandlungsfall abgerechnet.	

Feld Nr	Feld	Feld-art	Anzahl Stellen	Feldeigen-schaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
07	Leistungsbedarf der GOP	M	≤ 8	numerisch	Gibt den Leistungsbedarf der abgerechneten GOP nach EBM unter Berücksichtigung der Anzahl aus Feld 06 an.	
08	Kennzeichen Wert	M	1	numerisch	Kennzeichen für den Leistungsbedarf der GOP aus Feld 07: 1 = Zehntelpunkte 2 = Cent 3 = Sonst	
09	Art der Inanspruchnahme	M	1	numerisch	Die Art der Inanspruchnahme des Leistungsscheins: 1 = Ambulante Behandlung 2 = Überweisung 3 = Belegärztliche Behandlung 4 = Notfall	
10	LB_EURO_GO	M	≤ 8	numerisch	Leistungsbedarf der abgerechneten GOP nach regionaler Euro-Gebührenordnung in Cent unter Berücksichtigung der Anzahl aus Feld 06	

5.5 Satzart DS211 – Betriebsstättenverzeichnis

Dateiinhalt:
<p>Abgrenzung: Zu berücksichtigen sind Betriebsstätten des jeweiligen Abrechnungsquartals, die Leistungen für Versicherte bzw. betreute Personen (§ 264 Abs. 2 SGB V) der Geburtstagsstichprobe erbracht haben. Je Abrechnungsquartal wird zu solchen Betriebsstätten des jeweiligen KV-Bezirks ein Datensatz geliefert, für die Versichertenangaben aus Satzart DS201 erfolgreich Fallangaben aus Satzart DS202 zugeordnet werden konnten.</p> <p>Primärschlüssel: Die Kombination der Felder 01 bis 03 identifiziert einen Datensatz eindeutig.</p> <p>Anmerkung: Die einzelnen Datenfelder eines Datensatzes sind durch das Trennzeichen „#“ getrennt.</p>

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
00	Satzart	M	5	alphanum.	Konstant „DS211“	
01	Abrechnungsquartal	M	5	numerisch	Quartal im Format JJJQ	
02	Betriebsstättenpseudonym	M	≤8	numerisch	Ordnungsnummer eines Betriebsstättenpseudonyms, welches nach den Vorgaben aus den Abschnitten 1.2 und 1.5 erzeugt wurde	
03	KV-Nummer	M	2	alphanum.	Nummer der Kassenärztlichen Vereinigung am Ort der Betriebsstätte gemäß Schlüsselverzeichnis 2	
04	Abrechnungsgruppe	M	4	alphanum.	Abrechnungsgruppe der Betriebsstätte nach Tätigkeitsschwerpunkt gemäß Schlüsselverzeichnis 6	
05	HA-/FA-Kennzeichen	M	1	numerisch	Kennzeichen des Versorgungsbereiches der Betriebsstätte: 0 = Hausarzt 1 = Facharzt 2 = Versorgungsbereichsübergreifend	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
06	Praxistyp_VS	M	2	numerisch	Kennzeichen des Praxistyps: 10 = Einzelpraxis 11 = Örtliche Gemeinschaftspraxis / Berufsausübungsgemeinschaft: Fach- / schwerpunktgleich 12 = Örtliche Gemeinschaftspraxis / Berufsausübungsgemeinschaft: Fach- übergreifend 13 = Örtliche Gemeinschaftspraxis / Berufsausübungsgemeinschaft: Schwer- punktübergreifend 20 = Einrichtung nach § 311 SGB V 21 = MVZ 50 = Krankenhaus 90 = Sonstige	

5.6 Satzart DS213 – ICD-10-Stammdatei

Dateiinhalt:	
<p>Abgrenzung: Zu berücksichtigen sind ICD-10-Diagnosen, die im Rahmen der Abrechnung verwendet werden können und nach SGB V zulässig sind.</p> <p>Primärschlüssel: Die Kombination der Felder 01 und 02 identifiziert einen Datensatz eindeutig.</p> <p>Anmerkung: Die einzelnen Datenfelder eines Datensatzes sind durch das Trennzeichen #“ getrennt.</p>	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
00	Satzart	M	5	alphanum.	Konstant „DS213“	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
01	Gültigkeitszeitraum	M	4	numerisch	Jahr der Gültigkeit des Diagnose-Kodes	
02	ICD-10-Kode	M	≥ 3, ≤ 7	alphanum.	Gültiger ICD-10-Kode gemäß SGB V	Die Ersatzschlüsselnummer „UUU“ ist kein Bestandteil der vom BfArM veröffentlichten ICD-10-GM.
03	Text	M	≤ 255	alphanum.	Klartext des ICD-10-Kodes	
04	Notationskennzeichen	K	1	alphanum.	Erlaubte Zeichen: +, *, !	

5.7 Satzart DS215 – Gebührenordnungspositions-Stammdatei

Dateiinhalte:
<p>Abgrenzung: Für jede verschiedene Gebührenordnungsposition, die in einem Quartal des jeweiligen KV-Bezirks gültig ist, wird ein Datensatz geliefert.</p> <p>Primärschlüssel: Die Kombination der Felder 01 bis 04 identifiziert einen Datensatz eindeutig.</p> <p>Anmerkung: Die einzelnen Datenfelder eines Datensatzes sind durch das Trennzeichen „#“ getrennt. Die Wertangaben in den Feldern 05, 06, 08, 09 sind ganzzahlig zu übermitteln.</p>

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
00	Satzart	M	5	alphanum.	konstant „DS215“	
01	Leistungsquartal	M	5	numerisch	Quartal im Format JJJQ	
02	KV-Nummer	M	2	alphanum.	Nummer der Kassenärztlichen Vereinigung am Ort der Betriebsstätte gemäß Schlüsselverzeichnis 2	
03	GO-ART	M	1	numerisch	Beschreibt die Gebührenordnung, konstant „0“	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
04	GOP	M	≤ 6	alphanum.	Gebührenordnungsposition, linksbündig	
05	GOP-Wert ambulant	M	≤ 8	numerisch	Ambulanter Wert der GOP nach EBM in Zehntelpunkten oder Cent	
06	GOP-Wert stationär	M	≤ 8	numerisch	Stationärer Wert der GOP nach EBM in Zehntelpunkten oder Cent	
07	Kennzeichen Wert	M	1	numerisch	Kennzeichen für den Wert der GOP aus Feld 05 bzw. 06: 1 = Zehntelpunkte 2 = Cent 3 = Sonst	
08	LB_EURO_GO ambulant	M	≤ 8	numerisch	Ambulanter Wert der GOP nach regionaler Euro-Gebührenordnung in Cent	
09	LB_EURO_GO stationär	M	≤ 8	numerisch	Stationärer Wert der GOP nach regionaler Euro-Gebührenordnung in Cent	

5.8 Satzart DS219 – Kostenträgerverzeichnis (KT-Stamm)

Dateiinhalt:
Abgrenzung: In der Satzart DS219 erfolgt die Zuordnung eines KVK-IK zu einer Krankenkasse. Diese wird identifiziert durch ein eindeutiges Kassensitz-IK.
Primärschlüssel: Das Feld 01 identifiziert einen Datensatz eindeutig.
Anmerkung: Die einzelnen Datenfelder eines Datensatzes sind durch das Trennzeichen „#“ getrennt.

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
00	Satzart	M	5	alphanum.	konstant „DS219“	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
01	KVK-IK	M	9	alphanum.	Genau ein Datensatz für jedes dem GKV-SV bekannte Institutionskennzeichen	
02	Kassensitz-IK	M	9	alphanum.	Das für den Kostenträger ausgewählte Kassensitz-IK (ohne Berücksichtigung von Rechtsnachfolgern)	
03	Flag	M	1	numerisch	Falls im Rahmen der QS für dieses KVK-IK mindestens ein Kennzeichen gesetzt wurde: 1 = Kennzeichnung vorhanden (Auffälligkeiten) 0 = keine Kennzeichnung	Das Feld hat genau dann die Ausprägung „1“, wenn Versicherte dieser IK aufgrund „falsch pseudonymisierte Personen-ID“ markiert sind. Flags von Institutionskennzeichen beziehen sich auf dieses einzelne Objekt im Unterschied zu Kostenträgern, die zwar in der Satzart DS220 ebenfalls durch ein Institutionskennzeichen identifiziert sind, deren Flag sich aber über alle Institutionskennzeichen des Kostenträgers auswirken kann. Siehe dazu Abschnitt 2.3.4.

5.9 Satzart DS220 – Kostenträgerhistorie

Dateiinhalt:						
<p>Abgrenzung: Die Kostenträgerhistorie gibt Auskunft über die Zusammenlegung (Fusion) von Kostenträgern. Zu jedem in der Satzart DS219 gelieferten Kassensitz-IK wird in der Satzart DS220 der jeweils aktuelle Stand der Abbildung auf alle durch Kassenfusion entstandenen Rechtsnachfolger-IKs (geordnet nach Gültigkeitszeitraum) übermittelt. Mehrere Fusionen werden durch Angabe einer Fusionsstufe deutlich gemacht.</p> <p>Primärschlüssel: Das Feld 01 identifiziert einen Datensatz eindeutig.</p> <p>Anmerkung: Die einzelnen Datenfelder eines Datensatzes sind durch das Trennzeichen „#“ getrennt.</p>						

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
---------	------	---------	----------------	-----------------	--------------------	---------------------------

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
00	Satzart	M	5	alphanum.	konstant „DS220“	
01	Kassensitz-IK	M	9	alphanum.	Kassensitz-IK gemäß Satzart DS219	
02	Kostenträger-Name	M	< 101	alphanum.	Name des Kostenträgers aus Feld 01	
03	Kassenart	M	2	alphanum.	Kassenart des Kostenträgers aus Feld 01 gemäß Schlüsselverzeichnis 1	
04	Gültig ab	M	8	numerisch	Beginn der Gültigkeit der Zuordnung in diesem Datensatz, tagesgenau (Format JJJMMTT)	
05	Gültig bis	M	8	numerisch	Ende der Gültigkeit der Zuordnung in diesem Datensatz aufgrund einer nachfolgenden Rechtsnachfolge (Format JJJMMTT), „99991231“ falls aktuell gültig	
06	IK des direkten Rechtsnachfolgers	M	9	alphanum.	Kassensitz-IK des direkten Rechtsnachfolgers	
07	Name des direkten Rechtsnachfolgers	M	< 101	alphanum.	Name des direkten Rechtsnachfolgers	
08	Kassenart des direkten Rechtsnachfolgers	M	2	alphanum.	Kassenart des direkten Rechtsnachfolgers gemäß Schlüsselverzeichnis 1	
09	IK des aktuellen Rechtsnachfolgers	M	9	alphanum.	Kassensitz-IK des aktuellen Rechtsnachfolgers	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
10	Name des aktuellen Rechtsnachfolgers	M	< 101	alphanum.	Name des aktuellen Rechtsnachfolgers	
11	Kassenart des aktuellen Rechtsnachfolgers	M	2	alphanum.	Kassenart des aktuellen Rechtsnachfolgers gemäß Schlüsselverzeichnis 1	
12	Fusionsstufe	M	≤ 2	numerisch	<p>1: Die Kasse ist nicht fusioniert, es gibt keinen Rechtsnachfolger, in den Feldern 06-11 sind die Angaben für die Kasse selbst wiederholt.</p> <p>2: Die Kasse ist fusioniert, es gibt einen direkten Rechtsnachfolger. Die Felder 06-08 enthalten die Angaben zum direkten Rechtsnachfolger und stimmen mit den Angaben zum aktuellen Rechtsnachfolger in den Feldern 09-11 überein.</p> <p>3, 4, 5 usw.: Es gibt eine Fusionskette mit wiederholten Fusionen. Die Felder 06-08 enthalten die Angaben zum direkten Rechtsnachfolger. Die Felder 09-11 enthalten Angaben zum aktuellen Rechtsnachfolger. Die Zahl der Fusionen ergibt sich aus der Fusionsstufe minus 1.</p>	

Feld Nr	Feld	Feldart	Anzahl Stellen	Feldeigenschaft	Inhalt/Erläuterung	Anmerkungen für Empfänger
13	Kassensitz-IK-KV	m	2	alphanum.	Nummer der Kassenärztlichen Vereinigung am Ort des Kassensitzes gemäß Schlüsselverzeichnis 2, nur angegeben für den Datensatz eines gültigen aktuellen Rechtsnachfolgers mit der jeweils höchsten Fusionsstufe (1 entspricht der höchsten Stufe)	
14	Flag	M	1	numerisch	Falls im Rahmen der QS für dieses Kassensitz-IK mindestens ein Kennzeichen gesetzt wurde: 1 = Kennzeichnung vorhanden (Auffälligkeiten) 0 = keine Kennzeichnung	Die Ausprägung „1“ entsteht aufgrund der Auffälligkeit vom Typ „auffälliger Ausschöpfungsgrad des Leistungsbedarfs“. Flag-Ziel sind Versichertenverhältnisse (Satzart DS201), welche sich jeweils über die Satzart DS219 diesem Kassensitz-IK zuordnen lassen. Siehe dazu Abschnitt 2.3.5.

5.10 Festlegungen zur Datenübermittlung an die Gesamtvertragspartner

5.10.1 Arztseitig

Für die Satzarten DS201, DS202, DS203, DS210, DS211 und DS215 wird je Quartal und KV eine separate Datei mit der Dateinamenskonvention

Satzart_Quartal_KV_Lieferdatum.Endung

übermittelt, wobei folgende Formate verwendet werden:

- Satzart fünfstellig alphanumerisch (DS201, DS202, DS203, DS210, DS211, DS215)
- Quartal fünfstellig numerisch (20161, 20162, ..., 20184)
- KV zweistellig alphanumerisch gemäß Schlüsselverzeichnis 2
- Lieferdatum achtstellig numerisch (JJJJMMTT)
- Endung entweder .csv (vor Verschlüsselung) oder .zip (nach Verschlüsselung)

Die Satzart DS213 wird nur nach Jahren getrennt übermittelt, die Dateinamenskonvention lautet:

DS213_JJJJ0_00_Lieferdatum.Endung

Die Satzarten DS219 und DS220 werden je Lieferung nur einmal übermittelt, die Dateinamenskonventionen lauten:

DS219_00000_00_Lieferdatum.Endung bzw. DS220_00000_00_Lieferdatum.Endung

Auch dabei werden die folgenden Formate verwendet:

- Lieferdatum achtstellig numerisch (JJJJMMTT)
- Endung entweder .csv (vor Verschlüsselung) oder .zip (nach Verschlüsselung)

5.10.2 Kassenseitig

Für die Satzarten DS203, DS210, DS211 und DS215 wird je Quartal und KV eine separate Datei mit der Dateinamenskonvention

Satzart_Quartal_KV_Lieferdatum.Endung

übermittelt. Für die Satzarten DS201 und DS202 wird je Quartal, KV und Rechtsnachfolger-Kassenart eine separate Datei mit der Dateinamenskonvention

Satzart_Quartal_KV_Rechtsnachfolger-Kassenart_Lieferdatum.Endung

übermittelt. Dabei werden folgende Formate verwendet:

- Satzart fünfstellig alphanumerisch (DS201, DS202, DS203, DS210, DS211, DS215)
- Quartal fünfstellig numerisch (20161, 20162, ..., 20184)
- KV zweistellig alphanumerisch gemäß Schlüsselverzeichnis 2

- Rechtsnachfolger-Kassenart alphanumerisch (AOK, BKK, EK, IKK, KN, LKK)
- Lieferdatum achtstellig numerisch (JJJJMMTT)
- Endung entweder .csv (vor Verschlüsselung) oder .zip (nach Verschlüsselung)

Die Satzart DS213 wird nur nach Jahren getrennt übermittelt, die Dateinamenskonzvention lautet:

DS213_JJJJ_00_Lieferdatum.Endung

Die Satzarten DS219 und DS220 werden je Lieferung nur einmal übermittelt, die Dateinamenskonzventionen lauten:

DS219_00000_00_Lieferdatum.Endung bzw. DS220_00000_00_Lieferdatum.Endung

Auch dabei werden die folgenden Formate verwendet:

- Lieferdatum achtstellig numerisch (JJJJMMTT)
- Endung entweder .csv (vor Verschlüsselung) oder .zip (nach Verschlüsselung)

6 Geburtstagsstichprobe des Bewertungsausschusses (GSP)

In diesem Kapitel wird die der GSPR 2018 zugrunde liegende bundesweite Versichertenstichprobe (Geburtstagsstichprobe, GSP) beschrieben. Dazu wird kurz auf die Historie und die Beschlusslage des (Erweiterten) Bewertungsausschusses zur Geburtstagsstichprobe eingegangen (Abschnitt 6.1). Anschließend wird die Übermittlung der Daten der GSP an die Datenstelle des Bewertungsausschusses skizziert und kurz auf das Pseudonymisierungsverfahren eingegangen. Bestandteil dieses Kapitels ist außerdem eine Aufzählung der Satzarten der GSP (Abschnitt 6.2.3). Der Abschnitt 6.3 beschreibt detailliert den Zusammenführungsprozess der kassenseitig und arztseitig gelieferten Daten („Matching“). In Abschnitt 6.4 wird auf verbleibende Mängel in der GSP hingewiesen und der Umgang mit diesen erläutert. In Abschnitt 6.5 wird schließlich auf die Identifikation von Versicherten, die an Selektivverträgen teilnehmen, eingegangen.

6.1 Historie und Beschlusslage des (Erweiterten) Bewertungsausschusses

Geburtstagsstichprobe 2007–2008: Der Bewertungsausschuss hat in seiner 184. Sitzung eine Datenlieferung von den Trägerorganisationen des Bewertungsausschusses an die Datenstelle des Bewertungsausschusses beschlossen. Im Rahmen dieser Datenlieferung wurden pseudonymisierte Versichertenstammdaten, Kostenträgerstammdaten sowie vertragsärztliche Abrechnungsdaten übermittelt. Die Übermittlung bezieht sich dabei auf eine Versichertenstichprobe von versichertenvollständigen Abrechnungsdaten aus allen KV-Bezirken für die Quartale 2007/1 bis 2008/4. Grundlage für die Versichertenstichprobe 2007–2008 ist eine Geburtstagsstichprobe; dem Stichprobenverfahren liegen der 3., 10., 17. und 24. Kalendertag jedes Monats als Geburtsdatum des Versicherten

zugrunde. Die Daten der Geburtsstagsstichprobe 2007–2008 sind nicht Bestandteil der GSPR 2018.

Geburtsstagsstichprobe 2009: Der Erweiterte Bewertungsausschuss hat am 2. September 2009 in der 15. Sitzung die Fortschreibung der mit Beschluss des Bewertungsausschusses aus der 184. Sitzung erhobenen Daten zur bundesweiten Versichertenstichprobe beschlossen. Die Übermittlung bezieht sich auf die Quartale 2009/1 bis 2009/4. Grundlage für die Versichertenstichprobe 2009 ist eine Geburtsstagsstichprobe; dem Stichprobenverfahren liegen der 3., 10., 17. und 24. Kalendertag jedes Monats als Geburtsdatum des Versicherten zugrunde. Die Daten der Geburtsstagsstichprobe 2009 sind nicht Bestandteil der GSPR 2018.

Geburtsstagsstichprobe 2010: In der 237. Sitzung (schriftliche Beschlussfassung) hat der Bewertungsausschuss die Fortschreibung der mit Beschluss des Bewertungsausschusses aus der 184. Sitzung erhobenen Daten zur bundesweiten Versichertenstichprobe beschlossen. Die Übermittlung bezieht sich auf die Quartale 2010/1 bis 2010/4. Grundlage für die Versichertenstichprobe 2010 ist eine Geburtsstagsstichprobe; dem Stichprobenverfahren liegen der 3., 4., 10., 11., 17. und 24. Kalendertag jedes Monats als Geburtsdatum des Versicherten zugrunde. Die Daten der Geburtsstagsstichprobe 2010 sind nicht Bestandteil der GSPR 2018.

Geburtsstagsstichprobe 2011: In der 273. Sitzung (schriftliche Beschlussfassung), geändert durch Beschluss des Bewertungsausschusses in seiner 307. Sitzung (schriftliche Beschlussfassung), hat der Bewertungsausschuss die Fortschreibung der mit Beschluss des Bewertungsausschusses aus der 237. Sitzung erhobenen Daten zur bundesweiten Versichertenstichprobe beschlossen. Die Übermittlung bezieht sich auf die Quartale 2011/1 bis 2011/4. Grundlage für die Versichertenstichprobe 2011 ist eine Geburtsstagsstichprobe; dem Stichprobenverfahren liegen der 3., 4., 10., 11., 17., 18. und 24. Kalendertag jedes Monats als Geburtsdatum des Versicherten zugrunde. Die Daten der Geburtsstagsstichprobe 2011 sind nicht Bestandteil der GSPR 2018.

Geburtsstagsstichprobe 2012: In der 273. Sitzung (schriftliche Beschlussfassung), geändert durch Beschluss des Bewertungsausschusses in seiner 307. Sitzung (schriftliche Beschlussfassung), hat der Bewertungsausschuss die Fortschreibung der mit Beschluss des Bewertungsausschusses aus der 237. Sitzung erhobenen Daten zur bundesweiten Versichertenstichprobe beschlossen. Die Übermittlung bezieht sich auf die Quartale 2012/1 bis 2012/4. Grundlage für die Versichertenstichprobe 2012 ist eine Geburtsstagsstichprobe; dem Stichprobenverfahren liegen der 3., 4., 10., 11., 17., 18., 24. und 25. Kalendertag jedes Monats als Geburtsdatum des Versicherten zugrunde. Die Daten der Geburtsstagsstichprobe 2012 sind nicht Bestandteil der GSPR 2018.

Geburtsstagsstichprobe 2013: In der 273. Sitzung (schriftliche Beschlussfassung), geändert durch Beschluss des Bewertungsausschusses in seiner 307. Sitzung (schriftliche Beschlussfassung), hat der Bewertungsausschuss die Fortschreibung der mit Beschluss des Bewertungsausschusses aus der 237. Sitzung erhobenen Daten zur bundesweiten Versi-

chertenstichprobe beschlossen. Die Übermittlung bezieht sich auf die Quartale 2013/1 bis 2013/4. Grundlage für die Versichertenstichprobe 2013 ist eine Geburtstagsstichprobe; dem Stichprobenverfahren liegen der 4., 5., 11., 17., 18., 24. und 25. Kalendertag jedes Monats als Geburtsdatum des Versicherten zugrunde. Die Daten der Geburtstagsstichprobe 2013 sind nicht Bestandteil der GSPR 2018.

Geburtstagsstichprobe 2014: In der 348. Sitzung (schriftliche Beschlussfassung) hat der Bewertungsausschuss die Fortschreibung der mit Beschluss des Bewertungsausschusses aus der 273. Sitzung (schriftliche Beschlussfassung), zuletzt geändert durch Beschluss des Bewertungsausschusses in seiner 307. Sitzung (schriftliche Beschlussfassung), erhobenen Daten zur bundesweiten Versichertenstichprobe beschlossen. Die Übermittlung bezieht sich auf die Quartale 2014/1 bis 2014/4. Grundlage für die Versichertenstichprobe 2014 ist eine Geburtstagsstichprobe; dem Stichprobenverfahren liegen der 4., 5., 11., 12., 18., 24. und 25. Kalendertag jedes Monats als Geburtsdatum des Versicherten zugrunde. Die Daten der Geburtstagsstichprobe 2014 sind nicht Bestandteil der GSPR 2018.

Geburtstagsstichprobe 2015: In der 348. Sitzung (schriftliche Beschlussfassung) hat der Bewertungsausschuss die Fortschreibung der mit Beschluss des Bewertungsausschusses aus der 273. Sitzung (schriftliche Beschlussfassung), zuletzt geändert durch Beschluss des Bewertungsausschusses in seiner 307. Sitzung (schriftliche Beschlussfassung), erhobenen Daten zur bundesweiten Versichertenstichprobe beschlossen. Die Übermittlung bezieht sich auf die Quartale 2015/1 bis 2015/4. Grundlage für die Versichertenstichprobe 2015 ist eine Geburtstagsstichprobe; dem Stichprobenverfahren liegen der 4., 5., 11., 12., 18., 19. und 25. Kalendertag jedes Monats als Geburtsdatum des Versicherten zugrunde. Die Daten der Geburtstagsstichprobe 2015 sind nicht Bestandteil der GSPR 2018.

Geburtstagsstichprobe 2016: In der 348. Sitzung (schriftliche Beschlussfassung) hat der Bewertungsausschuss die Fortschreibung der mit Beschluss des Bewertungsausschusses aus der 273. Sitzung (schriftliche Beschlussfassung), zuletzt geändert durch Beschluss des Bewertungsausschusses in seiner 307. Sitzung (schriftliche Beschlussfassung), erhobenen Daten zur bundesweiten Versichertenstichprobe beschlossen. Die Übermittlung bezieht sich auf die Quartale 2016/1 bis 2016/4. Grundlage für die Versichertenstichprobe 2016 ist eine Geburtstagsstichprobe; dem Stichprobenverfahren liegen der 5., 11., 12., 18., 19., 25. und 26. Kalendertag jedes Monats als Geburtsdatum des Versicherten zugrunde. Die Daten der Geburtstagsstichprobe 2016 sind Bestandteil der GSPR 2018.

Geburtstagsstichprobe 2017: In der 348. Sitzung (schriftliche Beschlussfassung) hat der Bewertungsausschuss die Fortschreibung der mit Beschluss des Bewertungsausschusses aus der 273. Sitzung (schriftliche Beschlussfassung), zuletzt geändert durch Beschluss des Bewertungsausschusses in seiner 307. Sitzung (schriftliche Beschlussfassung), erhobenen Daten zur bundesweiten Versichertenstichprobe beschlossen. Die Übermittlung bezieht sich auf die Quartale 2017/1 bis 2017/4. Grundlage für die Versichertenstichprobe 2017 ist eine Geburtstagsstichprobe; dem Stichprobenverfahren liegen der 5., 6., 12., 18., 19., 25. und 26. Kalendertag jedes Monats als Geburtsdatum des Versicherten zu-

grunde. Die Daten der Geburtstagsstichprobe 2017 sind Bestandteil der GSPR 2018.

Geburtstagsstichprobe 2018: In der 437. Sitzung (schriftliche Beschlussfassung) hat der Bewertungsausschuss die Fortschreibung der mit Beschluss des Bewertungsausschusses aus der 348. Sitzung (schriftliche Beschlussfassung) erhobenen Daten zur bundesweiten Versichertenstichprobe beschlossen. Die Übermittlung bezieht sich auf die Quartale 2018/1 bis 2018/4. Grundlage für die Versichertenstichprobe 2018 ist eine Geburtstagsstichprobe; dem Stichprobenverfahren liegen der 5., 6., 12., 13., 19., 25. und 26. Kalendertag jedes Monats als Geburtsdatum des Versicherten zugrunde. Die Daten der Geburtstagsstichprobe 2018 sind Bestandteil der GSPR 2018.

6.2 Übermittlung der Daten an die Datenstelle des BA

6.2.1 Datenlieferung

Die Krankenkassen liefern für den entsprechenden Zeitraum die pseudonymisierten Versichertenstammdaten über den entsprechenden Dachverband der Kassenart oder einen für die Kassenart tätigen Dienstleister an den GKV-Spitzenverband. Dieser übermittelt die (nochmals) pseudonymisierten Versichertenstammdaten und Kostenträgerstammdaten an die Datenstelle des Bewertungsausschusses.

Die Kassenärztlichen Vereinigungen liefern für den entsprechenden Zeitraum die pseudonymisierten vertragsärztlichen Abrechnungsdaten an die Kassenärztliche Bundesvereinigung. Die Kassenärztliche Bundesvereinigung übermittelt die (nochmals) pseudonymisierten vertragsärztlichen Abrechnungsdaten anschließend an die Datenstelle des Bewertungsausschusses.

Die vom GKV-Spitzenverband und der Kassenärztlichen Bundesvereinigung an die Datenstelle des Bewertungsausschusses übermittelten Daten werden in der Datenstelle formal geprüft und der jeweilige Datenkörper einer umfangreichen, und zwischen Datengebern und Datenstelle abgestimmten, Qualitätssicherung unterzogen. Die Daten werden erst nach Freigabeerklärung durch die Trägerorganisationen in der Datenstelle weitergehend genutzt.

6.2.2 Pseudonymisierungsverfahren

Die Geburtstagsstichprobe umfasst pseudonymisierte Versichertenstammdaten der Krankenkassen und pseudonymisierte ambulante ärztliche Abrechnungsdaten von Patienten. Die Zusammenführung der Versichertenstammdaten mit den ambulanten ärztlichen Abrechnungsdaten in der Datenstelle erfolgt über die Attribute Institutionskennzeichen der Krankenkasse und pseudonymisierte Krankenversicherungsnummer (KVNR) von der Krankenversichertenkarte bzw. elektronischen Gesundheitskarte (eGK). Die Datenerhebung erfolgt derzeit für die Jahre 2016 bis 2018. Bereits für die Datenerhebung des Jahres 2010 erfolgte erstmals eine turnusmäßige Erweiterung der Stichprobenauswahl in

Verbindung mit einer Schlüsselerweiterung. Die Stichprobenauswahl wurde für die Jahre 2011 und 2012 erneut turnusmäßig erweitert. Seit 2013 ist die Anzahl der Geburtskalendertage in der Stichprobe konstant, wobei jedes Jahr ein Geburtskalendertag entfernt und durch einen neuen ersetzt wird.

Für weitergehende Informationen zum verwendeten Pseudonymisierungsverfahren sei auf den Beschluss des Bewertungsausschuss in seiner 414. Sitzung verwiesen.

6.2.3 Satzarten der Geburtstagsstichprobe

Alle Informationen, die in der GSPR 2018 enthalten sind, basieren auf den Daten der Geburtstagsstichprobe (GSP) sowie auf den Datenlieferungen zu Selektivverträgen. In der GSP erfolgt weder eine Trennung nach Wohnort-Prinzip noch wird eine Maskierung von kassenartfremden Institutionskennzeichen bzw. KV-fremden Betriebsstättennummern vorgenommen. Die Geburtstagsstichproben der Jahre 2016, 2017 und 2018 enthalten folgende Satzarten:

- Satzart 200: Versichertennummern (entfallen seit Berichtsjahr 2018)
- Satzart 201: Versichertenstammdaten
- Satzart 202: Ambulante Abrechnungen (KV-Fall)
- Satzart 203: Diagnosen der ambulanten Behandlung (KV-Fall-Diagnosen)
- Satzart 210: Gebührenordnungspositionen der ambulanten Behandlung
- Satzart 211: Betriebsstättenverzeichnis
- Satzart 213: ICD-10-Stammdatei
- Satzart 215: Gebührenordnungspositions-Stammdatei
- Satzart 219: Kostenträgerverzeichnis (KT-Stamm)
- Satzart 220: Kostenträgerhistorie

6.3 Zusammenführung der Datenlieferungen zur GSP – Matching

6.3.1 Definition

Der Begriff „Matching“ beschreibt die Zuordnung von Abrechnungsfällen der Satzart 202 zu Versicherten. Ein Versicherter steht dabei für eine natürliche – im Berichtszeitraum 2016–2018¹⁵ in der gesetzlichen Krankenversicherung versicherte – Person, deren pseudonymisierten Versichertenstammdaten und ggf. vertragsärztlichen Abrechnungsdaten in der Geburtstagsstichprobe vorhanden sind. Ziel des Matchingverfahrens ist es dabei,

¹⁵Die GSPR 2018 enthält nur Daten des Berichtszeitraums 2016–2018. Das beschriebene Matchingverfahren greift aber auch auf Daten älterer Geburtstagsstichproben zurück. Dies kann aufgrund des mehrstufigen Matchingverfahrens, bei dem einem Abrechnungsfall eine Personen-ID auch quartalsübergreifend zugeordnet werden kann, ggf. relevant sein.

allen Abrechnungsfällen der Satzart 202 eine eindeutige Personen-ID der Satzart 201 zuzuordnen.

Diejenigen Abrechnungsfälle der Satzart 202, denen keine Personen-ID der Satzart 201 aus *dem gleichen Kalenderjahr* zugeordnet werden kann, sind nicht Bestandteil des ausgelieferten Datenkörpers. Da die GSPR 2018 nur die Berichtsjahre 2016–2018 umfasst, werden auch keine Abrechnungsfälle weiter zurück liegender Berichtsjahre mit ausgeliefert. Es werden Daten zu GKV-Versicherten mit zugeordneten Abrechnungsfällen sowie Daten zu GKV-Versicherten ohne zugeordnete Abrechnungsfälle („Nichtinanspruchnehmer“) ausgeliefert.

Idealerweise lassen sich für jeden Abrechnungsfall der Satzart 202 eine Personen-ID und damit die Versichertenstammdaten der Satzart 201 quartalsbezogen zuordnen. Im Folgenden wird beschrieben, warum dies nicht in allen Fällen möglich ist (Abschnitt 6.3.2) und wie durch ein mehrstufiges Matchingverfahren versucht wird, die vorhandenen Zuordnungsprobleme zu beheben (Abschnitt 6.3.4).

Wie in Abschnitt 6.3.4 ausgeführt, erfolgt die Zuordnung der Personen-ID zum Abrechnungsfall ggf. quartalsübergreifend. Es ist also nicht ausgeschlossen, dass einem Abrechnungsfall in einem bestimmten Quartal eine Personen-ID zugeordnet wird, deren Versichertenstammdaten sich für das betreffende Quartal in der Satzart 201 nicht finden lassen. Es ist lediglich sichergestellt, dass sich für jeden in der GSPR 2018 ausgelieferten Abrechnungsfall in mindestens einem Quartal des betreffenden Jahres eine Personen-ID aus den Versichertenstammdaten finden lässt.

6.3.2 Aufgabenstellung

Sowohl in den Versichertenstammdaten als auch in den Daten zur Ambulanten Abrechnung ist für jeden Datensatz eine pseudonymisierte Versichertennummer und die Angabe des Kostenträgers zu finden. Idealerweise sollte die Kombination aus Pseudonym der Versichertennummer und Kostenträger also ausreichen, um jedem Abrechnungsfall der Satzart 202 eine Personen-ID der Satzarten 201 zuzuordnen.

Die Abrechnungsfälle der Satzart 202 entstammen der ärztlichen Rechnungslegung. In der ärztlichen Rechnungslegung liegen über den Versicherten die auf seiner vorgelegten Krankenversicherungskarte enthaltenen Informationen zum Krankenversicherungskarten-Institutionskennzeichen (KVK-IK) sowie die auf der Karte verzeichnete Versichertennummer vor.

Die Versichertennummern der Satzart 201 hingegen entstammen kassenseitigen Datenerlieferungen. Die Kostenträger übermitteln dem GKV-Spitzenverband quartals- und versichertenbezogen neben Stammdaten (z. B. Geschlecht und Geburtsjahr) ein Institutionskennzeichen (im Folgenden Melde-IK genannt), die für dieses Quartal aktuell gültigen, pseudonymisierten Versichertennummern sowie eine über den gesamten Berichtszeitraum hinweg gleich bleibende Personen-ID.

Die Kostenträger verfügen unter Umständen über mehrere Institutionskennzeichen (z. B. durch Fusionen oder aufgrund einer regionalen Gliederung des Kostenträgers) und liefern die Versichertenstammdaten nicht durchgängig mit dem auf der Krankenversicherungskarte verwendeten KVK-IK als Melde-IK an den GKV-Spitzenverband. Anders formuliert: Die KVK-IKs der arztseitigen Datenlieferungen und die Melde-IKs der kassenseitigen Rechnungslegung können sich unterscheiden, obwohl sich hinter beiden der gleiche Kostenträger verbirgt. Das einfache Zusammenführen arzt- und kassenseitig gelieferter Daten anhand der jeweils gemeldeten Versichertennummern und Institutionskennzeichen ist dann nicht möglich. Durch die vom GKV-Spitzenverband gelieferten Satzarten 219 und 220 können Melde-IKs und KVK-IKs allerdings in die entsprechenden Kassensitz-Institutionskennzeichen (Kassensitz-IK) überführt und Fusionen berücksichtigt werden.

Das Matching wird zusätzlich durch die Tatsache erschwert, dass in den Versichertenstammdaten ein zeitpunktbezogener Fusionsstand berücksichtigt ist. Der Datenbestand zum Lieferzeitpunkt entspricht dem Fusions- und Auflösungsstand der Krankenkassen zum jeweiligen Liefertermin. Bei Fusionen kann unklar sein, welche (dann Teil-)Kasse über die Versichertenstammdaten verfügt und in der Lieferverpflichtung steht.

6.3.3 Ziel des Matchingverfahrens

Ziel des Matchingverfahrens ist es, möglichst allen Abrechnungsfällen der Satzart 202 die für den Berichtszeitraum versichertenbezogen eindeutige Personen-ID richtig zuzuordnen. Die Abrechnungsfälle, für die keine Personen-ID gefunden wird, sind nicht Bestandteil des an die Trägerorganisationen des Bewertungsausschusses ausgelieferten und an die Gesamtvertragspartner weitergeleiteten Datenkörpers GSPR 2018.

Bei einem erfolgreichen Matching können den Abrechnungsfällen über die Satzart 201 personenbezogene Attribute zugeordnet werden.

6.3.4 Stufen des Matchingverfahrens

Für Leser, die mit der GSPR 2017 vertraut sind, sei als Erstes darauf hingewiesen, dass es keine Änderungen am Matchingverfahren gegeben hat. Insbesondere die Einteilung in Matchingstufen sowie deren grundsätzlicher Ablauf und deren Nummerierung haben sich im Vergleich zum Vorjahr nicht geändert.

Das Matchingverfahren erfolgt in fünf aufeinander folgenden Stufen. Die jeweils folgende Stufe wird nur dann angewendet, wenn die vorhergehende Stufe kein Ergebnis (kein Match) erbracht hat. Als Ergebnis wird dabei auch die Erkenntnis bezeichnet, dass sich zu einem Abrechnungsfall keine eindeutige Personen-ID zuordnen lässt (No-Match).

Die nicht zugeordneten Abrechnungsfälle werden nicht an die Trägerorganisationen und Gesamtvertragspartner übermittelt. Demzufolge sind die in den ausgelieferten Daten der

GSPR 2018 enthaltenen Fallzahlen, Diagnoseangaben, Gebührenordnungspositionshäufigkeiten und Leistungsbedarfe stets kleiner als die in den Daten der Geburtsstichprobe vollständig enthaltenen Abrechnungsdaten aller Kassenärztlichen Vereinigungen zusammen.

Stufe I: Matching über Personen-ID

Bei Abrechnungsfällen, denen die lebenslange Krankenversicherungsnummer zugrunde liegt (z. B. bei Vorliegen der elektronischen Gesundheitskarte (eGK)), wird aus dieser Nummer direkt die (pseudonymisierte) Personen-ID gebildet. Damit ist ein Matching direkt über die Personen-ID der Satzart 201 möglich.

Wenn also die pseudonymisierte Versicherungsnummer eines Abrechnungsfalles (Satzart 202) mit der Personen-ID eines Versichertenstammdatensatzes (Satzart 201) übereinstimmt, dann wird der Abrechnungsfall dieser Personen-ID zugeordnet. Die folgenden Matchingstufen kommen nicht zur Anwendung. Kann dagegen keine Personen-ID zugeordnet werden, so wird Stufe II des Matchingverfahrens durchgeführt. Auf der Stufe I werden also noch keine Datensätze als nicht zuordenbar identifiziert. Ein No-Match ist auf dieser Stufe damit nicht möglich.

Stufe II: Direktes Matching

Wenn es zu der Kombination aus KVK-IK und pseudonymisierter Versicherungsnummer eines Abrechnungsfalles aus der Satzart 202 mindestens einen Versichertenstammdatensatz (aus der Satzart 201 oder Kombination aus Satzart 201 und Satzart 200) gibt, bei dem

1. das KVK-IK aus Satzart 202 mit dem Melde-IK aus Satzart 201 bzw. aus Satzart 200 und
2. die (pseudonymisierte) Versicherungsnummer aus der Satzart 201 bzw. 200 mit derjenigen aus der Satzart 202 übereinstimmen,

dann wird diesem Abrechnungsfall die Personen-ID dieser Stammdatensätze zugeordnet, sofern alle Versichertenstammdatensätze dieselbe Personen-ID aufweisen.

Wenn es mehrere Versichertenstammdatensätze gibt, die die Bedingungen 1 und 2 erfüllen, diese aber nicht alle dieselbe Personen-ID aufweisen, so gilt dieser Abrechnungsfall als nicht zuordenbar (No-Match) und weitere Matchingstufen kommen für diesen Abrechnungsfall nicht mehr zur Anwendung. Der Abrechnungsfall ist damit nicht Bestandteil des an die Trägerorganisationen und Gesamtvertragspartner ausgelieferten Datenkörpers GSPR 2018.

Wenn es keinen Versichertenstammdatensatz gibt, der die Bedingungen 1 und 2 erfüllt, wird Stufe III des Matchingverfahrens durchgeführt.

Stufe III: Indirektes Matching

Stufe III kommt für diejenigen Datensätze zur Anwendung, bei denen die Stufen I und II kein Ergebnis geliefert haben. Die Stufe III wird für einen Abrechnungsfall erfolgreich abgeschlossen, wenn folgende zwei Bedingungen erfüllt sind:

1. Für den Abrechnungsfall wird mindestens ein Versichertenstammdatensatz der Satzart 201 bzw. der Kombination aus Satzart 201 und Satzart 200 gefunden, für den folgende zwei Eigenschaften gelten:
 - (a) Die (pseudonymisierte) Versichertennummer stimmt mit der des Abrechnungsfalls überein und
 - (b) das Kassensitz-IK, das dem Melde-IK der Satzart 201 über die Satzart 219 zugeordnet wurde, stimmt überein mit dem Kassensitz-IK, das dem KVK-IK der Satzart 202 über die Satzart 219 zugeordnet wurde.
2. Alle Versichertenstammdaten, die die Eigenschaften 1a und 1b erfüllen, weisen dieselbe Personen-ID aus.

Werden keine Versichertenstammdaten gefunden, die die Eigenschaften 1a und 1b erfüllen, so wird die Stufe IV des Matchingverfahrens durchgeführt.

Werden hingegen mehrere Versichertenstammdaten gefunden, die zwar alle die Eigenschaften 1a und 1b erfüllen, allerdings unterschiedliche Personen-IDs aufweisen, so gilt der Abrechnungsfall als nicht zuordenbar (No-Match) und weitere Matchingstufen kommen für diesen Abrechnungsfall nicht mehr zur Anwendung. Der Abrechnungsfall ist damit nicht Bestandteil des an die Trägerorganisationen und Gesamtvertragspartner ausgelieferten Datenkörpers GSPR 2018.

Stufe IV: Matching über Fusionen

Die Matchingstufe IV kommt nur zur Anwendung, wenn die Stufen I, II und III keine Ergebnisse geliefert haben. Durch diese Matchingstufe wird die Fusion von Kostenträgern berücksichtigt. Nach dem Zusammenschluss von Kostenträgern ist es z. B. möglich, dass Versicherte zwar ihre Versichertennummer und Krankenversicherungskarte behalten, aber unter einem anderen Melde-IK gemeldet werden. Dieses Melde-IK kann sowohl ein von den fusionierten Kostenträgern bereits vor der Fusion verwendetes Institutionskennzeichen als auch ein dem neu entstandenen Kostenträger zugewiesenes Institutionskennzeichen sein. Da die Festlegung der Versichertennummern vor dem Zusammenschluss zwischen den beteiligten Kostenträgern nicht zwangsläufig abgestimmt war, ergibt sich bei dieser Stufe eine höhere Gefahr, dass Versichertennummern miteinander kollidieren.

Die Berücksichtigung von Fusionen wird durch folgende recht komplexe Vorgehensweise erreicht. Zuerst wird ermittelt, ob zu einem Abrechnungsfall (Satzart 202) mindestens ein Versichertenstammdatensatz (Satzart 201 bzw. eine Kombination aus Satzart 201 und 200) existiert, für den gilt:

1. Die pseudonymisierte Versichertennummer stimmt mit der aus dem Abrechnungsfall

überein und

2. das dem Melde-IK über Satzart 219 zugeordnete Kassensitz-IK (im Folgenden als 201-Träger bezeichnet) und das dem KVK-IK über Satzart 219 zugeordnete Kassensitz-IK (im Folgenden als 202-Träger bezeichnet) erfüllen folgende drei Bedingungen:
 - (a) 201-Träger und 202-Träger haben einen gemeinsamen Rechtsnachfolger nach Satzart 220,
 - (b) der Beginn der Rechtsnachfolge zum 201-Träger liegt vor dem Ende des Abrechnungsquartals¹⁶ und
 - (c) der Beginn der Rechtsnachfolge zum 202-Träger liegt vor dem Ende des Abrechnungsquartals.

Weisen zudem alle Versichertenstammdatensätze (Satzart 201), die die obigen Bedingungen 1 und 2a erfüllen (ohne die Bedingungen 2b und 2c erfüllen zu müssen)¹⁷, dieselbe Personen-ID aus, dann wird der Abrechnungsfall dieser Personen-ID zugeordnet.

Weisen im Gegensatz dazu mehrere Versichertenstammdatensätze, die die obigen Bedingungen 1 und 2a erfüllen (ohne die Bedingungen 2b und 2c erfüllen zu müssen), unterschiedliche Personen-ID aus, und gibt es mindestens einen Versichertenstammdatensatz, der die Bedingungen 1 und 2 erfüllt, so gilt der Abrechnungsfall als nicht zuordenbar (No-Match) und der nächste Schritt des Matchingverfahrens kommt nicht zur Anwendung.

Existiert kein Versichertenstammdatensatz, der die Bedingungen 1 und 2 erfüllt, so wird für ihn der im Folgenden beschriebene abschließende Schritt des Matchingverfahrens durchgeführt.

Stufe V: Ergänzung durch Matching-Ergebnisse der Vorjahres-GSP

Zum Teil werden Versicherte von den Kostenträgern rückwirkend nur mit Daten zu neuen Versichertennummern gemeldet, obwohl sie im Berichtszeitraum noch im Besitz der alten Krankenversichertenkarte (KVK) waren und daher noch Fälle über die darauf gespeicherte alte Versichertennummer abgerechnet wurden. Solche Abrechnungsfälle können durch die Stufen I–IV des Matchingverfahrens unter Umständen nicht zugeordnet werden. Zudem kommt es ohnehin nach einem Wechsel der KVK zu Problemen bei der Zuordnung, wenn dieser Wechsel mit dem Beginn des Berichtszeitraumes der GSP zusammenfällt, da Abrechnungsfälle prozessbedingt zum Teil erst mit einem Quartal Verzögerung erfasst werden.

Aus diesen Gründen wird nach den Matchingstufen I–IV noch folgender ergänzender

¹⁶Das bedeutet, der Abrechnungsfall darf nicht zeitlich vor der Fusionierung stattgefunden haben.

¹⁷Die Bedingungen an einen akzeptierten Versichertenstammdatensatz sind somit strenger als an die Versichertenstammdatensätze, die auf eine abweichende Personen-ID hin untersucht werden. Auf diese Weise wird die Gefahr einer fälschlichen Zuordnung weiter reduziert. Insbesondere gilt, dass eine fälschliche Zuordnung eines Abrechnungsfalles in der Stufe IV ausgeschlossen ist, wenn ein Versichertenstammdatensatz mit der „richtigen“ Personen-ID einmal für ein beliebiges Quartal für einen Kostenträger bis zur höchsten Fusionsstufe für diesen Abrechnungsfall gefunden wurde.

Schritt durchgeführt: Wenn ein Abrechnungsfall (Satzart 202)

1. in den Matchingstufen I–IV keiner Personen-ID zugeordnet werden konnte,
2. in den Matchingstufen I–IV nicht als No-Match markiert wurde,
3. in den zugehörigen Abrechnungsdaten eine Kombination aus KVK-IK und KVK-Versichertennummer aufweist, der im Rahmen des Matchingverfahrens für die vorhergehende GSP in einem Quartal, welches nicht nach dem Leistungsquartal des Abrechnungsfalles liegt, eine Personen-ID zugeordnet wurde und
4. diese Personen-ID in den Versichertenstammdaten (Satzart 201) der aktuellen GSP vorkommt,

dann wird der Abrechnungsfall dieser Personen-ID zugeordnet. Als Matchingstufe des Abrechnungsfalles wird die Stufe eines Abrechnungsfalles aus der vorhergehenden GSP verwendet, der dieser Personen-ID zugeordnet ist. Gibt es mehr als einen solchen Abrechnungsfall in der vorhergehenden GSP, so wird der Fall mit der kleinsten Matchingstufe verwendet.

6.3.5 Matching-Quoten

Die Matching-Quote gibt den Anteil der in der Satzart 202 der Geburtstagsstichprobe gelieferten Abrechnungsfälle an, die durch das Matchingverfahren einem Versicherten zugeordnet werden konnten. In Tabelle 12 sind die Matching-Quoten pro Quartal für die einzelnen Matchingstufen angegeben. Dabei wird jeweils der Anteil der auf dieser Stufe zusätzlich zugeordneten Abrechnungsfälle ausgewiesen. Es fällt besonders auf, dass im Verlauf der Jahre 2016–2018 der Anteil der Abrechnungsfälle, die bereits auf Stufe I zugeordnet werden konnten, stark angestiegen ist. Tabelle 13 liefert die Matching-Quoten (über alle Stufen) pro Quartal auf KV-Ebene, wobei der angegebene KV-Bezirk der KV entspricht, über die der Fall abgerechnet wurde. In Tabelle 14 werden schließlich die Matching-Quoten (über alle Stufen) pro Quartal für die einzelnen Kassenarten angegeben.

Für Leser, die mit der GSPR 2017 vertraut sind, sei an dieser Stelle darauf hingewiesen, dass die Abweichungen zwischen den hier berichteten Matching-Quoten und denen in der Dokumentation der GSPR 2017 berichteten für die Quartale der Jahre 2016 und 2017 darin begründet sind, dass den Matching-Quoten der GSPR 2018 ein Matching mit den Daten der Jahre 2015–2018 zugrunde liegt, wohingegen die Matching-Quoten der GSPR 2017 auf einem Matching mit den Daten der Jahre 2014–2017 basieren.

Quartal	Anteil an Fällen					Gesamt
	Stufe I	Stufe II	Stufe III	Stufe IV	Stufe V	
2016/1	99,31 %	0,08 %	0,01 %	0,01 %	0,01 %	99,42 %
2016/2	99,37 %	0,07 %	0,01 %	0,01 %	0,01 %	99,46 %
2016/3	99,40 %	0,06 %	0,01 %	0,01 %	0,01 %	99,49 %
2016/4	99,42 %	0,05 %	0,01 %	0,01 %	0,00 %	99,49 %
2017/1	99,46 %	0,05 %	0,00 %	0,01 %	0,00 %	99,52 %
2017/2	99,47 %	0,04 %	0,00 %	0,01 %	0,00 %	99,52 %
2017/3	99,48 %	0,04 %	0,00 %	0,01 %	0,00 %	99,54 %
2017/4	99,54 %	0,04 %	0,00 %	0,01 %	0,00 %	99,59 %
2018/1	99,63 %	0,01 %	0,00 %	0,00 %	0,00 %	99,63 %
2018/2	99,63 %	0,00 %	0,00 %	0,00 %	0,00 %	99,64 %
2018/3	99,63 %	0,00 %	0,00 %	0,00 %	0,00 %	99,63 %
2018/4	99,63 %	0,00 %	0,00 %	0,00 %	0,00 %	99,64 %

Tabelle 12: KV-übergreifende Matching-Quoten pro Quartal auf den einzelnen Matchingstufen und insgesamt

KV	2016/1	2016/2	2016/3	2016/4	2017/1	2017/2	2017/3	2017/4	2018/1	2018/2	2018/3	2018/4
Baden-Württt.	99,33 %	99,35 %	99,39 %	99,41 %	99,42 %	99,44 %	99,45 %	99,45 %	99,58 %	99,58 %	99,58 %	99,60 %
Bayerns	99,09 %	99,11 %	99,18 %	99,14 %	99,18 %	99,14 %	99,16 %	99,53 %	99,60 %	99,62 %	99,60 %	99,60 %
Berlin	99,65 %	99,67 %	99,66 %	99,67 %	99,70 %	99,71 %	99,69 %	99,70 %	99,66 %	99,68 %	99,69 %	99,69 %
Brandenburg	99,87 %	99,86 %	99,87 %	99,87 %	99,86 %	99,87 %	99,87 %	99,86 %	99,86 %	99,86 %	99,87 %	99,86 %
Bremen	99,09 %	99,16 %	99,24 %	99,30 %	99,29 %	99,33 %	99,33 %	99,29 %	99,23 %	99,23 %	99,24 %	99,20 %
Hamburg	99,13 %	99,30 %	99,34 %	99,40 %	99,41 %	99,45 %	99,50 %	99,50 %	99,50 %	99,54 %	99,53 %	99,50 %
Hessen	99,48 %	99,58 %	99,58 %	99,58 %	99,58 %	99,59 %	99,60 %	99,59 %	99,49 %	99,50 %	99,52 %	99,52 %
Mecklenburg-Vorp.	99,79 %	99,80 %	99,78 %	99,81 %	99,80 %	99,80 %	99,80 %	99,84 %	99,84 %	99,83 %	99,81 %	99,82 %
Niedersachsen	99,55 %	99,58 %	99,59 %	99,61 %	99,61 %	99,63 %	99,64 %	99,64 %	99,66 %	99,64 %	99,63 %	99,65 %
Nordrhein	99,25 %	99,26 %	99,30 %	99,26 %	99,35 %	99,36 %	99,39 %	99,37 %	99,46 %	99,48 %	99,47 %	99,48 %
Rheinland-Pfalz	99,64 %	99,65 %	99,65 %	99,65 %	99,68 %	99,68 %	99,68 %	99,70 %	99,71 %	99,71 %	99,67 %	99,70 %
Saarland	99,59 %	99,63 %	99,64 %	99,65 %	99,65 %	99,65 %	99,64 %	99,63 %	99,64 %	99,55 %	99,63 %	99,64 %
Sachsen	99,89 %	99,90 %	99,90 %	99,90 %	99,90 %	99,90 %	99,91 %	99,90 %	99,91 %	99,91 %	99,91 %	99,90 %
Sachsen-Anhalt	99,89 %	99,89 %	99,88 %	99,89 %	99,89 %	99,89 %	99,89 %	99,90 %	99,89 %	99,90 %	99,89 %	99,90 %
Schleswig-Holstein	99,34 %	99,34 %	99,36 %	99,37 %	99,39 %	99,41 %	99,42 %	99,43 %	99,60 %	99,58 %	99,58 %	99,60 %
Thüringen	99,75 %	99,75 %	99,75 %	99,77 %	99,76 %	99,77 %	99,78 %	99,80 %	99,81 %	99,83 %	99,83 %	99,83 %
Westfalen-Lippe	99,28 %	99,41 %	99,46 %	99,51 %	99,55 %	99,58 %	99,60 %	99,61 %	99,67 %	99,66 %	99,65 %	99,66 %

Tabelle 13: Matching-Quoten pro Quartal für die einzelnen KV-Bezirke, über die die jeweiligen Fälle abgerechnet wurden.

Kassenart	2016/1	2016/2	2016/3	2016/4	2017/1	2017/2	2017/3	2017/4	2018/1	2018/2	2018/3	2018/4
AOK	99,42 %	99,46 %	99,48 %	99,47 %	99,51 %	99,52 %	99,53 %	99,58 %	99,63 %	99,63 %	99,62 %	99,63 %
BKK	99,46 %	99,50 %	99,53 %	99,53 %	99,57 %	99,57 %	99,58 %	99,66 %	99,72 %	99,72 %	99,72 %	99,72 %
EK	99,38 %	99,42 %	99,45 %	99,45 %	99,47 %	99,47 %	99,49 %	99,55 %	99,58 %	99,58 %	99,58 %	99,59 %
IKK	99,57 %	99,61 %	99,64 %	99,65 %	99,67 %	99,68 %	99,69 %	99,72 %	99,75 %	99,75 %	99,74 %	99,75 %
KN	99,46 %	99,55 %	99,60 %	99,61 %	99,62 %	99,66 %	99,67 %	99,68 %	99,69 %	99,70 %	99,71 %	99,72 %
LKK	99,56 %	99,63 %	99,66 %	99,66 %	99,67 %	99,68 %	99,70 %	99,75 %	99,77 %	99,78 %	99,77 %	99,79 %

Tabelle 14: Matching-Quoten pro Quartal für die einzelnen Kassenarten

6.4 Verbleibende Mängel der Datenqualität der zusammengeführten Daten

Nach der Zusammenführung der kassen- und arztseitigen Datenlieferungen (siehe Abschnitt 6.3) werden der entstandene Datenkörper geprüft und abgestimmte Qualitätssicherungsmaßnahmen (QS-Maßnahmen) durchgeführt.

Im Vordergrund der Qualitätssicherungs-Maßnahmen steht die Sicherstellung der Repräsentativität der Versichertenstichprobe im Vergleich zur Grundgesamtheit. Die Repräsentativität soll hoch sein und wenn möglich mit geeigneten Mitteln weiter erhöht werden. Ursächlich für eine eingeschränkte Repräsentativität können Probleme in den gelieferten Daten sowie Probleme bei der Datenzusammenführung sein. Darüber hinaus wird im Rahmen der Qualitätssicherung auch die Plausibilität der Daten geprüft.

Bei der Qualitätssicherung der Geburtstagsstichprobe auffällig gewordene Datensätze werden markiert. Die Markierung ist notwendig, da diese Datensätze unter Umständen Auswertungsergebnisse verzerren können. In den Abschnitten 2.3.2 und 2.3.3 wurden Gründe genannt, warum Datensätze der Satzart DS201 markiert werden. Wie dort beschrieben, wurden diese Gründe im Rahmen einer sorgfältigen Datenprüfung der Geburtstagsstichprobe identifiziert und die Markierungen von Datensätzen der Satzart DS201 in der GSPR 2018 entsprechen den Markierungen von Datensätzen der Satzart 201 in der Geburtstagsstichprobe. Allerdings lässt die Markierung eines Datensatzes in der GSPR 2018 keinen Rückschluss auf die Gründe zu, aus denen der Datensatz markiert wurde.

In der Geburtstagsstichprobe gibt es zudem zwei weitere Gründe für die Markierung eines Datensatzes: Erstens werden alle Datensätze von Versicherten markiert, für die die Angaben zu Geschlecht oder Geburtsquartal in den Versichertenstammdaten nicht eindeutig sind. Zweitens werden alle Datensätze von Versicherten markiert, deren Geburtsquartal nach dem Berichtszeitraum liegt. Die aus diesen beiden Gründen markierten Datensätze der Geburtstagsstichprobe werden nicht mit der GSPR 2018 ausgeliefert.

Für eine genaue Beschreibung der Gründe für Markierungen von Datensätzen der Satzart DS201 der GSPR 2018 bzw. der Gründe für das Markieren von Datensätzen der Satzart 201 der Geburtstagsstichprobe sei auf Abschnitt 2.3.2 verwiesen.

6.5 Identifizierung von Teilnehmern an Selektivverträgen

Zur quartalsbezogenen Feststellung, ob ein Versicherter in der Geburtstagsstichprobe an einem oder mehreren Selektivverträgen teilnimmt, werden im Rahmen der Qualitätssicherung der Daten der GSP neben den Angaben in der Satzart 201 zur Teilnahme an einem bereinigungsrelevanten Selektivvertrag auch Informationen aus den Datenlieferungen zu Selektivverträgen gemäß Beschluss des Erweiterten Bewertungsausschusses in seiner 51., 56. und 61. Sitzung (im Folgenden als Selektivvertragsdaten bezeichnet) herangezogen. Diese zusätzlichen Informationen umfassen:

1. Die Existenz eines Datensatzes zu dem Versicherten in der Satzart 004 der Selektivvertragsdaten (kassenseitige Lieferung der Teilnehmer an Selektivverträgen unter den Versicherten in der Geburtstagsstichprobe)
2. Die Existenz eines Datensatzes zu dem Versicherten in der Satzart 014 der Selektivvertragsdaten (arztseitige Lieferung der Teilnehmer an Selektivverträgen unter den Versicherten in der Geburtstagsstichprobe)
3. Das Datum des Beginns und des Endes der Teilnahme des Versicherten am jeweiligen Selektivvertrag laut den für ihn vorliegenden Datensätzen der Satzart 004 der Selektivvertragsdaten

Nach Feststellung der Teilnahme eines Versicherten an einem Selektivvertrag liefert die Satzart 001 dann in der großen Mehrzahl der Fälle auch die Art des Selektivvertrags. Die so zusammengetragenen Informationen werden im Feld 16 der Satzart DS201 der GSPR 2018 eingetragen und gegebenenfalls die Information zur Vollständigkeit des Versorgungsauftrages in Feld 18 angepasst¹⁸.

A Anhang

A.1 Glossar

Zur besseren Übersicht für den Leser wurden im Folgenden für einige der in diesem Dokument verwendeten zentralen Begriffe die bereits im Text gegebenen Definitionen zusammengetragen.

Ausschöpfungsgrad des Leistungsbedarfs: Als Ausschöpfungsgrad wird der Quotient aus dem durchschnittlichen Leistungsbedarf je Versicherten in der Geburtstagsstichprobe und dem durchschnittlichen Leistungsbedarf aller Versicherten bezeichnet (evtl. nach Abgrenzung von bestimmten Versichertengruppen wie z. B. alle Versicherten in einem KV-Bezirk). Der Ausschöpfungsgrad setzt damit den Leistungsbedarf pro Kopf der Geburtstagsstichprobe ins Verhältnis zum Leistungsbedarf pro Kopf aller Versicherten. Ein Ausschöpfungsgrad von kleiner (größer) als 100% zeigt damit an, dass der durchschnittliche Leistungsbedarf je Versicherten in der Geburtstagsstichprobe kleiner (größer) ist als der durchschnittliche Leistungsbedarf aller Versicherten.....18, 20, 30

Falschmatcher: Anhand von spezifischen Gebührenordnungspositionen und/oder bestimmten Diagnosen kann auf das vermeintliche Geschlecht des Versicherten geschlossen werden. Stimmt nun das in den Versichertenstammdaten angegebene Geschlecht des Versicherten nicht mit dem laut Abrechnungsfall zu vermutenden Geschlecht des Versicherten überein, so muss davon ausgegangen werden, dass dieser Versicherte dem Abrechnungsfall fälschlicherweise zugeordnet wurde. Ein

¹⁸Siehe Anmerkung zu Feld 18 in Abschnitt 5.1

- Versicherter mit mindestens einem in diesem Sinne fälschlich zugeordneten Abrechnungsfall wird als Falschmatcher bezeichnet. 17, 84, 85, 87
- Fusionskasse:** Krankenkasse, die zu einem bestimmten Stichtag existiert. In der GSPR 2018 wird hierfür der Stichtag 02.04.2020 verwendet, an dem insgesamt 105 Fusionskassen existierten. 6, 12
- Inanspruchnahmequote:** Zwar wird allen in der GSPR 2018 berichteten Abrechnungsfällen eine Personen-ID zugeordnet, aber selbstverständlich weist nicht jeder der in den Versichertenstammdaten berichteten Versicherten in jedem Quartal auch mindestens einen Abrechnungsfall auf. Die Inanspruchnahmequoten geben pro KV und Quartal den Anteil der Versicherten mit mindestens einem Abrechnungsfall an allen Stichprobenversicherten der GSPR 2018 an. In die Berechnung der Inanspruchnahmequote werden nur Versicherte einbezogen, die in dem Quartal nicht als auffällig markiert wurden. 20, 28
- Lieferquote der Versichertenstammdaten:** Als Lieferquote wird der Anteil der Versicherten in der Geburtstagsstichprobe, die nicht als auffällig markiert wurden, an der Menge aller GKV-Versicherten bezeichnet (evtl. nach Abgrenzung von bestimmten Versichertengruppen wie z. B. alle Versicherten in einem KV-Bezirk). Darüber hinaus wird auch die ursprüngliche Lieferquote betrachtet, in die auch die als auffällig markierten Versicherten einbezogen werden. Die Gesamtzahl der Versicherten einer Gruppe wird dabei auf Grundlage der von den Krankenkassen gelieferten ANZVER87a-Daten ermittelt. 20
- Matching-Quote:** Im vorliegenden Dokument beschreibt der Begriff Matching die Zuordnung von arztseitig gelieferten Abrechnungsfällen zu den kassenseitig gelieferten Versichertenstammdaten. Ziel des Matchingverfahrens ist es dabei, allen Abrechnungsfällen eine eindeutige Personen-ID der Versichertenstammdaten zuzuordnen. Diejenigen Abrechnungsfälle, denen keine Personen-ID aus dem entsprechenden Jahr zugeordnet werden kann, sind nicht Bestandteil der GSPR 2018. Die Matching-Quote gibt den Anteil der Abrechnungsfälle, denen eine Personen-ID zugeordnet werden kann, an allen arztseitig gelieferten Abrechnungsfällen an. Dem Matching liegen dabei grundsätzlich Daten der Berichtsjahre 2015–2018 zugrunde. 3, 77–80
- Nichtinanspruchnehmer:** Ein Versicherter wird in einem Quartal als Nichtinanspruchnehmer bezeichnet, wenn für ihn in diesem Quartal kein Abrechnungsfall vorliegt. 72
- Teilkasse:** Eine ehemals autonom existierende Krankenkasse, die im Berichtszeitraum oder auch davor mit einer anderen Krankenkasse fusioniert hat, wird in diesem

Dokument als Teilkasse bezeichnet. Darüber hinaus wird der Begriff Teilkasse in diesem Dokument auch für eine untergeordnete Einheit einer Krankenkasse verwendet, repräsentiert durch ein Institutionskennzeichen aus der Satzart DS219 (z. B. Geschäftsstellen-IK, Praxisnetz-IK). 7, 12

A.2 Geschlechtsspezifische Abrechnungsfälle

Im Rahmen der Datenprüfung und der QS-Maßnahmen wird der auf Quartal und Geburtskalendertag bezogene Anteil an Falschmatchern je Teilkasse bestimmt. Dabei gilt ein Versicherter als Falschmatcher, wenn bei mindestens einem dem Versicherten zugeordneten Abrechnungsfalle die gebührenordnungspositionen- bzw. diagnosenspezifische Geschlechtsangabe (Fallgeschlecht) im Widerspruch zu dem für den Versicherten angegebenen Geschlecht steht.

Zur Bestimmung des Fallgeschlechts werden die über Tabelle A1 ableitbaren geschlechtsspezifischen Diagnosen und die in Tabelle A2 aufgeführten geschlechtsspezifischen Gebührenordnungspositionen verwendet.

Erste Stellen des ICD-Codes	Geschlecht	Erste Stellen des ICD-Codes	Geschlecht	Erste Stellen des ICD-Codes	Geschlecht
A34	weiblich	F53	weiblich	Q978	weiblich
B260	männlich	I861	männlich	Q979	weiblich
B373	weiblich	I863	weiblich	Q98	männlich
C51	weiblich	L291	männlich	R86	männlich
C52	weiblich	L292	weiblich	R87	weiblich
C53	weiblich	L705	weiblich	S312	männlich
C54	weiblich	M800	weiblich	S313	männlich
C55	weiblich	M801	weiblich	S314	weiblich
C56	weiblich	M810	weiblich	S3732	männlich
C57	weiblich	M811	weiblich	S3733	männlich
C58	weiblich	M830	weiblich	S374	weiblich
C60	männlich	N4	männlich	S375	weiblich
C61	männlich	N5	männlich	S376	weiblich
C62	männlich	N7	weiblich	S3782	männlich
C63	männlich	N8	weiblich	S3783	männlich
C796	weiblich	N90	weiblich	S3784	männlich
D06	weiblich	N91	weiblich	S3980	männlich
D070	weiblich	N92	weiblich	T192	weiblich

Erste Stellen des ICD- Codes	Geschlecht	Erste Stellen des ICD- Codes	Geschlecht	Erste Stellen des ICD- Codes	Geschlecht
D071	weiblich	N93	weiblich	T193	weiblich
D072	weiblich	N94	weiblich	T833	weiblich
D073	weiblich	N95	weiblich	Z014	weiblich
D074	männlich	N96	weiblich	Z123	weiblich
D075	männlich	N97	weiblich	Z124	weiblich
D076	männlich	N98	weiblich	Z125	männlich
D176	männlich	N992	weiblich	Z301	weiblich
D25	weiblich	N993	weiblich	Z303	weiblich
D26	weiblich	O	weiblich	Z305	weiblich
D27	weiblich	P546	weiblich	Z311	weiblich
D28	weiblich	Q50	weiblich	Z312	weiblich
D29	männlich	Q51	weiblich	Z32	weiblich
D39	weiblich	Q52	weiblich	Z33	weiblich
D40	männlich	Q53	männlich	Z34	weiblich
E28	weiblich	Q54	männlich	Z35	weiblich
E29	männlich	Q55	männlich	Z39	weiblich
E894	weiblich	Q96	weiblich	Z4001	weiblich
E895	männlich	Q970	weiblich	Z437	weiblich
F524	männlich	Q971	weiblich	Z443	weiblich
F525	weiblich	Q972	weiblich	Z875	weiblich

Tabelle A1: Geschlechtsspezifische Diagnosen zur Identifizierung von Falschmatchern

GOP	Geschlecht	GOP	Geschlecht	GOP	Geschlecht
01730	weiblich	01829	weiblich	31302	weiblich
01731	männlich	01830	weiblich	31303	weiblich
01733	weiblich	01831	weiblich	31304	weiblich
01735	weiblich	01832	weiblich	31305	weiblich
01750	weiblich	01833	weiblich	31306	weiblich
01752	weiblich	01835	weiblich	31307	weiblich
01753	weiblich	01836	weiblich	31696	weiblich
01754	weiblich	01837	weiblich	31697	weiblich
01755	weiblich	01838	weiblich	31698	weiblich

GOP	Geschlecht	GOP	Geschlecht	GOP	Geschlecht
01756	weiblich	01839	weiblich	31699	weiblich
01757	weiblich	01840	weiblich	31700	weiblich
01758	weiblich	01854	männlich	31701	weiblich
01759	weiblich	01855	weiblich	31702	weiblich
01770	weiblich	01912	weiblich	33044	weiblich
01772	weiblich	01913	weiblich	33062	männlich
01773	weiblich	01915	weiblich	33064	männlich
01774	weiblich	08230	weiblich	33074	weiblich
01775	weiblich	08231	weiblich	01734M	männlich
01780	weiblich	08310	weiblich	01740M	männlich
01781	weiblich	08311	weiblich	01741M	männlich
01782	weiblich	08320	weiblich	01742M	männlich
01783	weiblich	08330	weiblich	01743M	männlich
01784	weiblich	08331	weiblich	01745M	männlich
01785	weiblich	08332	weiblich	01746M	männlich
01786	weiblich	08333	weiblich	01700V	weiblich
01787	weiblich	08334	weiblich	01701V	weiblich
01793	weiblich	08340	weiblich	32540V	weiblich
01800	weiblich	08341	weiblich	32541V	weiblich
01801	weiblich	08345	weiblich	32542V	weiblich
01802	weiblich	08410	weiblich	32543V	weiblich
01803	weiblich	08411	weiblich	32544V	weiblich
01804	weiblich	08412	weiblich	32545V	weiblich
01805	weiblich	08413	weiblich	32546V	weiblich
01806	weiblich	08414	weiblich	32550V	weiblich
01807	weiblich	08415	weiblich	32551V	weiblich
01808	weiblich	08416	weiblich	32552V	weiblich
01809	weiblich	08530	weiblich	32553V	weiblich
01810	weiblich	08531	weiblich	32554V	weiblich
01811	weiblich	08541	weiblich	32555V	weiblich
01812	weiblich	08542	weiblich	32556V	weiblich
01815	weiblich	08550	weiblich	33043V	weiblich
01816	weiblich	08551	weiblich	33044V	weiblich
01817	weiblich	08552	weiblich	33090V	weiblich

GOP	Geschlecht	GOP	Geschlecht	GOP	Geschlecht
01818	weiblich	08560	weiblich	01740P	weiblich
01822	weiblich	08561	weiblich	01740N	männlich
01825	weiblich	26310	männlich	01746P	weiblich
01826	weiblich	26311	weiblich	01746N	männlich
01827	weiblich	26341	männlich		
01828	weiblich	31301	weiblich		

Tabelle A2: Geschlechtsspezifische Gebührenordnungspositionen zur Identifizierung von Falschmatchern