

Projekt-Endbericht

OF-LÖG

Outcome-Forschung

basierend auf Leistungsdaten

des Österreichischen Gesundheitswesens

Arbeitsgemeinschaft

Medizinische Universität Wien

Univ.-Prof. DDr. Wolfgang DORDA

Ao. Univ.-Prof. Dr. Georg DUFTSCHMID

Ao. Univ.-Prof. Dr. Walter GALL

Universität Wien

Univ.-Prof. Dr. Wilfried GROSSMANN

Ao. Univ.-Prof. Dr. Karl FRÖSCHL

unter Mitarbeit von DI Dr. Milan Hronský

mit 53 Tabellen und 66 Abbildungen

Wien, im Juni 2011

Inhaltsübersicht

Einleitung	1
 Teilprojekt 1: Datenqualität Case-Mix GAP-DRG Daten und Vergleich mit Bevölkerungsdaten von Statistik Austria	
1. Beobachtungseinheiten	6
2. Bewertung der Qualität von Attributen der CASE-Mix Datei	17
3. Analysen mehrerer Variablen	29
Zusammenfassung	55
 Teilprojekt 2: Diabetes – Epidemiologische Analysen	
1. Ziel	58
2. Resultate	58
 Teilprojekt 3: Epidemiologie – Chronischer Schmerz	
1. Vorbemerkungen	81
2. Daten Hauptverband (HV)	83
3. Daten Gesundheitsbefragung (GB)	89
4. Vergleich Hauptverband – Gesundheitsbefragung (HV—GB)	93
5. Medikamentenkosten (HV)	96
6. Prävalenzen – Grafiken (HV—GB)	96
 Annex: Einige Anmerkungen zum GAP-DRG Datenmodell	110

Einleitung

Das Data Warehouse des Hauptverbands der Sozialversicherungsträger umfasst eine in Österreich einzigartige Sammlung patientenbezogener Leistungsdaten des Gesundheitswesens und bietet somit einen hervorragenden Ausgangspunkt für die epidemiologische Analyse im Sinne der *Evidence Based Medicine* (im Folgenden kurz: EBM). Gleichzeitig stellt sich natürlich auch die Frage, inwieweit die nach den internen Notwendigkeiten und Gegebenheiten des österreichischen Sozialversicherungssystems organisierten Leistungsdatenbestände den demographischen Erhebungen der Statistik Austria (insb. Zensusdaten) entsprechen bzw. welche Möglichkeiten sich für eine Verbesserung der Datenqualität aus einem methodischen Vergleich beider Datenkörper eröffnen. Zudem ist festzuhalten, dass das Data Warehouse des Hauptverbands seinerseits bereits aus einer Zusammenführung unterschiedlicher Datenquellen hervorgeht. Diese Quellen liefern zum überwiegenden Teil Transaktionsdaten, die durch inhaltliche administrative ebenso wie durch räumliche Kriterien definiert werden. Daneben spielen noch verschiedene Verzeichnisse im Sinne von Stammdaten eine wesentliche Rolle (wie etwa die Codierung der medizinischen Wirkstoffe ATC, Taxonomien zur Codierung medizinischer Einzelleistungen oder Listen medizinischer Leistungsträger, Vertragspartner usw.usf.). Es gehört zu den Eigenarten (auch) des österreichischen Gesundheitsversorgungssystems, dass die Datenbestände nach teilweise erheblich unterschiedlichen Gesichtspunkten erfasst und zusammengestellt werden. Hinzu kommt noch, dass einzelne Attribute z.B. von Patienten in verschiedenen Verrechnungskreisen auch mitunter unterschiedliche Ausprägungen annehmen können; bspw. kann die Angabe des politischen Bezirks zu einem Patienten aus sachlogischen Gründen – etwa infolge unterschiedlicher Verrechnungsstellen für medizinische Leistungen – in den einzelnen Datenquellen voneinander abweichend erfasst sein.

Für eine konsistente Nutzung dieses Datenbestands im Sinne der EBM ist daher die Herbeiführung eines *konsolidierten* Datenbestands aus diesen Quellen mit möglichst hohem Qualitätsstandard von vorrangigem Interesse. Den natürlichen Ausgangspunkt dieser Konsolidierung bildet dabei das Personen- bzw. Versichertenregister des Hauptverbands der Sozialversicherungsträger, das als wesentliche demographische Merkmale (Attribute) Angaben zu Geschlecht, Geburtsjahr, Staatsbürgerschaft, Familienstand sowie Ein- und Austrittszeiten bei einzelnen gesetzlichen Sozialversicherungsträgern umfasst. Als weiteres wichtiges analytisches Kriterium tritt die räumliche Gliederung dieser Personenstandsdaten hinzu.

Damit eignet sich dieses Register grundsätzlich zur Ermittlung sachlogisch abgrenzbarer stichtags- oder zeitraumbezogener Kollektive von Merkmalsträgern (Personen bzw. Patienten) als Basis weiterer Analysen im Sinne der EBM. Für derartige spezifische Fragestellungen werden die konsolidierten Registerdaten mit Daten aus weiteren Datenbeständen, die Auskunft über erbrachte medizinische Leistungen (z.B. Anstaltsaufenthalte, Medikationen, Krankenstandsmeldungen usw.) geben, je nach Verfügbarkeit und Anlassfall verknüpft, wobei in einer Reihe von Fällen vorteilhaft auf durchgehende Pseudonymisierungen der Patientendatensätze zurückgegriffen werden kann.

Der Aufbau eines derartigen konsolidierten Datenbestands setzt eine wohlüberlegte Planung der Integration der einzelnen Datenquellen voraus, die auch eine Be-

wertung der Qualität der einzelnen einbezogenen Datenquellen in Hinblick auf die Palette der Auswertungsziele miteinschließt. Erstrebenswert ist diesbzgl. ein *best practice model* – „good evidence practice“ – mit der Perspektive einer wiederholten Verwendbarkeit für unterschiedliche Arten von Fragestellungen. In diesem Sinne schlägt das Projekt OF-LÖG eine mehrstufige Vorgangsweise folgender Form vor:

- Den Ausgangspunkt bildet naturgemäß die im Hauptverband zur Verfügung stehende Menge der Datenquellen. Unbeschadet der bereits im GAP-DRG Datenschema vorgenommenen Zusammenführungen ist die Gewinnung einer Übersicht über die internen Beziehungen und Strukturen der Datenmodelle und Produktionsformen der einzelnen Datenquellen im Sinne einer *data provenance* für die anschließende Beurteilung der Datenqualität unverzichtbar. Eine solche *Inventur* versteht sich klarer Weise ergänzend zur bereits im Hauptverband vorliegenden Dokumentation, hält sich aber die Option zur Formulierung alternativer Datenmodelle und Integrationsansätze offen.
- Auf Grundlage der Analyse der *data provenance* vollzieht sich der Aufbau eines Datenschemas bzw. einer Datenbank unter Einbeziehung aller für die angepeilten EBM-Fragestellungen relevanten Merkmale, wie sie sich aus dem Integrationskonzept der verfügbaren Datenquellen ergeben. Formal gesehen realisiert sich die Integration über gemeinsame Schlüsselmerkmale der zu verbindenden Datenquellen, die entweder deterministisch (insb. über sog. *equi-joins*) oder über dazwischen geschaltete *record linkage*-Verfahren auf Datensatz-Ebene gekoppelt werden. Die solchermaßen abgeleiteten integrierten Kollektive stellen dann die Grundlage für die Aggregation dar, die in der Regel zu mehr oder weniger hochdimensionalen Elementar-Aggregationen (sog. „Daten-Würfel“) führt, aus denen je nach Bedarf und Datenverteilung zur Analyse und Präsentation geeignete Verdichtungen (Gruppierungen, Selektionen, Projektionen usw.) gewonnen werden. Die realisierte Methode der Datenintegration ist als wesentliche Maßnahme der Sicherung der Datenqualität (Nachvollziehbarkeit) im Detail zu dokumentieren. Im GAP-DRG Ansatz erfolgt die Verknüpfung personen-spezifischer Datensätze in aller Regel über die in den Datenquellen bereits verfügbaren Versicherten-Pseudonyme; im Falle der aus den Abrechnungen der Krankenanstalten stammenden sog. MBDS (minimum basic data set)-Datensätze ist ein komplexeres *record linkage*-Verfahren zur (Pseudo-)Reidentifikation der Zuordnung zu Versicherten unvermeidlich, das dem Projekt OF-LÖG aber bereits zur Verfügung gestanden ist.
- Der vorgenannte Schritt des Verfahrens zur Datenintegration bezieht sich auf eine vorerst formale Zusammenführung der Datenquellen, die per se noch keine Gewähr für die interne *Konsistenz* des Datenbestands leistet. Dies betrifft vornehmlich zwei Qualitätsaspekte:
 - *Fehlende Werte*: bedingt durch die Integration unterschiedlicher (und daher vielfach heterogener) Datenquellen ist eine statistische Analyse des Ursprungs fehlender Werte essentiell, um auf dieser Analyse aufbauend nach Möglichkeit Imputationsverfahren geeignet einsetzen zu können;
 - *Widersprüchliche Werte*: im Falle von in mehreren Datenquellen gleichzeitig auftretenden Merkmalen sind (Edit-)Regeln zur Auflösung inkonsistenter Ausprägungen (z.B. unterschiedliche Geschlechts- oder Al-

tersangaben) zu etablieren, wobei verschiedene Alternativen bestehen wie insb. Majoritätsregeln, Regeln aufgrund temporaler oder raumbezogener Präferenzen, oder auch Regeln auf Grundlage von Informationen aus der data provenance-Dokumentation.

Konsistenzprüfungen können u.a. zu Empfehlungen führen, wie einzelne (Teil-) Bereiche der Datenquellen und Datensysteme des Hauptverbands im Sinne der EBM vorteilhafter oder wirtschaftlicher als bisher administriert werden können.

- Die interne Konsistenz des integrierten Datenbestands bildet eine unverzichtbare Grundlage jeder sinnvollen Datennutzung; darüber hinaus stellt sich aber auch die Frage nach der *Kohärenz* der erreichten Konsolidierung. Angesichts der Verfügbarkeit einer Vielzahl überlappender Datenkörper medizinischer Relevanz bedeutet diese Kohärenz, dass in den Aggregaten (ab zumindest einer gewissen Aggregationshöhe) keine nennenswerten Diskrepanzen zutage treten. Mit Bezug auf das österreichische Kollektiv der Sozialversicherten ist insb. die Stimmigkeit mit den von der *Statistik Austria* erhobenen Bevölkerungsaggregaten hervorzuheben, die konkret ergänzt werden durch Stichprobenerhebungen (und darauf basierenden Hochrechnungen) aus der Gesundheitsbefragung sowie gelegentlich Datenerhebungen zu speziellen medizinischen Fragestellungen. Vor allem Zensusdaten sind naturgemäß besonders gut zur Beurteilung der Über- oder Unterdeckung der Zielpopulation geeignet, während bei spezifischeren Datenquellen i.A. lediglich die Plausibilität der Kohärenz für bestimmte Teilkollektive möglich ist. In jedem Fall sind Defizite in der Kohärenz Anlass für Ursachenforschung und, sofern möglich, die Vornahme qualitätssichernder Maßnahmen.
- Am Ende der Verfahrenskette schließlich steht die *Nutzung* des integrierten, konsolidierten und kohärenten Datenkörpers für eine Vielzahl an EBM-orientierten Analysen. Im Lichte der *statistisch-analytischen* Tönung dieser Datennutzung ist hier eine Datenschnittstelle wünschenswert, die auf möglichst komfortable Weise die Selektion statistischer Teilkollektive (merkmalsbasierte Filterung von Datensätzen) aus den (wenigen) integrierten Gesamtkollektiven – und hier in der Hauptsache wiederum das Gesamtkollektiv der Versicherten – ermöglicht. Da die betrachteten Kollektive allesamt durch eine analytisch stets relevante Dynamik gekennzeichnet sind, spielt die Nutzbarkeit temporaler Filterungen eine herausragende Rolle; Analoges gilt häufig auch für räumliche Abgrenzungen, sodass die Datenschnittstelle die interne Modellierung spatio-temporaler Merkmale tunlichst für die statistische Analytik *transparent* vorhalten sollte. Das bestehende GAP-DRG Datenmodell ist im Projekt OF-LÖG im betrachteten Datenausschnitt 2006/07 vor allem hinsichtlich dieser Kriterien kritisch zu hinterfragen gewesen.

Diesem skizzierten Verfahrensansatz und insb. den Fragen der Datenkohärenz widmete sich Teilprojekt 1 von OF-LÖG, wobei Fragen der Datenmodellierung – ohne abschließendes Fazit – im Annex dieses Berichts zusammengefasst sind. Die Teilprojekte 2 und 3 hingegen stellen konkrete Analysebeispiele – epidemiologische Analytik des Diabetes mellitus bzw. des chronischen Schmerzes – für die Nutzung des konsolidierten GAP-DRG Datenkörpers dar. Vor allem Teilprojekt 3 gibt einen

Eindruck für die Komplexität der Herleitung relevanter analytischer Kollektive, die von einer leistungsfähigen Datenschnittstelle zu bedienen sind.

Das Projekt OF-LÖG hat anhand konkreter Analysen in mehrfacher Hinsicht die Bedeutung eines konsolidierten und kohärenten Datenbestands für Fragestellungen der EBM sowie die Nützlichkeit einer effizienten Datenschnittstelle verdeutlicht und im gleichen Zug verschiedene Verbesserungsmöglichkeiten des bestehenden GAP-DRG Datenschemas sichtbar werden lassen. Die Autoren des Berichts geben daher der Hoffnung Ausdruck, mit den nachfolgenden Ausführungen und Anmerkungen einen hilfreichen Beitrag zur Verbesserung und Vereinfachung der Nutzung eines wichtigen epidemiologischen Datenkörpers und somit indirekt zur Förderung der EBM in Österreich geleistet zu haben. Aus den bisherigen Erfahrungen lässt sich schlussfolgern, dass – neben sicherlich zweckmäßigen Überarbeitungen des GAP-DRG Konzepts an sich – das Potential der Analysemöglichkeiten in keiner Weise ausgeschöpft werden konnte. Von längerfristiger Tragweite sind allerdings wohl jene aus dieser Studie ableitbaren Überlegungen, die zu einer Reorganisation der Erfassung medizinischer Leistungsdaten ab ovo Anlass geben, um eine Reihe immanenter Qualitätsdefizite der bestehenden Datenkörper hinkünftig bereits quellsseitig zu vermeiden.

Danksagungen

Das Projekt wäre in der vorliegenden Form nicht ohne die tatkräftige und kollegiale Unterstützung von Prof.es Dr. Edith Schober und Dr. Thomas Waldhör – im Bereich des Teilprojekts 2 zur Diabetes-Epidemiologie durch Auswertungen des juvenilen Diabetesregisters – sowie von Frau Dr. Gabriele Grögl-Aringer (Krankenanstaltenverbund) als Schmerz-Expertin – im Bereich des Teilprojekts 3 zur Epidemiologie des chronischen Schmerzes – über die Bühne gegangen. Die Projektgruppe dankt den genannten Personen ganz ausdrücklich für die engagierte Kooperation.

Ganz besonders gilt der Dank der Autoren Dr. Gottfried Endel (Hauptverband der Sozialversicherungsträger), der nicht nur in die Projektformulierung maßgeblich involviert war, sondern durch seine umfangreiche und profunde Sachkenntnis in medizinischer und EDV-technischer Hinsicht den Projektfortgang in erfolgsentscheidender Weise mitgestalt und mitgesteuert hat. Wir danken auch Hrn. Florian Endel für seine Hilfestellung in der Datenlogistik und zahlreiche Auskünfte betreffend die Struktur und Weiterentwicklung des GAP-DRG Datenmodells in der dem Projekt zugänglichen Implementierung an der TU Wien.

Begleitende Dokumentation

Zu allen drei Teilprojekten von OF-LÖG sind naturgemäß umfangreichere technische Dokumentationselemente (Datenbankauszüge und Rohtabellen, Datenbankabfragen und Verarbeitungs-Scripte, Grafiken in verschiedenen weiterverarbeitbaren Dateiformaten, umfangreiche Auswertungstabellen usw.) entstanden; diese sind jedoch in gesonderter Form gesammelt und der Einfachheit halber nicht in das vorliegende Print-Dokument miteinbezogen.

Teilprojekt 1

Datenqualität Case-Mix GAP-DRG Daten

und

**Vergleich mit Bevölkerungsdaten von
Statistik Austria**

1. Beobachtungseinheiten

1.1 Konzeptionelle Einheiten und Populationen

Die grundlegende Einheit von Interesse sind lebende Personen. Diese werden auf Grund von bestimmten Merkmalen zu den folgenden zwei Einheiten näher spezifiziert und für diese werden auch entsprechende Grundgesamtheiten definiert:

- *Wohnbevölkerungsperson*: Eine lebende Person ist eine Wohnbevölkerungsperson, wenn sie zu einem bestimmten Zeitpunkt oder während einer bestimmten Zeitperiode laut Bevölkerungsregister von Statistik Austria ihren Hauptwohnsitz in Österreich hat. Die daraus gebildete Population wird mit **PWB** bezeichnet.
- *Hauptverbandsperson*: Eine lebende Person ist eine Hauptverbandsperson, wenn für sie zu einem bestimmten Zeitpunkt oder während einer bestimmten Zeitperiode laut Aufzeichnungen des Hauptverbandes Leistungen erbracht/abgerechnet wurden. Die daraus gebildete Population wird mit **PHV** bezeichnet.

Im Rahmen dieser Untersuchung sind die Populationen **PWB** und **PHV** im Zeitraum 2006/07 von zentralem Interesse. Die Unterschiede zwischen diesen beiden Populationen sind primär durch die folgenden Tatsachen begründet:

- a) **PWB\PHV**: Die Untererfassung der Wohnbevölkerung im Hauptverband wird primär dadurch begründet, dass lebende in Österreich wohnhafte Personen keiner sozialversicherungs-gesetzlichen Krankenversicherung unterliegen: Dazu gehören Personen bestimmter Berufsgruppen wie z.B. Notare, Anwälte, Ziviltechniker, Ordensmitglieder, Präsenzdienler, zeitverpflichtete Soldaten, oder Versicherte von Krankenfürsorgeanstalten. Sozialhilfeempfänger sind im Hauptverband registriert.
- b) **PHV\PWB**: Die Übererfassung der Population des Hauptverbandes kann zu einem großen Teil durch die folgenden beiden Begriffe erfasst werden:
 - „Anspruchstouristen“: Urlauber, die in Österreich medizinische Leistungen in Anspruch nehmen (Personen werden im HV registriert, zählen aber nicht zur Wohnbevölkerung).
 - „Arbeitstouristen“: Personen die in Österreich keinen Hauptwohnsitz haben, auf Grund eines Beschäftigungsverhältnisses aber in Österreich sozialversichert sind (Personen werden im HV registriert, haben aber keinen Hauptwohnsitz in Österreich).

Neben diesen beiden Populationen spielt noch die Grundgesamtheit der Gesundheitsbefragung eine Rolle, die sich durch die Hochrechnung einer Stichprobe von 15.474 Personen über 15 Jahre im Zeitraum März 2006 – Februar 2007 durchgeführt wurde.

1.2 Realisierungen der Populationen in Registern

1.2.1 Wohnbevölkerungspopulation *PWB*

Da für Fragen der Evidenz basierten Medizin (nachfolgend kurz: EBM) nur Daten von gesamten Kohorten von Interesse sind, erfolgt die Realisierung der Population *PWB* im Zeitraum 2006/07 in dieser Untersuchung durch Daten der Statistik Austria, die auf der Homepage von Statistik Austria unter

<http://sdb.statistik.at/statistik.at/ext/superweb/loadDatabase.do?db=debevstand>

zur Verfügung stehen (Applikation Superstar). Als Referenzzeit wurde dabei das vierte Quartal 2006 gewählt. Dies bedeutet zwar eine gewisse Ungenauigkeit, scheint aber für die Zwecke eines ersten Vergleiches ausreichend. Ein weiterer Grund für die Wahl dieses Referenzquartals ist die Tatsache, dass die medizinisch relevanten Daten für die Population des Hauptverbandes in den weiteren Phasen des Projektes mit den Ergebnissen der Gesundheitsbefragung verglichen werden sollen. Diese Daten aus einer Stichprobe für ganz Österreich wurden im Zeitraum März 2006 bis Februar 2007 durchgeführt, so dass der Beginn des vierten Quartals etwa in der Mitte dieses Zeitraums liegt.

Die Daten der Bevölkerung stehen als Datenwürfel zur Verfügung, der durch die folgenden Dimensionen gebildet wird:

Alter × Geschlecht × Bundesland (NUTS 2)

Das Summary Attribut ist die Anzahl der lebenden Personen in jeder durch die Dimensionen definierten Zelle.

Die Gesamtzahl dieser Bevölkerung im Referenzzeitraum beträgt **8.276.466** Personen.

Um eine Vorstellung über die Schwankungen dieser Werte im Zeitverlauf zu bekommen, ist in der folgenden Tabelle 1 eine Übersicht über die Bevölkerungsentwicklung seit 2007 in den einzelnen Quartalen angegeben (Quelle Statistik Austria).

Tabelle 1: Entwicklung der Wohnbevölkerung seit 2007 nach Bundesländern

Bevölkerung zu Quartalsbeginn seit 2007 nach Staatsangehörigkeit und Bundesland

Quartal	Österreich	Burgenland	Kärnten	Niederösterreich	Oberösterreich	Salzburg	Steiermark	Tirol	Vorarlberg	Wien
Bevölkerung insgesamt										
1. Quartal 2007	8.282.984	280.062	559.393	1.588.567	1.403.663	526.048	1.202.483	697.253	364.269	1.661.246
2. Quartal 2007	8.292.023	280.110	559.721	1.590.402	1.404.317	526.301	1.203.077	697.898	364.804	1.665.393
3. Quartal 2007	8.299.411	280.548	560.268	1.593.091	1.405.608	526.252	1.203.712	697.947	365.211	1.666.774
4. Quartal 2007	8.311.573	281.038	560.615	1.596.063	1.407.043	526.932	1.204.647	698.721	365.503	1.671.011
1. Quartal 2008	8.318.592	281.185	560.322	1.596.538	1.406.664	527.452	1.204.947	700.707	365.868	1.674.909
2. Quartal 2008	8.327.174	281.472	560.324	1.598.530	1.407.908	528.045	1.205.380	701.675	366.409	1.677.431
3. Quartal 2008	8.334.059	282.146	560.648	1.600.999	1.409.416	528.174	1.205.752	701.515	366.674	1.678.735
4. Quartal 2008	8.348.027	282.950	560.899	1.604.310	1.410.614	528.565	1.207.433	702.495	367.288	1.683.473
1. Quartal 2009	8.355.260	283.118	560.605	1.605.122	1.410.403	529.217	1.207.479	704.472	367.573	1.687.271
2. Quartal 2009	8.356.874	283.122	560.054	1.605.011	1.410.193	529.305	1.206.909	704.690	367.791	1.689.799
3. Quartal 2009	8.360.219	283.417	560.287	1.606.515	1.411.164	529.053	1.207.263	704.044	367.844	1.690.632
4. Quartal 2009	8.369.878	283.931	559.922	1.608.422	1.411.988	529.391	1.208.289	704.781	368.374	1.694.780
1. Quartal 2010	8.375.290	283.965	559.315	1.607.976	1.411.238	529.861	1.208.372	706.873	368.868	1.698.822
2. Quartal 2010	8.382.602	283.916	559.054	1.608.410	1.411.450	530.472	1.208.498	707.321	369.286	1.704.195
3. Quartal 2010	8.387.142	284.263	559.237	1.609.672	1.412.534	530.332	1.209.121	706.590	369.535	1.705.858
4. Quartal 2010	8.402.549	284.922	558.997	1.612.155	1.413.936	531.155	1.211.054	707.538	369.889	1.712.903

Wie man erkennt, hat sich die Wohnbevölkerung nur geringfügig geändert. Vergleicht man beispielsweise die Wohnbevölkerung im 1. Quartal 2010 mit jener im ersten 1. Quartal 2007, so ergibt sich ein leichter Zuwachs von 1,1%, für Österreich. Die Werte schwanken etwas in den Bundesländern und sind in Wien mit einem Zuwachs von 2,26% am höchsten. Im Kontext dieser Untersuchung scheint dies aber vernachlässigbar.

1.2.2 Hauptverbandspopulation *PHV*

Für die Realisierung der Hauptverbandspopulation *PHV* wurde die GAP_DRG-Datenbank herangezogen. Diese Datenbank enthält Eintragungen über die Hauptverbandspopulation, die Einheit ist dabei eine Person. Zwei Tabellen dieser Datenbank geben Auskunft über die Basisdaten der Bevölkerung, die mit den Daten von Statistik Austria nach Aggregation verglichen werden kann. Es ist dies einerseits die Tabelle **gapdrg.personen** („Versicherte“) und die Tabelle **gapdrg.personendaten** („Versicherungsverhältnisse“) in der Version vom 30.10.2010. Die erste Tabelle enthält Basisdaten zu den einzelnen Personen, die durch die zentrale Partnerverwaltung geprüft wurden. Die folgenden Attribute aus dieser Tabelle wurden verwendet (vgl. auch semantische Beschreibung der Datenbank):

Tabelle 2: Attribute aus der Tabelle **gapdrg.personen**

P.PERS_ID	PK	Identifikationsnummer Person (numerischer Wert, aus vnump abgeleitet)
GEBJAHR		Geburtsjahr
TODDATUM		Datum des Todes
Gebjahr_mehrfach		Nur mit J befüllt, wenn beim FOKO Import mehr als ein Geburtsjahr eingetragen war, ansonsten (null)

In der zweiten Tabelle – Tabelle 3, nächste Seite – finden sich folgende Attribute (vgl. auch semantische Beschreibung der Datenbank).

In der Tabelle **gapdrg.personen** existiert für jede Person nur ein Eintrag. In der Tabelle **gapdrg.personendaten** finden sich für die einzelnen Personen mehrere Einträge, da diese Tabelle einer Aufzeichnung aller von Versicherungsträgern übermittelten Ausprägungen der Attribute von Sozialversicherungsverhältnissen einer Person entspricht.

Die Grundgesamtheit aller jener Fälle, die in der Tabelle **gapdrg.personen** enthalten sind und durch die zentrale Partnerverwaltung geprüft wurden, wird mit **PHV_0** bezeichnet. Die Größe dieser Population ist mit 11.338.569 Personen deutlich größer als die Population **PWB** von Statistik Austria. Anscheinend treten die in Abschnitt 1.1. angeführten Gründe für eine Übererfassung durch den Hauptverband wesentlich häufiger auf als die Gründe für eine Untererfassung.

Um aus diesen Daten ein Register der Hauptverbandspopulation für den Zeitraum 2006/07 zu erstellen ist zu beachten, dass diese Tabellen auch Daten von bereits verstorbene Personen enthält und auch von Personen, die nach dem 31.12. 2007 geboren wurden. Es ist daher eine Reihe von Filteroperationen notwendig, um die gewünschte Population **PHV** zu erhalten.

Tabelle 3: Attribute aus der Tabelle **gapdrg.personendaten**

ID	PK	Identifikationsnummer (Fortlaufende Nummer innerhalb der Tabelle), Autowert
PERS_ID	FK	Identifikationsnummer Person (Feld pers_id in Tabelle personen_gesamt)
VTR		Meldender Versicherungsträger Jener Träger, der diesen Datensatz gemeldet hat. Es ist nicht klar, ob dies Leistungszuständiger oder Abrechnender Versicherungsträger ist. Ist fachlich mit dem Feld Herkunft ident.
VNUMV		Pseudonym des Versicherten in Langform, ist meistens (null)
BEZIRK_ID	FK	Identifikationsnummer des Bezirks (Feld Bezirk_id in der Tabelle Bezirk)
VERSKAT	FK	Versicherungskategorie (Beschreibung siehe Tabelle Verskat)
REZGEB		Rezeptgebührenbefreiung (Ja/Nein)
VON		Von Datum (Gültigkeit der in dem Datensatz befüllten Ausprägung)
BIS		Bis Datum (Gültigkeit der in dem Datensatz befüllten Ausprägung)
HERKUNFT	PK	Jener Träger, der diesen Datensatz gemeldet hat. Ist fachlich mit dem Feld VTR ident.
GESL		Geschlecht, auch wenn mehrfach vorhanden
GEBJAHR		Geburtsjahr (wenn mehrfach vorhanden befüllt, ansonsten (null))

Bevor eine derartige Filterung durchgeführt werden kann, ist die Qualität der Sterbedaten zu untersuchen. Die folgende Tabelle 4 (nächste Seite) gibt Auskunft über die in **gapdrg.personen** eingetragenen Todesdaten nach Jahren und vergleicht diese Anzahlen mit der Anzahl der Sterbefälle von Statistik Austria seit 1991.

Wie man erkennt, sind die Todesfälle in den Daten des Hauptverbandes bis zum Jahr 2000 deutlich geringer als jene bei Statistik Austria. Danach tritt eine deutliche Verbesserung ein. Im Jahre 2006 wurde offensichtlich im HV eine Bereinigung durchgeführt. Die Anzahl der Todesfälle in den Jahren 2005, 2007 und 2008 scheinen im Vergleich mit den Daten von Statistik Austria plausibel, wenn man eine Übererfassung durch den Hauptverband unterstellt. Im Jahre 2009 wurde von Statistik Austria eine bessere Schätzung der im Ausland verstorbenen Personen durchgeführt, so dass sich das Verhältnis wieder umdreht und die Zahlen des Hauptverbandes weniger plausibel erscheinen.

Rechnet man noch die Todesfälle vor 1991 hinzu, so ergibt sich in den Daten des Hauptverbandes eine höhere Anzahl von Todesfällen. Allerdings sind hier Fälle mit Todesjahren 1916, 1921 und 1928 aufklärungsbedürftig.

Auf Grund der offensichtlichen Unsicherheit über die Sterbefälle im Jahr 2006 wurde in einer ersten Näherung die folgende Annahme zur Bestimmung der Population **PHV** für 2006/07 getroffen:

*Es werden im Folgenden nur jene Personen aus **PHV_0** betrachtet, die während des gesamten Zeitraumes 2006/07 gelebt haben (kein eingetragenes Todesdatum oder Datum des Todes nach dem 31.12.2007).*

Wir bezeichnen diese Population mit **PHV_1**. Diese umfasst 9.847.957 Personen.

Tabelle 4: Sterbefälle in **gapdrg.personen** im Vergleich zu Sterbefällen nach Statistik Austria

Todesfälle gapdrg.personen		Todesfälle Statistik Austria		Differenz
1991	34.315	1991	83.428	-49.113
1992	34.936	1992	83.162	-48.226
1993	33.981	1993	82.517	-48.536
1994	34.191	1994	80.684	-46.493
1995	35.157	1995	81.171	-46.014
1996	35.532	1996	80.790	-45.258
1997	41.364	1997	79.432	-38.068
1998	42.788	1998	78.339	-35.551
1999	49.660	1999	78.200	-28.540
2000	54.528	2000	76.780	-22.252
2001	73.314	2001	74.767	-1.453
2002	75.215	2002	76.131	-916
2003	76.978	2003	77.209	-231
2004	74.480	2004	74.292	188
2005	79.464	2005	75.189	4.275
2006	278.333	2006	74.295	204.038
2007	77.239	2007	74.625	2.614
2008	80.850	2008	75.083	5.767
2009	72.863	2009	77.381	-4.518
Gesamt	1.285.188		1.483.475	-198.287
Todesfälle vor 31.12.1990	304.940			
Total	1.590.128			

Da wir an medizinischen Interventionen interessiert sind, wird für die weiteren Einschränkungen die Information aus der Tabelle **gapdrg.personendaten** verwendet. Ein Matching der Daten aus den beiden Tabellen **gapdrg.personen** und **gapdrg.personendaten** auf Basis der Personenidentifikation zeigt, dass nur 9.838.895 Personen in beiden Tabellen zu finden sind. Die Grundgesamtheit dieser Personen wird mit **PHV_2** bezeichnet.

Eine nähere Analyse der 9.062 Fälle, die zwar in **gapdrg.personen** existieren, in **gapdrg.personendaten** aber keine Einträge haben zeigt, dass von diesen nur 361 als verstorben registriert sind. Der Todeszeitpunkt dieser Personen liegt aber nach dem 01.01.2008 und daher wurden sie im ersten Schritt nicht gefiltert.

Für jene 9.062 Personen in **PHV_1**, die nicht zu **PHV_2** gehören wurde das Geburtsjahr analysiert. Abbildung 1 zeigt Histogramme der Geburtsjahre gegliedert nach vorhandenem Eintrag eines Todesdatums.

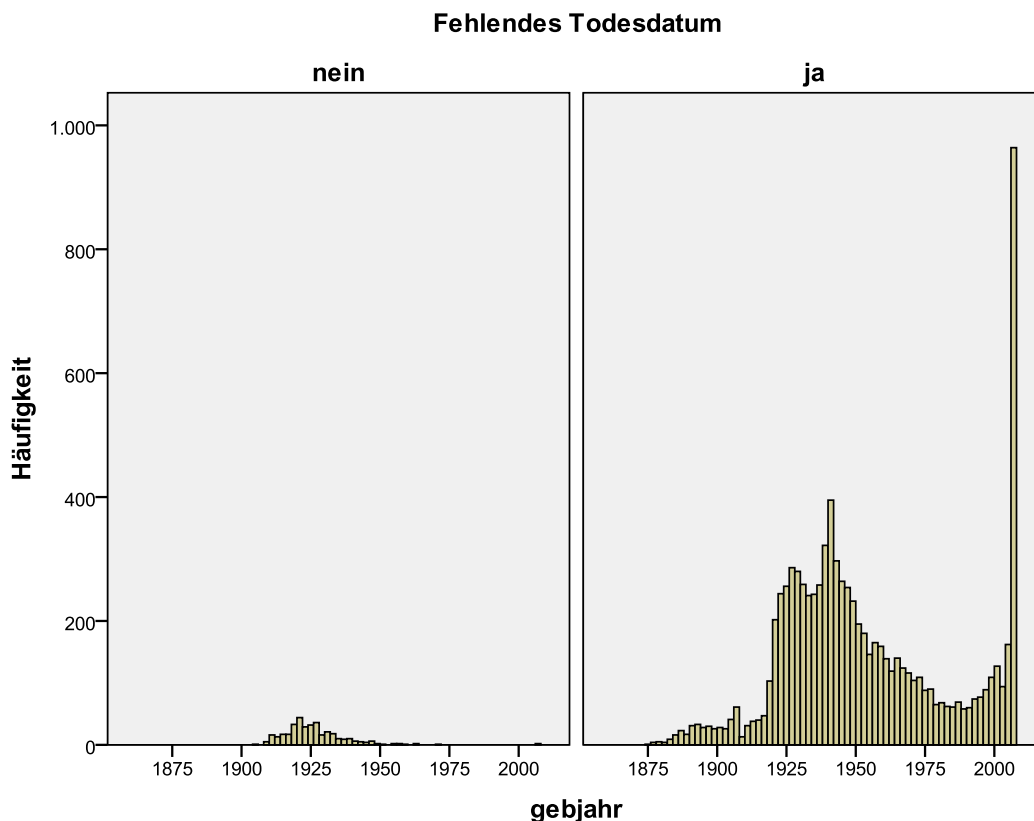


Abbildung 1: Verteilung der Geburtsjahre für Personen ohne Versicherungsverhältnisse (PHV_1\PHV_2) gegliedert nach existierendem Todesdatum

Die folgende Tabelle 5 zeigt dazu noch die Kennzahlen der Verteilungen der Geburtsjahre in den beiden Gruppen.

Tabelle 5: Basisstatistiken für die Personen ohne Leistungsdaten gegliedert nach vorhandenem Todesdatum (obere Tabelle) und nicht vorhandenem Todesdatum (untere Tabelle)

Deskriptive Statistik^a

	N	Minimum	Maximum	Mittelwert	Standardabweichung
gebjahr	361	1905	2007	1925,76	11,816
Gültige Werte (Listenweise)	361				

a. Fehlendes Todesdatum = nein

Deskriptive Statistik^a

	N	Minimum	Maximum	Mittelwert	Standardabweichung
gebjahr	8701	1875	2007	1954,27	30,867
Gültige Werte (Listenweise)	8701				

a. Fehlendes Todesdatum = ja

Wie man erkennt, gibt es eine Reihe von lebend registrierten Personen mit Geburtsjahren vor 1900. Es besteht also der Verdacht, dass unter diesen Personen einige bereits verstorben sind und daher keine Einträge in den Leistungsdaten vorhanden

sind. Auch die große Anzahl von Personen mit einem Geburtsjahr 2006 und 2007 (d.h., etwa 10%) scheint problematisch. Wenn man die Gegebenheiten des österreichischen Gesundheitssystems berücksichtigt, so scheint es nicht sehr plausibel, dass für Kinder keine Sozialversicherung besteht.

Es erschien daher gerechtfertigt, diese 9.062 Personen von allen weiteren Überlegungen auszuschließen.

Unter den verbleibenden Personen in der Population **PHV_2** gibt es auch Fälle für die zwar von einem Versicherungsträger Stammdaten eingetragen wurden, im betrachteten Zeitraum 2006/07 aber keine Leistungen erbracht wurden. Dies könnte folgende Ursachen haben:

- (1) Die Person ist bereits verstorben, es wurde aber kein Todesdatum eingetragen.
- (2) Die Person war einmal Arbeitstourist oder Anspruchstourist, im Zeitraum 2006/07 wurden aber keine Leistungen (mehr) bezogen.

Daher wurden die Geburtsdaten dieser 68.616 Personen aus der Tabelle **gapdrg-personen (PVH_0)** einer Plausibilitätsüberprüfung unterzogen. Die Basisstatistiken sind in Tabelle 6 dargestellt, ein Histogramm der Verteilung der Geburtsjahre ist in Abbildung 2 zu finden.

Aus diesen Darstellungen wird deutlich, dass beide Ursachen für Personen in dieser Gruppe möglich sind. Einerseits gibt es extrem alte Personen, bei denen vermutlich kein Todesdatum eingetragen wurde, andererseits gibt es auch Kinder, für die üblicherweise Leistungen der Sozialversicherung in Anspruch genommen werden.

*Tabelle 6: Basisstatistiken für das Geburtsdatum für Personen aus **PVH_2** ohne Stammdateneinträge im Zeitraum 2006/07*

Statistik	Wert	Standardfehler
Mittelwert	1974,84	,083
95% Konfidenzintervall des Mittelwerts	Untergrenze Obergrenze	1974,68 1975,01
5% getrimmtes Mittel	1976,71	
Median	1980,00	
Varianz	409,701	
Standardabweichung	20,241	
Minimum	1877	
Maximum	2007	
Spannweite	130	
Interquartilbereich	20	

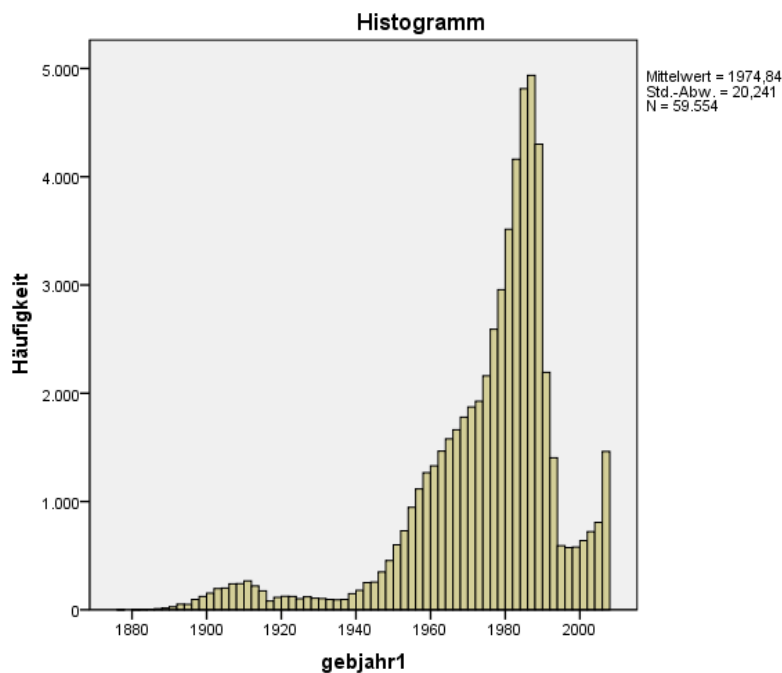


Abbildung 2: Verteilung der Geburtsjahre von Personen aus **PVH_2** ohne Stammdaten im Zeitraum 2006/07

Da das Geburtsjahr auch in der Tabelle **gapdrg.personendaten** enthalten ist, schien es sinnvoll zu überprüfen, ob diese Geburtsdaten für die auszuscheidenden Personen übereinstimmen. Zunächst einmal ist festzustellen, dass die Information über das Geburtsdatum in der Tabelle **gapdrg.personendaten** offensichtlich lückenhaft ist. Für fast zwei Drittel der Personen (62,8%) ist in **gapdrg.personendaten** kein Geburtsdatum eingetragen.

Bemerkenswert ist auch ein Vergleich zwischen den Eintragungen der Geburtsjahre in den beiden Tabellen **gapdrg.personen** und **gapdrg.personendaten**. Tabelle 7 zeigt, dass für etwa 6% der Fälle ein Fehler in der Hunderterstelle des Geburtsdatums vorliegt. In **gapdrg.personendaten** ist das Geburtsdatum um 100 Jahre größer.

Tabelle 7: Unterschiede in den Geburtsdaten der beiden Tabellen *gapdrg.personen* und *gapdrg.personendaten*

		gebdiff			
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	,00	20880	30,4	94,2	94,2
	100,00	1285	1,9	5,8	100,0
	Gesamt	22165	32,3	100,0	
Fehlend	System	46451	67,7		
Gesamt		68616	100,0		

Versucht man die Quellen dieser Unterschiede auszumachen, so stellt man fest, dass die Verteilung sehr unterschiedlich bezüglich der einzelnen Versicherungs-

träger –VTR– ist. Für die Unterschiede in den Geburtsjahren ist praktisch nur der VTR 13 (GKK Burgenland) verantwortlich:

Tabelle 8: Herkunft der unterschiedlichen Geburtsdaten nach Versicherungsträger

		gebdiff		Gesamt
		,00	100,00	
vtr	Eisenbahnen und Bergbau	199	0	199
	Öffentliche Bedienstete	283	0	283
	Wiener GKK	5748	2	5750
	Burgenland GKK	2316	1279	3595
	Oberösterreich GKK	3113	0	3113
	Steiermark GKK	2478	4	2482
	Kärnten GKK	1073	0	1073
	Neusiedler Betriebskrankenkasse	1	0	1
	Sozialvers. Gewerbl. Wirtschaft	5669	0	5669
Gesamt		20880	1285	22165

Auch die fehlenden Geburtsdaten verteilen sich sehr unterschiedlich auf die einzelnen VTR, wie die folgende Tabelle zeigt: Bei der Niederösterreichischen und der Tiroler Gebietskrankenkasse fehlt das Geburtsdatum vollständig, bei den Öffentlichen Bediensteten zum Teil, bei VTR 7 (BVA) zum größten Teil. (Die Anzahl der Eintragungen in **gapdrg.personendaten** ist dabei größer als die Anzahl der Personen, da ja für eine Person im betrachteten Zeitraum mehr als ein Versicherungsverhältnis bestehen kann.)

Tabelle 9: Herkunft der fehlenden Geburtsdaten nach VTR

		gebmiss		Gesamt
		Geburtsdatum fehlt in gapdrg.personendaten	Geburtsdatum vorhanden	
vtr	Eisenbahnen und Bergbau	99	249	348
	Öffentliche Bedienstete	1881	431	2312
	Wiener GKK	24	6105	6129
	Niederösterreich GKK	42967	0	42967
	Burgenland GKK	2	3631	3633
	Oberösterreich GKK	21	3212	3233
	Steiermark GKK	8	2572	2580
	Kärnten GKK	1	1119	1120
	Tirol GKK	15474	0	15474
	Neusiedler Betriebskrankenkasse	0	1	1
	Sozialvers. Gewerbl. Wirtschaft	17	5850	5867
Gesamt		60494	23170	83664

Da offensichtlich die Datenqualität für jene Personen in **PHV_2**, die im betrachteten Zeitraum keine Stammdateneinträge haben, nicht sehr gut ist, spricht vieles dafür,

dass es sich bei diese Fällen um Personen handelt die entweder 2006/07 schon verstorben waren oder aber in diesem Zeitraum nicht mehr Anspruchs- oder Arbeitstouristen in Österreich waren.

Somit werden auch diese Personen ausgeschlossen und wir erhalten eine Grundgesamtheit **PHV_3**, die insgesamt 9.779.341 Fälle enthält. Diese wird im Weiteren als Hauptverbandspopulation **PHV** angesehen. Der Prozess der Bestimmung der Population des Hauptverbandes ist in Abbildung 3 schematisch dargestellt.

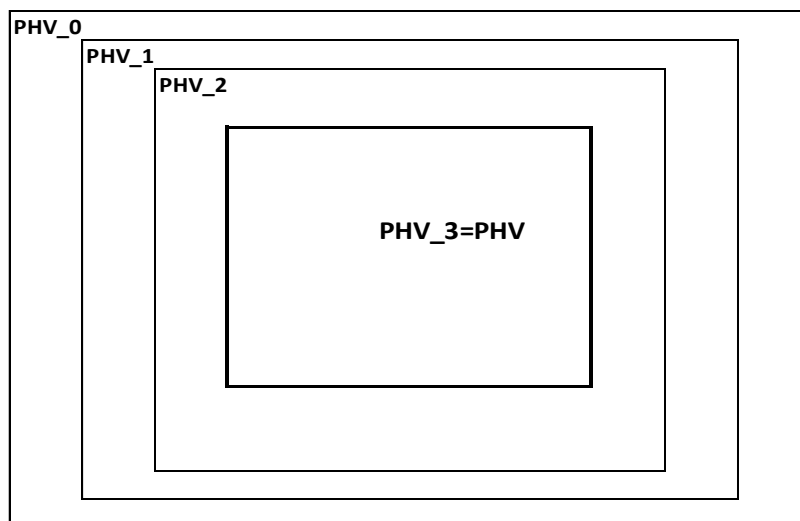


Abbildung 3: Schematische Darstellung der Bestimmung der Hauptverbandspopulation für den Zeitraum 2006/07

- **PHV_0**: Population aller Personen in **gapdrg.personen**, die durch die zentrale Partnerverwaltung bestätigt sind: **11.338.569 Personen**.
- **PHV_1**: Population von **9.847.975** Personen mit folgenden Eigenschaften:
 - Person ist vor dem 1.1.2008 geboren;
 - Eintrag in **gapdrg.personen** durch ZPV bestätigt;
 - Kein Todesdatum eingetragen oder ein Todesdatum nach dem 31.12.2007.
- **PHV_2**: Personen aus PHV_1, die in **gapdrg.personendaten** eingetragen sind. Dabei handelt es sich um **9.838.895** Personen.
- **PHV_3**: Personen in PHV_1, für die im Zeitraum 2006/07 ein gültiger Stammdateneintrag existiert. Es handelt sich dabei um **9.779.341** Personen.

Es verbleiben somit 9.779.341 Eintragungen von Personen für die Population **PHV**. Der Unterschied zur Wohnbevölkerung **PWB** ist mit 1.502.875 Personen oder 18,15% beträchtlich. Die in Abschnitt 1 angesprochene Untererfassung der Wohnbevölkerung ist also weit geringer als die Erweiterung der Population durch Phänomene wie Arbeits- und Anspruchstouristen.

1.2.3 Population der Gesundheitsbefragung

Über den subjektiven Gesundheitszustand der Bevölkerung im Zeitraum 2006/07 geben die Daten der Gesundheitsbefragung Auskunft. Diese Erhebung wurde im

Zeitraum von März 2006 bis Februar 2007 von Statistik Austria durchgeführt. Dabei wurden insgesamt 15.474 Personen, die älter als 15 Jahre sind, in ganz Österreich befragt. Die Ergebnisse dieser Befragung geben einen Überblick über den subjektiven Gesundheitszustand der Bevölkerung. In den Abbildungen 4 und 5 werden die relativen Abweichungen der hochgerechneten Bevölkerungsdaten aus der Gesundheitsbefragung von der Wohnbevölkerung **PWB** in Fünfjahreskohorten für Männer und Frauen dargestellt.

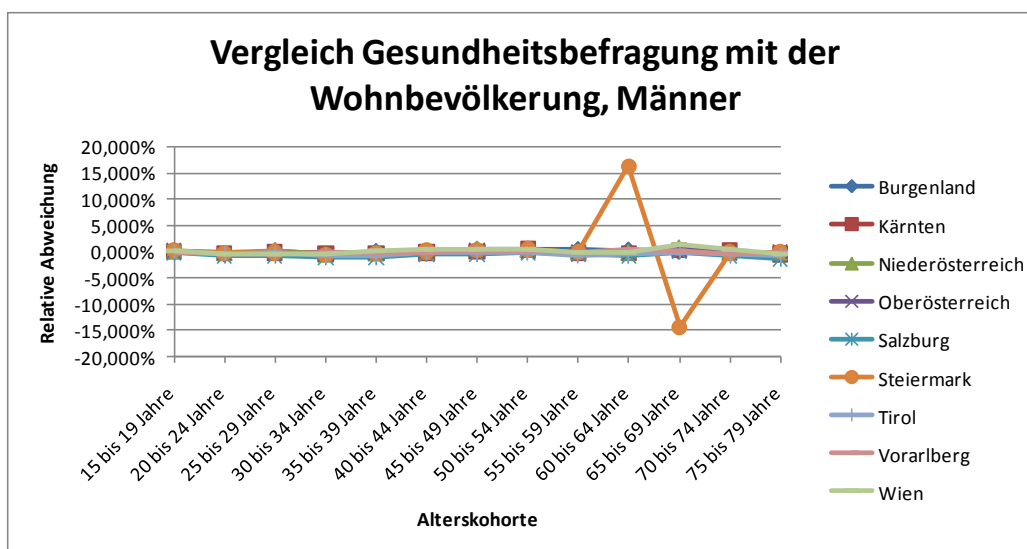


Abbildung 4: Vergleich der hochgerechneten Bevölkerungszahlen der Gesundheitsbefragung mit der Wohnbevölkerung in Fünfjahreskohorten in den Bundesländern für Männer

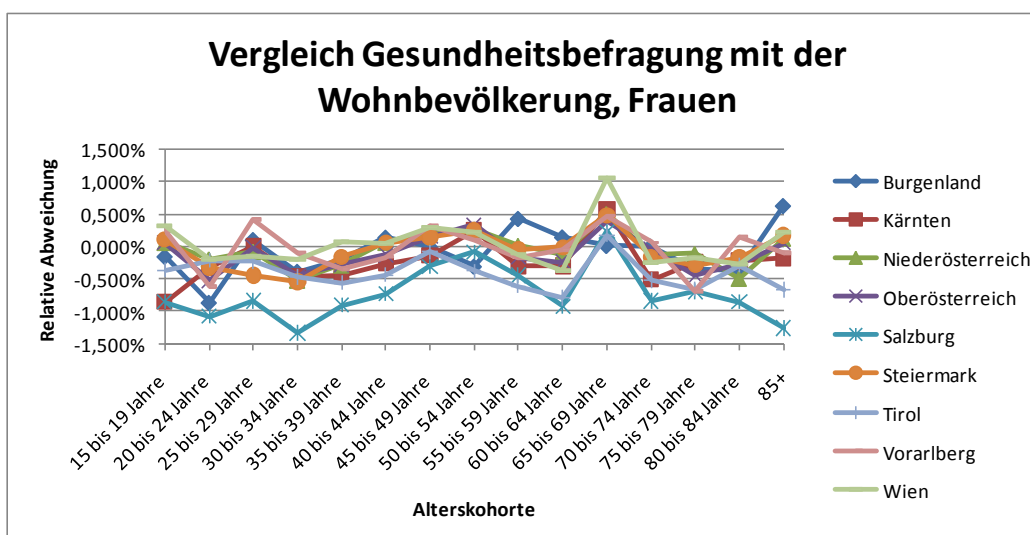


Abbildung 5: Vergleich der hochgerechneten Bevölkerungszahlen der Gesundheitsbefragung mit der Wohnbevölkerung in Fünfjahreskohorten in den Bundesländern für Frauen

Wie man erkennt, ist der Unterschied bei den Frauen in allen Bundesländern fast durchgehend unter 1%. Etwas problematischer ist das Ergebnis bei den Männern. In der Steiermark scheint ein Problem mit den Alterskohorten 60 – 64 Jahre und 65 – 69 Jahre vorzuliegen. In Summe gleichen sich diese allerdings aus. Die Alters-

kohorten über 80 Jahre wurden bei den Männern nicht dargestellt, da hier die relativen Abweichungen sehr hoch sind.

Zusammenfassend lässt sich aber feststellen, dass die hochgerechneten Bevölkerungszahlen der Gesundheitsbefragung mit wenigen Ausnahmen gut mit der Wohnbevölkerung im dritten Quartal 2006 übereinstimmen. Die Basisdaten der Grundgesamtheiten **PWB**, **PHV** und der Gesundheitsbefragung sind in folgender Tabelle 10 nochmals zusammengefasst.

Tabelle 10: Basisdaten für die Populationen **PWB**, **PHV** und Gesundheitsbefragung

Population	PWB	PHV (PHV_3)	Gesundheitsbefragung
Zeitbezug	3. Quartal 2006	2006 – 2007	März 2006 – Februar 2007
Operative Referenz (Datenabzug)	18.12.2010	30.10.2010	10.6.2009
Größe (Personen)	8.276.466	9.779.341	6.991.892
	Alle Altersjahrgänge	Alle Altersjahrgänge	Personen älter als 15 Jahre

2. Bewertung der Qualität von Attributen der CASE-Mix Datei

Alle in diesem Abschnitt dargestellten Ergebnisse beziehen sich auf den Zustand der „Urdaten“, d.h. der im vorigen Abschnitt definierten Hauptverbandspopulation **PHV**. Dabei wird vom Stand der Daten vom 30.10.2010 mit den in Abschnitt 1 beschriebenen Bereinigungen ausgegangen. Die Datei umfasst somit 9.779.341 Personen. Im Folgenden werden die einzelnen Attribute der Tabelle **gapdrg.personendaten** analysiert.

2.1 Konzentration der Versicherungsverhältnisse

Zunächst einmal betrachten wir die Verteilung der einzelnen Versicherungsverhältnisse. Tabelle 11 (nächste Seite) gibt eine Übersicht über die Anzahl der Personen mit 1, 2, 3, ... Stammdatensätzen (Konzentration der Stammdatensätze auf Personen).

Tabelle 11: Anzahl der Stammdatensätze pro Person

Anzahl der Einträge pro Person	Anzahl Personen	Gesamtzahl der Eintragungen
1	2.266.262	2.266.262
2	2.161.118	4.322.236
3	1.396.435	4.189.305
4	1.176.130	4.704.520
5	874.579	4.372.895
6	630.736	3.784.416
7	427.503	2.992.521
8	312.155	2.497.240
9	197.465	1.777.185
10	129.028	1.290.280
11	83.771	921.481
12	51.483	617.796
13	30.839	400.907
14	18.130	253.820
15	10.422	156.330
16	5.850	93.600
17	3.301	56.117
18	1.846	33.228
19	1.030	19.570
20	587	11.740
21	320	6.720
22	172	3.784
23	91	2.093
24	42	1.008
25	18	450
26	12	312
27	10	270
28	3	84
29	3	87
Summe	9.779.341	34.776.257

Somit liegen im Zeitraum 2006/2007 insgesamt 34.776.257 Stammdatensätze vor. Damit ergeben sich pro Person im Durchschnitt 3,5 Eintragungen innerhalb des Zeitraumes 2006/2007; diese Zählung unterscheidet nicht zwischen synchronen und aufeinanderfolgenden Versicherungsverhältnissen.

2.2 Analyse der Merkmale „vtr“ und „herkunft“

Diese beiden Merkmale geben Auskunft über die Versicherungsträger, von denen die Informationen in der GAP-DRG-Datenbank geliefert werden. Laut Beschreibung sollten diese beiden Merkmale gleich sein. Die folgende Tabelle 12 zeigt, dass diese Bedingung erfüllt ist.

Tabelle 12: Übereinstimmung der Merkmale ‚vtr‘ und ‚herkunft‘

vtr = herkunft	34.776.257
vtr≠ herkunft	0
vtr fehlend	0
herkunft fehlend	0

Nachstehende Tabelle 13 gibt eine Übersicht über die Meldungen der einzelnen Versicherungsträger.

Tabelle 13: Übersicht über den Datenumfang der einzelnen Versicherungsträger

ID	VTR	Personen	Stammdaten	Stammdaten/ Person
5	Eisenbahnen und Bergbau	358.527	703.532	1,96
7	Öffentliche Bedienstete	1.355.441	1.823.416	1,35
11	Wiener GKK	3.433.123	7.052.689	2,05
12	Niederösterreich GKK	1.864.716	4.801.527	2,57
13	Burgenland GKK	325.605	543.149	1,67
14	Oberösterreich GKK	1.693.369	3.191.028	1,88
15	Steiermark GKK	1.688.681	4.008.218	2,37
16	Kärnten GKK	558.475	1.217.796	2,18
17	Salzburg GKK	1.117.560	2.572.169	2,30
18	Tirol GKK	1.232.954	1.670.735	1,36
19	Vorarlberg GKK	399.785	777.367	1,94
22	Wiener VB Betriebskrankenkasse	24.843	37.644	1,52
24	Neusiedler Betriebskrankenkasse	4.064	5.897	1,45
25	VOEST Alpine Donawitz Betriebskrankenkasse	16.639	27.438	1,65
28	Kapfenberg Betriebskrankenkasse	12.258	22.673	1,85
40	Sozialvers. Gewerbl. Wirtschaft	1.761.323	5.420.133	3,08
50	Sozialvers. Bauern	585.129	900.846	1,54
	Gesamt	16.432.492	34.776.257	2,12

Die höhere Anzahl der Personen erklärt sich dadurch, dass eine Person bei mehreren Versicherungsträgern versichert sein kann. Dementsprechend ist die Durchschnittliche Anzahl der Einträge pro Person bei einem Versicherungsträger etwas geringer als in der Tabelle der Anzahl der Stammdaten pro Person

2.3 Verteilung der Bezirke

Tabelle 14 gibt Auskunft über die Häufigkeit von verschiedenen Bezirks_IDs für die einzelnen Personen:

Tabelle 14: Verteilung der Anzahl der Bezirke pro Person

Anzahl der Bezirke / Person	Personen	Prozent
Kein Eintrag	641.221	6,56%
1	7.565.610	77,36%
2	1.310.566	13,40%
3	220.877	2,26%
4	34.863	0,36%
5	5.260	0,05%
6	788	0,01%
7	123	0,00%
8	26	0,00%
9	6	0,00%
10	1	0,00%
	9.779.341	100,00%

Etwa 3/4 aller Personen wechselt also nicht den Bezirk im betrachteten Zeitraum. Wie man weiter erkennt sind bei etwa 6,6% aller Personen keine Informationen über eine räumliche Zuordnung vorhanden. Über die Verteilung von Eintragungen mit fehlender Bezirks-ID auf die einzelnen Versicherungsträger gibt folgende Tabelle 15 Auskunft.

Tabelle 15: Verteilung der fehlenden Bezirke auf die einzelnen Vertragsträger

ID	Vertragsträger	Bezirksangabe vorhanden	Bezirksangabe fehlt	Gesamt	Prozent fehlend
5	Eisenbahnen und Bergbau	631.382	72.150	703.532	10,26%
7	Öffentliche Bedienstete	1.337.592	485.824	1.823.416	26,64%
11	Wiener GKK	6.722.404	330.285	7.052.689	4,68%
12	Niederösterreich GKK	4.742.753	58.774	4.801.527	1,22%
13	Burgenland GKK	499.747	43.402	543.149	7,99%
14	Oberösterreich GKK	3.071.231	119.797	3.191.028	3,75%
15	Steiermark GKK	3.263.848	744.370	4.008.218	18,57%
16	Kärnten GKK	1.210.572	7.224	1.217.796	0,59%
17	Salzburg GKK	1.905.053	667.116	2.572.169	25,94%
18	Tirol GKK	1.344.215	326.520	1.670.735	19,54%
19	Vorarlberg GKK	737.165	40.202	777.367	5,17%
22	Wiener VB Betriebskrankenkasse	37.459	185	37.644	0,49%
24	Neusiedler Betriebskrankenkasse	5.871	26	5.897	0,44%
25	VOEST Alpine Donawitz BKK	27.337	101	27.438	0,37%
28	Kapfenberg Betriebskrankenkasse	22.655	18	22.673	0,08%
40	Sozialvers. Gewerbl. Wirtschaft	5.376.576	43.557	5.420.133	0,80%
50	Sozialvers. Bauern	898.307	2.539	900.846	0,28%
	Gesamt	31.834.167	2.942.090	34.776.257	8,46%

Damit wird deutlich, dass die Hauptverantwortlichen für die fehlenden Bezirkseintragungen die Sonderversicherungsträger (Eisenbahn und BVA) und die GKKs von Salzburg, Tirol und der Steiermark sind.

2.4 Berechnungen der Versicherungsdauer

Bei der Betrachtung der Versicherungsdauern können zwei Klassen gebildet werden: Versicherungsverträge, die im Laufe der betrachteten Zeitperiode beendet wurden und Vertragsverhältnisse die zum Ende der betrachteten Zeitperiode, also am 31.12.2007, noch aktiv waren. Tabelle 16 (nächste Seite) stellt die wesentlichen Kennzahlen der Vertragsdauern in diesen beiden Gruppen dar.

Tabelle 16 (nächste Seite) zeigt deutlich, dass abgeschlossene Verträge eine merklich kürzere Vertragsdauer haben als jene Verträge, die am Ende der betrachteten Zeitperiode noch laufen. Dies ist auch ein Hinweis, dass in den Daten ein erheblicher Anteil von Anspruchstouristen enthalten ist.

Tabelle 16: Verteilung der Versicherungsdauern in Vierteljahren

Vertragsdauer in Vierteljahren				
		abgeschlossene Verträge	offene Verträge	Gesamt
N	Gültig	26.057.879	8.718.378	34.776.257
	Fehlend	0	0	0
Mittelwert		5,09	9,83	6,28
Standardfehler des Mittelwertes		0,0011	,0031	0,0012
Median		3,00	5,00	4,00
Modus		1,00	4,00	1,00
Standardabweichung		5,54	9,17	6,95
Minimum		1,00	1,00	1,00
Maximum		40,00	40,00	40,00
Perzentile	10	1,00	2,00	1,00
	20	1,00	4,00	1,00
	30	1,00	4,00	2,00
	40	2,00	5,00	3,00
	50	3,00	5,00	4,00
	60	4,00	8,00	4,00
	70	5,00	11,00	6,00
	80	8,00	16,00	10,00
	90	14,00	26,00	17,00

Betrachtet man die Vertragsdauern bei den einzelnen Vertragspartnern, so zeigt sich, dass diese Unterschiede in den mittleren Vertragsdauern bei fast allen Versicherungsträgern auftreten. Alle Gebietskrankenkassen zeigen die Unterschiede in den mittleren Vertragsdauern.

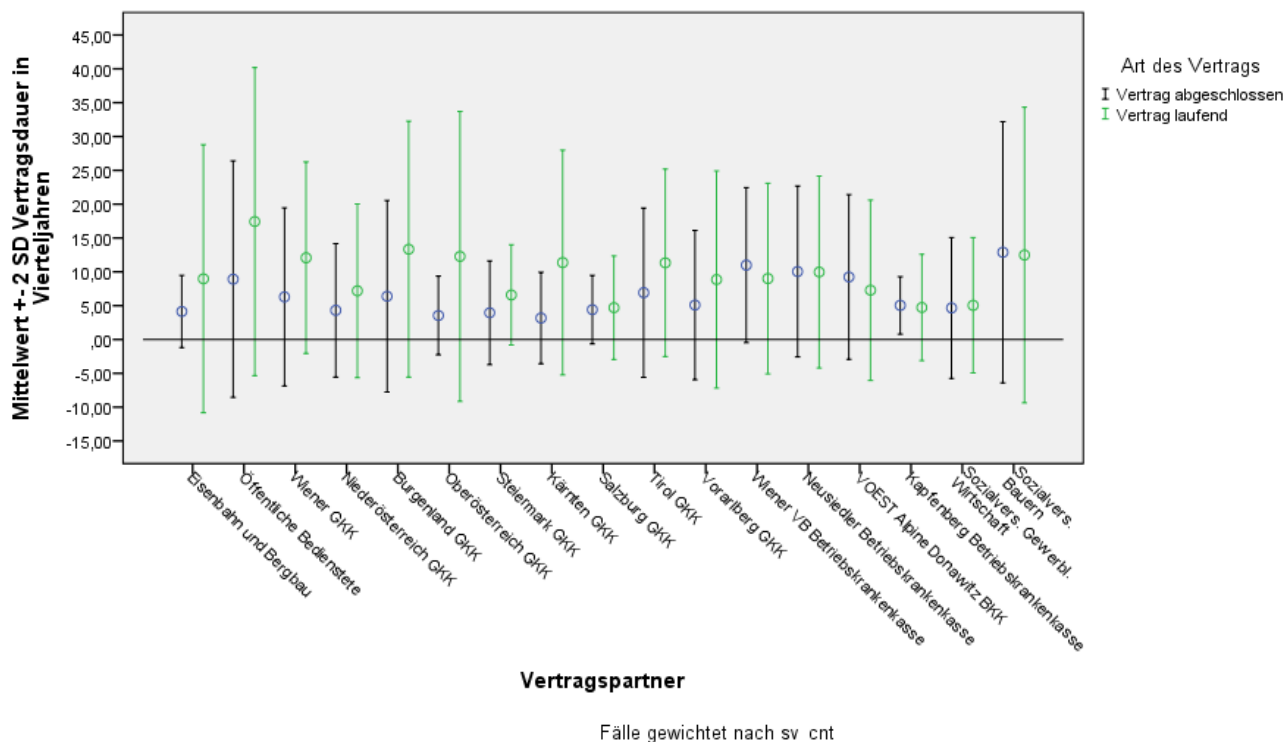


Abbildung 6: Mittlere Versicherungsdauer bei den einzelnen Versicherungsträgern und Schwankungsbereiche

Etwas anders liegen die Verhältnisse bei den Betriebskrankenkassen, bei der Sozialversicherung der Bauern und bei der Sozialversicherung der gewerblichen Wirtschaft. Hier sind keine deutlichen Unterschiede zwischen abgeschlossenen und laufenden Verträgen festzustellen, bei den Betriebskrankenkassen der Wiener Verkehrsbetriebe und der Voest-Alpine Donawitz sind die laufenden Verträge im Mittel kürzer als die abgeschlossenen Verträge. Abbildung 6 (vorige Seite) stellt die Ergebnisse grafisch dar. Auch diese Ergebnisse sprechen für die Erklärung, dass die höhere Personenanzahl in PHV durch die Einbeziehung von Anspruchstouristen erklärt werden muss.

2.5 Rezeptgebühren

Die Daten bezüglich der Rezeptgebühr zeigen, dass rund 81% aller Personen nie von der Rezeptgebühr befreit wurden. Ein Wechsel zwischen Befreiung und Rezeptgebührenpflicht kommt in etwa 4% aller Fälle vor. Bei den Personen mit Versicherungsverhältnissen mit fehlender Eintragung zum Status der Rezeptgebührenbefreiung ist die Kombination ‚Befreiung und fehlend‘ die häufigste.

Tabelle 17: Verteilung der verschiedenen Kombinationen von Eintragungen zur Rezeptgebühr

Kategorie	Anzahl	Prozent
Personen immer befreit von Rezeptgebühr	230912	2,36%
Personen befreit und fehlend	1195436	12,22%
Personen nie befreit von Rezeptgebühr	7891974	80,70%
Personen nie befreit und fehlend	7486	0,08%
Personen mit wechselnder Befreiung	378986	3,88%
Personen mit wechselnder Befreiung und fehlend	42678	0,44%
Personen mit immer fehlenden Eintragungen	31869	0,33%
Gesamt	9779341	100,00%

Tabelle 18 (nächste Seite) gibt Auskunft über die Befreiung von der Rezeptgebühr bei den einzelnen Versicherungsträgern.

Hier ist zunächst einmal auffallend, dass nur bei vier Versicherungsträgern fehlende Einträge vorkommen. Bei der der Salzburger GKK fehlt diese Information in 50% aller Einträge, bei Eisenbahn und Bergbau und bei der GKK Vorarlberg zu knapp 20% bzw. 17%. Bei der GKK Burgenland ist der Anteil mit 0,12% vernachlässigbar klein.

Tabelle 18: Fehlende Einträge bei Rezeptgebühren bei den einzelnen Vertragsträger

ID	Vertragsträger	Eintrag Rezeptgebühr	Fehlender Eintrag Rezeptgebühr	Gesamt	Prozent fehlend
5	Eisenbahnen und Bergbau	565.783	137.749	703.532	19,58%
7	Öffentliche Bedienstete	1.823.416	0	1.823.416	0,00%
11	Wiener GKK	7.052.689	0	7.052.689	0,00%
12	Niederösterreich GKK	4.801.527	0	4.801.527	0,00%
13	Burgenland GKK	542.514	635	543.149	0,12%
14	Oberösterreich GKK	3.191.028	0	3.191.028	0,00%
15	Steiermark GKK	4.008.217	1	4.008.218	0,00%
16	Kärnten GKK	1.217.796	0	1.217.796	0,00%
17	Salzburg GKK	1.272.151	1.300.018	2.572.169	50,54%
18	Tirol GKK	1.670.735	0	1.670.735	0,00%
19	Vorarlberg GKK	642.937	134.430	777.367	17,29%
22	Wiener VB Betriebskrankenkasse	37.644	0	37.644	0,00%
24	Neusiedler Betriebskrankenkasse	5.897	0	5.897	0,00%
25	VOEST Alpine Donawitz BKK	27.438	0	27.438	0,00%
28	Kapfenberg Betriebskrankenkasse	22.673	0	22.673	0,00%
40	Sozialvers. Gewerbl. Wirtschaft	5.420.133	0	5.420.133	0,00%
50	Sozialvers. Bauern	900.846	0	900.846	0,00%
	Gesamt	33.203.424	1.572.833	34.776.257	4,52%

2.6 Geschlecht

Die Analyse der Eintragungen vom Geschlecht zeigt, dass es für 21,5% der Personen keine Eintragungen für das Feld Geschlecht gibt. Die Anzahl der Personen mit gemischtem Geschlecht ist mit knapp 2 Promille vernachlässigbar.

Tabelle 19: Verteilung der verschiedenen Kombinationen von Eintragungen zum Geschlecht

Kategorie	Anzahl Personen
Geschlecht immer männlich	3.890.966
Geschlecht immer weiblich	3.772.049
Geschlecht weiblich und fehlend	894.775
Geschlecht männlich und fehlend	880.260
Geschlecht immer fehlend	325.592
Geschlecht mehrdeutig	11.652
Geschlecht mehrdeutig und fehlend	4.047
Gesamt	9.779.341

Betrachtet man die Einträge, die zu fehlendem Geschlecht geführt haben, und ordnet sie den verschiedenen Versicherungsträgern zu, so zeigt sich auch hier wieder eine Konzentration auf wenige Träger: GKK Salzburg, GKK Oberösterreich und die SVA-Gewerbliche Wirtschaft sind für die Probleme verantwortlich; vgl. Tabelle 20 (nächste Seite).

Tabelle 20: Verteilung der fehlenden Werte beim Geschlecht auf die einzelnen Vertragsträger

ID	Vertragsträger	Geschlecht vorhanden	Geschlecht fehlend	Gesamt	Prozent fehlend
5	Eisenbahnen und Bergbau	703.532	0	703.532	0,00%
7	Öffentliche Bedienstete	1.823.416	0	1.823.416	0,00%
11	Wiener GKK	7.052.688	1	7.052.689	0,00%
12	Niederösterreich GKK	4.801.527	0	4.801.527	0,00%
13	Burgenland GKK	543.149	0	543.149	0,00%
14	Oberösterreich GKK	2.485.243	705.785	3.191.028	22,12%
15	Steiermark GKK	4.008.218	0	4.008.218	0,00%
16	Kärnten GKK	1.217.796	0	1.217.796	0,00%
17	Salzburg GKK	1.252.104	1.320.065	2.572.169	51,32%
18	Tirol GKK	1.670.735	0	1.670.735	0,00%
19	Vorarlberg GKK	777.367	0	777.367	0,00%
22	Wiener VB Betriebskrankenkasse	37.644	0	37.644	0,00%
24	Neusiedler Betriebskrankenkasse	5.897	0	5.897	0,00%
25	VOEST Alpine Donawitz BKK	27.438	0	27.438	0,00%
28	Kapfenberg Betriebskrankenkasse	22.673	0	22.673	0,00%
40	Sozialvers. Gewerbl. Wirtschaft	3.423.915	1.996.218	5.420.133	36,83%
50	Sozialvers. Bauern	900.846	0	900.846	0,00%
Gesamt		30.754.188	4.022.069	34.776.257	

2.7 Alter

Das Attribut Alter ist auf Grund der Bereinigungen durch die zentrale Partnerverwaltung überall vorhanden. Tabelle 21 (nächste und übernächste Seiten) zeigt die Altersverteilung in der Tabelle **gapdrg.personen** im Vergleich mit der Altersverteilung der Wohnbevölkerung **PWB**.

Auch hier wird die Problematik der alten Personen in der Hauptverbandspopulation deutlich. Die extremen Abweichungen im Alter 60, 69 und ab 80 Jahren machen deutlich, dass hier einerseits Personen nicht ausgeschieden wurden, andererseits vermutlich falsche Altersangaben eingesetzt wurden. Die extremen Abweichungen bei Personen über 95 Jahre dürften primär auf nicht eingetragene Sterbedaten zurück zu führen sein. Es gibt etwa in **PHV** eine Person mit einem Alter von 179 Jahren. Die folgende Grafik zeigt die relativen Abweichungen der Altersverteilung in der Hauptverbandsbevölkerung **PHV** gegenüber der Wohnbevölkerung **PWB** von Statistik Austria. Die Gruppe der Personen über 95 Jahre wird aus Gründen der Übersichtlichkeit nicht dargestellt.

Tabelle 21 (a): Altersstruktur der Hauptverbandspopulation im Vergleich zur Altersstruktur in der Wohnbevölkerung

Alter in Jahren	Häufigkeit Hauptverband	Prozent	Häufigkeit Statistik Austria	Prozent	Differenz HV-StatA absolut	Differenz HV-StatA relativ
0	76.331	0,78	77.434	,9	-1.103	-1,42%
1	79.816	0,82	80.155	1,0	-339	-0,42%
2	81.225	0,83	80.422	1,0	803	1,00%
3	82.922	0,85	80.444	1,0	2.478	3,08%
4	81.837	0,84	80.614	1,0	1.223	1,52%
5	84.114	0,86	79.926	1,0	4.188	5,24%
6	81.741	0,84	81.990	1,0	-249	-0,30%
7	85.043	0,87	83.020	1,0	2.023	2,44%
8	85.403	0,87	85.492	1,0	-89	-0,10%
9	89.061	0,91	90.813	1,1	-1.752	-1,93%
10	92.073	0,94	91.924	1,1	149	0,16%
11	96.604	0,99	92.870	1,1	3.734	4,02%
12	96.558	0,99	96.461	1,2	97	0,10%
13	99.772	1,02	98.847	1,2	925	0,94%
14	102.465	1,05	99.876	1,2	2.589	2,59%
15	104.228	1,07	100.348	1,2	3.880	3,87%
16	106.210	1,09	98.124	1,2	8.086	8,24%
17	105.252	1,08	98.877	1,2	6.375	6,45%
18	106.513	1,09	98.648	1,2	7.865	7,97%
19	110.268	1,13	97.978	1,2	12.290	12,54%
20	113.497	1,16	100.087	1,2	13.410	13,40%
21	119.040	1,22	102.103	1,2	16.937	16,59%
22	124.288	1,27	104.196	1,3	20.092	19,28%
23	129.130	1,32	106.614	1,3	22.516	21,12%
24	131.691	1,35	109.845	1,3	21.846	19,89%
25	138.478	1,42	110.357	1,3	28.121	25,48%
26	139.162	1,42	106.072	1,3	33.090	31,20%
27	138.470	1,42	102.951	1,2	35.519	34,50%
28	131.770	1,35	101.588	1,2	30.182	29,71%
29	130.878	1,34	102.138	1,2	28.740	28,14%
30	130.165	1,33	104.564	1,3	25.601	24,48%
31	131.902	1,35	109.919	1,3	21.983	20,00%
32	135.820	1,39	111.845	1,4	23.975	21,44%
33	138.869	1,42	114.253	1,4	24.616	21,55%
34	138.861	1,42	121.306	1,5	17.555	14,47%
35	146.837	1,50	125.824	1,5	21.013	16,70%
36	151.948	1,55	130.408	1,6	21.540	16,52%
37	156.382	1,60	138.549	1,7	17.833	12,87%
38	164.993	1,69	141.978	1,7	23.015	16,21%
39	169.018	1,73	141.016	1,7	28.002	19,86%
40	168.489	1,72	141.768	1,7	26.721	18,85%
41	168.391	1,72	144.200	1,7	24.191	16,78%
42	170.377	1,74	144.232	1,7	26.145	18,13%
43	171.125	1,75	144.230	1,7	26.895	18,65%
44	169.876	1,74	140.730	1,7	29.146	20,71%
45	166.779	1,71	134.300	1,6	32.479	24,18%
46	162.742	1,66	132.424	1,6	30.318	22,89%
47	158.498	1,62	126.825	1,5	31.673	24,97%
48	153.555	1,57	121.141	1,5	32.414	26,76%
49	148.082	1,51	118.574	1,4	29.508	24,89%

Tabelle 21 (b): Altersstruktur der Hauptverbandspopulation im Vergleich zur Altersstruktur in der Wohnbevölkerung

Alter in Jahren	Häufigkeit Hauptverband	Prozent	Häufigkeit Statistik Austria	Prozent	Differenz HV-StatA absolut	Differenz HV-StatA relativ
50	145.312	1,49	116.142	1,4	29.170	25,12%
51	145.665	1,49	108.490	1,3	37.175	34,27%
52	137.538	1,41	103.315	1,2	34.223	33,12%
53	133.288	1,36	101.266	1,2	32.022	31,62%
54	129.946	1,33	99.296	1,2	30.650	30,87%
55	130.138	1,33	96.306	1,2	33.832	35,13%
56	122.890	1,26	97.568	1,2	25.322	25,95%
57	126.947	1,30	101.932	1,2	25.015	24,54%
58	124.407	1,27	101.948	1,2	22.459	22,03%
59	126.227	1,29	104.218	1,3	22.009	21,12%
60	124.423	1,27	76.618	,9	47.805	62,39%
61	104.744	1,07	76.793	,9	27.951	36,40%
62	85.953	0,88	92.467	1,1	-6.514	-7,04%
63	106.719	1,09	88.847	1,1	17.872	20,12%
64	104.854	1,07	93.094	1,1	11.760	12,63%
65	103.811	1,06	102.747	1,2	1.064	1,04%
66	115.498	1,18	110.280	1,3	5.218	4,73%
67	122.751	1,26	98.166	1,2	24.585	25,04%
68	116.483	1,19	67.569	,8	48.914	72,39%
69	82.251	0,84	63.178	,8	19.073	30,19%
70	72.609	0,74	61.323	,7	11.286	18,40%
71	70.623	0,72	59.999	,7	10.624	17,71%
72	67.733	0,69	59.302	,7	8.431	14,22%
73	66.632	0,68	59.517	,7	7.115	11,95%
74	66.571	0,68	60.714	,7	5.857	9,65%
75	66.948	0,68	59.703	,7	7.245	12,14%
76	64.826	0,66	58.457	,7	6.369	10,90%
77	65.245	0,67	56.329	,7	8.916	15,83%
78	61.057	0,62	52.976	,6	8.081	15,25%
79	58.031	0,59	50.974	,6	7.057	13,84%
80	53.812	0,55	49.453	,6	4.359	8,81%
81	52.035	0,53	47.526	,6	4.509	9,49%
82	49.703	0,51	44.048	,5	5.655	12,84%
83	45.952	0,47	41.481	,5	4.471	10,78%
84	43.007	0,44	38.694	,5	4.313	11,15%
85	39.516	0,40	33.314	,4	6.202	18,62%
86	35.040	0,36	29.277	,4	5.763	19,68%
87	30.090	0,31	16.566	,2	13.524	81,64%
88	20.902	0,21	10.846	,1	10.056	92,72%
89	12.765	0,13	9.696	,1	3.069	31,65%
90	10.239	0,10	8.397	,1	1.842	21,94%
91	9.050	0,09	9.145	,1	-95	-1,04%
92	9.162	0,09	8.491	,1	671	7,90%
93	9.800	0,10	6.837	,1	2.963	43,34%
94	8.155	0,08	5.047	,1	3.108	61,58%
95 und älter	53.444	0,55	9.814	,1	43.630	444,57%
Gesamt	9.779.341	100,00	8.276.466	100,0	1.502.875	18,16%

Wie man erkennt, ist der Grad der Übereinstimmung in der Altersgruppe bis 15 Jahre akzeptabel; danach wird deutlich, dass mehr Personen in der Grundgesamtheit des Hauptverbandes sind (vgl. Abbildung 7, nächste Seite) Dieser Überhang ist bis zu einem Alter von etwa 60 Jahren in einer Größenordnung von 20%, mit einigen Abweichungen nach oben in den Altersgruppen 25 – 30 Jahre und 50 – 60 Jahre. Ab einem Alter von 60 Jahren scheinen die Werte in der Population des

Hauptverbandes nur mehr beschränkt mit den Werten in der Wohnbevölkerung vergleichbar.

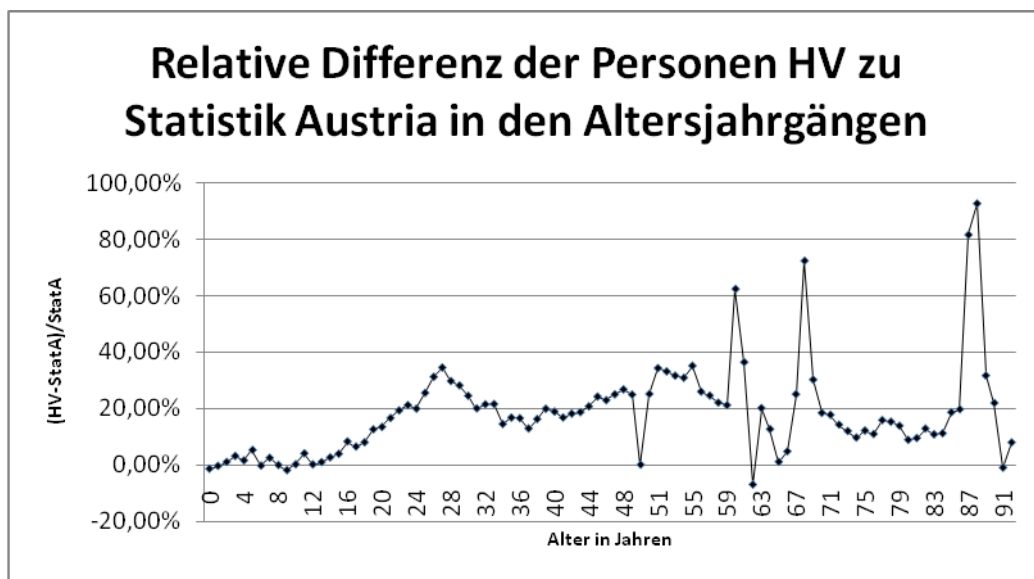


Abbildung 7: Relative Abweichungen der Altersverteilung der Bevölkerung im Hauptverband (PHV) und Statistik Austria (PWB)

Abbildung 8 zeigt die Zuordnung der Verträge von Personen über 80 Jahre zu den einzelnen Versicherungsträgern im Verhältnis zu den Zuordnungen für alle Verträge.

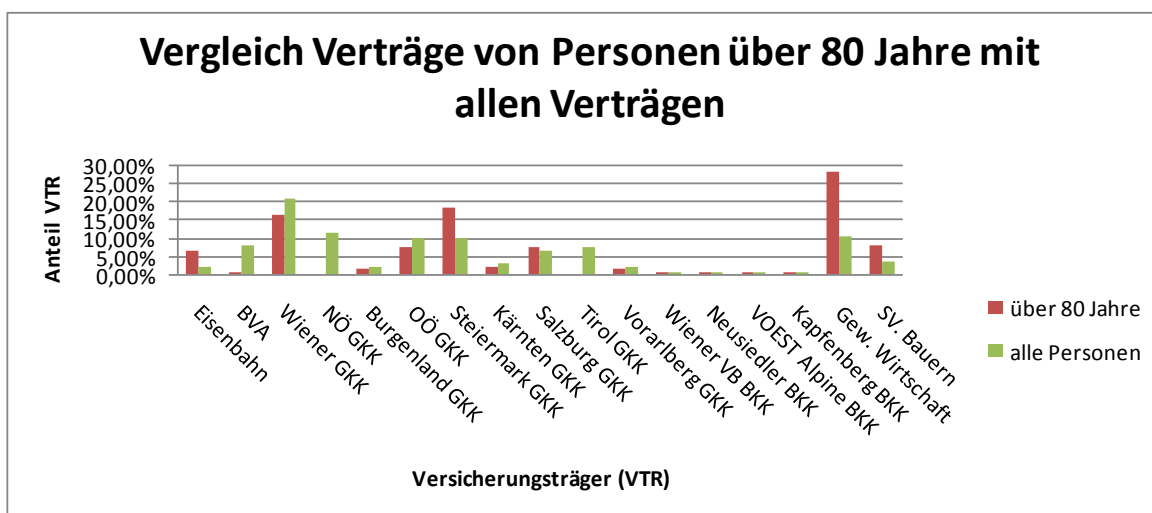


Abbildung 8: Anteil der Verträge von Personen über 80 Jahre in den einzelnen VTRs im Vergleich zu allen Verträgen

Überproportional viele Vertragsverhältnisse mit Personen über 80 Jahre gibt es bei der Gewerblichen Wirtschaft, bei der steirischen Gebietskrankenkasse, bei den Eisenbahnern und bei der Sozialversicherung der Bauern. Die Erklärung dürfte auch hier der fehlende Todeszeitpunkt sein.

Die bisher betrachtete Altersverteilung beruht auf den Daten in der Tabelle **gapdrg.personen**. Das Geburtsjahr ist aber auch in der Tabelle **gapdrg.personendaten**

zu finden. Aus den Analysen zur Definition der Population **PHV** ist bekannt, dass es Fehler in den Eintragungen der Geburtsjahre gibt, insbesondere können sich Geburtsjahre um einen Wert von 100 unterscheiden. Tabelle 22 zeigt die Verteilung der Differenzen in den Geburtsjahren für die einzelnen Versicherungsträger.

Tabelle 22: Verteilung der unterschiedlichen Geburtsjahre in **gapdrg.personen** und **gapdrg.personendaten** auf die einzelnen VTR

Versicherungsträger	Differenz 100	gleich	andere Wert	fehlend	Summe
Eisenbahnen und Bergbau	163	554.178		149.191	703.532
Öffentliche Bedienstete	11	41.029		1.782.376	1.823.416
Wiener GKK	6.220	7.012.692		33.777	7.052.689
Niederösterreich GKK				4.801.527	4.801.527
Burgenland GKK	95	381.029		162.025	543.149
Oberösterreich GKK	817	3.181.655		8.556	3.191.028
Steiermark GKK	28.580	3.970.175		9.463	4.008.218
Kärnten GKK	82	1.039.092		178.622	1.217.796
Salzburg GKK	5.182	2.463.481		103.506	2.572.169
Tirol GKK				1.670.735	1.670.735
Vorarlberg GKK	66	677.323		99.978	777.367
Wiener VB Betriebskrankenkasse	16	29.261		8.367	37.644
Neusiedler Betriebskrankenkasse		3.815		2.082	5.897
VOEST Alpine Donawitz BKK	12	21.471		5.955	27.438
Kapfenberg Betriebskrankenkasse	4	17.769		4.900	22.673
Sozialvers. Gewerbl. Wirtschaft	11.350	5.353.281		55.502	5.420.133
Sozialvers. Bauern	4	11.063	882.849	6.930	900.846
Summe	52.602	24.757.314	882.849	9.083.492	34.776.257

Ein um hundert Jahre größeres Geburtsjahr tritt insgesamt sehr selten auf (1,5 Promille aller Verträge) und konzentriert sich vornehmlich auf die Steirische GKK (54% aller dieser Fälle), die Sozialversicherung der gewerblichen Wirtschaft (22% aller dieser Fälle), Wiener GKK (12%) und die Salzburger GKK (10%). Fehlende Geburtsjahre sind vor allem bei den öffentlichen Bediensteten, der Niederösterreichischen GKK, und der Tiroler GKK zu finden. Diese drei Versicherungsträger sind für etwa 90% der fehlenden Werte verantwortlich. Einen Sonderfall stellt die Sozialversicherung der Bauern dar. Hier treten praktisch alle Werte von unterschiedlichen Eintragungen im Alter zwischen den beiden Tabellen **gapdrg.personen** und **gapdrg.personendaten** auf.

3. Analysen mehrerer Variablen

Von zentralem Interesse für viele Fragen der EBM ist die Gliederung der Bevölkerung nach Alter, Geschlecht und Region. In diesem Abschnitt wird daher die Struktur der Hauptverbandspopulation **PHV** nach diesen drei Merkmalen dargestellt und mit der Wohnbevölkerung der Statistik Austria **PWB** verglichen. Von diesen drei Merkmalen ist nur das Alter vollständig vorhanden, beim Geschlecht und beim Bundesland gibt es fehlende Werte. Es werden daher nicht nur die Tabellen der gemeinsamen Verteilungen betrachtet, sondern auch die Struktur der fehlenden Werte bezüglich der anderen Variablen.

3.1 Alter und Geschlecht

Betrachtet man die Verteilung von Alter und Geschlecht, so ist zu nächst das fehlende Geschlecht zu berücksichtigen.

Abbildung 9 zeigt fehlendes Geschlecht in den einzelnen Alterskohorten (Jahren). Es wird dabei deutlich, dass besonders in der Gruppe unter 15 Jahren das Geschlecht oft fehlt. Dies könnte ein Problem von Kindern mit nicht-österreichischer Staatsangehörigkeit sein. Weiter ist das häufig fehlende Geschlecht bei alten Personen auffallend. Dies lässt sich als ein Indiz für nicht eingetragene Sterbedaten deuten.

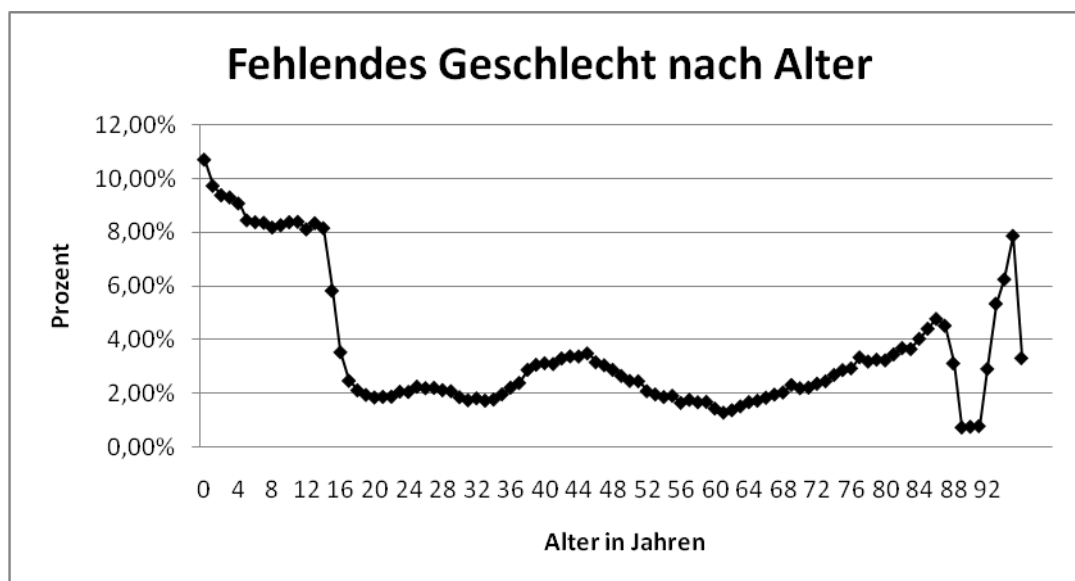


Abbildung 9: Anteil von fehlendem Geschlecht in den einzelnen Altersjahrgängen

Neben fehlendem Geschlecht gibt es auch noch jene Fälle, in denen eine Person mehrere Versicherungsverhältnisse mit unterschiedlich angegebenem Geschlecht hat. Um die Effekte eines fehlenden oder nicht eindeutigen Geschlechts in der Darstellung einer Bevölkerungspyramide auszugleichen, wurden folgende Regeln zur Imputation des Geschlechts angewendet:

- bei fehlendem Geschlecht wurde dieses nach dem Zufallsprinzip imputiert (männlich 50%, weiblich 50%);

- bei mehrdeutigem Geschlecht wurde ein Majoritätsprinzip angewendet (das am häufigsten genannte Geschlecht gilt, bei gleicher Anzahl wird das Zufallsprinzip angewendet).

Damit ergibt sich mit den imputierten Werten eine Verteilung von Alter und Geschlecht, die in Abbildung 10 als Bevölkerungspyramide dargestellt ist.

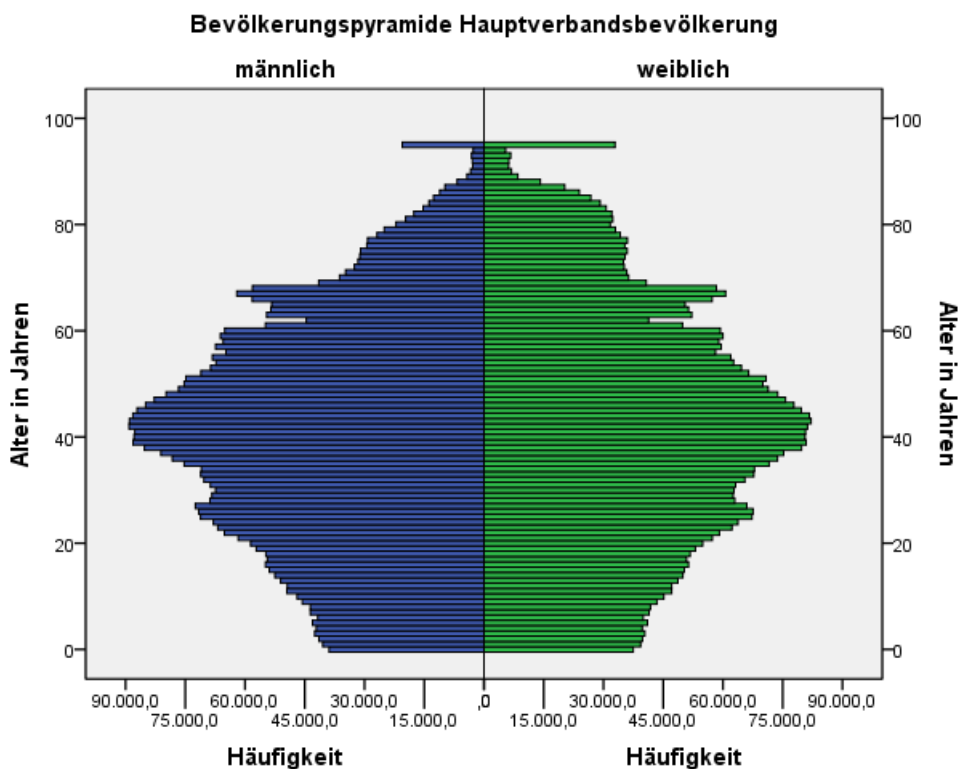


Abbildung 10 (a): Bevölkerungspyramide für Österreich der Hauptverbandspopulation **PHV**

Abbildung 11 (nächste Seite) entspricht Abbildung 7 in Abschnitt 2.7 und zeigt die relativen Abweichungen zwischen der Hauptverbandsbevölkerung **PHV** und der Wohnbevölkerung **PWB** bezogen auf die Wohnbevölkerung für Männer und Frauen getrennt.

Aus diesen Grafiken wird deutlich, dass die Übereinstimmung der Bevölkerungen in den Alterskohorten bis 15 Jahre relativ gut ist. Dabei ist aber zu beachten, dass gerade in dieser Gruppe sehr viele Werte für das Geschlecht imputiert werden mussten. Danach steigt bis etwa 20 Jahre die relative Differenz sowohl bei Männern als auch bei Frauen an und pendelt sich dann auf etwa 20% bei den Männern ein, bei den Frauen auf etwas unter 20%. Ab etwa 44 Jahren steigt der Überhang im Hauptverband wieder an und erreicht bei Männern einen Wert von 40%, bei Frauen von etwa 30%. Ab einem Alter von etwa 60 Jahren treten einige sehr große Differenzen auf, die Normwerte selbst liegen wieder bei etwa 20% bei den Männern, bei den Frauen unter 10%.

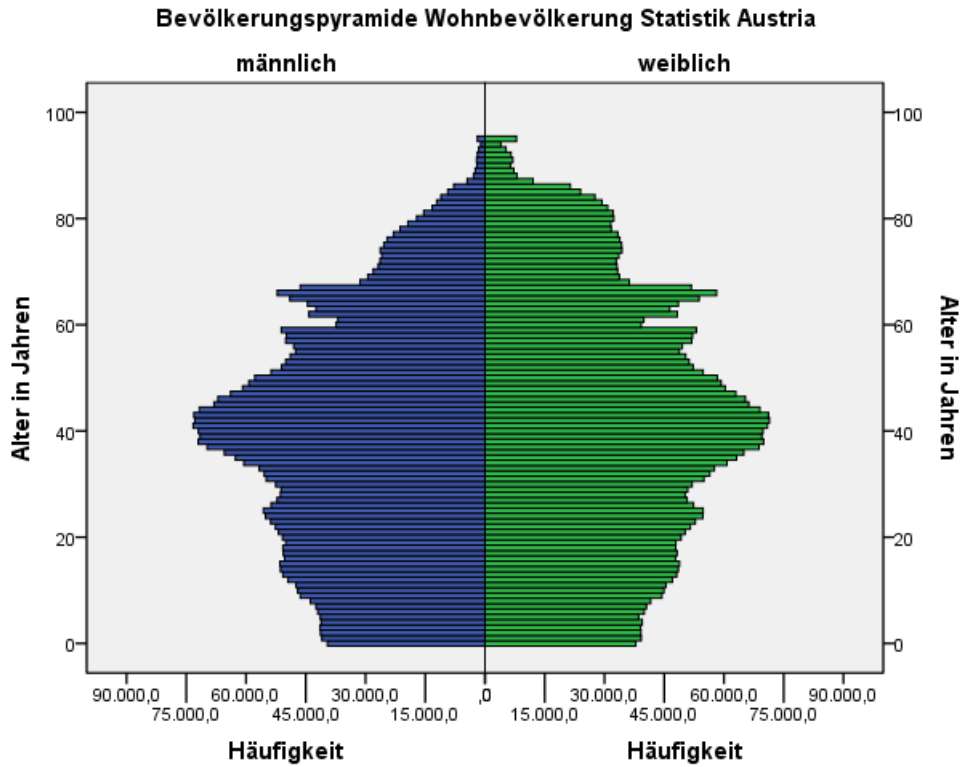


Abbildung 10 (b): Bevölkerungspyramide für Österreich der Wohnbevölkerung von Statistik Austria **PWB**

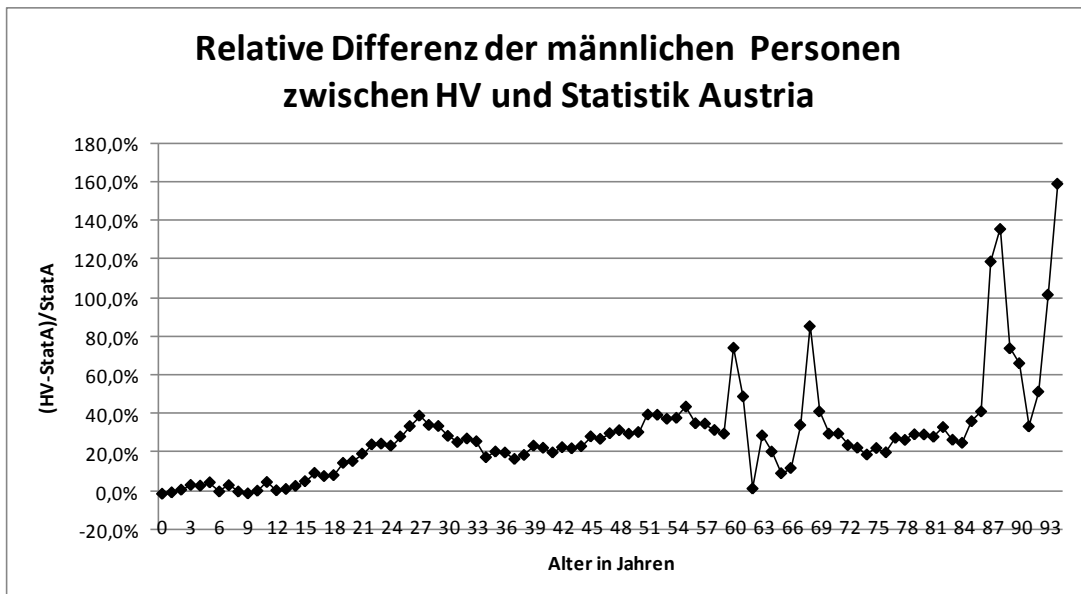


Abbildung 11 (a): Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung **PWB** von Statistik Austria für Männer in den einzelnen Jahrgängen

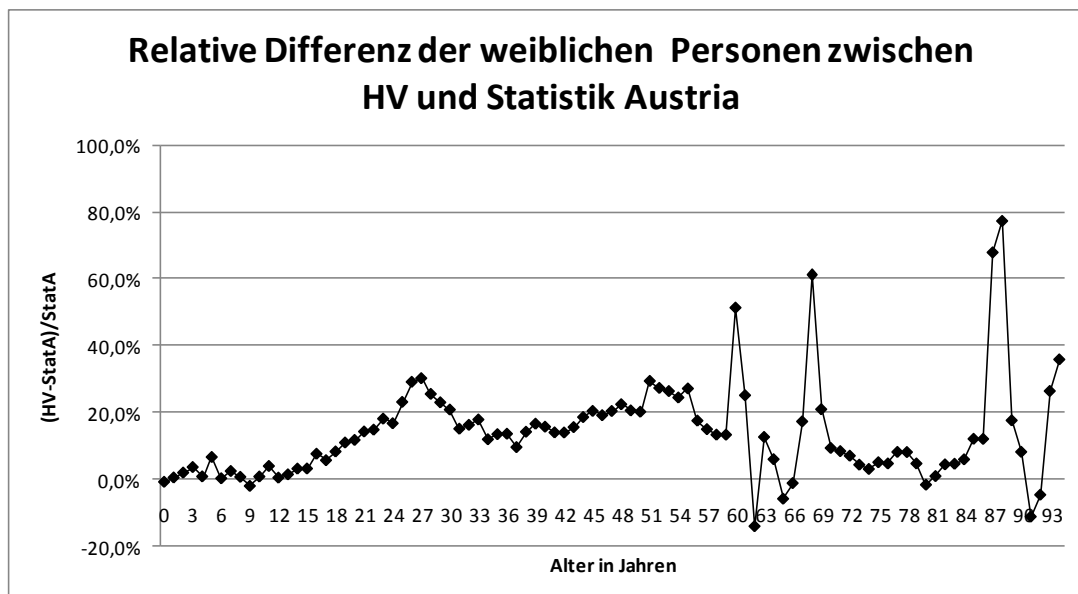


Abbildung 11 (b): Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung **PWB** von Statistik Austria für Frauen in den einzelnen Jahrgängen

Zur Erklärung der Unterschiede bieten sich zwei Argumente an:

- (1) In der Altersgruppe der Berufstätigen (20 – 60 Jahre) ist der Anteil der Arbeitstouristen (siehe Kapitel 1) bei den Männern höher als bei den Frauen.
- (2) Die extrem großen relativen Differenzen in den hohen Altersjahrgängen dürften zu einem Gutteil durch die mangelnde Qualität der Daten über das Sterbedatum zu erklären sein.

3.2 Alter und Bundesland

In der Analyse der einzelnen Merkmale wurde der Raumbezug nur auf Ebene des Bezirks (Merkmal BEZIRK_ID; vgl. Tabelle 14) analysiert. Zum Vergleich mit den Daten von der Statistik Austria wurden die Bezirke auf Ebene der Bundesländer aggregiert und Datensätze mit fehlender Bundesländer-Angabe mit dem Wert 10 markiert. Die fehlenden Daten für den Raumbezug könnten auf Ebene der Bundesländer zum Teil durch den Versicherungsträger imputiert werden, doch wurde dies bisher nicht gemacht. Abbildung 12 (nächste Seite) zeigt den Anteil der fehlenden Bundesländer in den einzelnen Altersjahrgängen. Ab etwa 55 Jahren ist die Qualität der Daten wieder deutlich schlechter, was auch auf die Problematik der nicht korrekt ausgeschiedenen hohen Altersjahrgänge hindeutet.

Tabelle 23 (nächste Seite) zeigt die Verteilung der Altersjahrgänge für Personen über 95 Jahre. Bei etwa 28.500 Personen fehlen in dieser Alterskohorte die Angaben zum Bundesland, was auch ein Indiz dafür ist, dass hier Personen nicht durch Todesdatum ausgeschieden wurden. Auffallend ist auch, dass 40% der Personen in dieser Altersgruppe mit einem eingetragenen Bundesland aus Wien stammen.

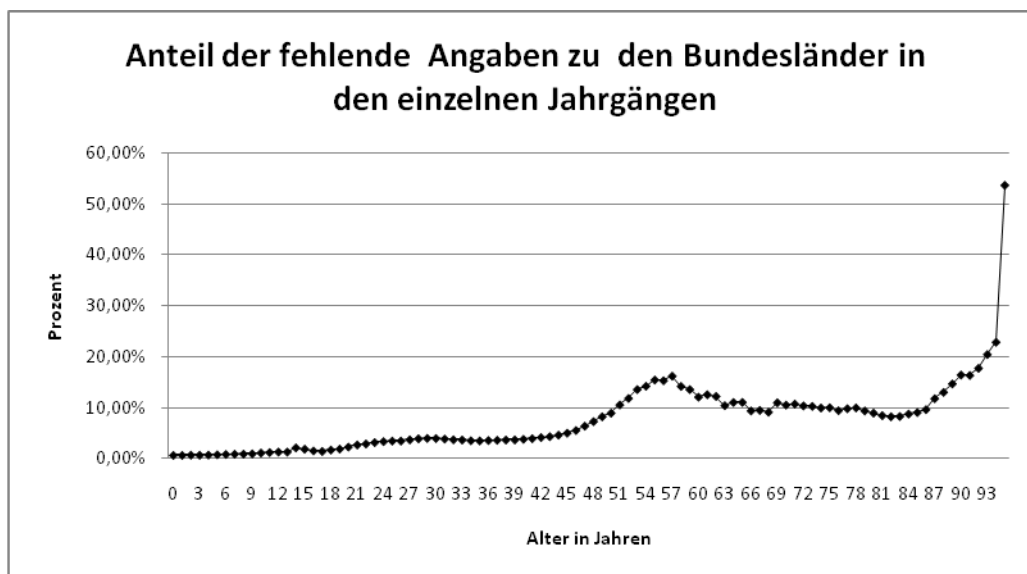


Abbildung 12: Fehlende Angaben zu Bundesländern in den einzelnen Jahrgängen

Tabelle 23: Verteilung der Bundesländer bei Personen über 95 Jahre

		Bundesland			
		Häufigkeit	Prozent	Gültige	Kumulierte
		t		Prozente	Prozente
Gültig	Burgenland	435	,8	1,8	1,8
	Kärnten	1165	2,2	4,7	6,4
	Niederösterreich	2317	4,3	9,3	15,8
	Oberösterreich	3078	5,8	12,4	28,2
	Salzburg	1599	3,0	6,4	34,6
	Steiermark	2793	5,2	11,3	45,9
	Tirol	2888	5,4	11,6	57,5
	Vorarlberg	428	,8	1,7	59,3
	Wien	10108	18,9	40,7	100,0
	Gesamt	24811	46,4	100,0	
Fehlend	10	28633	53,6		
Gesamt		53444	100,0		

Den Vergleich der Altersverteilung nach Bundesländern der beiden Populationen findet man in Abschnitt 3.4.

3.3 Bundesland und Geschlecht

Die Tabellen 24 und 25 zeigen die Verteilung des Geschlechts gegliedert nach Bundesländern in der Hauptverbandspopulation **PHV** bzw. in der Wohnbevölkerung **PWB**. In Tabelle 24 (nächste Seite) wurde das fehlende Geschlecht entsprechend den Ausführungen in Abschnitt 3.1 imputiert. Entsprechend dem höheren Anteil an

Männern in der Hauptverbandspopulation **PHV** ergeben sich leicht unterschiedliche Geschlechtsparitäten. Besonders im Fall des fehlenden Bundeslandes überwiegt der Anteil der Männer mit 60% deutlich.

Das fehlende Bundesland könnte ein Indikator für Anspruchstouristen sein, die nicht zur Wohnbevölkerung zählen. Für die Interpretation der Daten und die Beurteilung der Datenqualität wäre es sehr hilfreich, wenn die *Staatsangehörigkeit* als weiteres Gliederungsmerkmal verwendet werden könnte.

*Tabelle 24: Verteilung des Geschlechts nach Bundesländern in der Hauptverbandspopulation **PHV***

Verteilung des Geschlechts in den Bundesländern Hauptverbandspopulation

			Geschlecht		Gesamt
			männlich	weiblich	
Bundesland	Burgenland	Anzahl	133772	137081	270853
		% innerhalb von Bundesland	49,4%	50,6%	100,0%
	Kärnten	Anzahl	282832	290454	573286
		% innerhalb von Bundesland	49,3%	50,7%	100,0%
	Niederösterreich	Anzahl	831234	826831	1658065
		% innerhalb von Bundesland	50,1%	49,9%	100,0%
	Oberösterreich	Anzahl	723037	720502	1443539
		% innerhalb von Bundesland	50,1%	49,9%	100,0%
	Salzburg	Anzahl	289222	300632	589854
		% innerhalb von Bundesland	49,0%	51,0%	100,0%
	Steiermark	Anzahl	622307	639954	1262261
		% innerhalb von Bundesland	49,3%	50,7%	100,0%
	Tirol	Anzahl	407211	407057	814268
		% innerhalb von Bundesland	50,0%	50,0%	100,0%
	Vorarlberg	Anzahl	185930	187687	373617
		% innerhalb von Bundesland	49,8%	50,2%	100,0%
	Wien	Anzahl	1080069	1072308	2152377
		% innerhalb von Bundesland	50,2%	49,8%	100,0%
	Bundesland fehlend	Anzahl	386047	255174	641221
		% innerhalb von Bundesland	60,2%	39,8%	100,0%
Gesamt		Anzahl	4941661	4837680	9779341
		% innerhalb von Bundesland	50,5%	49,5%	100,0%

Tabelle 25: Verteilung des Geschlechts nach Bundesländern in der Wohnbevölkerung **PWB**

			Geschlecht		Gesamt
			männlich	weiblich	
Bundesland	Burgenland	Anzahl	137020	143065	280085
		% innerhalb von Bundesland	48,9%	51,1%	100,0%
	Kärnten	Anzahl	271033	288689	559722
		% innerhalb von Bundesland	48,4%	51,6%	100,0%
	Niederösterreich	Anzahl	777719	809645	1587364
		% innerhalb von Bundesland	49,0%	51,0%	100,0%
	Oberösterreich	Anzahl	688289	715459	1403748
		% innerhalb von Bundesland	49,0%	51,0%	100,0%
	Salzburg	Anzahl	254885	270614	525499
		% innerhalb von Bundesland	48,5%	51,5%	100,0%
	Steiermark	Anzahl	586636	615657	1202293
		% innerhalb von Bundesland	48,8%	51,2%	100,0%
	Tirol	Anzahl	339910	355569	695479
		% innerhalb von Bundesland	48,9%	51,1%	100,0%
	Vorarlberg	Anzahl	179588	184209	363797
		% innerhalb von Bundesland	49,4%	50,6%	100,0%
	Wien	Anzahl	792109	866370	1658479
		% innerhalb von Bundesland	47,8%	52,2%	100,0%
Gesamt		Anzahl	4027189	4249277	8276466
		% innerhalb von Bundesland	48,7%	51,3%	100,0%

3.4 Alter, Bundesland und Geschlecht

Zum Vergleich der Bevölkerungen **PHV** und **PWB** bezüglich der drei Gliederungsmerkmale Alter, Geschlecht und Bundesland werden in diesem Abschnitt für jedes Bundesland zwei Arten von Grafiken verwendet. Einerseits sind das Bevölkerungspyramiden für die Populationen **PHV** und **PWB** in den einzelnen Bundesländern, andererseits die relativen Abweichungen der Hauptverbandspopulationen von den Wohnbevölkerungspopulationen in den Bundesländern, gegliedert jeweils nach Geschlecht. Die Abbildungen 13 – 30 (folgende Seiten) entsprechen den Abbildungen 10 bzw. 11 aus Abschnitt 3.1 für die einzelnen Bundesländer.

Eine Interpretation der Daten kann am einfachsten durch die Grafiken der relativen Abweichungen erfolgen. Es zeigt sich, dass die Unterschiede der beiden Populationen, wie in Abschnitt 3.1 beschrieben, regional auf der Ebene des Bundeslandes sehr stark variieren. Wegen der Datenprobleme in den hohen Altersjahr-

gängen scheint eine Interpretation für die Altersjahrgänge nur bis 60 Jahre sinnvoll.

- Im *Burgenland* sind die relativen Abweichungen am kleinsten, es scheint sogar eine kleine Untererfassung in der Hauptverbandspopulation vorzuliegen.
- In *Kärnten* spiegelt das Muster der relativen Abweichungen die Abweichungen für ganz Österreich wider, allerdings auf einem etwas niedrigeren Niveau.
- In *Niederösterreich* ist bei der männlichen Bevölkerung eine extreme Spitze in den Altersgruppen von 20 – 30 Jahren festzustellen, bei den Frauen ist die Abweichung geringer als im österreichischen Durchschnitt.
- In *Oberösterreich* liegen die Abweichung bei Männern und Frauen ebenfalls unter dem gesamtösterreichischen Niveau.
- In *Salzburg* liegen bei den Männern Abweichungen vor, die mit jenen von ganz Österreich vergleichbar sind; bei den Frauen hingegen zeigt sich im Alter von 25 – 30 Jahren eine überdurchschnittliche Abweichung.
- In der *Steiermark* liegen die Abweichungen bei Männern und Frauen in einer Größenordnung von 10%, unter dem österreichischen Niveau.
- *Tirol* zeigt ein deutlich unterschiedliches Verhalten. Hier sind Spitzen in den Altersgruppen von 20 Jahren bis etwa 40 Jahren extrem hoch.
- In *Vorarlberg* sind mit Ausnahme der Jahrgänge von 20 – 30 Jahren die Differenzen eher gering.
- *Wien* unterscheidet sich sowohl bei Männern als auch bei Frauen deutlich vom österreichischen Durchschnitt. Auffallend sind die großen Unterschiede in den Jahrgängen von 45 – 55 Jahren bei den Männern. Das Niveau bei den Frauen ist in den Jahrgängen von 25 – 55 Jahren mit fast 40% auch deutlich höher als im österreichischen Durchschnitt.

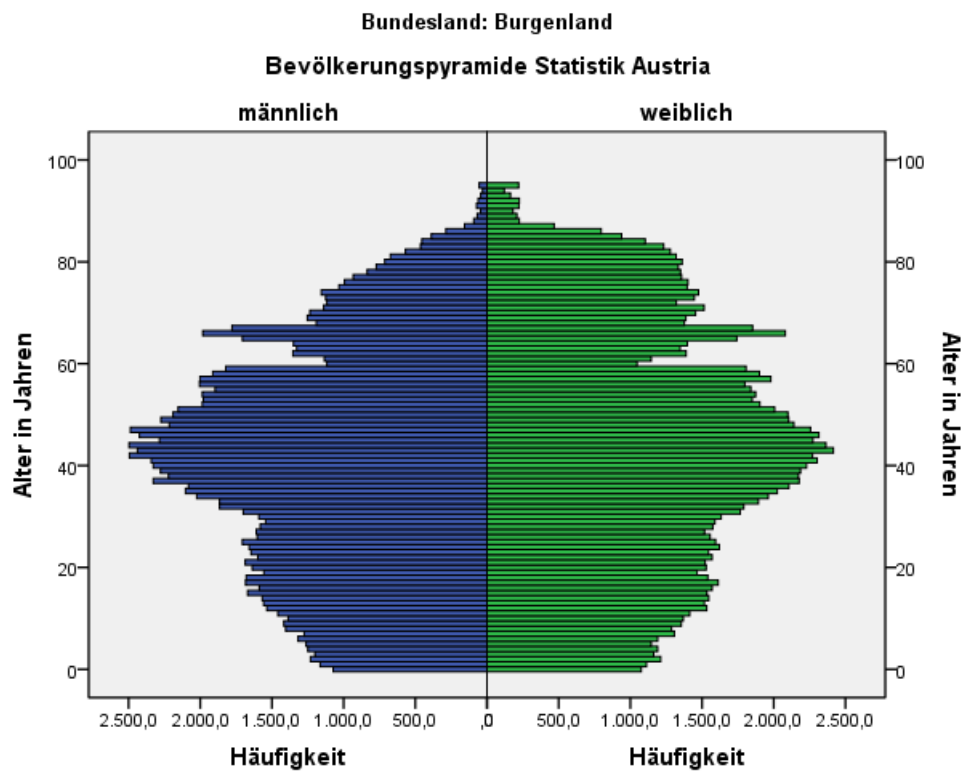
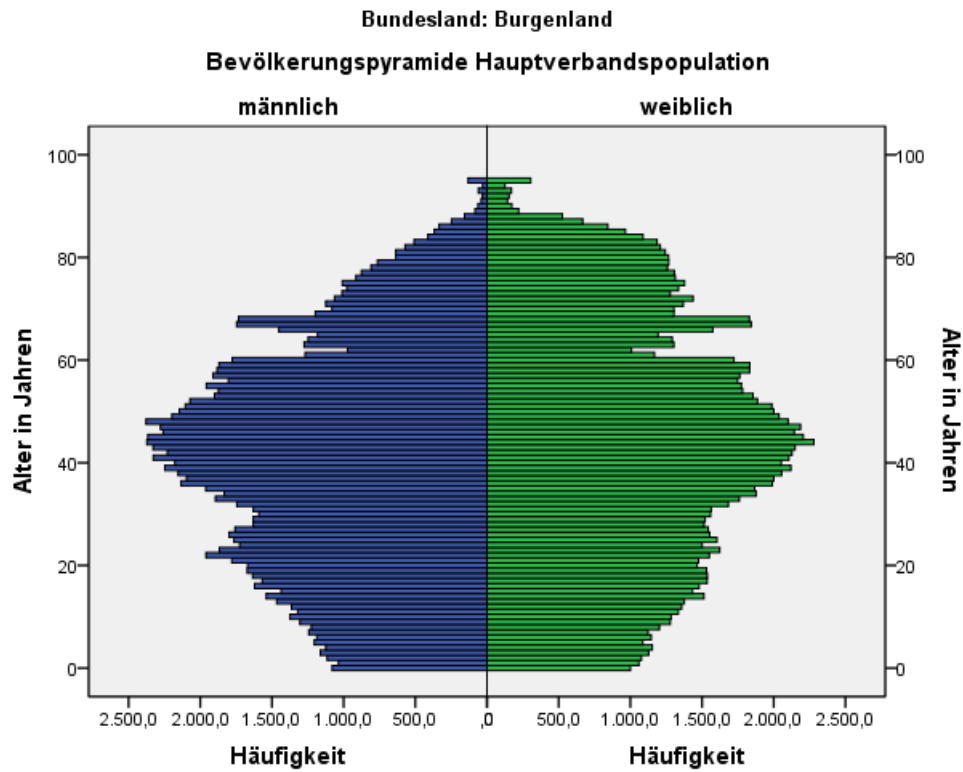


Abbildung 13: Bevölkerungspyramide für Burgenland der Hauptverbandspopulation **PHV** (oben) und der Wohnbevölkerung von Statistik Austria **PWB** (unten)

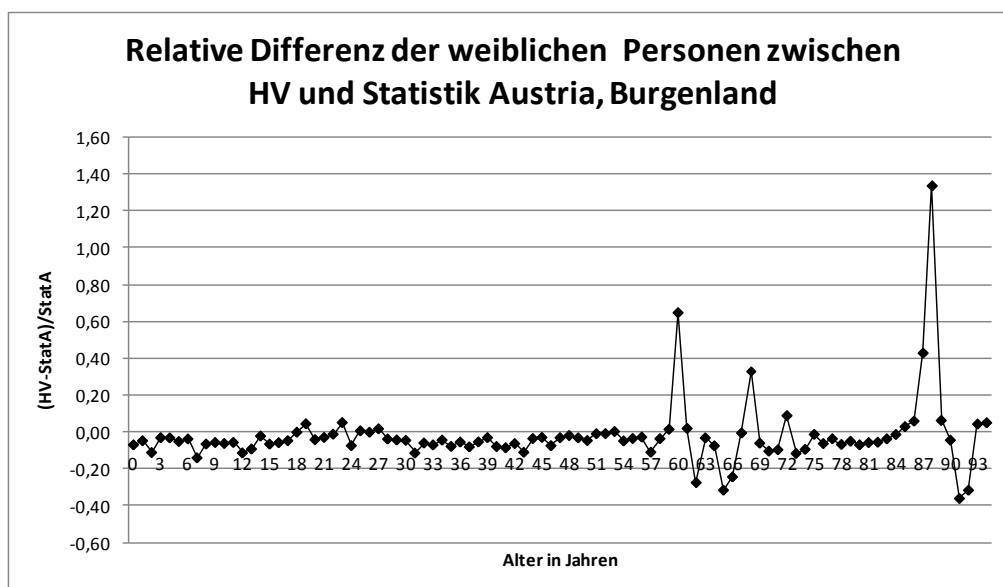
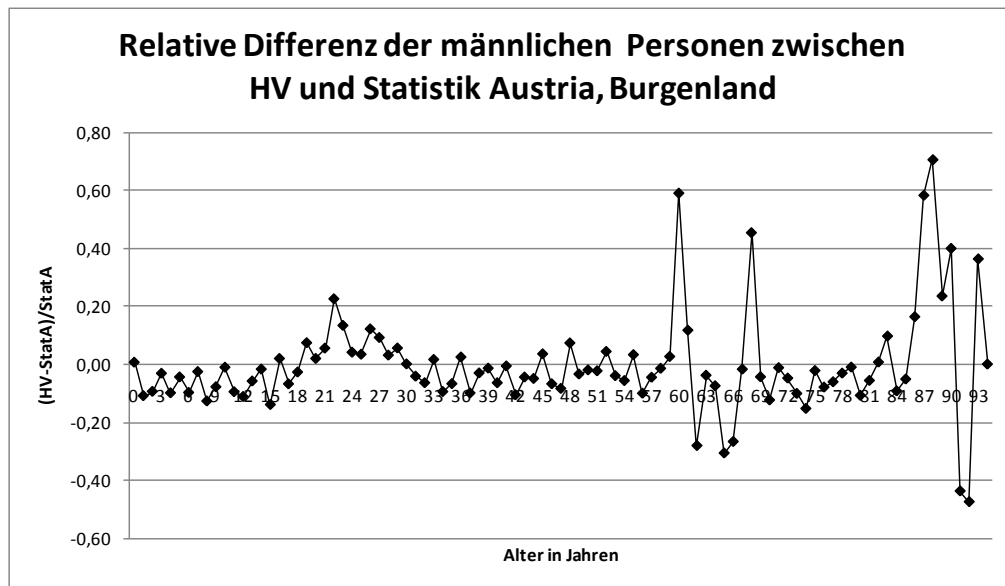


Abbildung 14: Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung von Statistik Austria **PWB** für Männer und Frauen in den einzelnen Jahrgängen im Burgenland

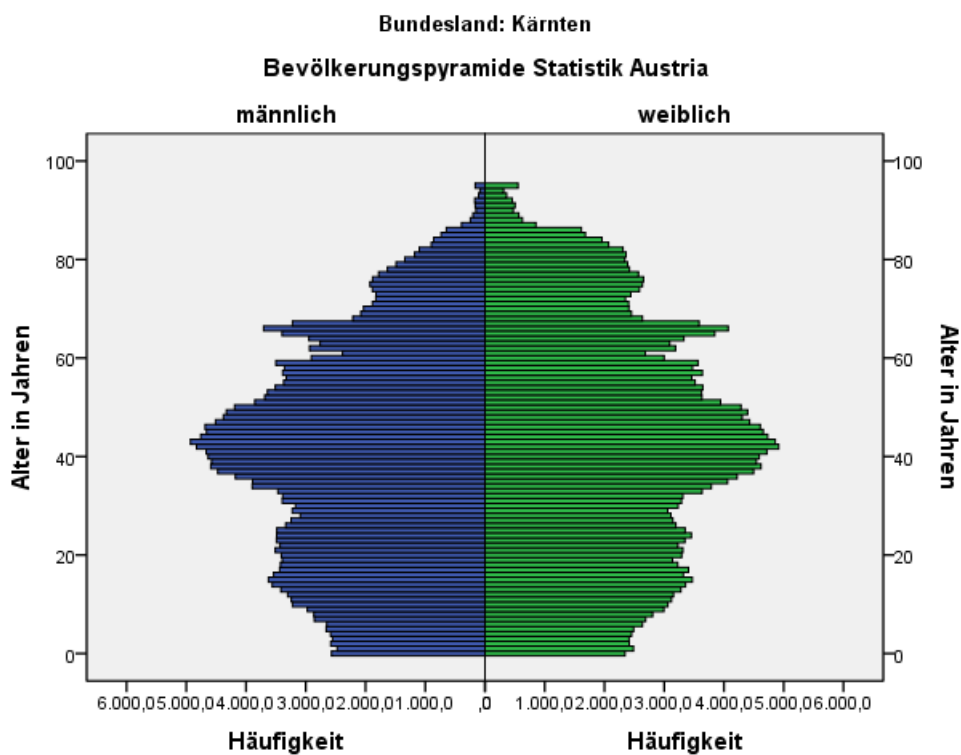
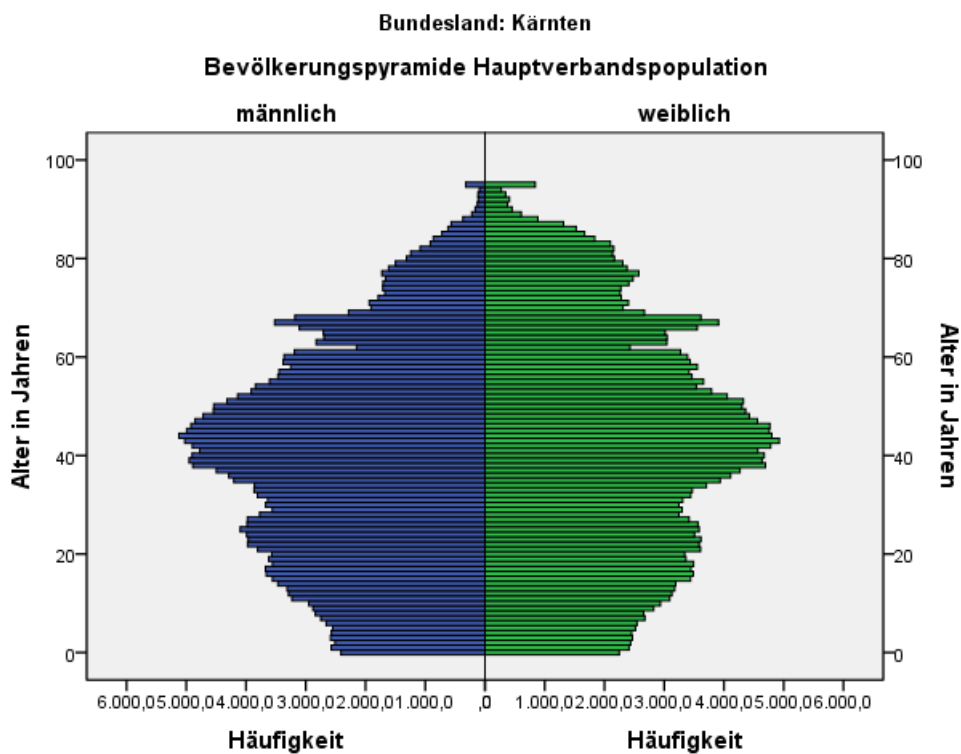


Abbildung 15: Bevölkerungspyramide für Kärnten der Hauptverbandspopulation **PHV** (oben) und der Wohnbevölkerung von Statistik Austria **PWB** (unten)

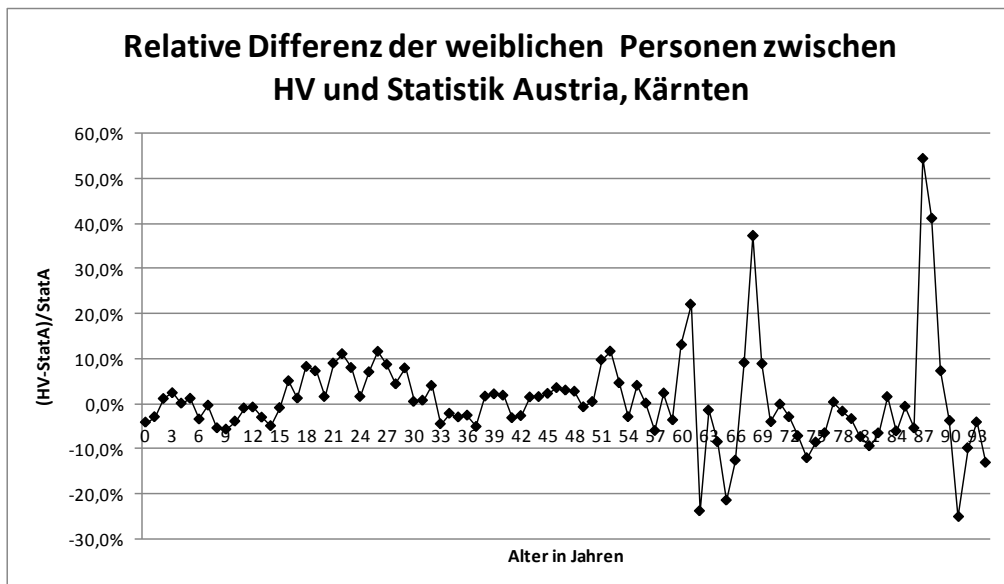
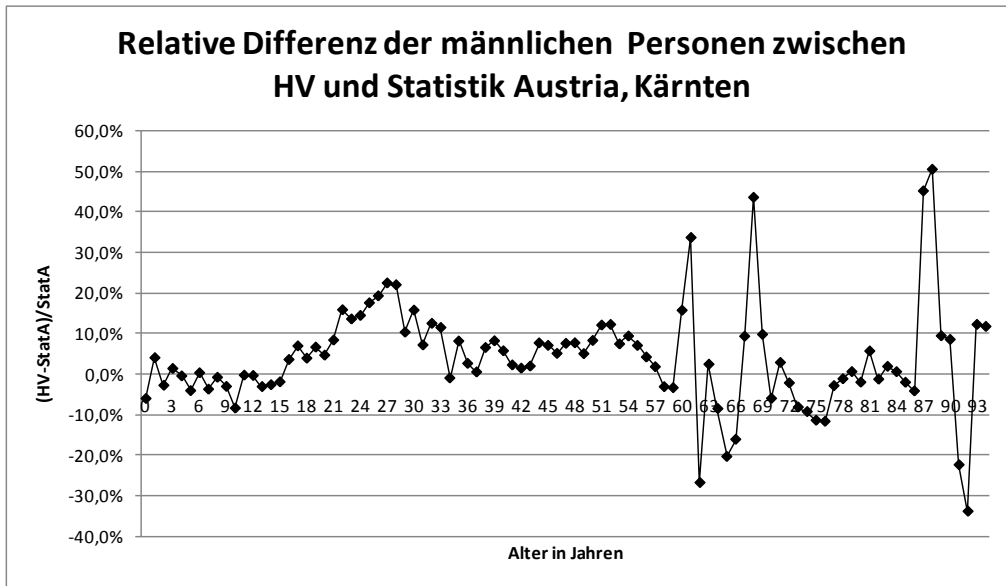


Abbildung 16: Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung von Statistik Austria **PWB** für Männer und Frauen in den einzelnen Jahrgängen in Kärnten

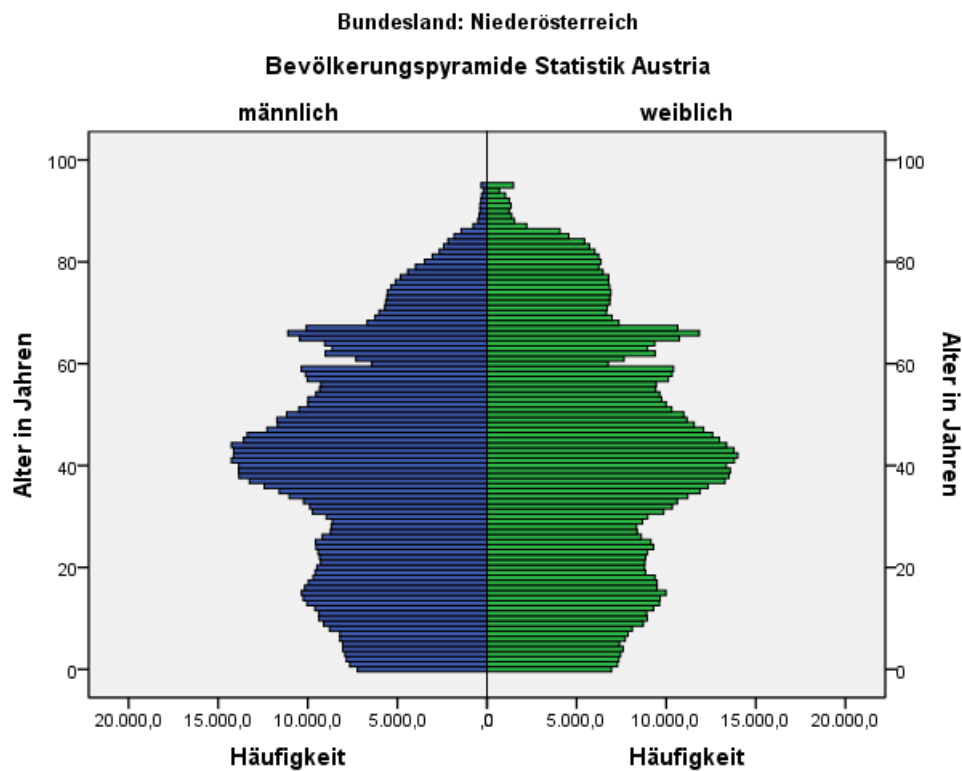
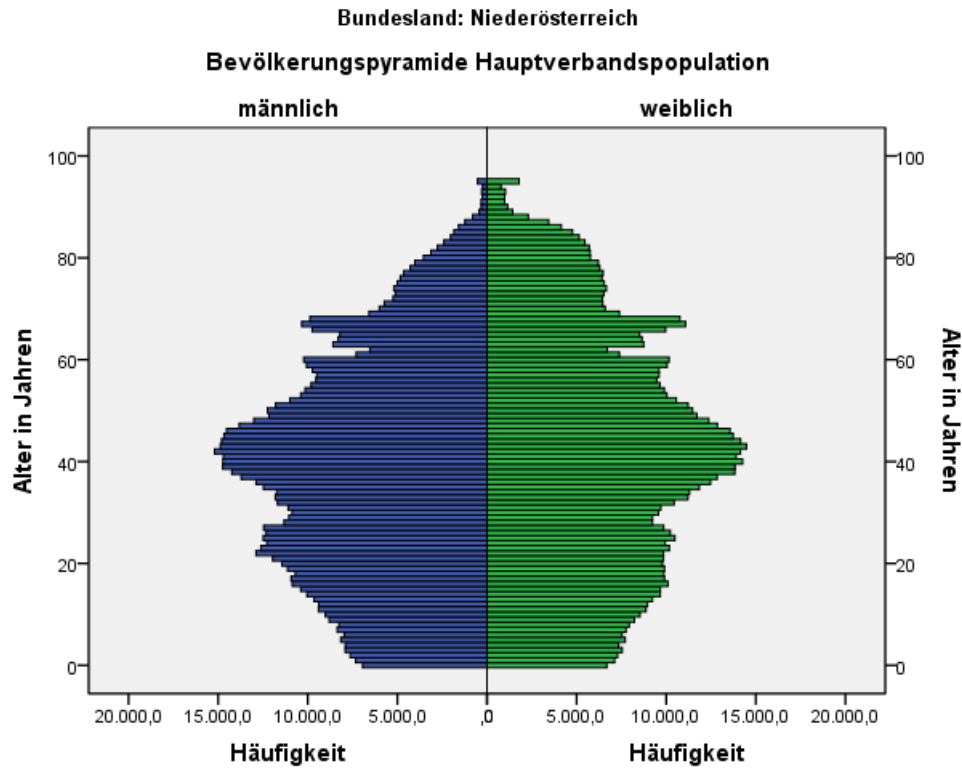


Abbildung 17: Bevölkerungspyramide für Niederösterreich der Hauptverbandspopulation **PHV** (oben) und der Wohnbevölkerung von Statistik Austria **PWB** (unten)

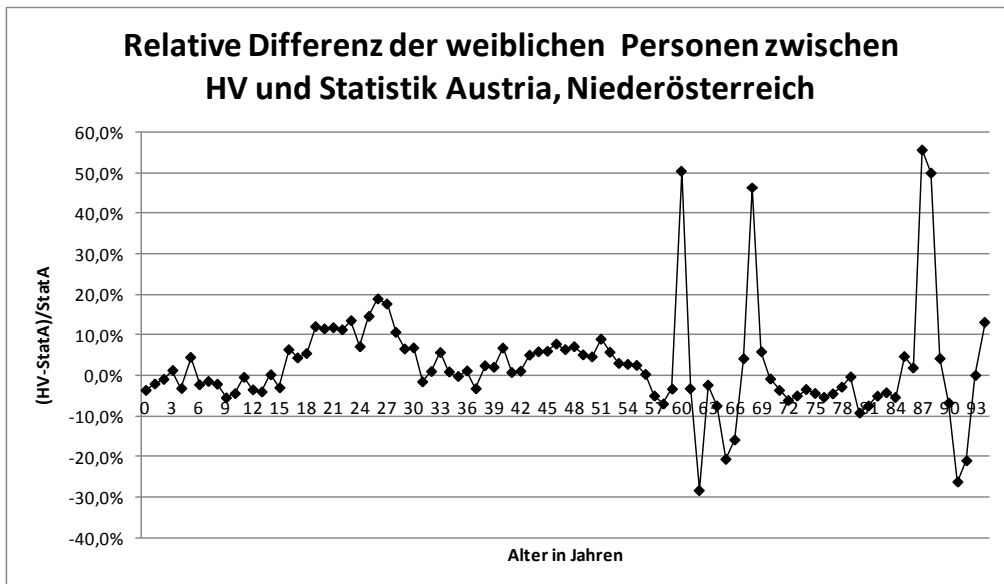
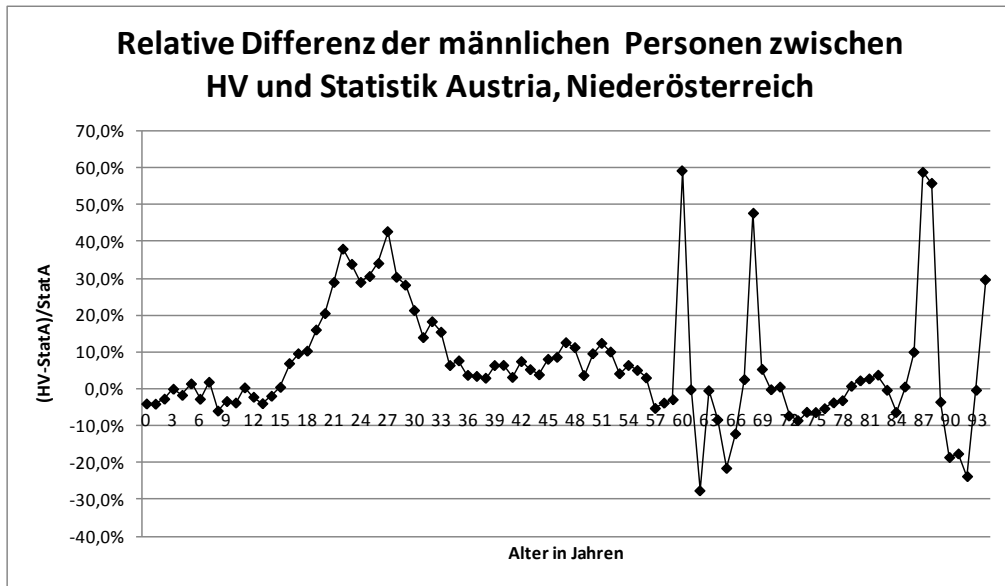


Abbildung 18: Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung von Statistik Austria **PWB** für Männer und Frauen in den einzelnen Jahrgängen in Niederösterreich

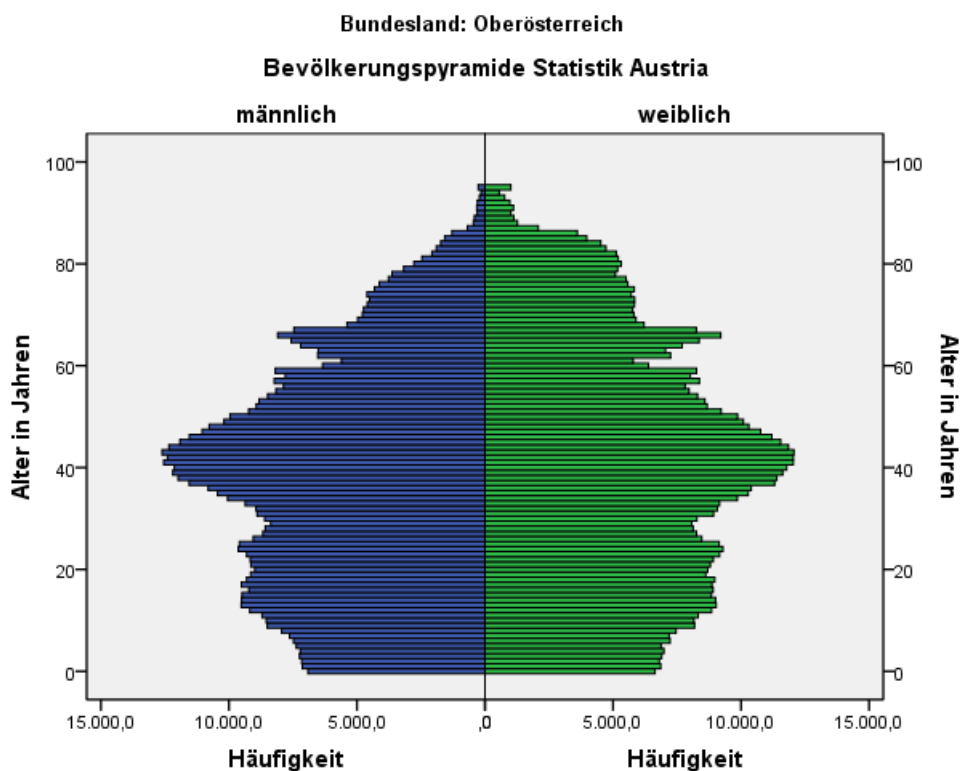
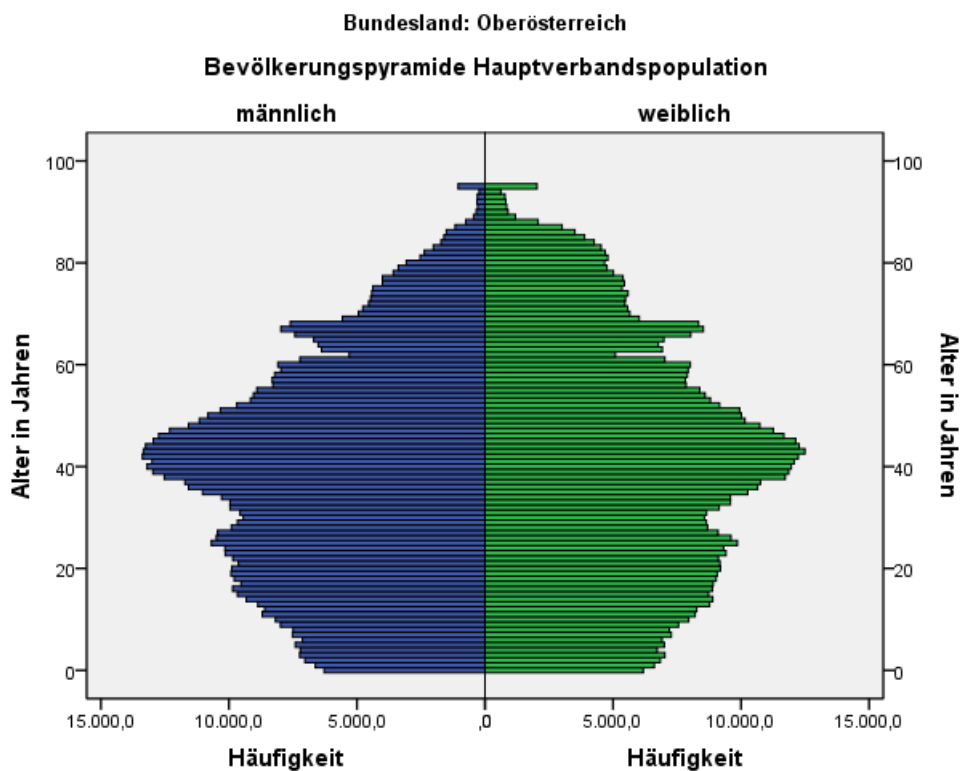


Abbildung 19: Bevölkerungspyramide für Oberösterreich der Hauptverbandspopulation **PHV** (oben) und der Wohnbevölkerung von Statistik Austria **PWB** (unten)

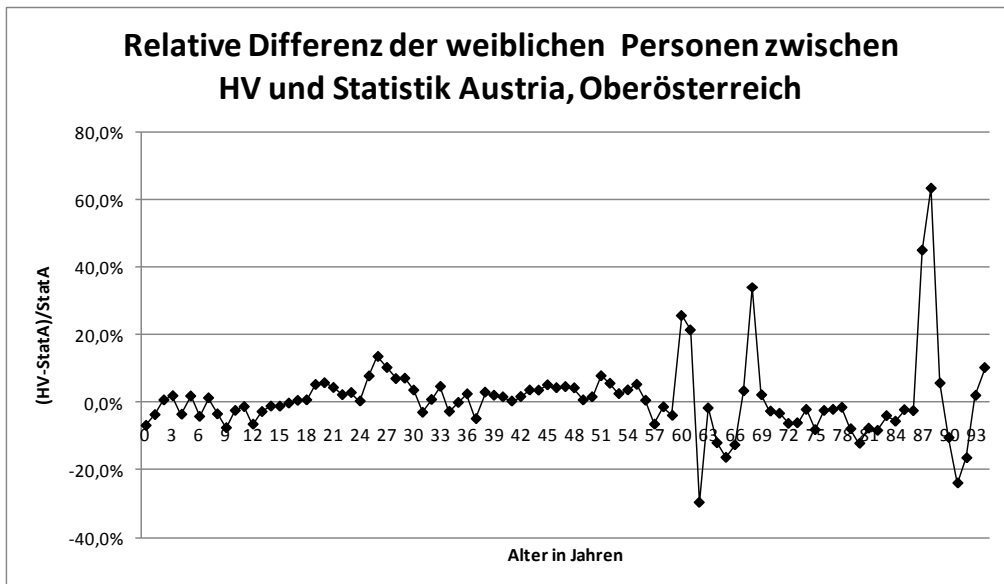
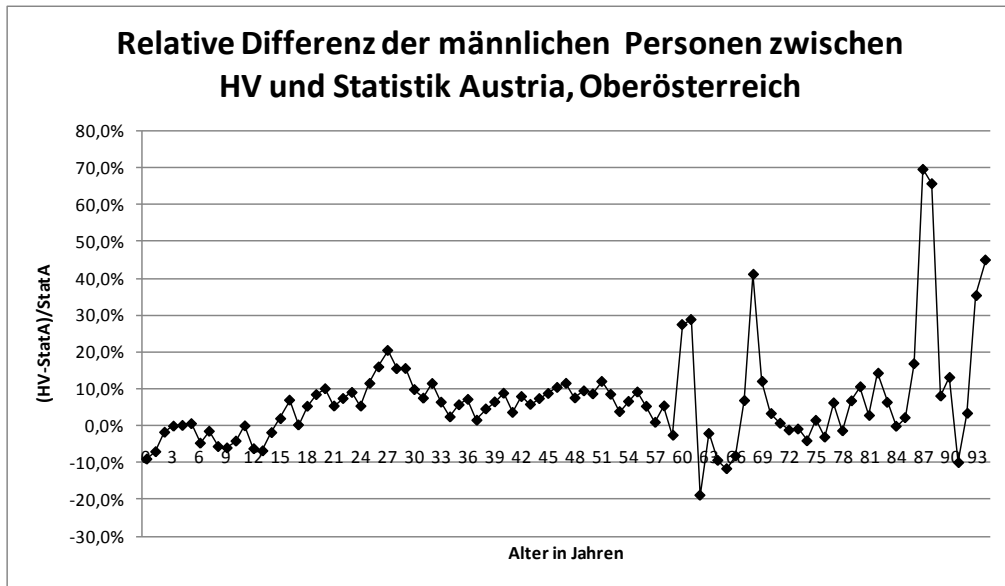


Abbildung 20: Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung von Statistik Austria **PWB** für Männer und Frauen in den einzelnen Jahrgängen in Oberösterreich

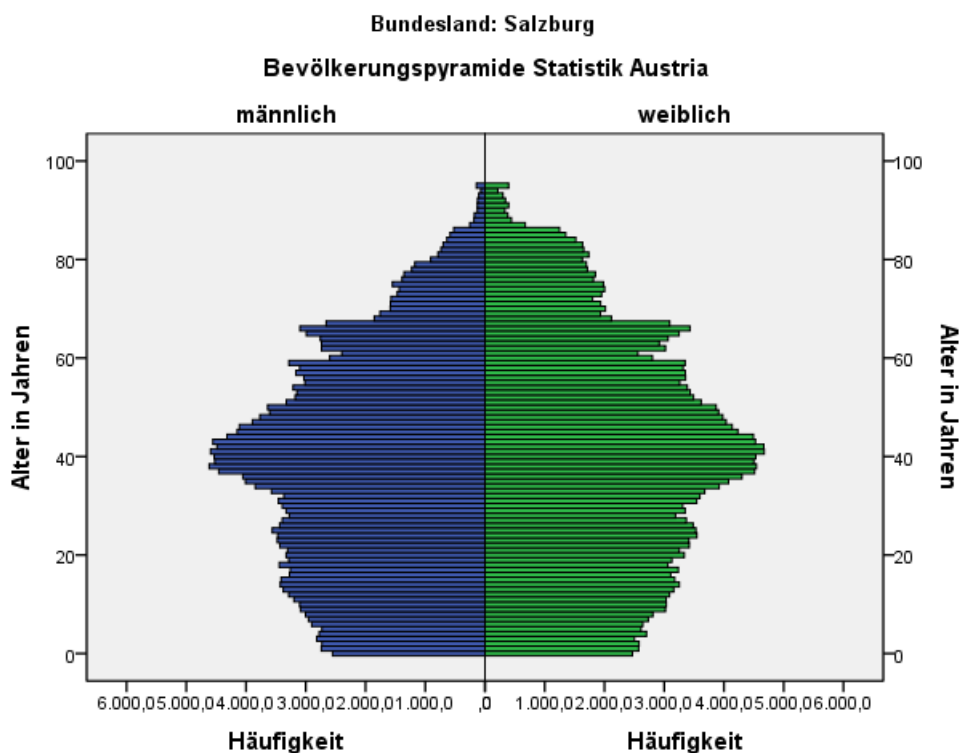
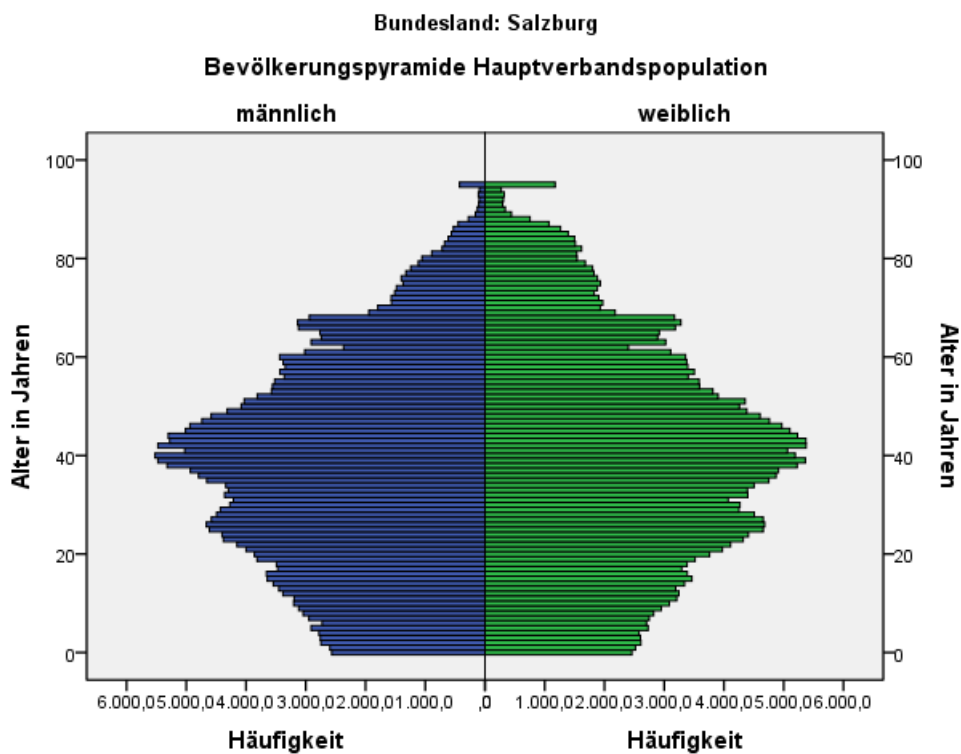


Abbildung 21: Bevölkerungspyramide für Salzburg der Hauptverbandspopulation **PHV** (oben) und der Wohnbevölkerung von Statistik Austria **PWB** (unten)

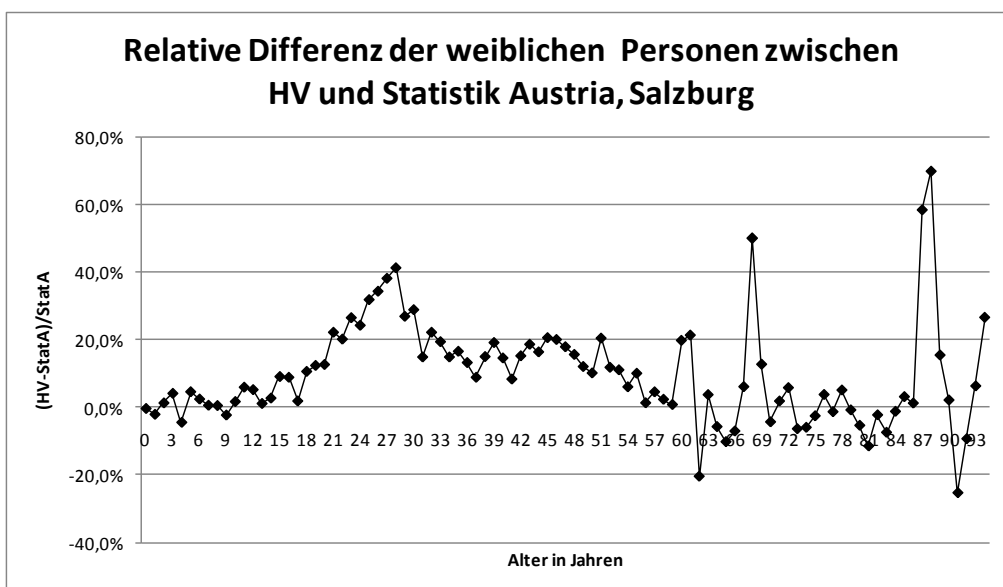
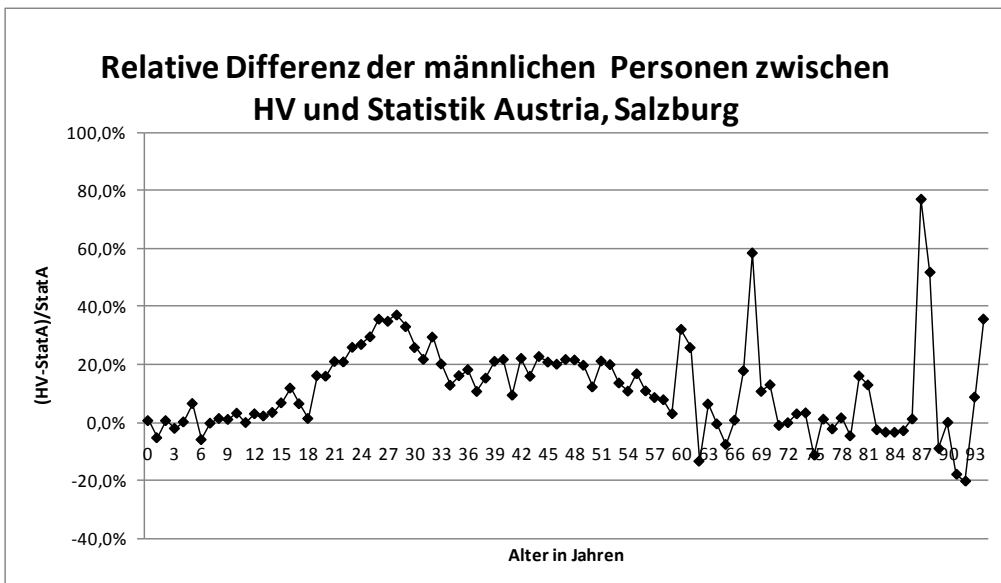


Abbildung 22: Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung von Statistik Austria **PWB** für Männer und Frauen in den einzelnen Jahrgängen in Salzburg

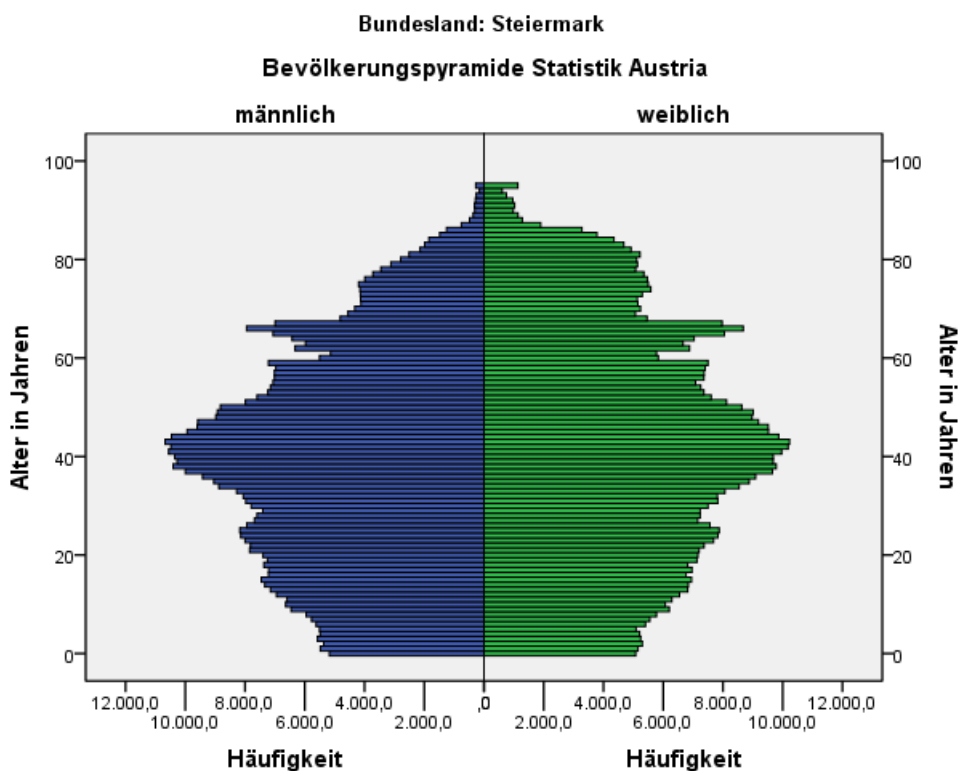
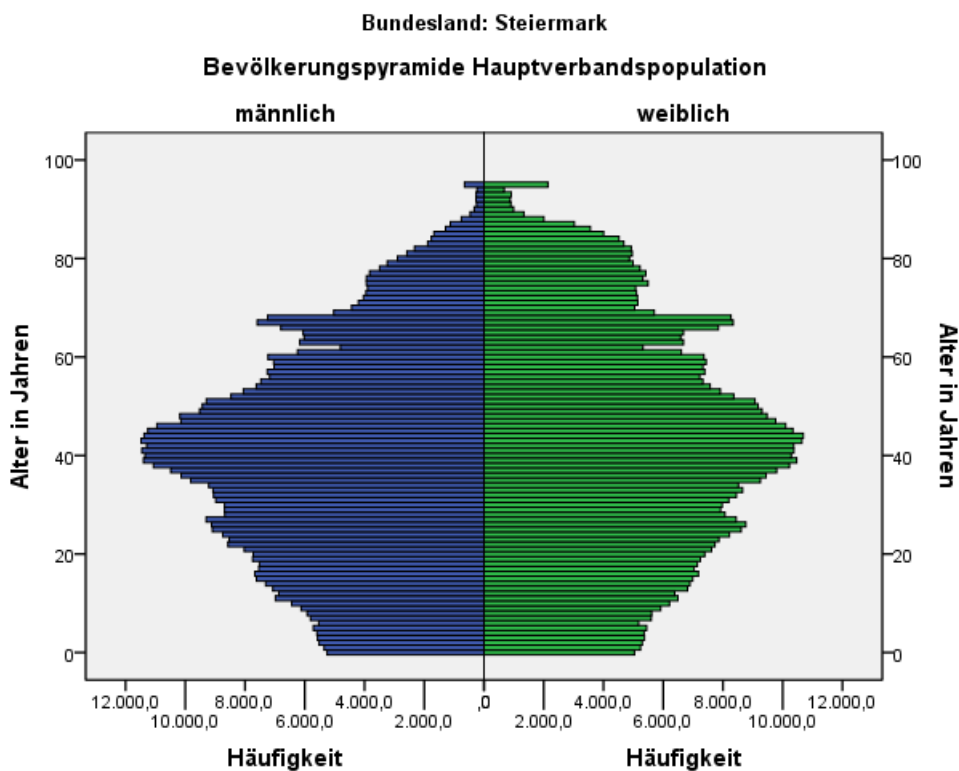


Abbildung 23: Bevölkerungspyramide für die Steiermark der Hauptverbandspopulation **PHV** (oben) und der Wohnbevölkerung von Statistik Austria **PWB** (unten)

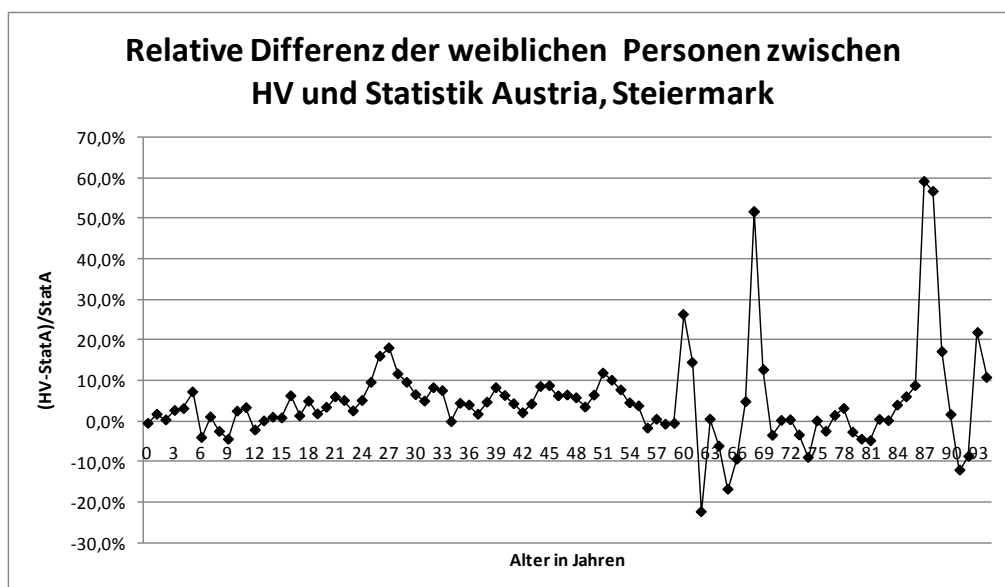
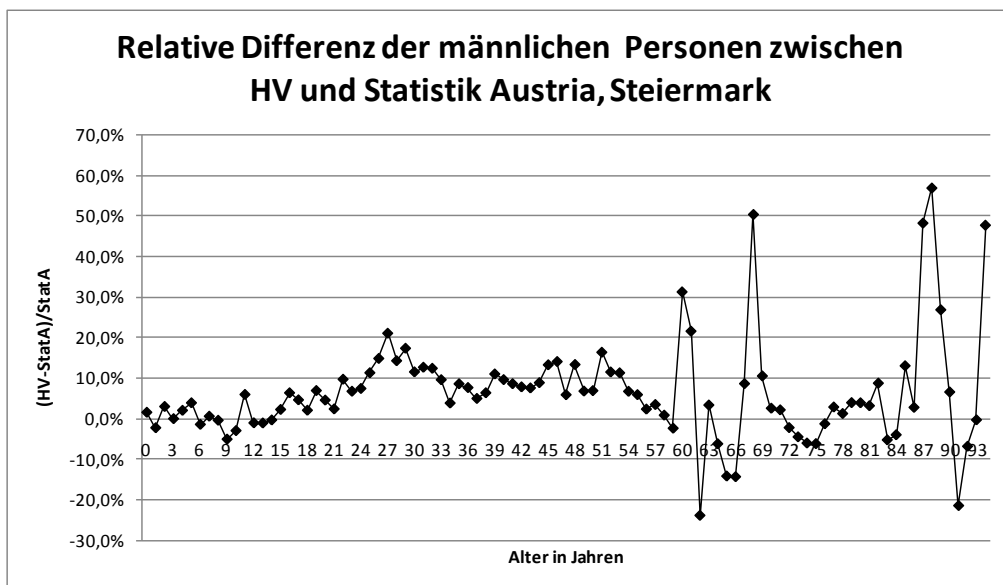


Abbildung 24: Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung von Statistik Austria **PWB** für Männer und Frauen in den einzelnen Jahrgängen in der Steiermark

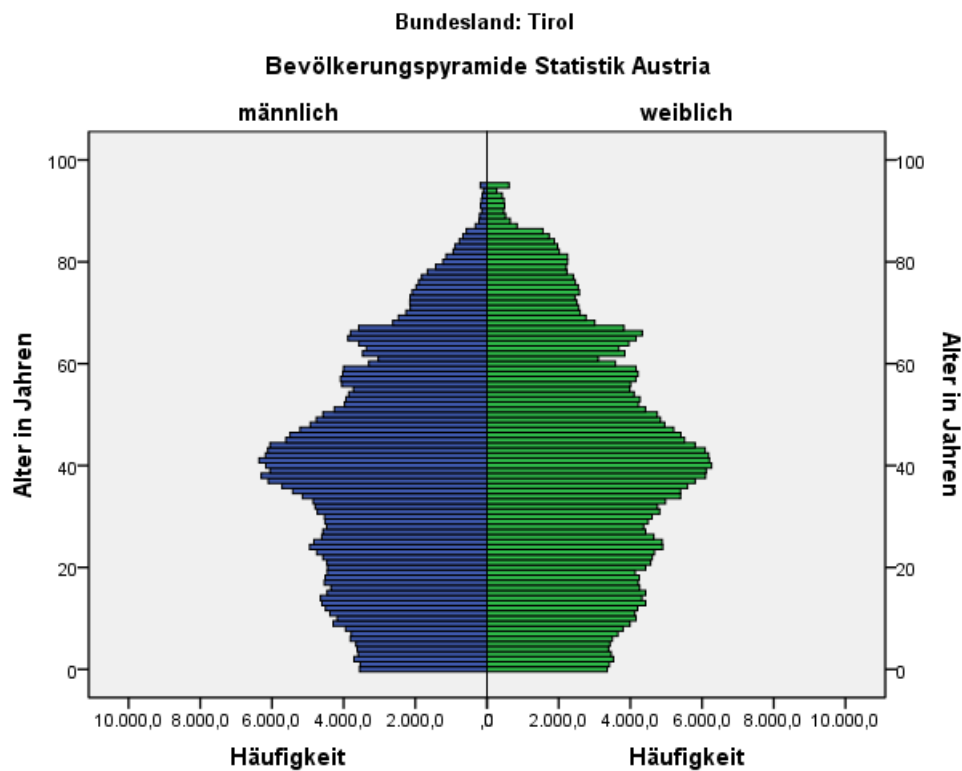
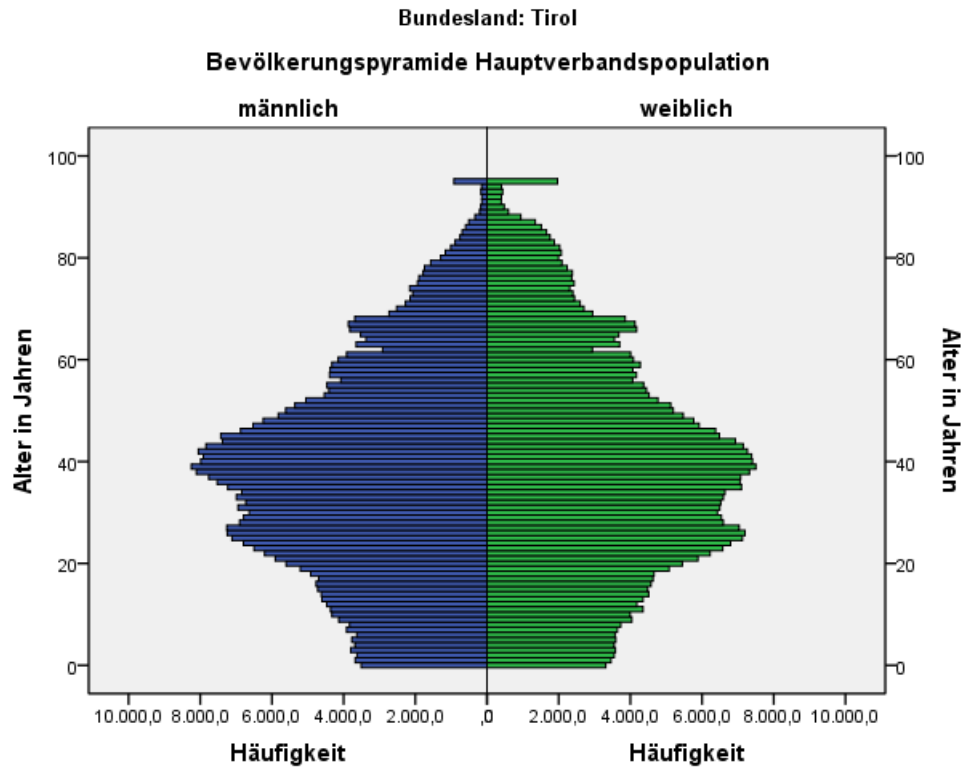


Abbildung 25: Bevölkerungspyramide für Tirol der Hauptverbandspopulation **PHV** (oben) und der Wohnbevölkerung von Statistik Austria **PWB** (unten)

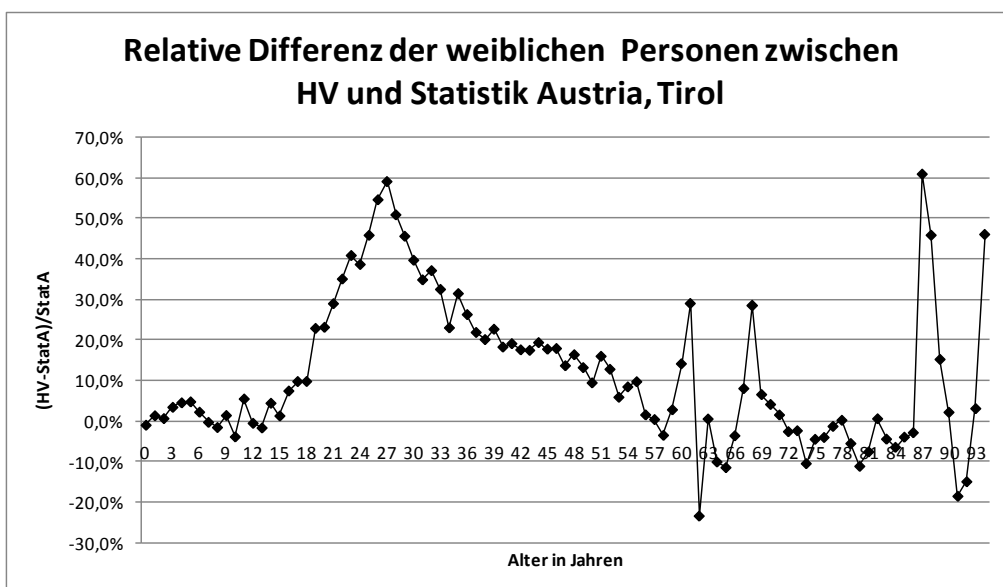
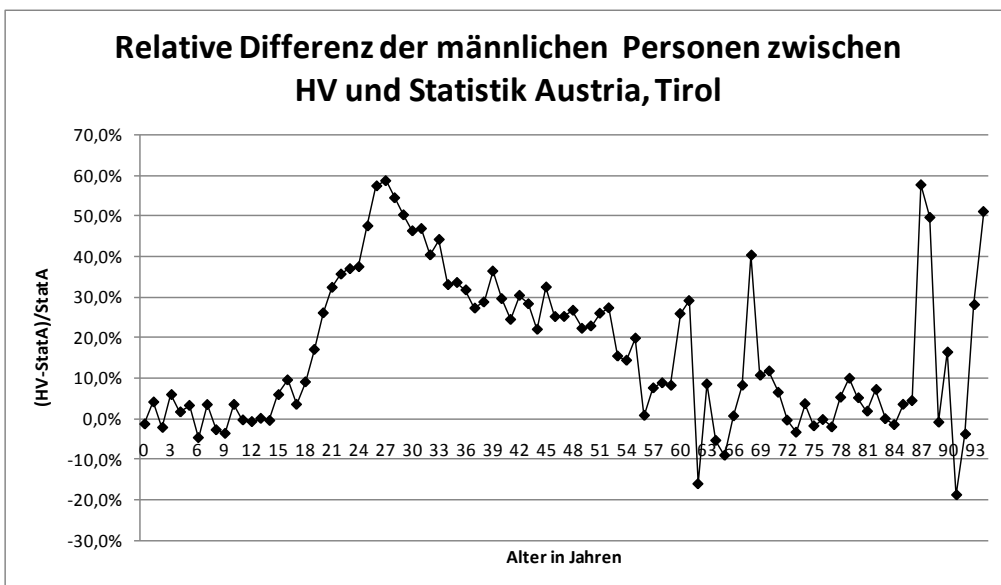


Abbildung 26: Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung von Statistik Austria **PWB** für Männer und Frauen in den einzelnen Jahrgängen in Tirol

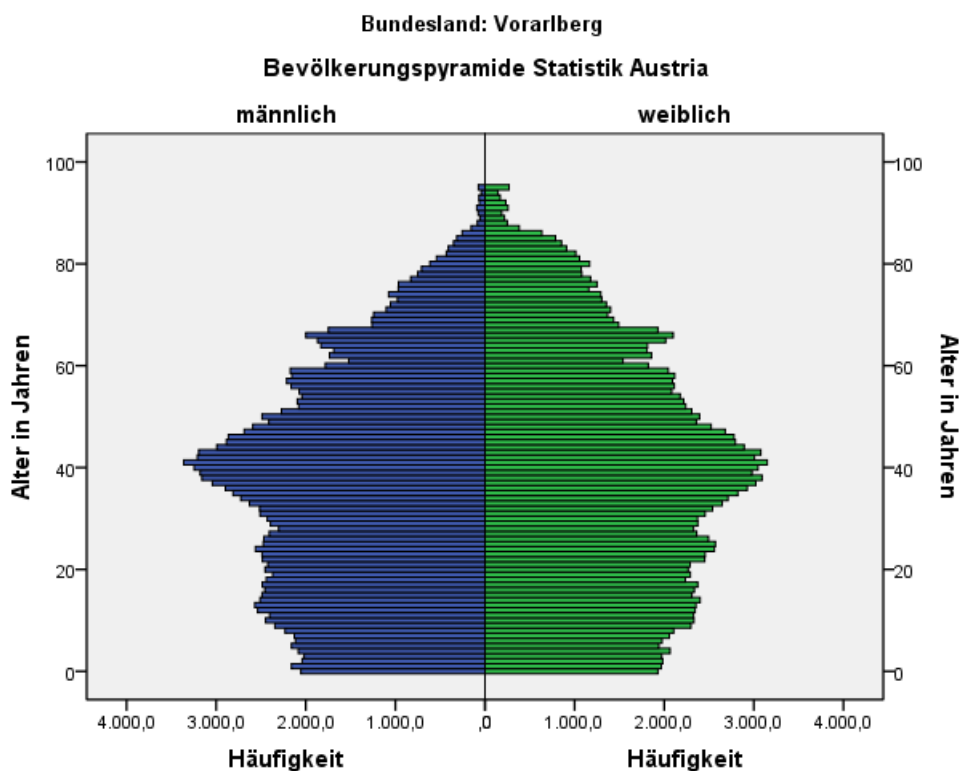
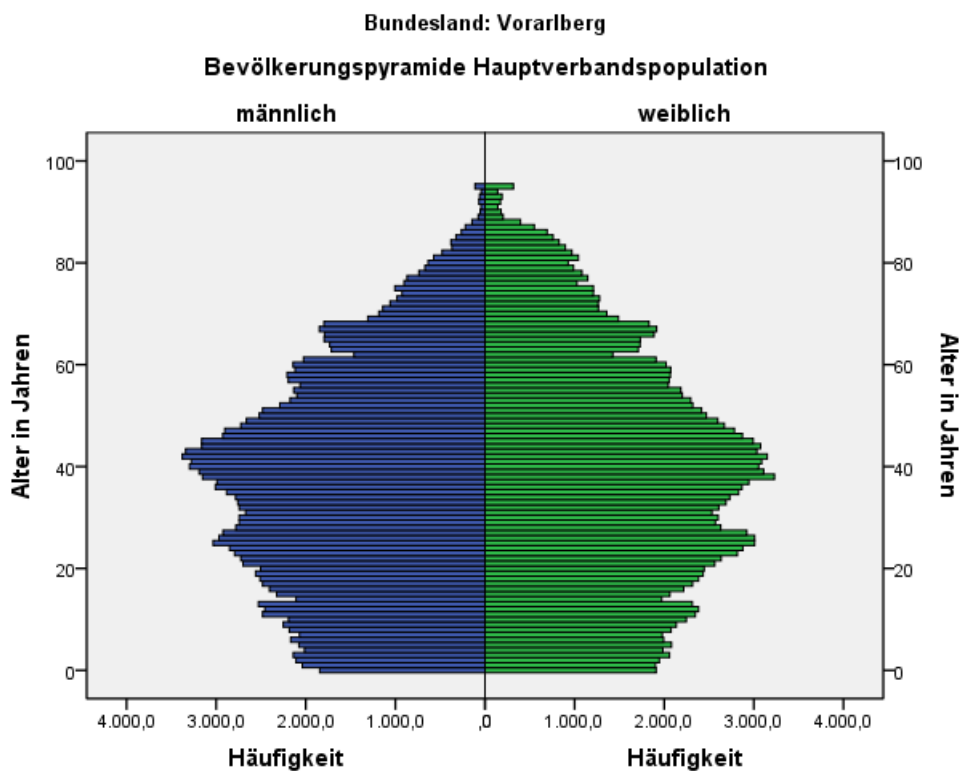


Abbildung 27: Bevölkerungspyramide für Vorarlberg der Hauptverbandspopulation **PHV** (oben) und der Wohnbevölkerung von Statistik Austria **PWB** (unten)

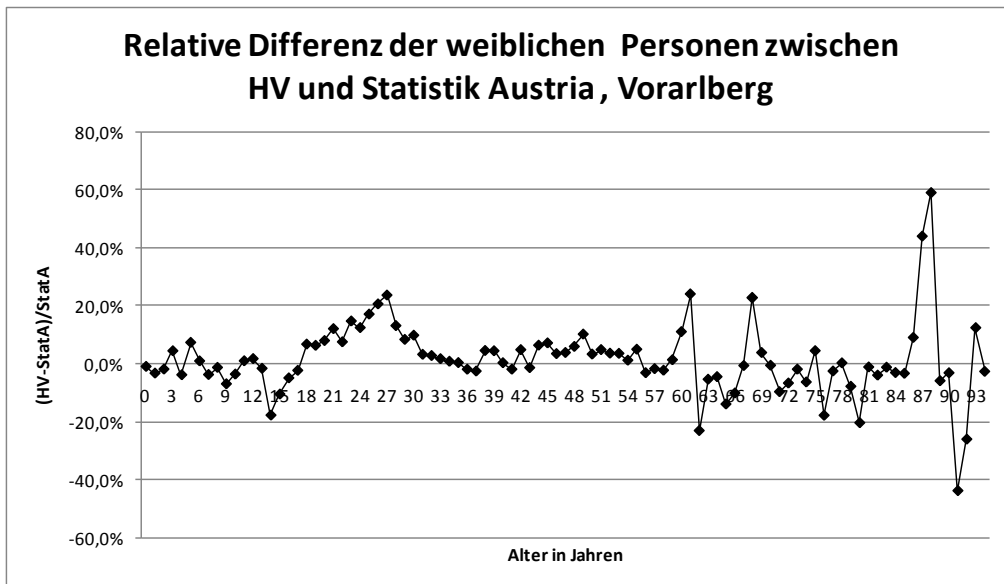
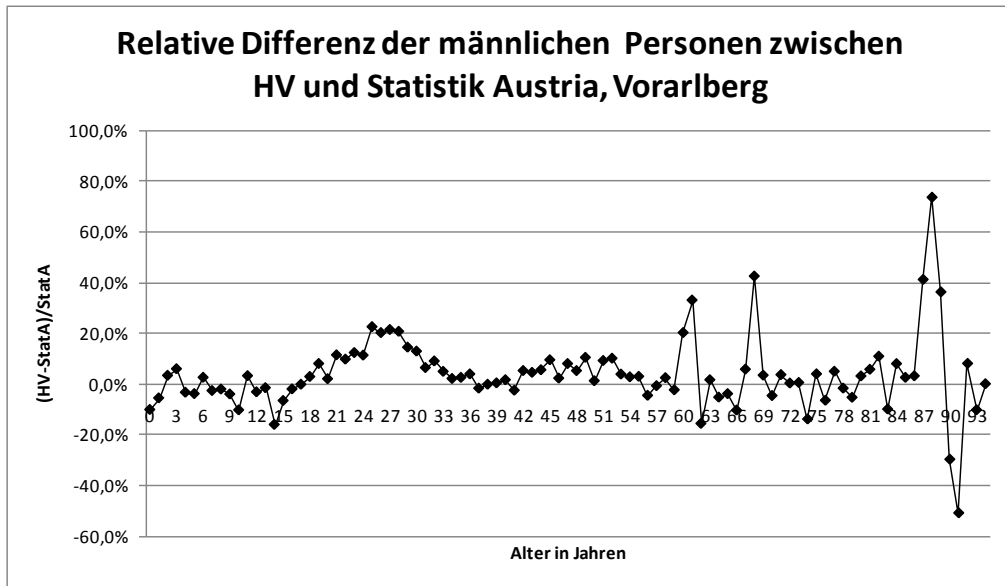


Abbildung 28: Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung von Statistik Austria **PWB** für Männer und Frauen in den einzelnen Jahrgängen in Vorarlberg

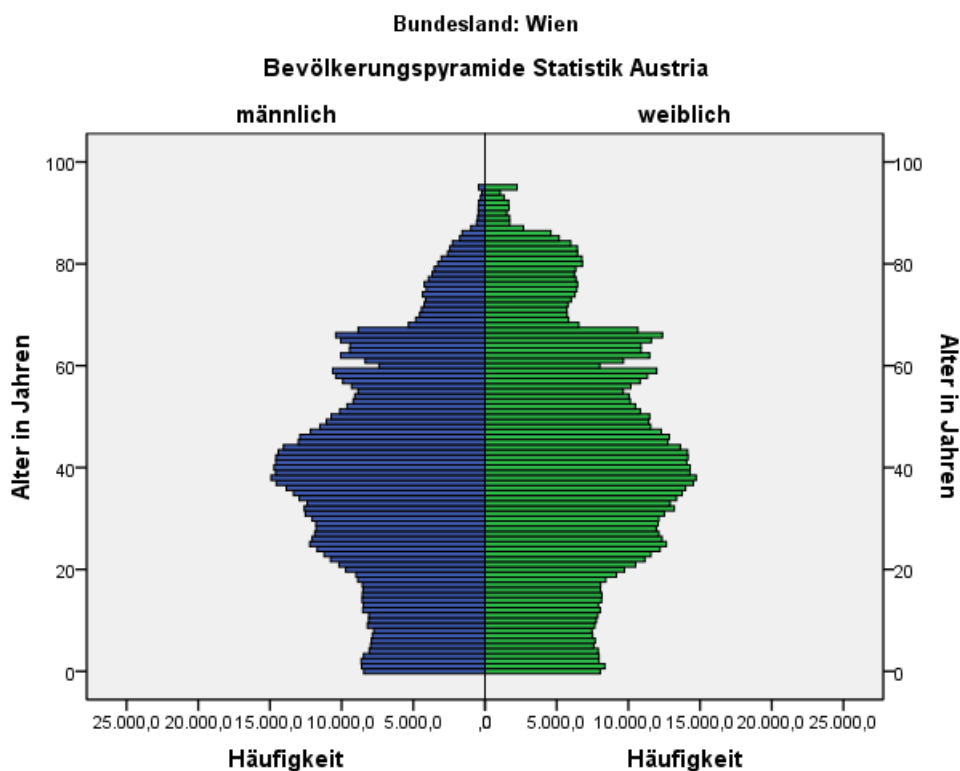
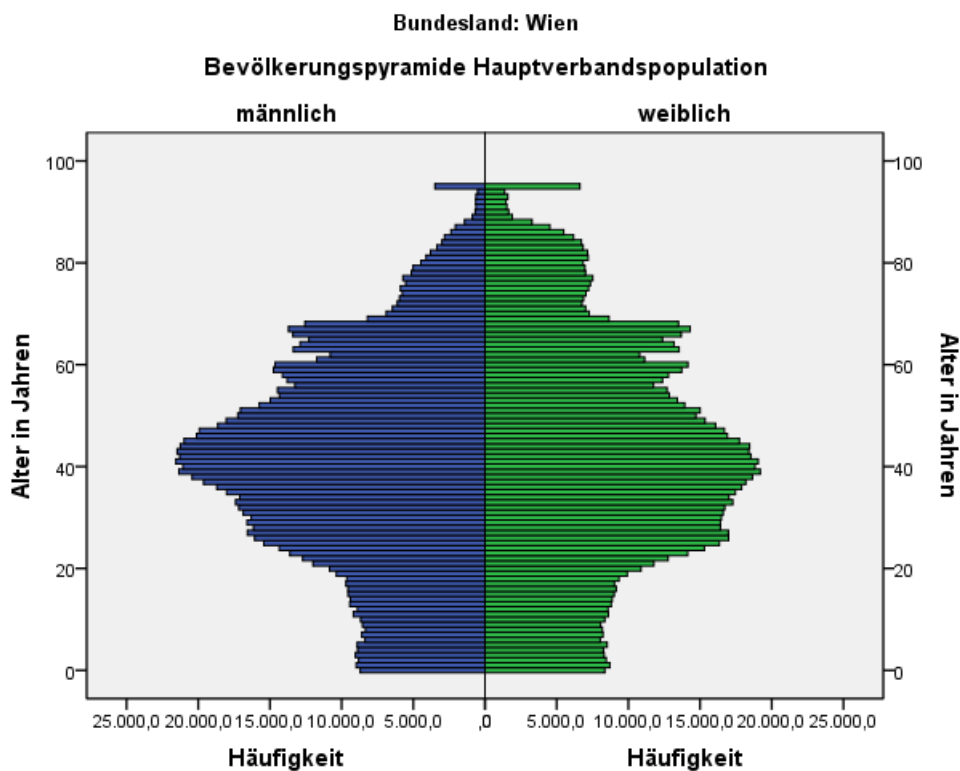


Abbildung 29: Bevölkerungspyramide für Wien der Hauptverbandspopulation **PHV** (oben) und der Wohnbevölkerung von Statistik Austria **PWB** (unten)

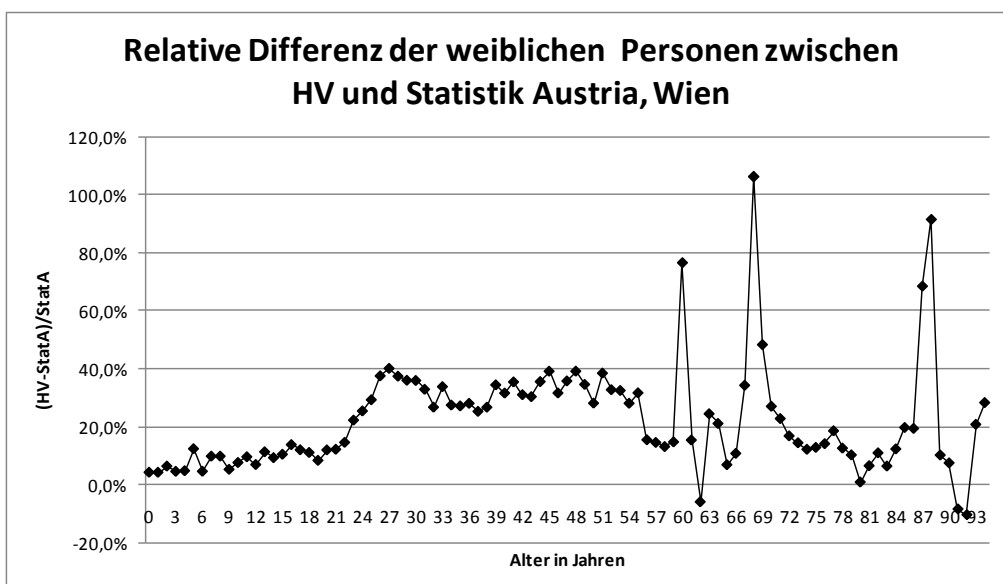
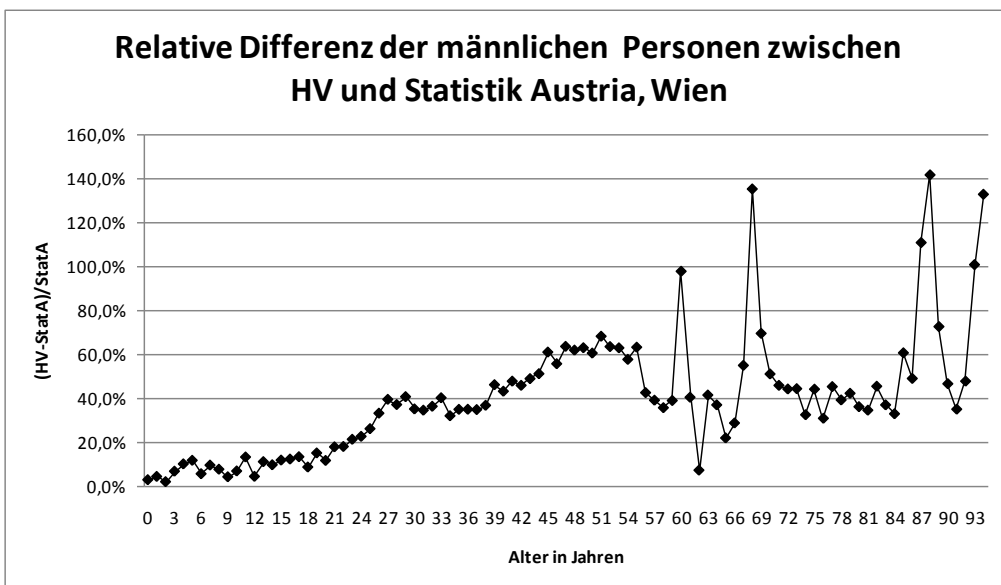
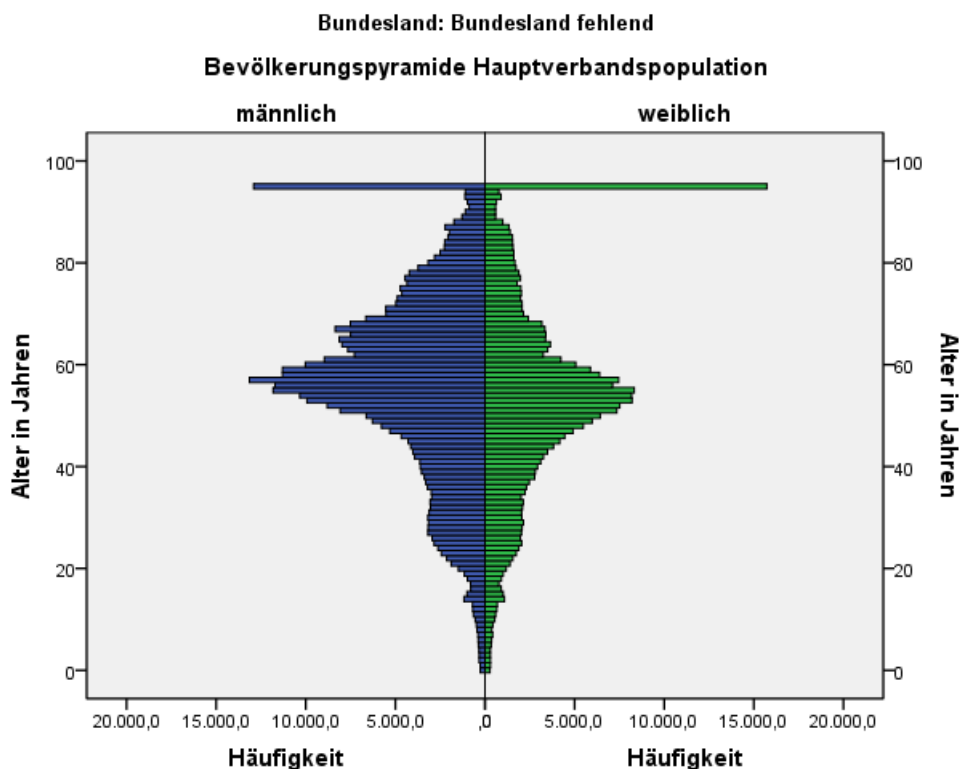


Abbildung 30: Relative Differenz der Hauptverbandspopulation **PHV** zur Wohnbevölkerung von Statistik Austria **PWB** für Männer und Frauen in den einzelnen Jahrgängen in Wien



*Abbildung 31: Bevölkerungspyramide für fehlendes Bundesland der Hauptverbandspopulation **PHV***

Die in Abbildung 31 dargestellte Bevölkerungspyramide macht die schon einige Male angesprochene Datenproblematik noch einmal deutlich. Das fehlende Bundesland tritt vor allem in den Jahrgängen 55 – 60 Jahre auf, besonders bei den Männern. Die extrem große Zahl an fehlenden Bundesländern in der Altersklasse 95 Jahre und älter deutet wieder auf die Problematik der Erfassung der Sterbedaten hin.

Zusammenfassung

Aufgrund der in diesem Bericht dargestellten Ergebnisse lässt sich über die Verwendung der Merkmale in den Tabellen **gapdrg.personen** und **gapdrg.personendaten** der GAP-DRG-Datenbank für die Grundgesamtheit **PHV** folgendes sagen:

Geburtsdatum bzw. Alter der Person: Hier sollten die Werte aus der Tabelle **gapdrg.personen** verwendet werden. Unterschiede zu den Altersangaben in der Tabelle **gapdrg.personendaten** treten vor allem bei der Sozialversicherung der Bauern auf, fehlende Altersangaben sind vor allem bei der BVA, der Niederösterreichischen GKK und der Tiroler GKK im hohem Maße zu finden. Bei einigen Vertragsträgern gibt es offensichtlich falsche Jahrhundertwerte, vor allem bei der Steirischen GKK.

Todesdatum: Das Todesdatum wurde offensichtlich 2006 bereinigt, es scheint aber noch immer Fälle zu geben, wo das Todesdatum nicht eingetragen wurde. Daher ist besonders bei den Alterskohorten über 80 Jahren mit großen Verzerrungen zu rechnen. Die Alterskohorten zwischen 60 und 80 Jahren weisen ebenfalls Verzerrungen auf.

Geschlecht: Besonders schlecht ist die Qualität der Angaben zum Geschlecht bei der Oberösterreichischen Gebietskrankenkasse, bei der Salzburger Gebietskrankenkasse und bei der Sozialversicherung der Gewerblichen Wirtschaft. Hier kann man sich aber mit einer Imputation nach der Geschlechtsparität in der Bevölkerung helfen.

Bezirks_ID: Angaben zum Bezirk fehlen in 6,7% aller Einträge in der Tabelle **gapdrg.personendaten**. Besonders kritisch sind die überregionalen Vertragsträger Eisenbahn und Bergbau und Öffentliche Bedienstete, und die GKK Steiermark, Salzburg und Tirol. Ist man nur an einer regionalen Gliederung in Bundesländer interessiert, so könnte man bei den Gebietskrankenkassen eine Imputation nach der Region der Gebietskrankenkasse durchführen. Bisher wurden fehlende Angaben zum Raumbezug durch ein fiktives zehntes Bundesland ersetzt.

Rezeptgebühr: Angaben zur Rezeptgebühr fehlen vor allem bei den Vertragsträgern Eisenbahn und Bergbau, und den Gebietskrankenkassen in Salzburg und Vorarlberg.

Vertragsdauern: Von den 34.776.257 Verträgen der Tabelle **gapdrg.personendaten** waren am 31.12. 2007 nur 8.718.387 aktuell laufend. Dies deutet darauf hin, dass in der Grundgesamtheit **PHV** sehr viele Anspruchstouristen enthalten sind, also Personen die sich im Zeitraum 2006 – 2007 nur kurzfristig in Österreich aufgehalten haben.

Hinsichtlich der Übereinstimmung der Grundgesamtheit **PHV** mit der österreichischen Wohnbevölkerung **PWB** lässt sich feststellen, dass diese bis zu einem Alter von etwa 60 Jahren bei den Frauen recht gut wiedergegeben wird. Bei den Männern zeigen sich bei den Alterskohorten zwischen 20 und 30 Jahren eine Überrepräsentation in der Grundgesamtheit **PHV**.

Die Hochrechnung der Gesundheitsbefragung auf die österreichische Wohnbevölkerung stimmt mit eine Ausnahme bei den Männern im Alter von 60 – 70 Jahren in der Steiermark bis zu einem Alter von 80 Jahren recht gut mit der Wohnbevölkerung überein.

Teilprojekt 2

Diabetes

Epidemiologische Analysen

1. Ziel

Ziel dieses Teilprojekts ist es, auf Basis der Daten der GAP-DRG Datenbank des Hauptverbands der österreichischen Sozialversicherung das (pseudonymisierte) Kollektiv der in Österreich therapierten Diabetes mellitus (DM)-PatientInnen hinsichtlich verschiedener Charakteristika zu analysieren.

Als erster Schritt soll dabei zunächst überprüft werden, wie präzise sich das DM-Kollektiv in der GAP-DRG Datenbank identifizieren lässt. Zu diesem Zweck wird das in der GAP-DRG Datenbank identifizierte Kollektiv zwei vergleichbaren Kollektiven gegenübergestellt, die aus einem Diabetes-Register bzw. der österreichischen Gesundheitsbefragung 2006/07 stammen. In weiterer Folge soll das aus der GAP-DRG Datenbank extrahierte DM-Kollektiv hinsichtlich der Verteilung in folgenden Dimensionen analysiert werden:

- Bundesland
- Alter
- Geschlecht
- Behandlung mittels
 - o nur Insuline und Analoga
 - o nur orale Antidiabetika oder andere Antidiabetika
 - o Kombination aus obigen beiden Therapiegruppen

Weiters sollen auf Basis des aus der GAP-DRG Datenbank extrahierten DM-Kollektivs Prävalenzen für Diabetes pro Bundesland, Altersgruppe und Geschlecht bestimmt werden.

2. Resultate

2.1 Identifikation des DM-Kollektivs in der GAP-DRG Datenbank

Zunächst wurde das Kollektiv der DM-PatientInnen in der GAP-DRG Datenbank über die rezeptierten Medikationen ermittelt. Dabei wurde nach allen PatientInnen gesucht, die im Zeitraum 2006-2007 mindestens einmal ein Medikament der ATC-Gruppen „A10A – Insuline und Analoga“, „A10B – Orale Antidiabetika“ oder „A10X – Andere Antidiabetika“ verordnet bekommen haben. Dies ist über die Attribute ‚atc_e1‘, ‚atc_e2‘ und ‚atc_e3‘ der Tabelle **gapdrg-heilmittel** eruierbar, die eine nähere Beschreibung zu jeder der in der Tabelle **gapdrg.rezeptdaten** gespeicherten Verordnungen der PatientInnen beinhaltet. Dabei wurden 174.761 männliche und 176.943 weibliche (insgesamt 351.704) DM-PatientInnen identifiziert (siehe Tabelle 26, nächste Seite). Dieses Kollektiv bildet die Basis der weiteren Arbeiten und wird im Folgenden als **DM(GAPDRG)** bezeichnet.

Bei 6.638 (1,9%) dieser Personen wurde aufgrund fehlender Daten in der GAP-DRG Datenbank das Geschlecht künstlich inputiert. Bei 717 (0,2%) der Personen ist der Wohnbezirk nicht dokumentiert. Das Geburtsjahr ist bei allen DM-PatientInnen dokumentiert. Die Analyse der Tabelle **gapdrg.rezeptdaten** zeigte, dass für den Zeitraum 2006–2007 keine Verordnung der ATC-Gruppe ‚A10X‘ in der GAP-DRG Datenbank existiert; die Personen des Kollektivs **DM(GAPDRG)** wurden also ausschließlich mit Medikamenten der ATC-Gruppen ‚A10A‘ und ‚A10B‘ therapiert.

Tabelle 26: DM-PatientInnen auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007 gegliedert nach Bundesland und Geschlecht

Bundesland	männlich	weiblich	Summe
Burgenland	7185	8139	15324
Kärnten	10109	10525	20634
Niederösterreich	37090	37509	74599
Oberösterreich	26791	25829	52620
Salzburg	8870	8563	17433
Steiermark	23686	25834	49520
Tirol	10054	10499	20553
Vorarlberg	5666	5665	11331
Wien	44995	43978	88973
fehlender Wohnort	315	402	717
Summe	174761	176943	351704

Zu beachten ist, dass das Kollektiv **DM(GAPDRG)** nur medikamentös therapierte DM-PatientInnen beinhaltet, andere Therapieformen sind nicht berücksichtigt.

2.2 Konsistenzprüfung des DM-Kollektivs laut GAP-DRG Datenbank

Die inhaltliche Konsistenz des Kollektivs **DM(GAPDRG)** soll geprüft werden, in dem es mit Kollektiven von DM-PatientInnen anderer Datenquellen verglichen wird. Für Personen ab 15 Jahren eignet sich dafür ein entsprechendes aus der österreichischen Gesundheitsbefragung 2006/07¹ gewonnenes Kollektiv. Für Personen unter 15 Jahren kann mit dem an der Klinik für Kinder- und Jugendheilkunde der Medizinischen Universität Wien geführten juvenilen Diabetesregister verglichen werden.

2.2.1 Vergleich GAPDRG versus juveniles Diabetesregister

Es wurde ein Vergleich der Anzahl der DM-PatientInnen für den Stichtag 31.12.2006 durchgeführt. Die Auswertungen des juvenilen Diabetesregisters wurden von Prof. Dr. Edith Schober und Prof. Dr. Thomas Waldhör im Rahmen einer sehr effizienten Kooperation durchgeführt. Für den Vergleich wurden folgende Personen des in Kap. 0 beschriebenen Kollektivs **DM(GAPDRG)** nicht berücksichtigt:

- Personen, die am 31.12.2006 15 Jahre oder älter waren. Auch im juvenilen Diabetesregister sind nur Patientinnen unter 15 Jahren erfasst.
- Personen mit einem Todesdatum nicht später als 31.12.2006; es sollen ja die am 31.12.2006 lebenden Patienten verglichen werden.
- Personen, die erst nach dem 31.12.2006 ein DM-spezifisches Medikament bezogen, bis zum 31.12.2006 jedoch keines. Für diese PatientInnen ist nicht gesichert, dass sie am 31.12.2006 bereits an DM litten.

Aus dem juvenilen Diabetesregister wurde ein vergleichbares Kollektiv ermittelt, das im Folgenden als **DM(JDR)** bezeichnet wird. Laut **DM(GAPDRG)** ergibt sich ein Stand von 1.257 DM-PatientInnen, davon 641 männliche und 616 weibliche. Laut

¹ Siehe

http://www.statistik.at/web_de/dynamic/statistiken/gesundheit/publdetail?id=4&listid=4&detail=457

DM(JDR) ergibt sich ein Stand von 1.316 DM-PatientInnen, davon 672 männliche und 644 weibliche (siehe Tabelle 17).

Tabelle 17: Vergleich der DM-PatientInnen unter 15 Jahren laut Kollektiv DM(GAPDRG) bzw. laut Kollektiv DM(JDR) für den Stichtag 31.12.2006, gegliedert nach Bundesland und Geschlecht

Bundesland	männlich		weiblich		gesamt	
	DM(GAPDRG)	DM(JDR)	DM(GAPDRG)	DM(JDR)	DM(GAPDRG)	DM(JDR)
Burgenland	24	29	24	26	48	55
Kärnten	49	48	35	38	84	86
Niederösterreich	133	144	134	138	267	282
Oberösterreich	121	123	127	133	248	256
Salzburg	36	43	28	32	64	75
Steiermark	87	86	81	94	168	180
Tirol	49	55	53	52	102	107
Vorarlberg	32	30	27	27	59	57
Wien	109	112	106	102	215	214
Fehlender Wohnort	1	2	1	2	2	4
Summe	641	672	616	644	1257	1316
	Delta= 5%		Delta= 5%		Delta= 5%	

Damit ergibt sich eine gleichmäßige Untererfassung von 5% seitens **DM(GAPDRG)**. Als eine potentielle Erklärung für einen Teil dieser Untererfassung kann der Umstand in Betracht gezogen werden, dass die Daten von Personen (im vorliegenden Fall wohl jene Angehörigen, bei denen die PatientInnen mitversichert sind), die über Krankenfürsorgeanstalten versichert sind bzw. von Personen bestimmter Berufsgruppen (z.B. Notare, Anwälte, Ziviltechniker, Ordensmitglieder, Präsenzdienler oder zeitverpflichtete Soldaten) nicht in der GAP-DRG Datenbank enthalten sind.

2.2.2 Vergleich GAPDRG versus Gesundheitsbefragung 2006/07

In der Gesundheitsbefragung 2006/07 wurde in Frage S4.1d eruiert „Hatten Sie jemals Diabetes?“ und als spezifische Unterfrage dazu in Frage S4.4a „Haben Sie aufgrund dieses Problems in den letzten 12 Monaten Medikamente eingenommen oder wurden Sie deswegen behandelt?“. Da das Kollektiv **DM(GAPDRG)** auch auf Basis der diabetesspezifischen Medikamente gebildet wird, ist von einer Vergleichbarkeit mit dem Kollektiv jener Personen auszugehen, die in der Gesundheitsbefragung 2006/07 auf die Fragen S4.1d und S4.4a jeweils mit ‚Ja‘ geantwortet haben. Letzteres wird im Folgenden als DM(GB) bezeichnet.

Aus dem in Kap. 0 beschriebenen Kollektiv **DM(GAPDRG)** wurden jene Personen entfernt, die am 31.12.2006 unter 15 Jahren waren. Dies ist dadurch begründet, dass in der Gesundheitsbefragung nur Personen ab 15 Jahren befragt wurden. Da die Gesundheitsbefragung auf einer Hochrechnung basiert, ist im Gegensatz zum Vergleich mit dem juvenilen Diabetesregister die Entfernung der Teilkollektive „Personen mit einem Todesdatum nicht später als 31.12.2006“ sowie „Personen, die erst nach dem 31.12.2006 ein DM-spezifisches Medikament bezogen“, nicht relevant.

Der Vergleich der DM-PatientInnen laut Kollektiv **DM(GAPDRG)** bzw. laut Kollektiv **DM(GB)** ist in Tabelle 28 dargestellt. Es ergibt sich ein Stand laut **DM(GAPDRG)**

von insgesamt 350.041 (173.905 männliche und 176.136 weibliche) bzw. laut **DM(GB)** von insgesamt 353.039 (158.050 männliche und 194.989 weibliche) Personen. Das ergibt in **DM(GAPDRG)** eine Übererfassung von 9% bei den männlichen und eine Untererfassung von 11% bei den weiblichen PatientInnen, insgesamt eine Untererfassung von 1%.

Tabelle 28: Vergleich der DM-PatientInnen ab 15 Jahren laut Kollektiv DM(GAPDRG) bzw. laut Kollektiv DM(GB) gegliedert nach Bundesland und Geschlecht

Bundesland	männlich		weiblich		Gesamt	
	DM(GAPDRG)	DM(GB)	DM(GAPDRG)	DM(GB)	DM(GAPDRG)	DM(GB)
Burgenland	7153	7233	8111	6563	15264	13796
Kärnten	10046	9443	10482	13624	20528	23067
Niederösterreich	36907	26235	37347	39333	74254	65568
Oberösterreich	26632	16803	25667	30821	52299	47624
Salzburg	8828	7925	8525	7930	17353	15855
Steiermark	23577	25381	25734	27608	49311	52989
Tirol	9984	11748	10428	12098	20412	23846
Vorarlberg	5628	4886	5627	6488	11255	11374
Wien	44836	48396	43814	50524	88650	98920
Fehlender Wohnort	314		401		715	
Summe	173905	158050	176136	194989	350041	353039
	Delta= -9%		Delta= 11%		Delta= 1%	

Als eine potentielle Erklärung für einen Teil dieser Untererfassung kann der Umstand in Betracht gezogen werden, dass die Daten von Personen, die über Krankenfürsorgeanstalten versichert sind bzw. von Personen bestimmter Berufsgruppen (z.B. Notare, Anwälte, Ziviltechniker, Ordensmitglieder, Präsenzdienler oder zeitverpflichtete Soldaten) nicht in der GAP-DRG Datenbank enthalten sind.

2.3 Verteilung der DM-PatientInnen laut GAPDRG hinsichtlich Bundesland, Alter, Geschlecht und Medikation

Im Folgenden wird die Verteilung des Kollektivs **DM(GAPDRG)** auf die verschiedenen Medikations-Varianten beleuchtet, wobei folgende Medikations-Varianten unterschieden werden:

- Therapie mittels Insulinen und Analoga (entspricht ATC-Gruppe ‚A10A‘) oder oralen Antidiabetika (entspricht ATC-Gruppe ‚A10B‘). Dies beinhaltet auch PatientInnen mit Kombinationstherapien, d.h. Patienten die Medikamente aus mehr als einer der zwei ATC-Gruppen erhalten. Diese Therapievariante fasst im Kontext dieser Arbeit das Gesamtkollektiv der DM-PatientInnen zusammen; über alle Altersklassen summiert entspricht dies der in Tabelle 26 dargestellten Verteilung.
- Therapie nur mittels Insulinen und Analoga (entspricht ATC-Gruppe ‚A10A‘). PatientInnen dieses Kollektivs erhielten im Betrachtungszeitraum keine Therapie mittels oralen Antidiabetika (‚A10B‘).
- Therapie nur mittels oraler Antidiabetika (‚A10B‘). PatientInnen dieses Kollektivs erhielten im Betrachtungszeitraum keine Therapie mittels Insulinen und Analoga (entspricht ATC-Gruppe ‚A10A‘).

- Kombinationstherapie. PatientInnen dieses Kollektivs erhielten im Betrachtungszeitraum mindestens eine Therapie mit Insulinen und Analoga („A10A“) als auch mindestens eine Therapie mit oralen Antidiabetika („A10B“).

Tabelle 29 zeigt die Verteilung des Gesamtkollektivs der DM-PatientInnen aller Altersklassen auf die genannten Medikationsvarianten.

Tabelle 29: Gesamtkollektiv der DM-PatientInnen aller Altersklassen auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation, Bundesland und Geschlecht

	Insuline oder orale Antidiabetika			nur Insuline			nur orale Antidiabetika			Kombinations-therapie		
	M	w	ges	m	w	ges	m	w	ges	m	w	Ges
BGL	7185	8139	15324	795	762	1557	5787	6541	12328	603	836	1439
KTN	10109	10525	20634	1411	1293	2704	7846	8190	16036	852	1042	1894
NO												
E	37090	37509	74599	4672	4344	9016	28959	29192	58151	3459	3973	7432
OO												
E	26791	25829	52620	3584	2993	6577	20633	20005	40638	2574	2831	5405
SB												
G	8870	8563	17433	1087	901	1988	6874	6648	13522	909	1014	1923
ST												
M	23686	25834	49520	4091	3932	8023	17091	18733	35824	2504	3169	5673
TIR	10054	10499	20553	1753	1526	3279	7093	7579	14672	1208	1394	2602
VB												
G	5666	5665	11331	765	624	1389	4218	4295	8513	683	746	1429
						1013						
W	44995	43978	88973	5254	4880	4	34913	34017	68930	4828	5081	9909
	315	402	717	40	42	82	251	321	572	24	39	63
Σ	17476	17694	35170	2345	2129	4474	13366	13552	26918	1764	2012	3776
	1	3	4	2	7	9	5	1	6	4	5	9

Abbildung (nächste Seite) stellt die Verteilung des Gesamtkollektivs aller Altersklassen auf die Medikationsvarianten graphisch dar. Die Tabellen 30 bis 35 (folgende Seiten) zeigen die Verteilung der jeweiligen Teilkollektive pro Altersklasse auf die Medikationsvarianten; die Abbildungen 32 bis 38 (folgende Seiten) stellen die Verteilung der jeweiligen Teilkollektive pro Altersklasse auf die Medikationsvarianten graphisch dar. Die Abbildungen 39 bis 42 zeigen sodann jeweils für eine Medikationsvariante, wie sich die Anzahl der DM-PatientInnen über die Altersklassen ändert.

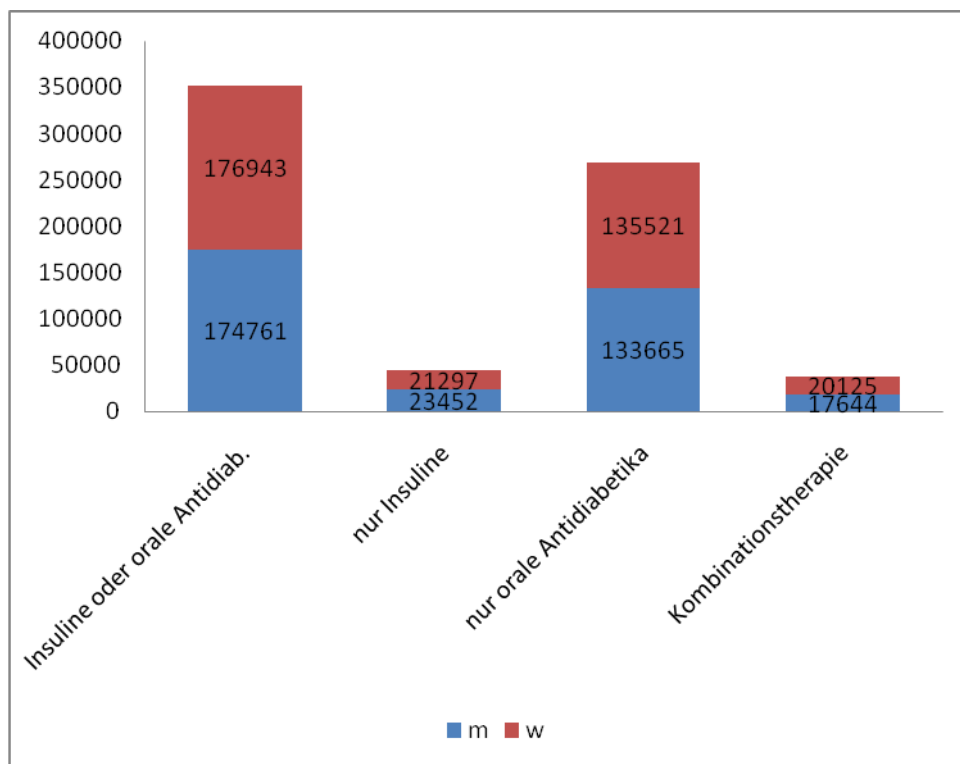


Abbildung 32: Ansicht der DM-PatientInnen aller Altersklassen auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation und Geschlecht

Tabelle 30: Kollektiv der DM-PatientInnen der Altersklasse 0 bis 14 Jahre auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation, Bundesland und Geschlecht

	Insuline oder orale Antidiab.			nur Insuline			nur orale Antidiabetika			Kombinations-therapie		
	m	w	ges	m	w	ges	m	w	ges	m	w	ges
BGL	32	28	60	28	24	52	3	3	6	1	1	2
KTN	63	43	106	61	42	103	2	1	3	0	0	0
NOE	183	162	345	169	149	318	12	11	23	2	2	4
OOE	159	162	321	151	148	299	7	13	20	1	1	2
SBG	42	38	80	41	35	76	1	3	4	0	0	0
STM	109	100	209	102	93	195	7	5	12	0	2	2
TIR	70	71	141	64	64	128	5	5	10	1	2	3
VBG	38	38	76	35	34	69	3	4	7	0	0	0
W	159	164	323	135	131	266	23	31	54	1	2	3
	1	1	2	1	1	2	0	0	0	0	0	0
Σ	856	807	1663	787	721	1508	63	76	139	6	10	16

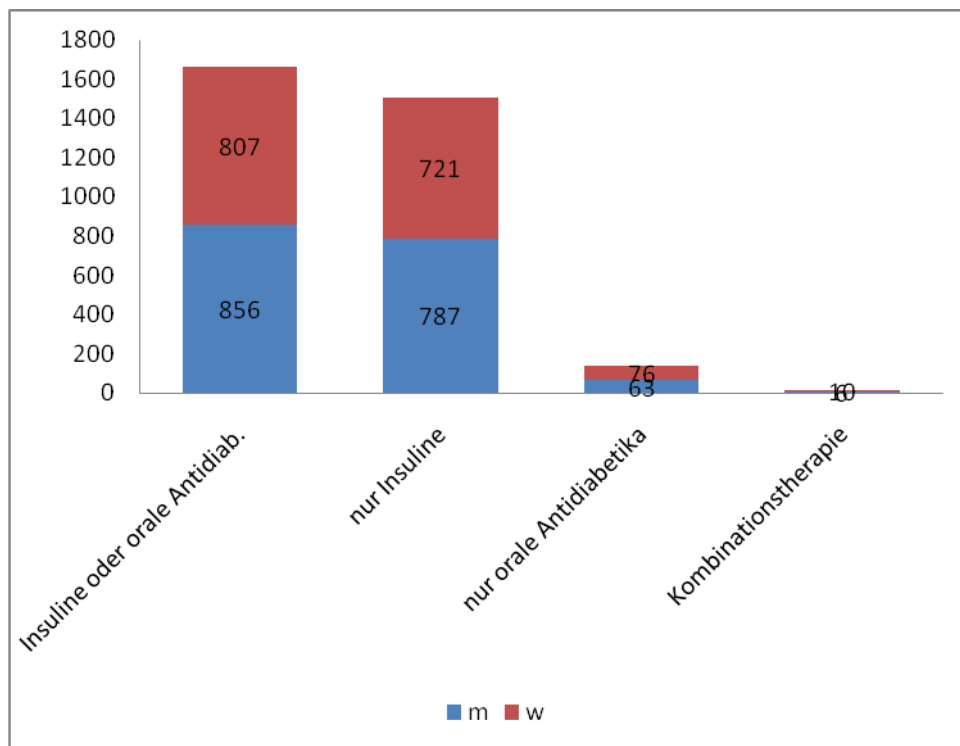


Abbildung 33: Ansicht der DM-PatientInnen der Altersklasse 0 bis 14 Jahre auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation und Geschlecht

Tabelle 31: Kollektiv der DM-PatientInnen der Altersklasse 15 bis 29 Jahre auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation, Bundesland und Geschlecht

	Insuline oder orale Antidiab.			nur Insuline			nur orale Antidiabetika			Kombinations-therapie		
	m	w	ges	m	w	ges	m	w	ges	m	w	ges
BGL	72	93	165	59	44	103	11	47	58	2	2	4
KTN	138	212	350	126	109	235	9	98	107	3	5	8
NOE	431	535	966	369	290	659	54	225	279	8	20	28
OOE	397	544	941	305	289	594	79	230	309	13	25	38
SBG	144	170	314	122	101	223	16	60	76	6	9	15
STM	332	392	724	287	246	533	37	133	170	8	13	21
TIR	196	248	444	163	141	304	29	98	127	4	9	13
VBG	89	156	245	77	79	156	9	71	80	3	6	9
W	501	930	1431	370	514	884	107	384	491	24	32	56
	2	3	5	1	3	4	0	0	0	1	0	1
Σ	2302	3283	5585	1879	1816	3695	351	1346	1697	72	121	193

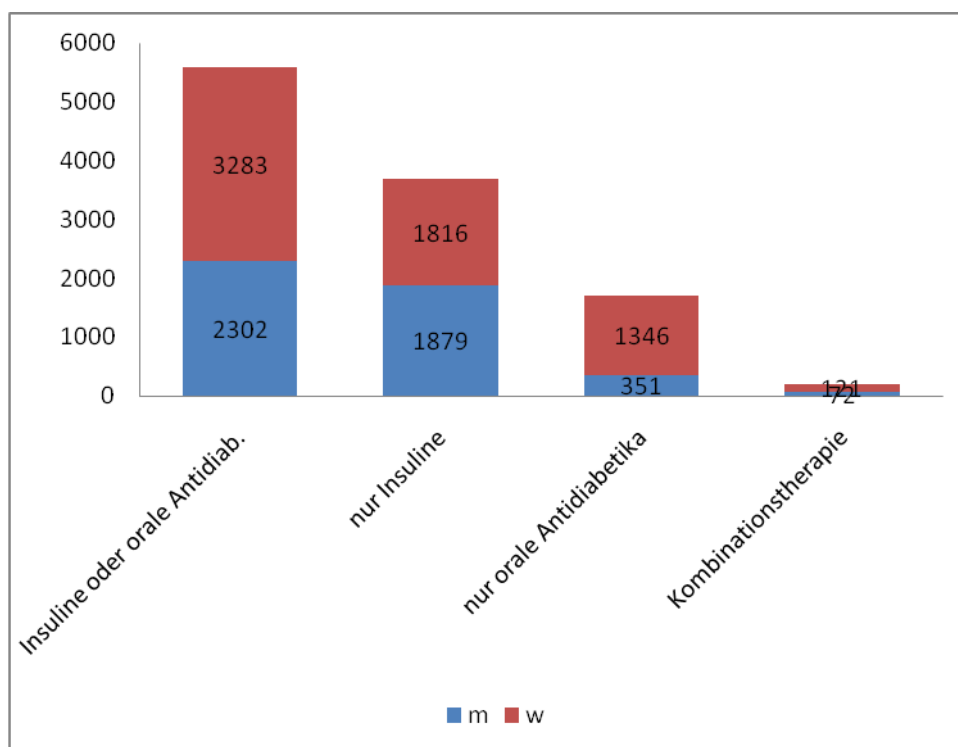


Abbildung 34: Ansicht der DM-PatientInnen der Altersklasse 15 bis 29 Jahre auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation und Geschlecht

Tabelle 32: Kollektiv der DM-PatientInnen der Altersklasse 30 bis 44 Jahre auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation, Bundesland und Geschlecht

	Insuline oder orale Antidiab.			nur Insuline			nur orale Antidiabetika			Kombinations-therapie		
	M	w	ges	m	w	ges	m	w	ges	m	w	ges
BGL	346	265	611	132	73	205	191	171	362	23	21	44
KTN	557	474	1031	243	182	425	276	265	541	38	27	65
NOE	2112	1641	3753	742	585	1327	1233	926	2159	137	130	267
OOE	1740	1336	3076	651	444	1095	957	780	1737	132	112	244
SBG	570	458	1028	203	186	389	324	226	550	43	46	89
STM	1474	1059	2533	680	451	1131	668	524	1192	126	84	210
TIR	735	592	1327	329	245	574	340	300	640	66	47	113
VBG	393	372	765	135	123	258	210	209	419	48	40	88
W	3157	2878	6035	862	971	1833	2032	1654	3686	263	253	516
	13	10	23	3	5	8	10	5	15	0	0	0
Σ	11097	9085	20182	3980	3265	7245	6241	5060	11301	876	760	1636

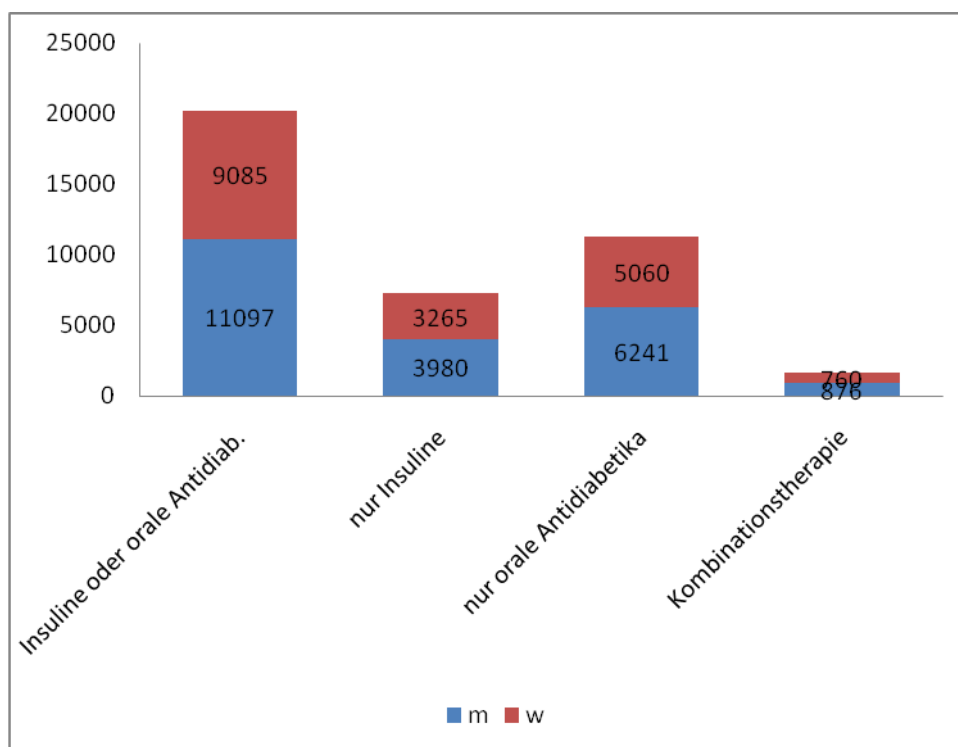


Abbildung 35: Ansicht der DM-PatientInnen der Altersklasse 30 bis 44 Jahre auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation und Geschlecht

Tabelle 33: Kollektiv der DM-PatientInnen der Altersklasse 45 bis 59 Jahre auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation, Bundesland und Geschlecht

	Insuline oder orale Antidiab.			nur Insuline			nur orale Antidiabetika			Kombinations-therapie		
	M	w	ges	m	w	ges	m	w	ges	m	w	ges
BGL	1890	1342	3232	203	130	333	1554	1075	2629	133	137	270
KTN	2815	1710	4525	377	218	595	2197	1318	3515	241	174	415
NOE	9973	6748	16721	1257	674	1931	7773	5342	13115	943	732	1675
OOE	7999	4734	12733	1007	508	1515	6172	3656	9828	820	570	1390
SBG	2615	1765	4380	316	185	501	1998	1336	3334	301	244	545
STM	6159	3850	10009	1041	560	1601	4471	2804	7275	647	486	1133
TIR	2838	1829	4667	433	197	630	2036	1347	3383	369	285	654
VBG	1714	1250	2964	222	98	320	1243	966	2209	249	186	435
W	12618	9132	21750	1318	802	2120	9896	7198	17094	1404	1132	2536
	104	74	178	12	8	20	83	59	142	9	7	16
Σ	48725	32434	81159	6186	3380	9566	37423	25101	62524	5116	3953	9069

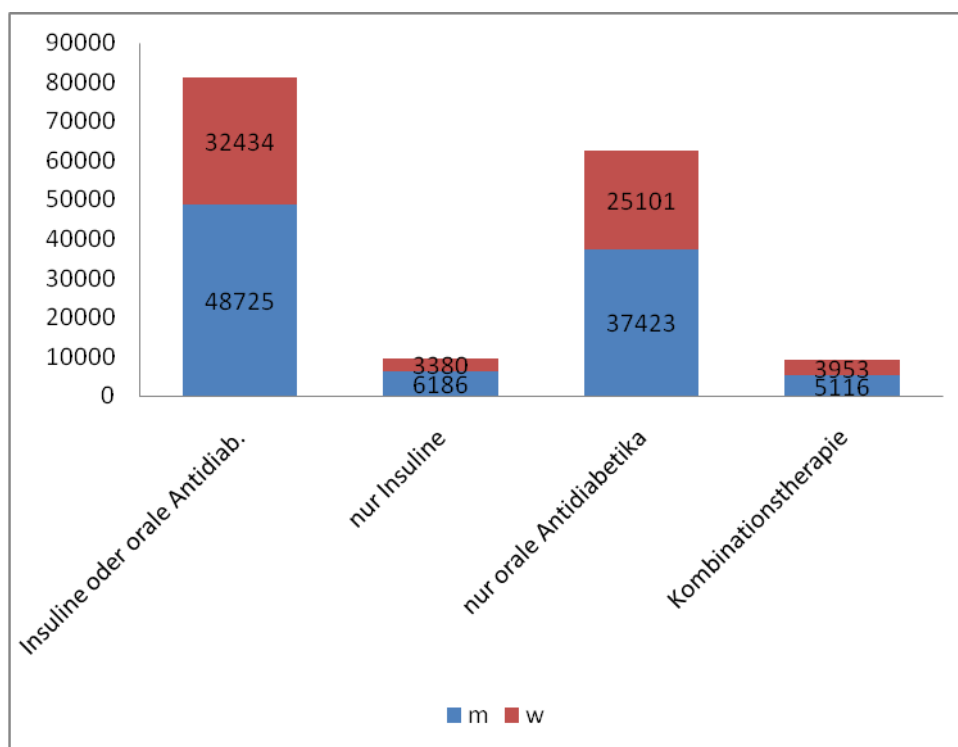


Abbildung 36: Ansicht der DM-PatientInnen der Altersklasse 45 bis 59 Jahre auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation und Geschlecht

Tabelle 34: Kollektiv der DM-PatientInnen der Altersklasse 60 bis 74 Jahre auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation, Bundesland und Geschlecht

	Insuline oder orale Antidiab.			nur Insuline			nur orale Antidiabetika			Kombinations-therapie					
	M	w	ges	m	w	ges	m	w	ges	m	w	ges			
BGL	3214	3226	6440	229	197	426	2689	2661	5350	296	368	664			
KTN	4272	3825	8097	401	351	752	3500	3048	6548	371	426	797			
NOE	1680	1481	31621	1440	1213	2653	1363	1182	25455	1729	1784	3513			
OOE	3	8	1138	976	766	1742	4	1	9271	7796	17067	1139	1202	2341	
SBG	6	9764	21150	271	173	444	9271	7796	17067	1139	1202	2341			
STM	3773	3120	6893	1024	9526	19773	1272	1013	2285	7800	7206	15006	1175	1307	2482
TIR	7	9526	19773	1272	1013	2285	7800	7206	15006	1175	1307	2482			
VBG	4213	3663	7876	485	362	847	3167	2757	5924	561	544	1105			
W	2435	2142	4577	218	142	360	1928	1686	3614	289	314	603			
	1936	1646	35837	1675	1076	2751	1544	1328	28727	2252	2107	4359			
	9	8	337	13	14	27	2	5	274	14	22	36			
Σ	7584	6675	14260	6980	5307	12287	6064	5298	11362	8222	8468	16690			

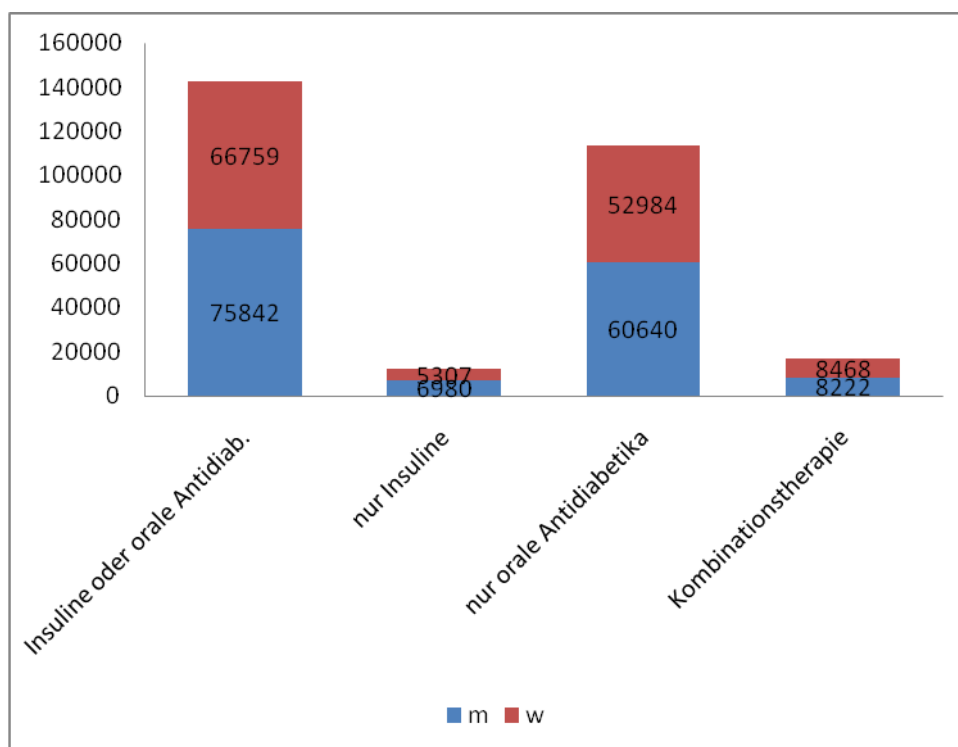


Abbildung 37: Ansicht der DM-PatientInnen der Altersklasse 60 bis 74 Jahre auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation und Geschlecht

Tabelle 35: Kollektiv der DM-PatientInnen der Altersklasse 75 Jahre und älter auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation, Bundesland und Geschlecht

	Insuline oder orale Antidiab.			nur Insuline			nur orale Antidiabetika			Kombinations-therapie		
	M	w	ges	m	w	ges	m	w	ges	m	w	ges
BGL	1631	3185	4816	144	294	438	1339	2584	3923	148	307	455
KTN	2264	4261	6525	203	391	594	1862	3460	5322	199	410	609
NOE	7588	13605	21193	695	1433	2128	6253	10867	17120	640	1305	1945
OOE	5110	9289	14399	494	838	1332	4147	7530	11677	469	921	1390
SBG	1726	3012	4738	134	221	355	1429	2470	3899	163	321	484
STM	5365	10907	16272	709	1569	2278	4108	8061	12169	548	1277	1825
TIR	2002	4096	6098	279	517	796	1516	3072	4588	207	507	714
VBG	997	1707	2704	78	148	226	825	1359	2184	94	200	294
W	9191	14406	23597	894	1386	2280	7413	11465	18878	884	1555	2439
	65	107	172	10	11	21	55	86	141	0	10	10
Σ	35939	64575	100514	3640	6808	10448	28947	50954	79901	3352	6813	10165

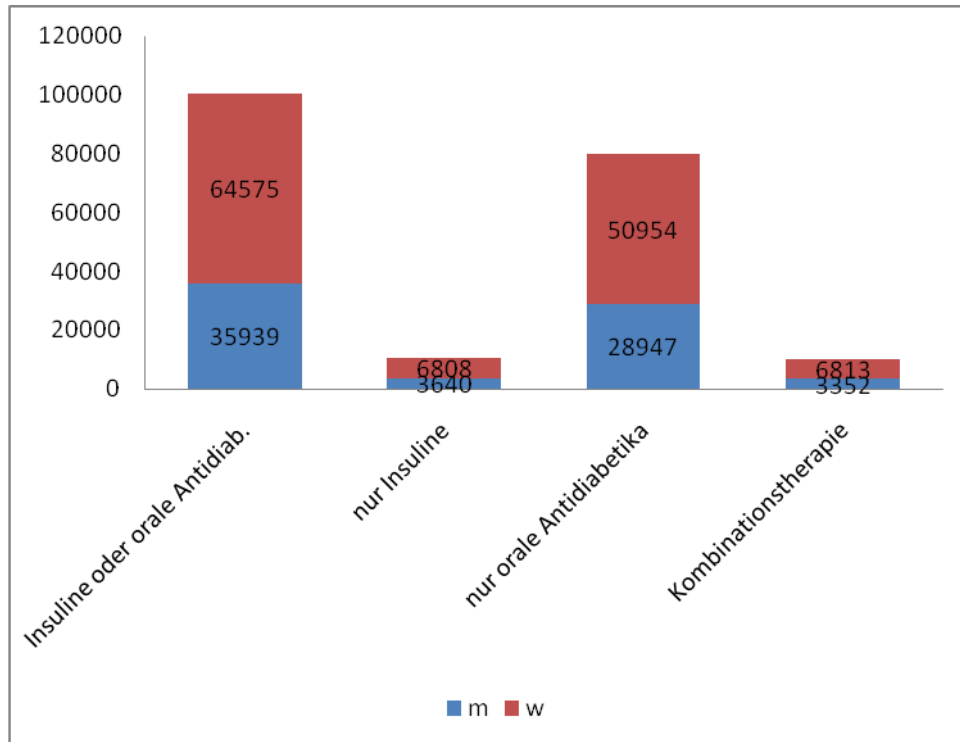


Abbildung 38: Ansicht der DM-PatientInnen der Altersklasse 75 Jahre und älter auf Basis der GAP-DRG Datenbank im Zeitraum 2006 und 2007, gegliedert nach Art der Medikation und Geschlecht

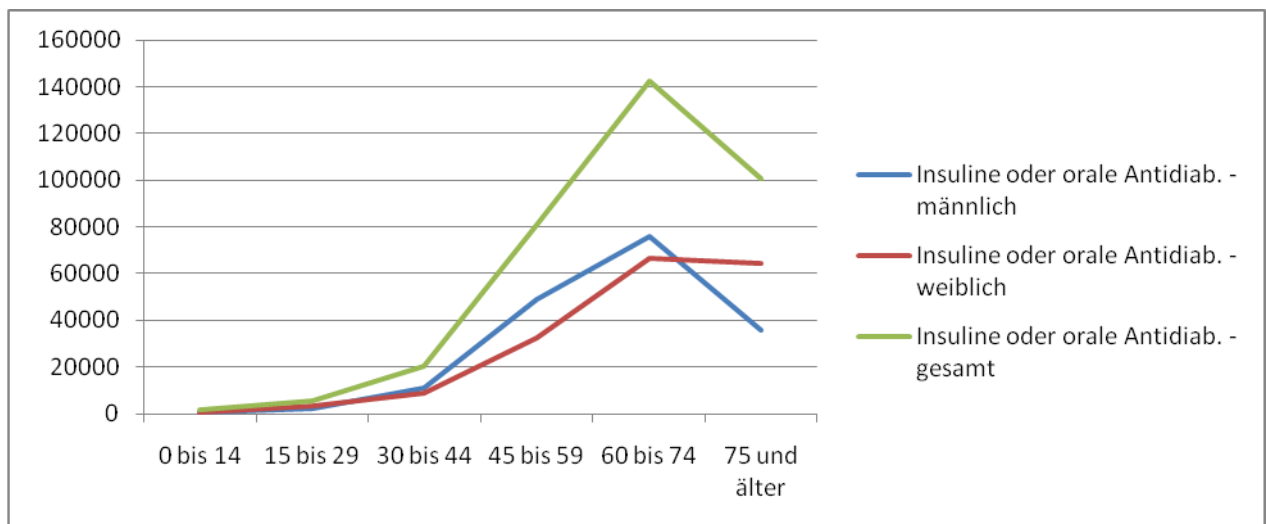


Abbildung 39: Ansicht der DM-PatientInnen, die mit Insulinen oder oralen Antidiabetika therapiert werden, nach Altersklassen und Geschlecht

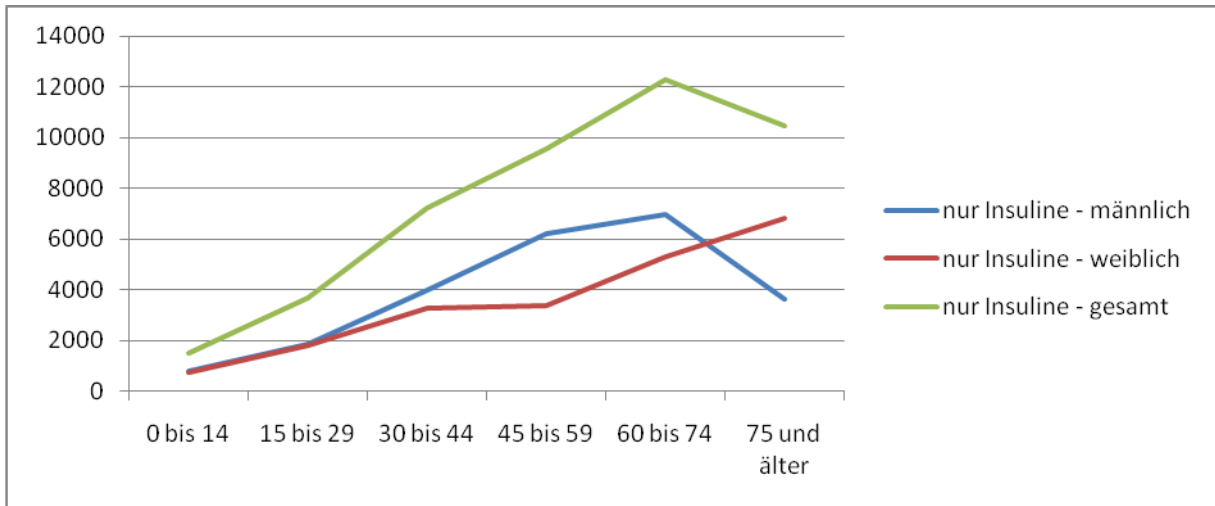


Abbildung 40: Ansicht der DM-PatientInnen, die nur mit Insulinen therapiert werden, nach Altersklassen und Geschlecht

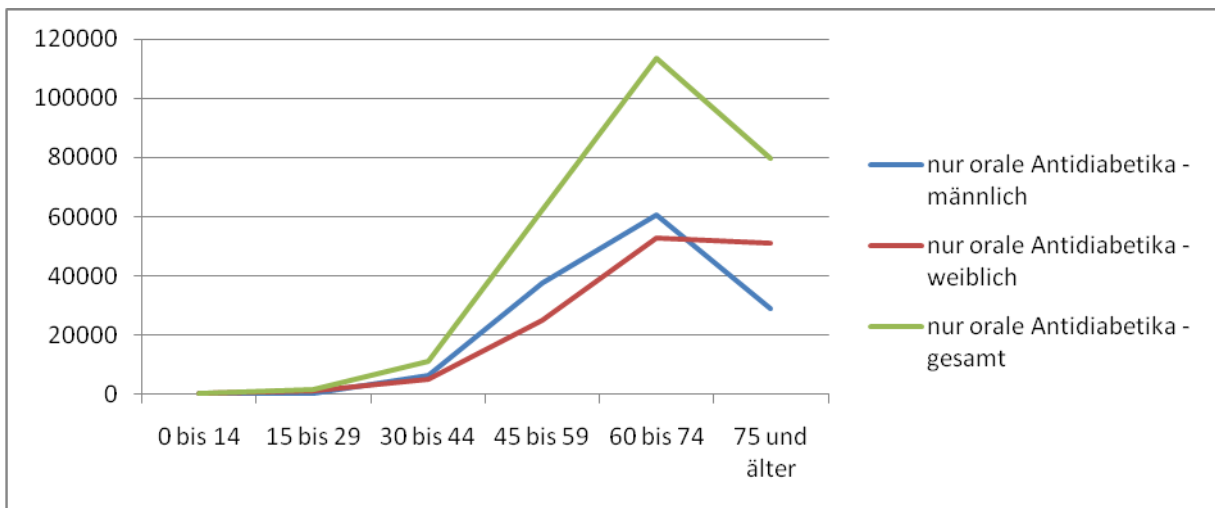


Abbildung 41: Ansicht der DM-PatientInnen, die nur mit oralen Antidiabetika therapiert werden, nach Altersklassen und Geschlecht

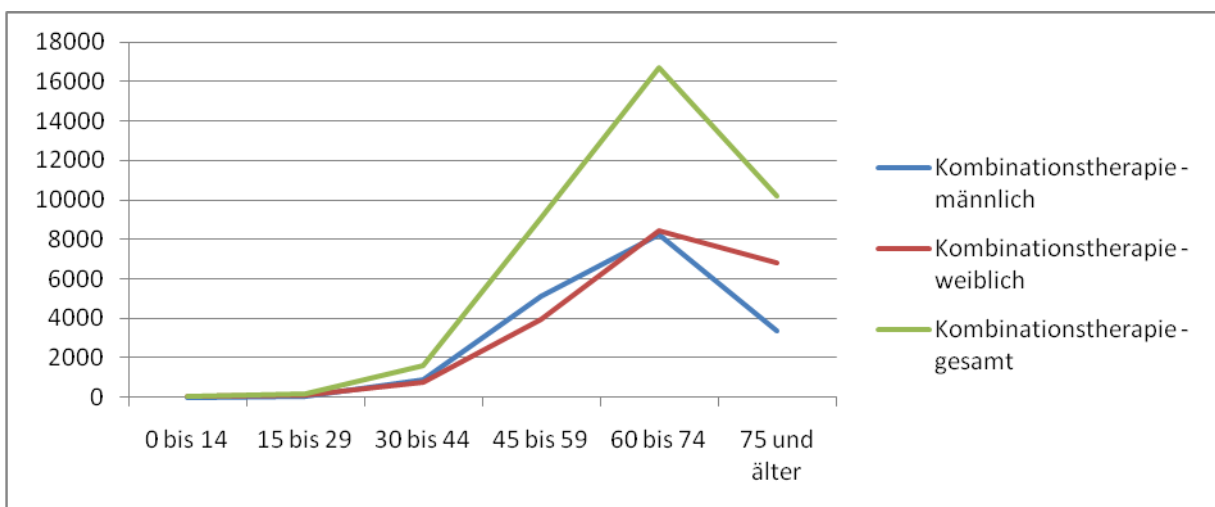


Abbildung 42: Ansicht der DM-PatientInnen, die mit einer Kombinationstherapie therapiert werden, nach Altersklassen und Geschlecht

2.4 Diabetes-Prävalenz in Österreich laut GAP-DRG Datenbank

Im Folgenden wird der Anteil der im Zeitraum 2006 bis 2007 in Österreich therapierten DM-PatientInnen an der österreichischen Wohnbevölkerung dargestellt. Beim DM-Kollektiv wird von dem in Tabelle 26 dargestellten Kollektiv **DM(GAPDRG)** ausgegangen, wobei für die Prävalenz-Berechnung nur Personen aus **DM(GAPDRG)** berücksichtigt werden, die vor dem 01.01.2008 geboren wurden und mindestens bis zum 31.12.2007 gelebt haben. Dies entspricht dem Teilkollektiv der DiabetikerInnen aus dem in Teilprojekt 1 definierten Kollektiv **PHV_3**. Durch diese Fokussierung reduziert sich die Gesamtzahl der betrachteten DM-PatientInnen gegenüber dem Kollektiv **DM(GAPDRG)** von 351.704 auf 328.120 Patienten, was einer Reduzierung um 6,7% entspricht.

Als Referenzpopulation wird die mittlere Wohnbevölkerungspopulation der Quartale 1/2006 bis 4/2007 laut Statistik Austria (SA) herangezogen. Dies ergibt eine Anzahl von 8,280.711 Personen. Die Wahl der Wohnbevölkerungspopulation als Referenzpopulation ist im Falle des Diabetes insofern sinnvoll, als die in der GAP-DRG Datenbank identifizierten DM-Patienten zum überwiegenden Anteil der österreichischen Wohnbevölkerung zuzurechnen sind – nur für 0,2% des Kollektivs ist der (österreichische) Wohnbezirk nicht dokumentiert (siehe Abschnitt 2.1). Dies scheint plausibel, da bei chronischen Erkrankungen wie dem Diabetes im Gegensatz zu spontan auftretenden Erkrankungen primär mit einer Behandlung im Heimatland, und daher auch nur mit einem minimalen Anteil von Anspruchstouristen und Arbeitstouristen an den in der GAP-DRG Datenbank erfassten, in Österreich therapierten DiabetikerInnen zu rechnen ist.

Abbildungen 43 und 44 (nächste und übernächste Seite) zeigen die Prävalenz der DM-PatientInnen aller Altersklassen je Bundesland in geschlechtsübergreifender Darstellung sowie nach Geschlecht getrennt. Die Abbildungen 45 bis 50 (folgende Seiten) zeigen dann die Prävalenz der DM-PatientInnen je Bundesland pro Altersklasse und nach Geschlecht getrennt.

Prevalence of Diabetes Mellitus (HVB/SA)

average = 3.96%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

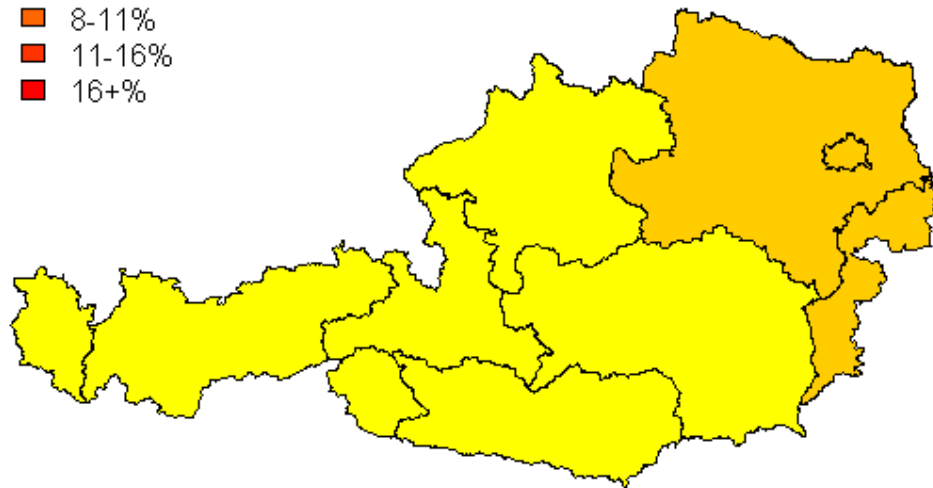
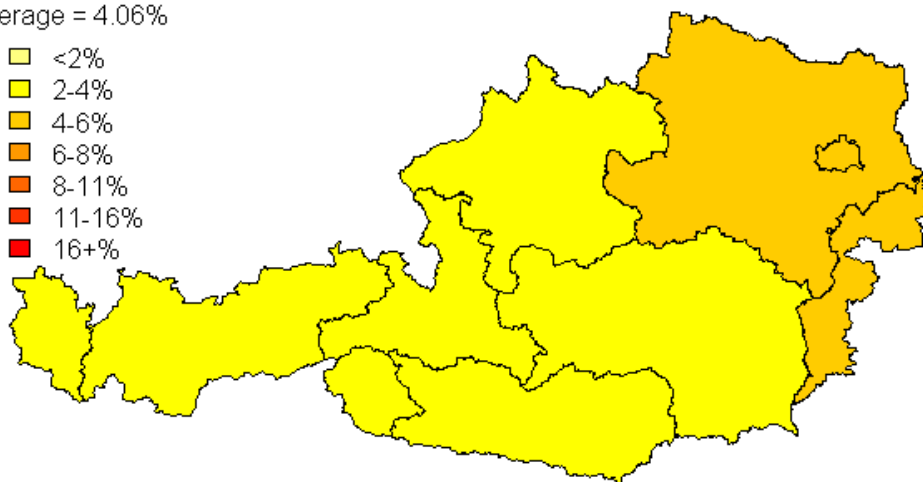


Abbildung 43: Prävalenz der DM-PatientInnen aller Altersklassen je Bundesland auf Basis der GAP-DRG Datenbank im Zeitraum 2006 bis 2007 bezogen auf die österreichische Wohnbevölkerung laut Statistik Austria

Prevalence of Diabetes Mellitus Males (HVB/SA)

average = 4.06%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%



Prevalence of Diabetes Mellitus Females (HVB/SA)

average = 3.87%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

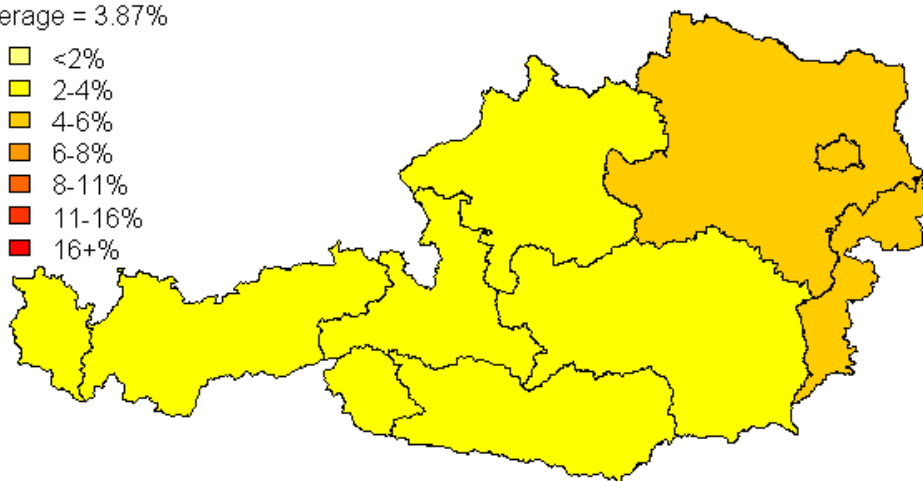
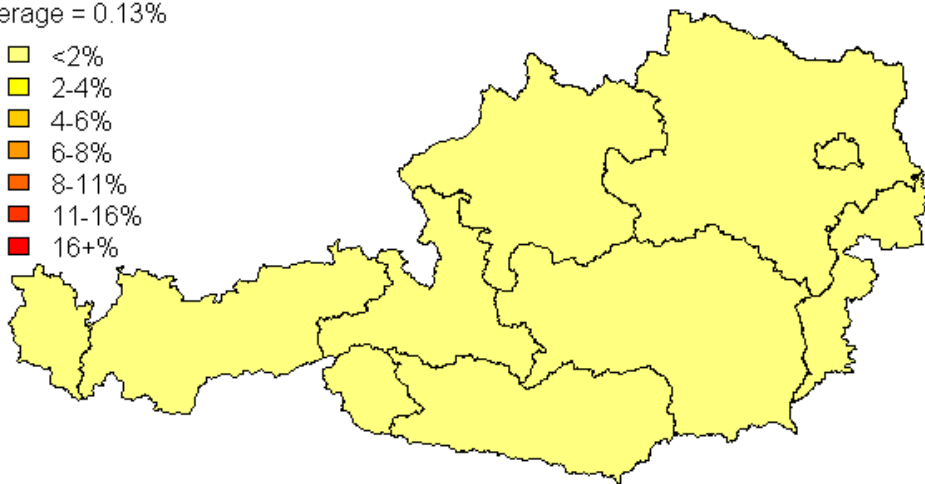


Abbildung 44: Prävalenz der DM-PatientInnen aller Altersklassen je Bundesland auf Basis der GAP-DRG Datenbank im Zeitraum 2006 bis 2007 bezogen auf die österreichische Wohnbevölkerung laut Statistik Austria, getrennt nach Geschlecht

Prevalence of Diabetes Mellitus Males 0-14 years (HVB/SA)

average = 0.13%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

**Prevalence of Diabetes Mellitus Females 0-14 years (HVB/SA)**

average = 0.13%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

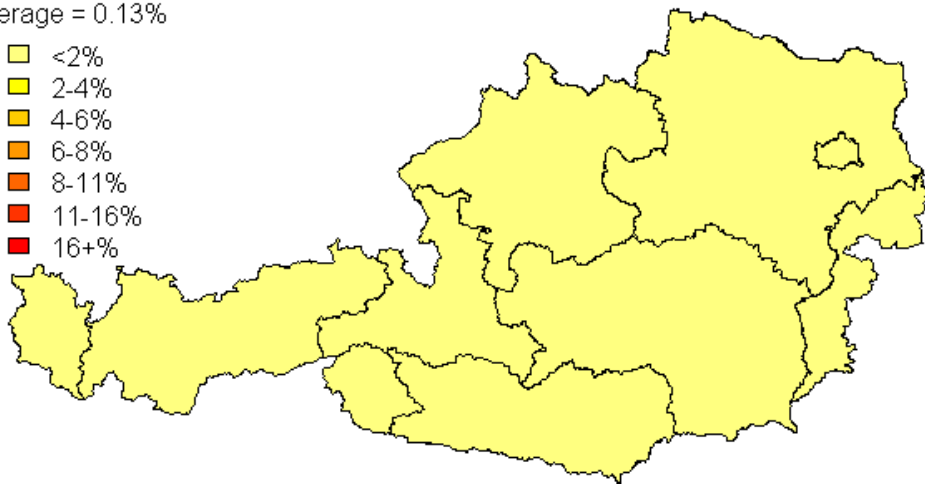
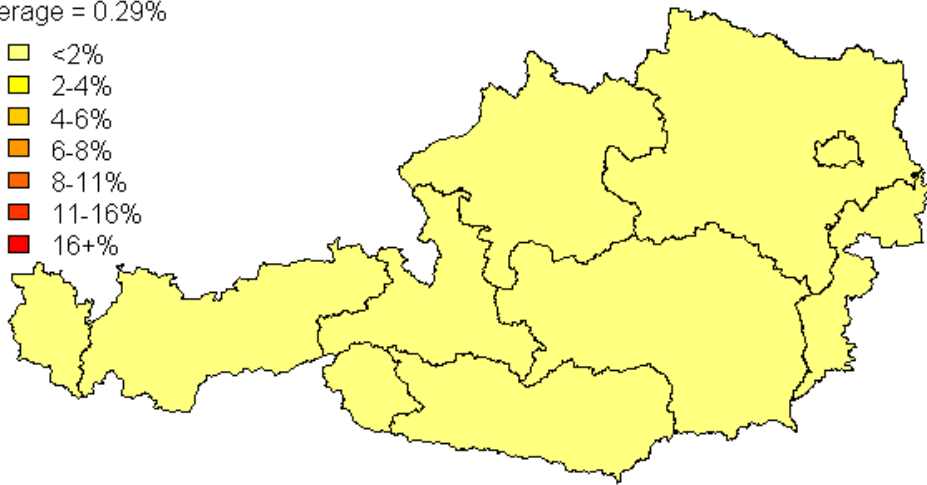


Abbildung 45: Prävalenz der DM-PatientInnen für die Altersklasse 0 bis 14 Jahre je Bundesland auf Basis der GAP-DRG Datenbank im Zeitraum 2006 bis 2007 bezogen auf die österreichische Wohnbevölkerung laut Statistik Austria, getrennt nach Geschlecht

Prevalence of Diabetes Mellitus Males 15-29 years (HVB/SA)

average = 0.29%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

**Prevalence of Diabetes Mellitus Females 15-29 years (HVB/SA)**

average = 0.43%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

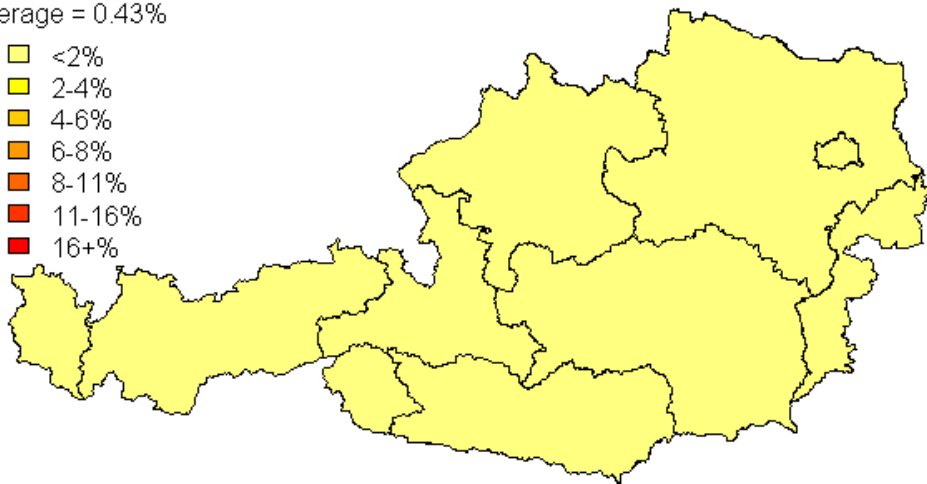
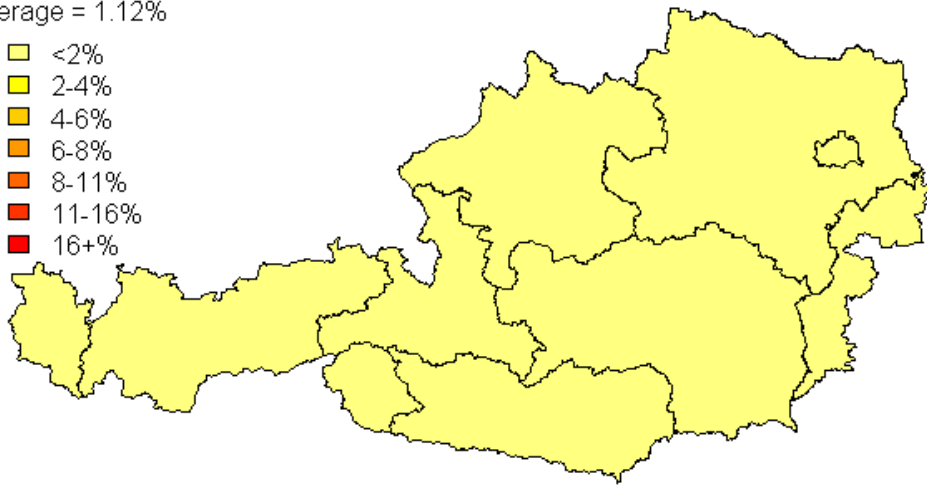


Abbildung 46: Prävalenz der DM-PatientInnen für die Altersklasse 15 bis 29 Jahre je Bundesland auf Basis der GAP-DRG Datenbank im Zeitraum 2006 bis 2007 bezogen auf die österreichische Wohnbevölkerung laut Statistik Austria, getrennt nach Geschlecht

Prevalence of Diabetes Mellitus Males 30-44 years (HVB/SA)

average = 1.12%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%



Prevalence of Diabetes Mellitus Females 30-44 years (HVB/SA)

average = 0.93%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

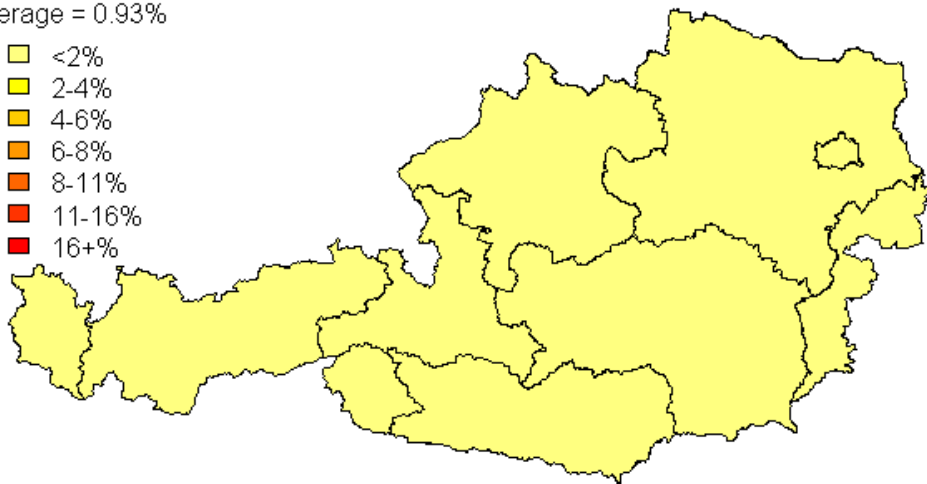
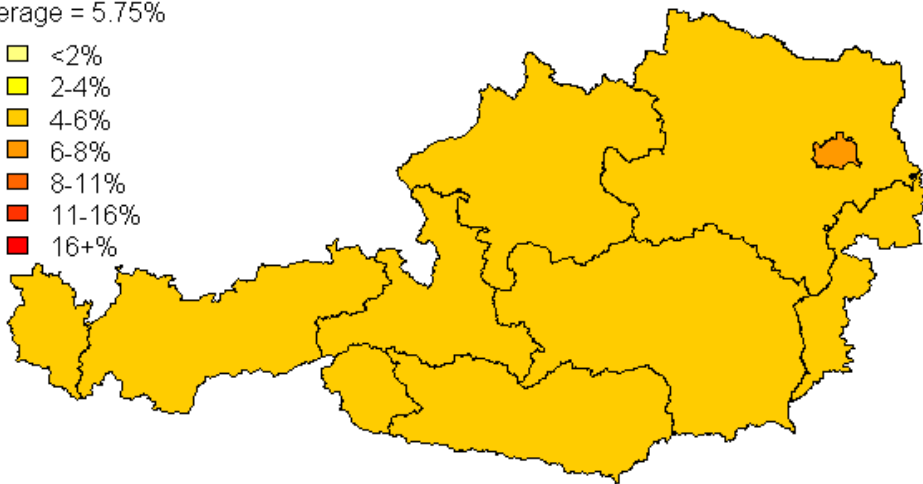


Abbildung 47: Prävalenz der DM-PatientInnen für die Altersklasse 30 bis 44 Jahre je Bundesland auf Basis der GAP-DRG Datenbank im Zeitraum 2006 bis 2007 bezogen auf die österreichische Wohnbevölkerung laut Statistik Austria, getrennt nach Geschlecht

Prevalence of Diabetes Mellitus Males 45-59 years (HVB/SA)

average = 5.75%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%



Prevalence of Diabetes Mellitus Females 45-59 years (HVB/SA)

average = 3.82%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

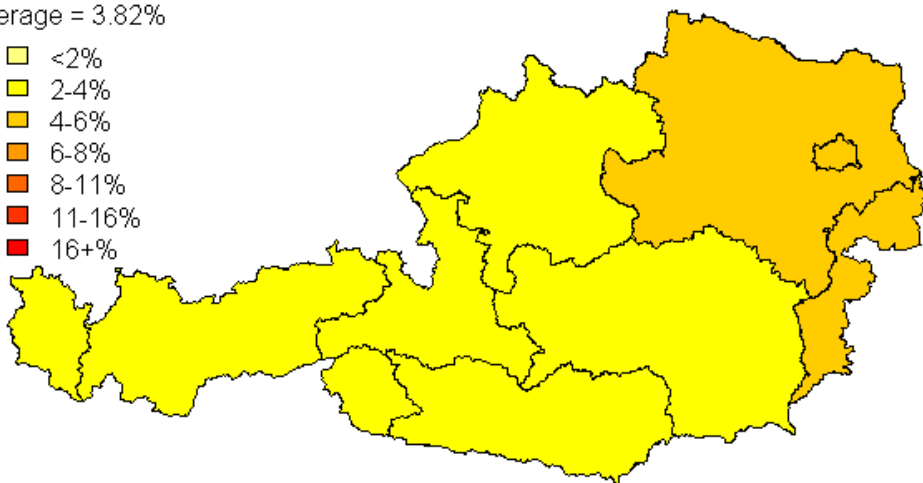
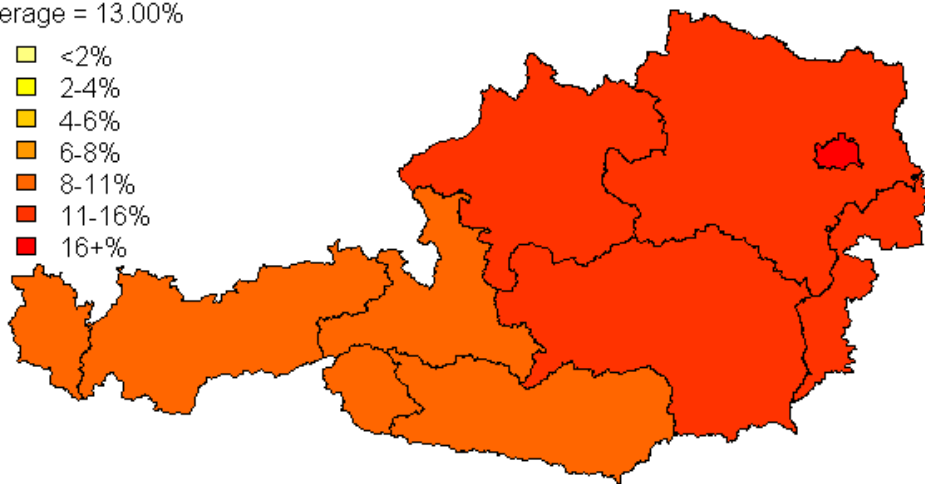


Abbildung 48: Prävalenz der DM-PatientInnen für die Altersklasse 45 bis 59 Jahre je Bundesland auf Basis der GAP-DRG Datenbank im Zeitraum 2006 bis 2007 bezogen auf die österreichische Wohnbevölkerung laut Statistik Austria, getrennt nach Geschlecht

Prevalence of Diabetes Mellitus Males 60-74 years (HVB/SA)

average = 13.00%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%



Prevalence of Diabetes Mellitus Females 60-74 years (HVB/SA)

average = 10.28%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

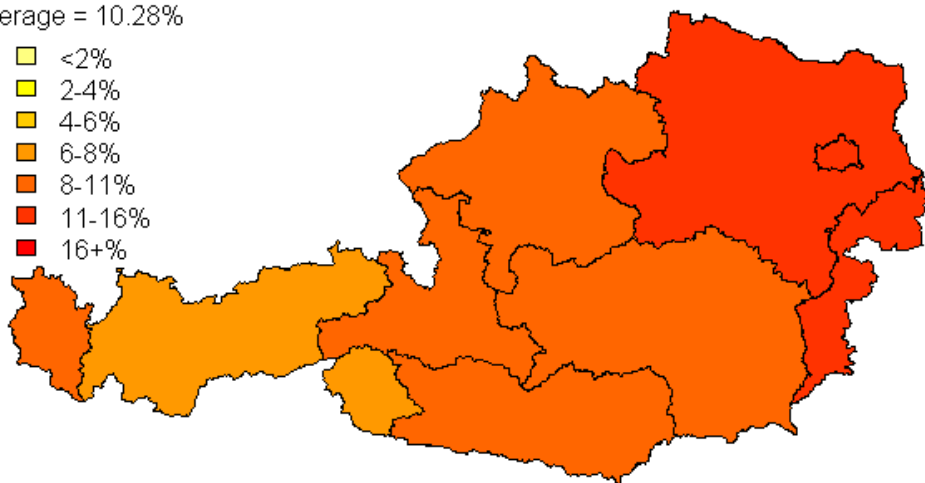
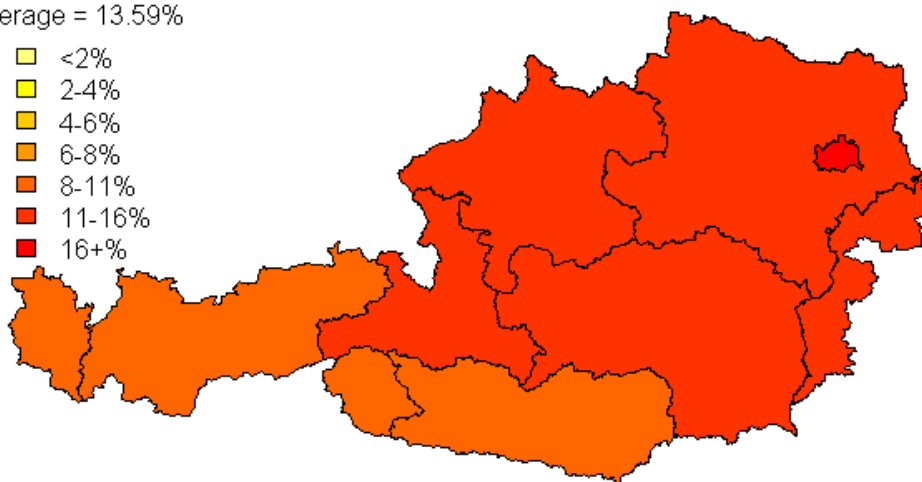


Abbildung 49: Prävalenz der DM-PatientInnen für die Altersklasse 60 bis 74 Jahre je Bundesland auf Basis der GAP-DRG Datenbank im Zeitraum 2006 bis 2007 bezogen auf die österreichische Wohnbevölkerung laut Statistik Austria, getrennt nach Geschlecht

Prevalence of Diabetes Mellitus Males 75+ years (HVB/SA)

average = 13.59%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

**Prevalence of Diabetes Mellitus Females 75+ years (HVB/SA)**

average = 12.92%

- <2%
- 2-4%
- 4-6%
- 6-8%
- 8-11%
- 11-16%
- 16+%

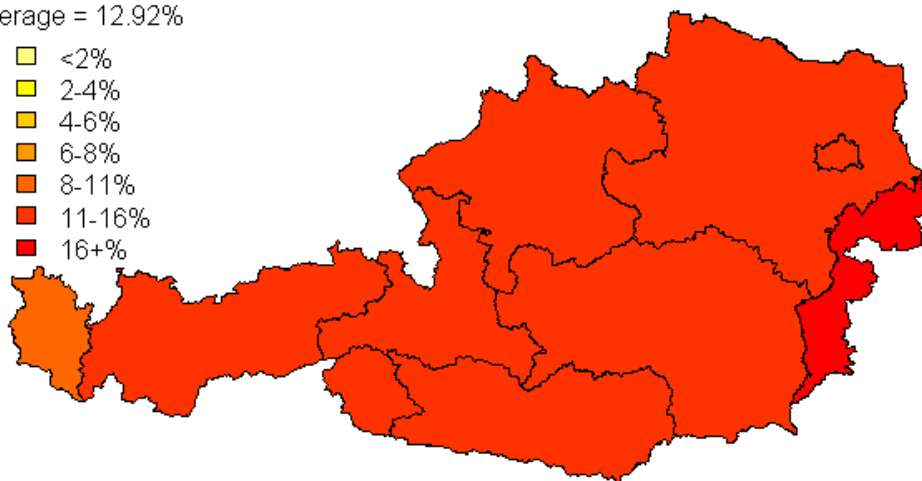


Abbildung 50: Prävalenz der DM-PatientInnen für die Altersklasse 75 Jahre und älter je Bundesland auf Basis der GAP-DRG Datenbank im Zeitraum 2006 bis 2007 bezogen auf die österreichische Wohnbevölkerung laut Statistik Austria, getrennt nach Geschlecht

Teilprojekt 3

Epidemiologie

Chronischer Schmerz

1. Vorbemerkungen

1.1 Ziel

Ziel des dritten Teilprojekts ist die Analyse des Auftretens von chronischem Schmerz in Österreich. Dabei sollen auch unterschiedliche Definitionen dieses Krankheitsbildes bezüglich der Schmerzdauer als Basis herangezogen und die dadurch bedingten epidemiologischen Unterschiede dargestellt werden.

Basierend auf den Leistungsdaten des Hauptverbandes der Österreichischen Sozialversicherungen aus den Jahren 2006 und 2007 wird die Anzahl und Verteilung der Personen mit einer Behandlung mittels Schmerzmedikamenten dargestellt und mit den Ergebnissen der Gesundheitsbefragung der Statistik Austria aus diesem Zeitraum verglichen.

1.2 Definition „chronischer Schmerz“

Schmerz wird als *chronisch* bezeichnet, wenn er die Dauer von 3 bis maximal 6 Monaten überschreitet. Nach dieser Zeitspanne finden auf unterschiedlichen – somatischen und psychosozialen – Ebenen Chronifizierungsvorgänge statt, die eine sekundäre Kausalkette für die weitere Aufrechterhaltung des Schmerzes darstellen. Durch die schrittweise Verselbstständigung des Schmerzgeschehens von seiner ursprünglich auslösenden Ursache verliert der chronische Schmerz seine Warnfunktion. Er ist nicht mehr Hinweis auf eine zugrundeliegende Verletzung oder Erkrankung, sondern eine eigenständige Erkrankung geworden. Die Schmerzintensität ist dabei weder zum organmedizinischen Befund noch zur subjektiven Beeinträchtigung bzw. Behinderung progressiv linear.²

Risikofaktoren für die Entstehung chronischer Schmerzen laut *International Association for the Study of Pain* sind:

- **Geschlecht:** die Inzidenz chronischer Schmerzen ist bei Frauen größer als bei Männern. Weibliche Geschlechtshormone scheinen für die stärker ausgeprägte nozizeptive Antwort bei Frauen eine bedeutende Rolle zu spielen, was auch in Tierexperimenten gezeigt werden konnte. Diese Theorie wird ebenso durch den Umstand bekräftigt, dass der geschlechtsspezifische Unterschied in der Reaktionsstärke auf nozizeptive Reize erst nach der Pubertät auftritt, und nach der Menopause nicht mehr festzustellen ist.
- **Alter:** die Prävalenz chronischer Schmerzen nimmt mit dem Alter zu. Die Ursache dafür liegt einerseits in altersbedingten degenerativen Veränderungen, vor allem des Bewegungsapparates, andererseits kommt es zur Abnahme der Effektivität des endogenen schmerzhemmenden Kontrollsystems.
- **Genetische Prädisposition:** ihr liegt der genetisch determinierte Unterschied in der Schmerzverarbeitung zugrunde (z.B.: Opioidrezeptor-Polymorphismus, Neuropeptid- Metabolismus). Ebenso findet sich eine große Variabilität hinsichtlich der individuellen Pharmakodynamik und Pharmakokinetik von Analgetika.

² Waddell G. Biopsychosocial analysis of low back pain. *Baillières Clin.Rheumatol* 1992; 6: 523-557.

- **Umweltfaktoren:** negative Erfahrungen mit schmerzhaften Erkrankungen und Behandlungen können als Prädiktoren für die Entwicklung eines chronischen Schmerzsyndroms angesehen werden. Frühgeborene Kinder, die wiederholt schmerzhaften Behandlungen ausgesetzt waren, weisen in der späteren Kindheit ein Defizit in der Effektivität des endogenen schmerzhemmenden Kontrollsystems auf.
- **Psychosoziale Faktoren:** Angst, Depression, Selbstwertmangel, übermäßiger Leistungsanspruch, erhöhte Prävalenz psychischer Traumatisierungen in der Vorgeschichte oder lebensverändernde Ereignisse müssen als Prädiktoren für eine Schmerzchronifizierung angesehen werden.

1.3 Variablennamen

In den nachfolgenden Darstellungen werden folgende Namen und Abkürzungen verwendet:

A. Daten des Hauptverbands der Sozialversicherungsträger (HV)

- HV-do: HV-Personen mit Schmerzmedikation
- HV-do-ch3: Personen mit 3 Monaten kontinuierlicher Schmerzmedikation
- HV-do-ch6: Personen mit 6 Monaten kontinuierlicher Schmerzmedikation
- HV-do-ch3-mk: Durchschnittliche Medikamentenkosten der Personen mit 3 Monaten kontinuierlicher Schmerzmedikation in den Jahren 2006-2007
- HV-do-ch6-mk: Durchschnittliche Medikamentenkosten der Personen mit 6 Monaten kontinuierlicher Schmerzmedikation in den Jahren 2006-2007
- HV-do-vm-s1m: Personen mit Schmerzmedikation „im letzten Monat“ (Simulation)

B. Daten der Gesundheitsbefragung (GB)

- GB-p: Gesamtpopulation Gesundheitsbefragung
- GB-do-p: Personen mit Schmerzen
- GB-do-ch-p: Personen mit Schmerzen länger als 3 Monate
- GB-m-d-p: Personen die in den letzten 2 Wochen Schmerzmedikamente eingenommen haben
- GB-vm-do-p: Personen die in den letzten 2 Wochen vom Arzt verordnete Schmerzmedikamente eingenommen haben
- GB-nvm-do-p: Personen die in den letzten 2 Wochen rezeptfreie Schmerzmedikamente eingenommen haben

C. Daten der Statistik Austria (SA)

- SA-ges: Meldebevölkerung als Referenzpopulation

2. Daten Hauptverband (HV)

Als Basis für alle Analysen wurde aus der GAP-DRG Datenbank das Kollektiv der Personen mit rezeptierten Schmerzmedikamenten gebildet.

2.1 ATC-Gruppen

Es wurde eine Liste von Arzneispezialitäten definiert, welche in der Schmerztherapie eingesetzt werden und entsprechend ihrer ATC-Codes in vier Gruppen eingeteilt; vgl. Tabelle 36.

Das „T“ in der Gruppe „N03A-T“ soll darauf aufmerksam machen, dass aus der Gruppe der Antiepileptika („N03A“) nur drei Unter-codes verwendet werden.

Tabelle 36: ATC-Gruppen der Schmerzmedikamente (HV)

Gruppe	ATC	Beschreibung	
M01-M02-M03	M01	Antiphlogistika und Antirheumatika	
	M02	Topische Mittel gegen Gelenk und Muskelschmerzen	
	M03	Muskelrelaxanzien, Peripher wirkende Mittel	
N02	N02	Analgetika	
N03A-T	N03AF01	Carbamazepin	(Antiepileptika)
	N03AX12	Gabapentin	
	N03AX16	Pregabalin	
N06AA	N06AA	Nichtselektive Monoamin-Wiederaufnahmehemmer (Trizykl. Antidepressiva)	

2.2 Patientenkollektiv

Aus der Hauptverbandspopulation (Population **PHV**, siehe Teilprojekt 1) wurden die Personen selektiert, die im Zeitraum 2006-2007 mindestens einmal ein Schmerzmedikament der vier Gruppen bekommen haben. Dabei wurden 3,166.889 Personen identifiziert. Dieses Kollektiv bildet die Basis für die weiteren Analysen.

2.3 Rezepte

Abbildung 51 und Tabelle 37 (jeweils nächste Seite) zeigen die Anzahl der Rezeptpositionen mit Schmerzmittel pro Person. Die Abbildungen 52 und 53 (übernächste Seite) zeigen die Anzahl der Pharmanummern bzw. die Anzahl der Einlösdaten je Person.

Die insgesamt 17 Millionen Rezeptpositionen enthalten 19 Millionen ATC-Codes (eine Position kann mehrere Packungen enthalten), die sich aus 949 unterschiedlichen Pharmanummern zusammensetzen.

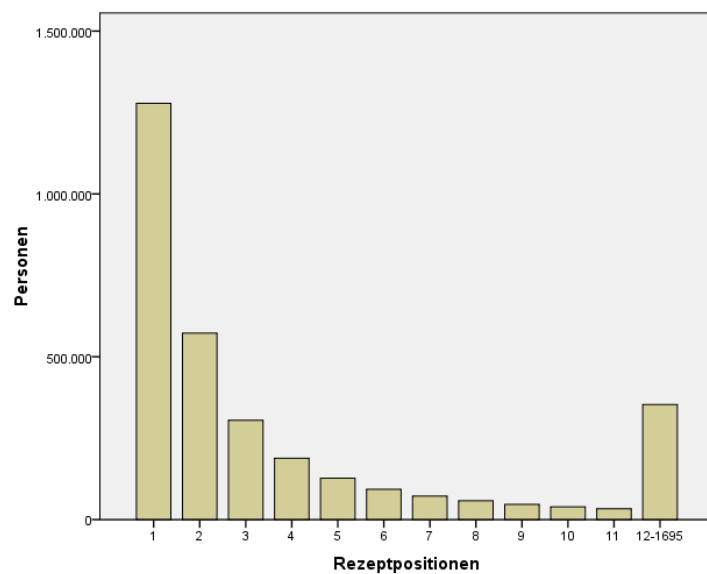


Abbildung 51: Anzahl Rezeptpositionen pro Person (HV)

Tabelle 37: Anzahl Rezeptpositionen pro Person nach Altersklassen und gesamt (HV)

Alter [Jahre]	Rezeptpositionen pro Person						Alle Rezept-Positionen	Personen (SA-ges)
	1	2	3	4	5	Max		
0-14	89.941	25.546	11.956	7.099	4.690	259	367.782	1.297.626
15-29	272.057	91.219	35.433	16.246	8.669	665	951.792	1.542.962
30-44	354.615	152.827	74.957	41.500	25.781	842	2.549.749	1.949.686
45-59	292.753	150.372	85.292	54.516	36.578	1695	4.292.840	1.665.231
60-74	189.014	104.846	65.212	44.873	32.589	522	4.764.423	1.177.221
75+	79.824	47.898	32.136	24.081	18.877	516	4.281.885	647.985
Summe	1.278.204	572.708	304.986	188.315	127.184		17.208.471	8.280.711

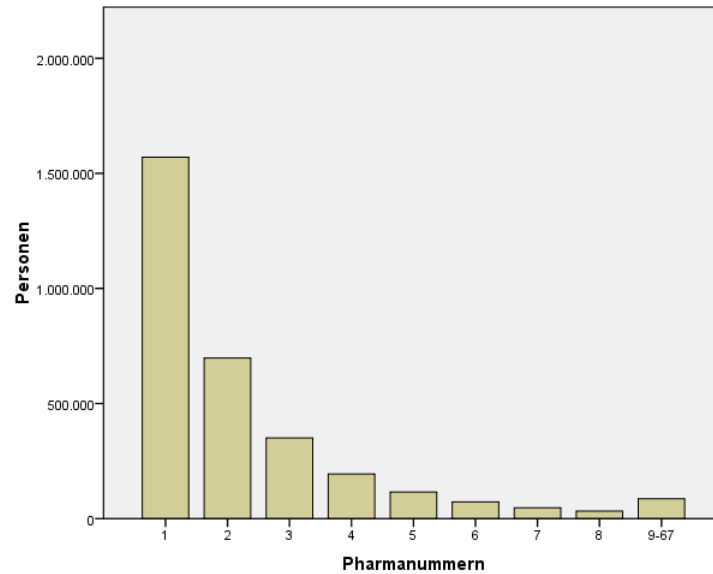


Abbildung 53: Anzahl Pharmanummern pro Person (HV)

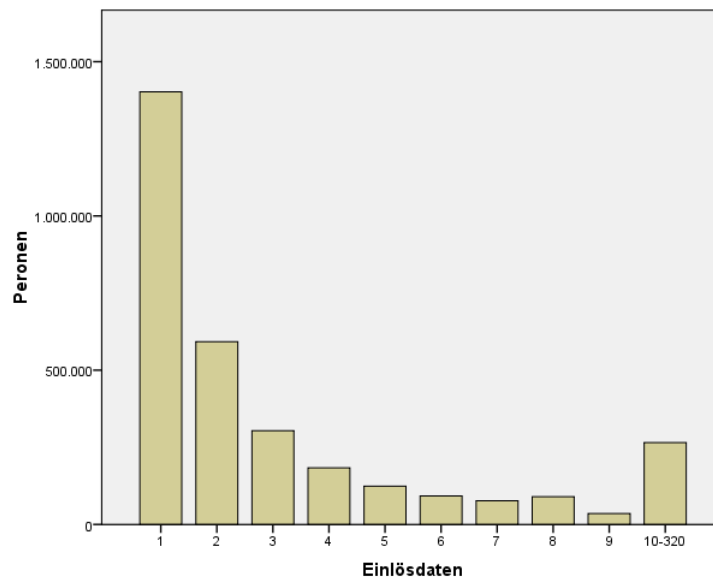


Abbildung 53: Anzahl Einlösdaten pro Person (HV)

2.4 Kontinuität der Medikation

Tabelle 38 (nächste Seite) zeigt die über (Kalender-)Monate kontinuierliche Medikation mit Schmerzmitteln. Zum Beispiel haben in den Jahren 2006–2007 130.392 Personen kontinuierlich Schmerzmittel über mindestens sechs Monate bekommen. Diese 130.392 Personen sind auch in den Häufigkeiten mit mindestens zwei kontinuierlichen Monaten (682.128) und in der Gesamtzahl der Personen, die (mindestens in einem Monat) Schmerzmedikamente erhielten, enthalten. Bei mehrfachen kontinuierlichen Episoden einer Person wird diese Person nur einmal gezählt. Man

© 2011 MU Wien/Univ. Wien ▶ 85

erkennt, dass eine Verkürzung in der Definition „chronischer Schmerz“ von sechs auf drei kontinuierliche Monate die Anzahl der Personen (345.474) mehr als verdoppelt. Tabelle 39 zeigt die erhaltenen Medikamente in diesen kontinuierlichen Monaten nach den vier Medikamentengruppen aufgeschlüsselt

Tabelle 18: Kontinuierliche Medikationsdauer (HV)

Mindest-kont. Dauer [Kalender-Monate]	Personen
1	3.166.889
2	682.128
3	345.474
4	224.981
5	165.416
6	130.392
7	106.049

Tabelle 39: Kontinuierliche Medikationsdauer nach Medikamentengruppen (HV)

Medikamenten Gruppen	Mindest-Dauer [Kalender-Monate]				
	3	4	5	6	7
M01-M02-M03	324835	210209	153699	120534	97692
N02	206961	150392	117089	95727	79928
N03A-T	41934	32005	25606	21211	17895
N06AA	33018	24897	19901	16423	13848

2.5 Verordnete Medikamentengruppen und Kombinationen

Tabelle 40 zeigt die Häufigkeiten der verordneten ATC-Codes nach den vier Medikamentengruppen. Insgesamt wurden 19 Millionen Schmerzmedikamente verordnet.

Tabelle 40: Häufigkeiten der ATC-Codes

Gruppe	n
M01-M02-M03	13.118.381
N02	5.446.259
N03A-T	707.325
N06AA	438.597
	Σ 19.710.562

Die Tabellen 41 und 42 zeigen die Häufigkeiten der Kombinationen von Schmerzmedikamenten für die Personen mit 3 bzw. 6 Monaten Kontinuität. Die drei häufigsten Ränge sind in den beiden Gruppen gleich.

Tabelle 41: Kombination der Schmerzmedikamente mit 3 Monaten Kontinuität (HV)

Medikamentengruppe				Personen (3 Monate)
M01-M02-M03	N02	N03A-T	N06AA	
1	1			151.796
1	1	1		18.836
1	1		1	15.928
1		1		10.462
1			1	7.826
1	1	1	1	5.116
	1	1		2.012
1		1	1	1.160
	1		1	891
	1	1	1	391
		1	1	309
1				113.711
	1			11.991
		1		3.648
			1	1.397
				345.474

Tabelle 42: Kombination der Schmerzmedikamente mit 6 Monaten Kontinuität (HV)

Medikamentengruppe				Personen (6 Monate)
M01-M02-M03	N02	N03A-T	N06AA	
1	1			64.243
1	1	1		10.863
1	1		1	8.823
1	1	1	1	3.612
1		1		3.518
1			1	2.234
	1	1		1.132
1		1	1	513
	1		1	487
	1	1	1	269
		1	1	117
1				26.728
	1			6.298
		1		1.187
			1	368
				130.392

2.6 Diagnosen

Durch die vom Hauptverband durchgeführte ACD-ICD Zuordnung ist es möglich von einer Medikamentenverordnung (pro Person, ATC-Code und Datum) auf eine Diagnose zu schließen.

Tabelle 43 zeigt die zugeordneten Diagnosen (in ICD-9 Gruppen) für die Personen mit chronischem Schmerz (3 Monate). Z.B. waren „Dorsopathien“ 27.021 mal die häufigste Diagnose pro Person und 31.061 mal die zweithäufigste. Die Diagnosegruppen sind nach den *häufigsten Diagnosen pro Person* geordnet. Es gibt Diagnosegruppen (markiert) die eine hohe Häufigkeit haben, deren „Häufigkeit pro Patient“ aber nicht an vorderer Stelle liegt.

Tabelle 43: Häufigste Diagnosen nach ICD-9 Gruppen der Personen mit chronischem Schmerz (HV)

IDC-9 Gruppe	Bezeichnung	Häufigste Diagnosen pro Person					Summe 1. bis 10.
		1.	2.	3.	4.	5.	
710 - 719	ARTHOPATHIEN UND VERWANDTE AFFEKTIONEN	43.517	33.146	24.172	18.957	17.134	180.724
720 - 724	DORSOPATHIEN	27.021	31.061	27.210	22.940	18.966	183.762
401 - 405	HYPERTONIE	22.173	12.661	12.021	10.892	9.809	104.873
730 - 739	OSTEOPATHIEN, CHONDROPATHIEN UND ERWORBERNE DEFORMITÄTEN DES MUSKELSKELETTSYSTEMS	11.991	12.278	12.278	10.690	11.111	96.094
360 - 379	AFFEKTIONEN DES AUGES UND SEINER ANHANGSGEBILDE	11.393	8.578	7.982	6.714	5.733	70.038
295 - 299	ANDERE PSYCHOSEN	9.616	11.037	8.590	7.661	6.833	72.904
410 - 414	ISCHÄMISCHE HERZKRANKHEITEN	8.352	10.096	8.188	7.407	7.140	77.688
530 - 537	KRANKHEITEN DES ÖSOPHAGUS, DES MAGENS UND DES DUODENUMS	6.849	10.161	9.833	10.202	10.152	98.498
290 - 294	ORGANISCHE PSYCHOSEN	6.765	6.166	5.311	4.565	4.504	49.052
460 - 466	AKUTE INFEKTIONEN DER ATMUNGSORGANE	6.543	5.454	4.876	5.395	4.948	48.503
340 - 349	ANDERE KRANKHEITEN DES ZENTRALNERVENSYSTEMS	6.535	5.410	4.251	4.674	4.079	44.943
300 - 316	NEUROSEN, PERSÖNLICHKEITSTÖRUNGEN (PSYCHOPATHIEN) UND ANDERE NICHTPSYCHOTISCHE PSYCHISCHE STÖRUNGEN	6.273	9.000	9.430	7.605	6.284	64.999
350 - 359	AFFEKTIONEN DES PERIPHEREN NERVENSYSTEMS	5.446	4.960	6.655	9.951	11.387	90.704
420 - 429	ANDERE FORMEN VON HERZKRANKHEITEN	4.745	6.705	9.107	7.478	8.160	70.869
250 - 259	KRANKHEITEN SONSTIGER ENDOKRINER DRÜSEN	3.721	2.451	2.087	2.154	2.274	24.808
451 - 459	KRANKHEITEN DER VENEN UND LYMPHGEFÄSSE SOWIE SONSTIGE KRANKHEITEN DES KREISLAUFSYSTEMS	3.016	2.297	2.496	3.072	3.296	35.672
270 - 279	SONSTIGE STOFFWECHSELKRANKHEITEN UND STÖRUNGEN IM IMMUNSYSTEM	2.963	2.012	2.567	2.880	3.028	28.705
160 - 165	MALIGNNE NEOPLASMIEN DER ATMUNGS- UND INTRATHORAKALEN ORGANE	2.728	2.567	1.417	1.400	1.594	16.525
330 - 337	HEREDITÄRE UND DEGENERATIVE KRANKHEITEN DES ZENTRALNERVENSYSTEMS	2.646	3.105	2.254	3.068	3.038	27.722
570 - 579	ANDERE KRANKHEITEN DER VERDAUUNGSORGANE	2.430	4.094	3.028	3.075	3.123	33.815
490 - 496	CHRONISCHE OBSTRUKTIVE LUNGENKRANKHEITEN UND VERWANDTE AFFEKTIONEN	2.396	2.663	2.762	2.921	3.705	38.756
700 - 709	ANDERE KRANKHEITEN DER HAUT UND DER SUBCUTIS	2.145	1.734	3.545	4.741	3.801	38.015
725 - 729	RHEUMATISMUS, EXKL. DES RÜCKENS	1.592	5.625	10.207	13.740	14.290	109.510

3. Daten Gesundheitsbefragung (GB)

3.1 Fragen

Aus der Gesundheitsbefragung wurden aus 2 Gruppen folgende Fragen ausgewählt:

S5 – Schmerzen:

- S5.1: „Hatten Sie während der letzten zwölf Monate **erhebliche** Schmerzen in einer oder mehreren Körperregionen?“
- S5.2: „Bitte zeigen Sie mir, an welcher Stelle des Körpers die Schmerzen aufgetreten sind“ (Ein Bild mit 14 Körperregionen wird vorgelegt).
- S5.7: „Bestehen diese Schmerzen schon **länger als drei Monate?**“

C4 - Medikamente:

- C4.1: „Haben Sie in den letzten beiden Wochen von einem Arzt verordnete Medikamente eingenommen?“
- C4.2 „Ich gehe mit Ihnen die Liste durch, bitte sagen Sie mir jeweils, ob Sie solche Medikamente innerhalb der letzten beiden Wochen eingenommen haben.“

Aus der Liste mit 18 Möglichkeiten (C4.A) wurden 3 für die Analyse ausgewählt:

- e) Gelenksschmerzen (Arthrose, Arthritis)
- f) Kopfschmerzen oder Migräne
- g) Andere Schmerzen

- C4.3: „Haben Sie in den letzten beiden Wochen Medikamente eingenommen, die nicht von einem Arzt verordnet waren (rezeptfreie Medikamente)?“
- C4.4 „Ich gehe mit Ihnen die Liste durch, bitte sagen Sie mir jeweils, ob Sie solche Medikamente innerhalb der letzten beiden Wochen eingenommen haben.“

Aus der Liste mit 8 Möglichkeiten (C4.B) wurden 3 für die Analyse ausgewählt:

- a) Gelenksschmerzen (Arthrose, Arthritis)
- b) Kopfschmerzen oder Migräne
- c) Andere Schmerzen

Eventuell für weitere vergleichende Analysen geeignete Fragen der Gesundheitsbefragung sind:

- S5.3: „Waren Sie wegen dieser Schmerzen in den letzten zwölf Monaten im Krankenstand?“
- S5.4: „Wie viele Tage?“
- S5.5: „Hatten Sie während der letzten sieben Tage auch Schmerzen in dieser Region?“
- S5.6: „Bitte geben Sie die durchschnittliche Stärke Ihrer Schmerzen der letzten sieben Tage mit einer Zahl zwischen 1 und 10 an. 1 bedeutet geringe Schmerzen, 10 bezeichnet den stärksten vorstellbaren Schmerz.“

Aus den Antwortlisten der zwei Fragen nach den Schmerzmedikamenten (C4.A, C4.B) könnte man zusätzlich zu den je drei verwendeten noch weitere berücksichtigen (z.B. Depression, Schlaftabletten, andere Krankheiten).

3.2 Übersicht

Tabelle 44 zeigt die Häufigkeiten der Personen mit erheblichen Schmerzen (GB-do-p) und mit Schmerz länger als 3 Monate (GB-do-ch-p). Der Konsum von Schmerzmedikamenten („in den letzten 14 Tagen“) wird in Summe (GB-m-do-p) und aufgeteilt in die Einnahme verordneten (GB-vm-do-p) und rezeptfreier (GB-do-nvm-p) Medikamente dargestellt.

Tabelle 44: Schmerzempfinden und eingenommene Schmerzmedikamente (GB)

Alter [Jahre]	GB-p	Schmerzempfinden		Schmerzmedikamente		
		GB-do-p	GB-do-ch-p	GB-m-do-p	GB-vm-do-p	GB-nvm-do-p
15-29	1.545.534	351.883	141.255	373.082	152.827	253.105
30-44	1.957.747	649.324	338.206	519.489	291.852	279.738
45-59	1.665.726	765.854	521.350	571.122	421.493	224.768
60-74	1.172.619	578.398	437.993	474.422	416.514	119.741
75+	650.266	352.085	298.680	342.343	316.441	70.839
Summe	6.991.892	2.697.544	1.737.481	2.280.458	1.599.127	948.191

3.3 Körperregionen

In der Gesundheitsbefragung wurde bei erheblichen Schmerzen „in den letzten 12 Monaten“ nach der Körperregion gefragt. Folgende 14 Körperregionen standen zur Verfügung:

1. Kopfschmerzen und Migräne
2. Schmerzen im Gesicht, in den Kaumuskeln, im Kiefergelenk oder im Ohrenbereich
3. Nackenschmerzen (im Bereich der Halswirbelsäule)
4. Schmerzen in den Schultern
5. Schmerzen in den Oberarmen/Ellenbogen/Unterarmen
6. Schmerzen in den Fingern und Händen
7. Schmerzen im Brustkorb
8. Schmerzen im Bauch oder Magen
9. Schmerzen im Rücken (im Bereich der Brustwirbelsäule)
10. Kreuzschmerzen (im Bereich der Lendenwirbelsäule)
11. Schmerzen im Unterleib
12. Schmerzen in den Hüften
13. Schmerzen in den Oberschenkeln/Knieen/Unterschenkeln
14. Schmerzen in den Füßen oder Zehen

In Tabelle 45 werden die Körperregionen in denen Schmerzen „länger als 3 Monate“ bestehen nach ihrer Häufigkeit aufgelistet. Kombiniertes Auftreten von chronischen Schmerzen an mehreren Körperregionen beginnt erst ab der 11. Stelle (Schmerzen in den Lendenwirbeln kombiniert mit Schmerzen im Nackenbereich). Die drei

häufigsten Körperregionen (1. Kreuzschmerzen Lendenbereich, 2. Beine, 3. Kopfschmerzen und Migräne) werden in Tabelle 46 nach Altersgruppen gezeigt.

Tabelle 45: Körperregionen chronischer Schmerzen (GB)

Personen	Rang	Körperregion													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
301.505	1.										1				
146.412	2.													1	
101.475	3.	1													
81.968	4.			1											
64.490	5.									1					
64.257	6.				1										
58.899	7.														1
51.771	8.												1		
42.906	9.						1								
32.178	10.								1						
29.615	K 1.			1							1				
27.521	K 2.										1			1	
26.999	11.					1									
21.592	K 3.									1	1				
19.498	12.							1							
18.312	K 4.				1						1				
...															
1.737.481															

Tabelle 46: Drei häufigsten Körperregionen mit Schmerzen mehr als 3 Monate (GB)

Alter [Jahre]	GB-p	GB-do-ch-p	Schmerzort		
			Lendenwirbelsäule	Beine	Kopf
15-29	1.545.534	141.255	41.557	14.276	56.442
30-44	1.957.747	338.206	141.767	52.226	63.356
45-59	1.665.726	521.350	226.858	118.105	82.975
60-74	1.172.619	437.993	173.641	141.363	41.783
75+	650.266	298.680	130.944	111.547	19.744
Summe	6.991.892	1.737.481	714.767	437.517	264.300

3.4 Kombinationen Schmerzmedikamente

Die Tabellen 47 und 48 enthalten die Personen die „in den letzten beiden Wochen“ verordnete bzw. rezeptfreie Medikamente der 3 ausgewählten Schmerzmedikamente eingenommen haben.

Tabelle 47: Kombinationen verordneter Schmerzmedikamente (GB)

Personen	Verordnetes Schmerzmedikament		
	Gelenks- schmerzen (e)	Kopfschmerzen, Migräne (f)	Andere Schmerzen (g)
520.508			1
365.701	1		
208.677	1		1
205.506		1	
118.848	1	1	1
105.607		1	1
74.280	1	1	
1.599.127	Summe		
1.751.070	Personen mit anderen Medikamenten als e) f) und g)		
3.641.695	Personen ohne verordnete Medikamente		
6.991.892	Alle Personen		

Tabelle 48: Kombinationen rezeptfreier Schmerzmedikamente (GB)

Personen	Verordnetes Schmerzmedikament		
	Gelenks- schmerzen (a)	Kopfschmerzen, Migräne (b)	Andere Schmerzen (c)
522.383		1	
195.135			1
74.066		1	1
71.766	1		
31.150	1		1
30.641	1	1	
23.048	1	1	1
948.189	Summe		
720.243	Personen mit anderen rezeptfreien Medikamenten als a) b) und c)		
5.323.459	Personen ohne rezeptfreie Medikamente		
6.991.892	Alle Personen		

4. Vergleich Hauptverband – Gesundheitsbefragung (HV–GB)

Für den Vergleich der Daten des Hauptverbandes mit der Gesundheitsbefragung wurde das Kollektiv des Hauptverbandes auf Personen größer als 14 Jahre eingeschränkt, da in der Gesundheitsbefragung diese Personengruppe bis 14 Jahren nicht befragt wurde.

4.1 Personen mit Schmerzen

In Tabelle 49 (gesamt bzw. nach Geschlecht) werden die Personen die Schmerzmedikamente erhielten und die Personen mit chronischen Schmerzen (3-Monate) gezeigt. Die Anzahlen der chronischen Schmerzpatienten werden auch als Anteile ausgewiesen wobei als Referenzpopulation die Meldebevölkerung laut Statistik Austria verwendet wird.

Tabelle 49 (a): Vergleich HV-GB – gesamt

Bundesland	SA-ges	GB-do	HV-do	GB-do-ch	HV-do-ch3	GB-do-ch[%]	HV-do-ch3 [%]
Burgenland	240.582	100.940	98.411	66.109	4.793	27,48	1,99
Kärnten	474.770	175.901	200.914	118.750	32.773	25,01	6,90
Niederösterr.	1.335.701	444.542	562.781	269.770	83.132	20,20	6,22
Oberösterr.	1.169.717	477.026	499.462	308.685	67.755	26,39	5,79
Salzburg	439.054	157.029	192.633	102.054	24.602	23,24	5,60
Steiermark	1.025.281	420.739	445.037	283.158	19.465	27,62	1,90
Tirol	580.073	262.147	248.764	163.054	7.119	28,11	1,23
Vorarlberg	298.341	96.684	117.258	53.828	1.818	18,04	0,61
Wien	1.419.566	562.536	641.699	372.073	97.577	26,21	6,87
Fehlt	-		6.332		390		
Österreich	6.983.085	2.697.544	3.013.291	1.737.481	339.424	24,88	4,86

Tabelle 49 (b): Vergleich HV-GB – Frauen

Bundesland	SA-ges	GB-do	HV-do	GB-do-ch	HV-do-ch3	GB-do-ch-[%]	HV-do-ch3 [%]
Burgenland	123.674	57.870	57.706	38.016	3.076	30,74	2,49
Kärnten	247.081	95.641	117.017	66.898	22.116	27,08	8,95
Niederösterr.	687.294	234.741	322.084	157.471	56.251	22,91	8,18
Oberösterr.	601.207	266.660	274.199	188.116	44.317	31,29	7,37
Salzburg	228.646	85.383	108.542	58.896	16.131	25,76	7,06
Steiermark	529.439	238.038	255.526	170.348	12.755	32,18	2,41
Tirol	299.411	136.405	136.650	88.970	4.401	29,72	1,47
Vorarlberg	152.406	49.913	64.870	27.764	1.062	18,22	0,70
Wien	749.740	340.394	379.181	231.902	66.279	30,93	8,84
Fehlt	-		3.682		263		
Österreich	3.618.898	1.505.045	1.719.457	1.028.381	226.651	28,42	6,26

Tabelle 49 (c): Vergleich HV-GB – Männer

Bundesland	SA-ges	GB-do	HV-do	GB-do-ch	HV-do-ch3	GB-do-ch- [%]	HV-do-ch3 [%]
Burgenland	116.908	43.070	40.705	28.094	1.717	24,03	1,47
Kärnten	227.689	80.260	83.897	51.851	10.657	22,77	4,68
Niederösterreich	648.407	209.800	240.697	112.301	26.881	17,32	4,15
Oberösterreich	568.510	210.366	225.263	120.569	23.438	21,21	4,12
Salzburg	210.408	71.645	84.091	43.160	8.471	20,51	4,03
Steiermark	495.842	182.701	189.511	112.811	6.710	22,75	1,35
Tirol	280.662	125.741	112.114	74.082	2.718	26,40	0,97
Vorarlberg	145.935	46.771	52.388	26.065	756	17,86	0,52
Wien	669.826	222.141	262.518	140.170	31.298	20,93	4,67
Fehlt			2.650		127		
Österreich	3.364.187	1.192.495	1.293.834	709.103	112.773	21,08	3,35

Von den 6.000 in den HV-Daten enthaltenen Personen mit fehlendem Bundesland entfallen auf die chronischen Schmerzpatienten 390 Personen.

4.2 Aktuelle Schmerzmedikationen (Simulation)

In Tabelle 50 (gesamt bzw. nach Geschlecht ausgewiesen) werden die Anzahlen der HV-Personen, die „im letzten Monat“ Schmerzmedikamente erhielten (HV-do-vm-s1m) und die GB-Personen, die „in den letzten 14 Tagen“ verordnete Schmerzmedikamente erhielten (GB-vm-do-p) gezeigt.

Die Werte des Hauptverbandes wurden durch eine Simulation ermittelt, bei der für alle Personen mit Schmerzmedikamenten über den Zeitraum 1. Quartal 2006 bis 1. Quartal 2007 (Zeitraum der Gesundheitsbefragung) ein Zeitpunkt randomisiert ermittelt wurde.

Tabelle 50 (a): Vergleich der Verordnungen (Simulation) - gesamt

Bundesland	HV-do-ch3	HV-do-vm-s1m	GB-vm-do-p	GB-do-ch
Burgenland	4.793	13.016	56.021	66.109
Kärnten	32.773	32.706	120.837	118.750
Niederösterreich	83.132	85.826	269.000	269.770
Oberösterreich	67.755	72.966	266.392	308.685
Salzburg	24.602	27.080	96.747	102.054
Steiermark	19.465	57.063	299.197	283.158
Tirol	7.119	27.522	146.710	163.054
Vorarlberg	1.818	12.250	45.042	53.828
Wien	97.577	101.128	299.181	372.073
Fehlt	390	696		
Summe	339.424	430.253	1.599.127	1.737.481

Tabelle 50 (b): Vergleich der Verordnungen (Simulation) - Frauen

Bundesland	HV-do-ch3	HV-do-vm-s1m	GB-vm-do-p	GB-do-ch
Burgenland	3.076	8.362	35.845	38.015
Kärnten	22.116	21.222	79.951	66.899
Niederösterreich	56.251	55.209	165.184	157.469
Oberösterreich	44.317	44.904	168.592	188.116
Salzburg	16.131	16.835	59.354	58.895
Steiermark	12.755	36.325	189.582	170.347
Tirol	4.401	16.660	83.924	88.972
Vorarlberg	1.062	7.348	27.147	27.763
Wien	66.279	66.031	197.525	231.903
Null	263	462		
Summe	226.651	273.358	1.007.104	1.028.381

Tabelle 50 (c): Vergleich der Verordnungen (Simulation) - Männer

Bundesland	HV-do-ch3	HV-do-vm-s1m	GB-vm-do-p	GB-do-ch
Burgenland	1.717	4.654	20.176	43.070
Kärnten	10.657	11.484	40.886	80.260
Niederösterreich	26.881	30.617	103.816	209.800
Oberösterreich	23.438	28.062	97.800	210.366
Salzburg	8.471	10.245	37.393	71.645
Steiermark	6.710	20.738	109.614	182.701
Tirol	2.718	10.862	62.786	125.741
Vorarlberg	756	4.902	17.895	46.771
Wien	31.298	35.097	101.656	222.141
Null	127	234		
Summe	112.773	156.895	592.022	1.192.495

5. Medikamentenkosten (HV)

Tabelle 51 zeigt die durchschnittlichen Kosten der Schmerzmedikamente der HV-Personen mit drei und sechs Monaten chronischen Schmerzen die in den Jahren 2006-2007 verordnet wurden. Die Gesamtkosten der Schmerzmedikamente belaufen sich für die 345.000 Personen mit mindestens 3 Monaten chronischem Schmerz auf 120 Millionen €. Davon sind für die 130.000 Personen mit mindestens sechs Monaten chronischem Schmerz 87 Millionen € enthalten.

Tabelle 51: Durchschnittliche Kosten der Schmerzmedikamente (HV)

Alter [Jahre]	HV-do-ch3	HV-do-ch3-mk	HV-do-ch6	HV-do-ch6-mk
0-14	6.050	40	637	134
15-29	9.659	233	2.039	649
30-44	37.430	340	10.463	828
45-59	81.324	348	27.479	719
60-74	106.176	331	40.183	619
75+	104.835	393	49.591	653
Summe	345.474		130.392	

6. Prävalenzen – Grafiken (HV-GB)

Nachfolgend werden die Prävalenzen der „HV-Personen“ mit (mindestens) drei Monaten chronischem Schmerz dargestellt. Die Werte beziehen sich immer auf die Referenzpopulation der Meldebevölkerung laut Statistik Austria. Die Prävalenzen für ganz Österreich sind in den Abbildungen 54 bis 65 (folgende Seiten) immer mit „average“ gekennzeichnet.

Zuerst werden in Abschnitt 6.1 die Prävalenzen auf Ebene der Bundesländer für alle Altersklassen gesamt dargestellt. Als Vergleich werden auch die Prävalenzen der Personen mit chronischem Schmerz der Gesundheitsbefragung gezeigt.

In Abschnitt 6.2 werden die Prävalenzen nach Altersklassen und Geschlecht auf Ebene der Bundesländer dargestellt.

In Abschnitt 6.3 werden die Prävalenzen nach Altersklassen und Geschlecht auf Ebene der Bezirke für Personen ab 30 Jahre dargestellt.

6.1 Chronische Schmerzpatienten aller Altersklassen (HV+GB)

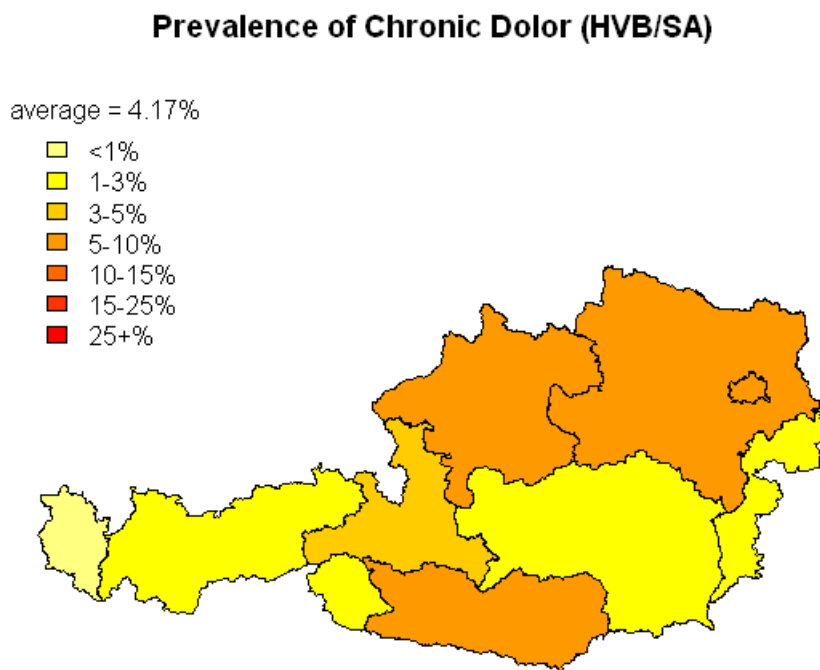
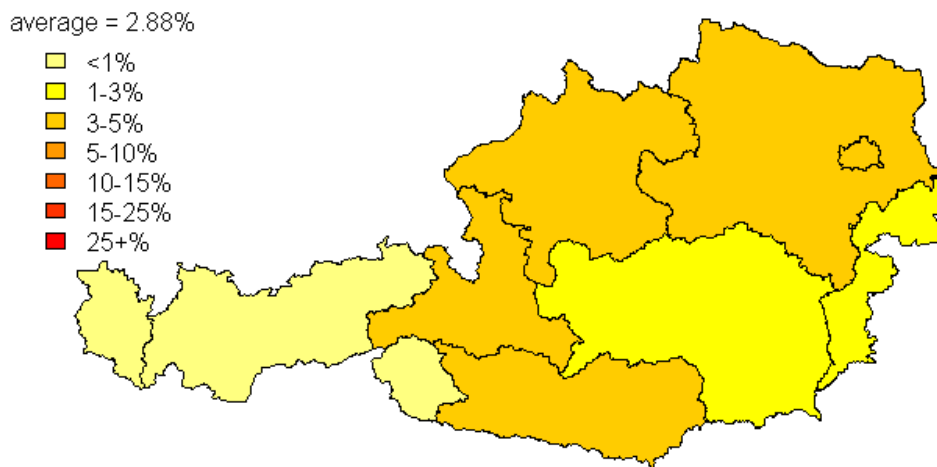


Abbildung 54 (a): Prävalenz der HV-Personen mit chronischem Schmerz – gesamt

Prevalence of Chronic Dolor Males (HVB/SA)



Prevalence of Chronic Dolor Females (HVB/SA)

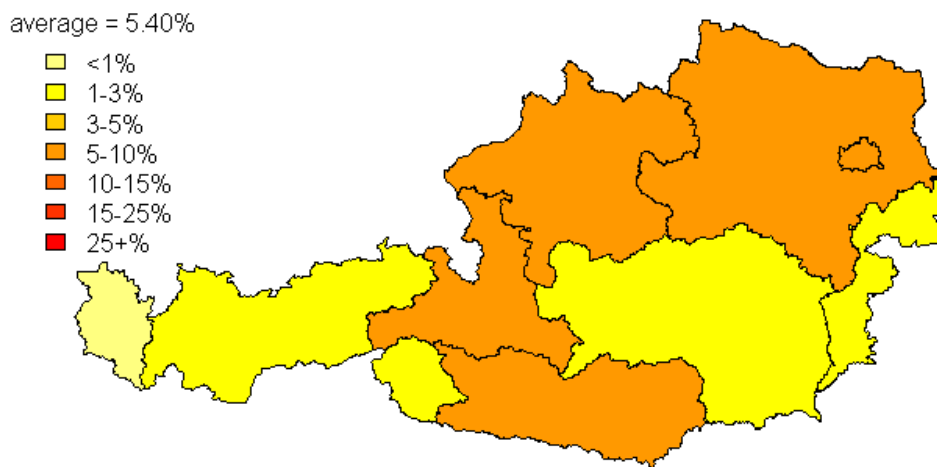
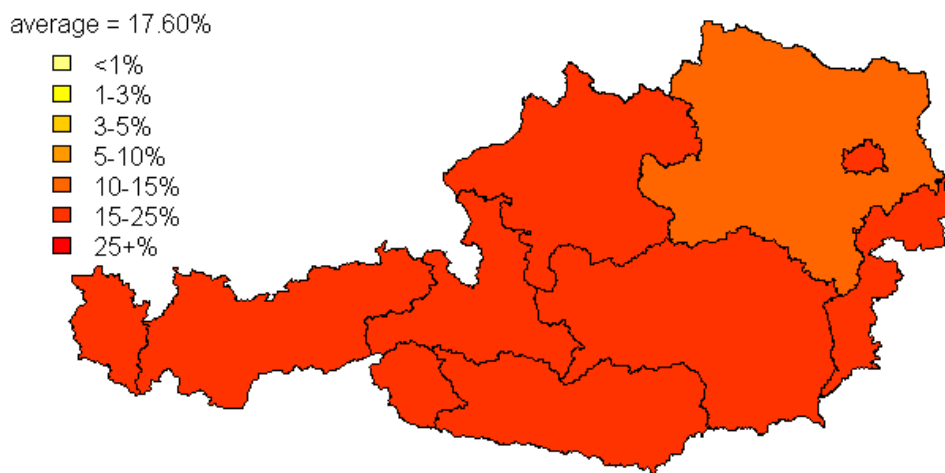


Abbildung 54 (b): Prävalenz der HV-Personen mit chronischem Schmerz - nach Geschlecht

Prevalence of Chronic Dolor Males (GB/SA)



Prevalence of Chronic Dolor Females (GB/SA)

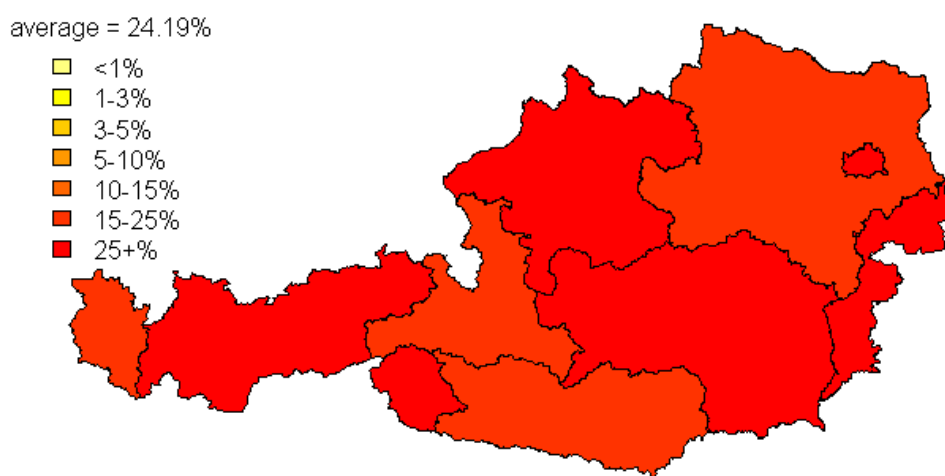


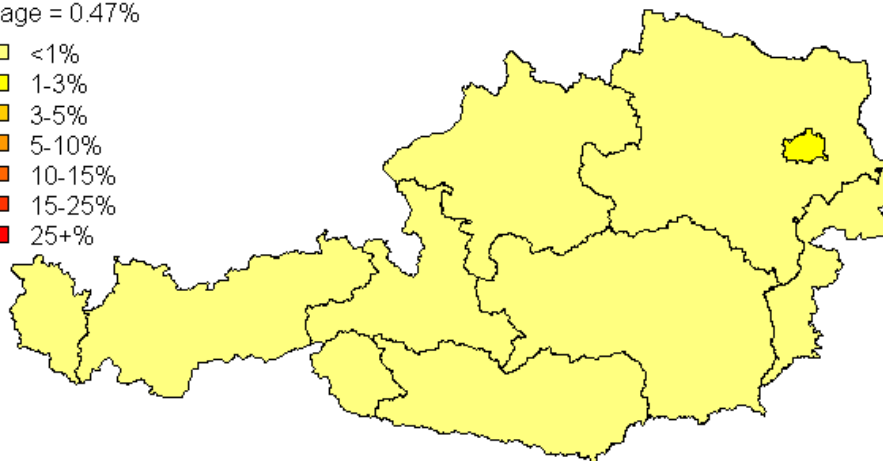
Abbildung 55: Prävalenz der GB-Personen mit Schmerzen länger als 3 Monate

6.2 Chronische Schmerzpatienten nach Altersgruppen und Geschlecht auf Ebene Bundesland (HV)

Prevalence of Chronic Dolor Males 0-14 years (HVB/SA)

average = 0.47%

- <1%
- 1-3%
- 3-5%
- 5-10%
- 10-15%
- 15-25%
- 25+%



Prevalence of Chronic Dolor Females 0-14 years (HVB/SA)

average = 0.46%

- <1%
- 1-3%
- 3-5%
- 5-10%
- 10-15%
- 15-25%
- 25+%

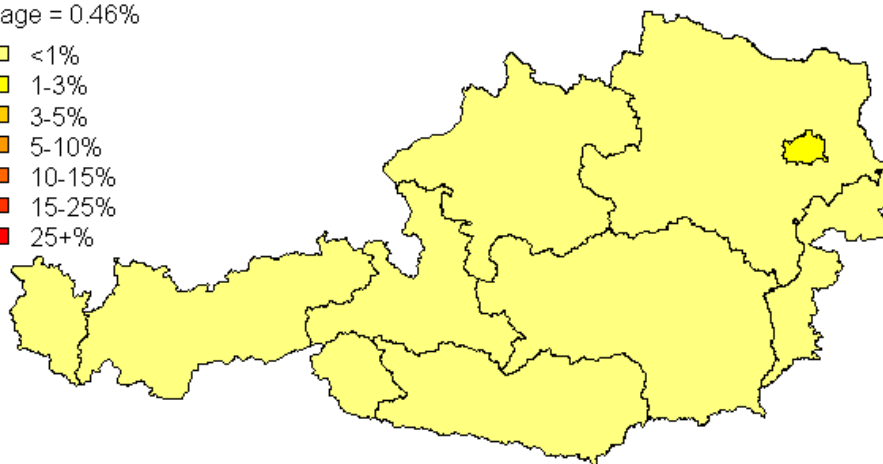
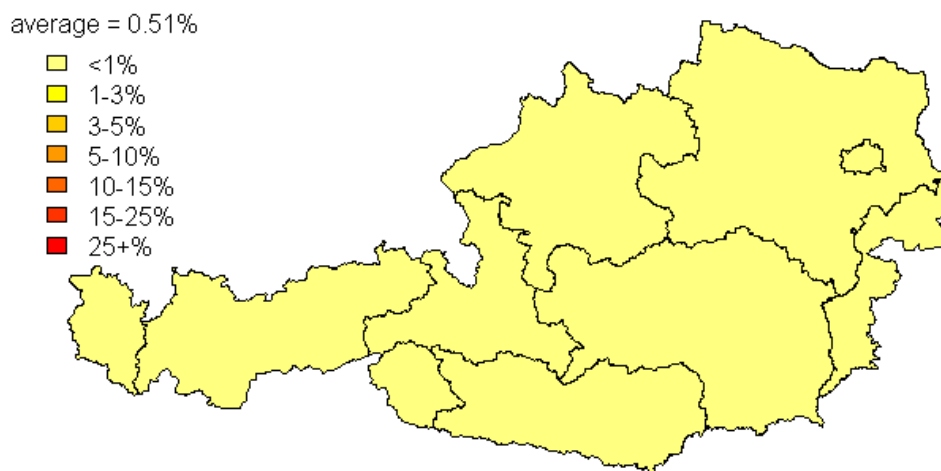


Abbildung 56: Prävalenz der HV-Personen mit chronischem Schmerz – Alter 0-14 nach Geschlecht auf Ebene Bundesland

Prevalence of Chronic Dolor Males 15-29 years (HVB/SA)



Prevalence of Chronic Dolor Females 15-29 years (HVB/SA)

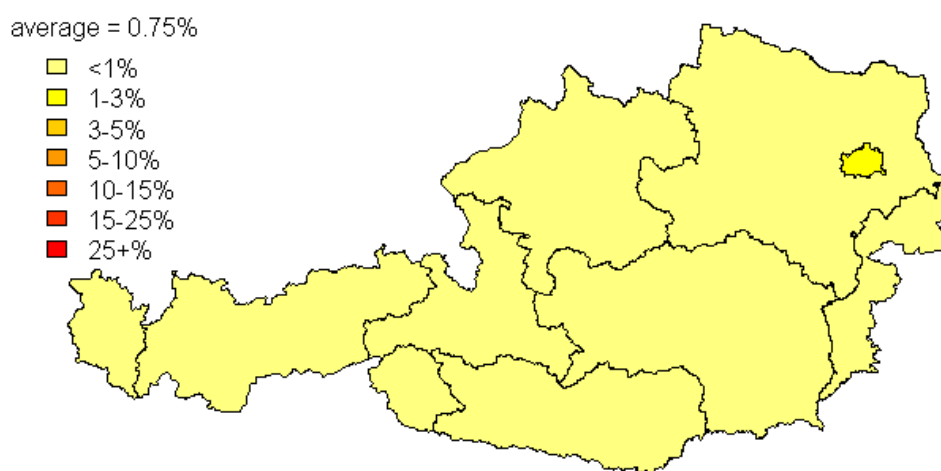
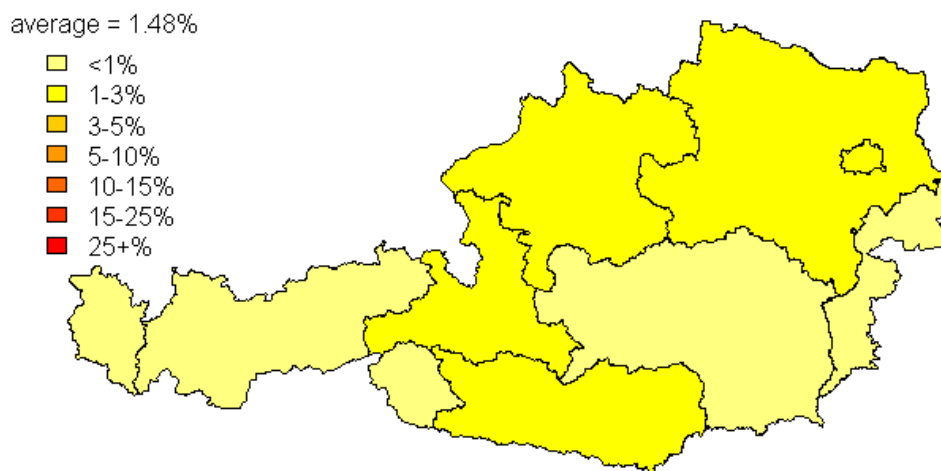


Abbildung 57: Prävalenz der HV-Personen mit chronischem Schmerz - Alter 15-29 nach Geschlecht auf Ebene Bundesland

Prevalence of Chronic Dolor Males 30-44 years (HVB/SA)



Prevalence of Chronic Dolor Females 30-44 years (HVB/SA)

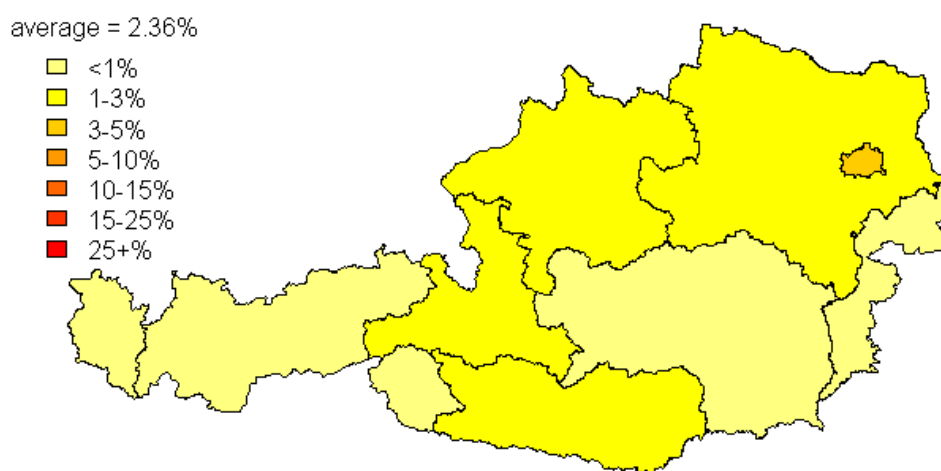
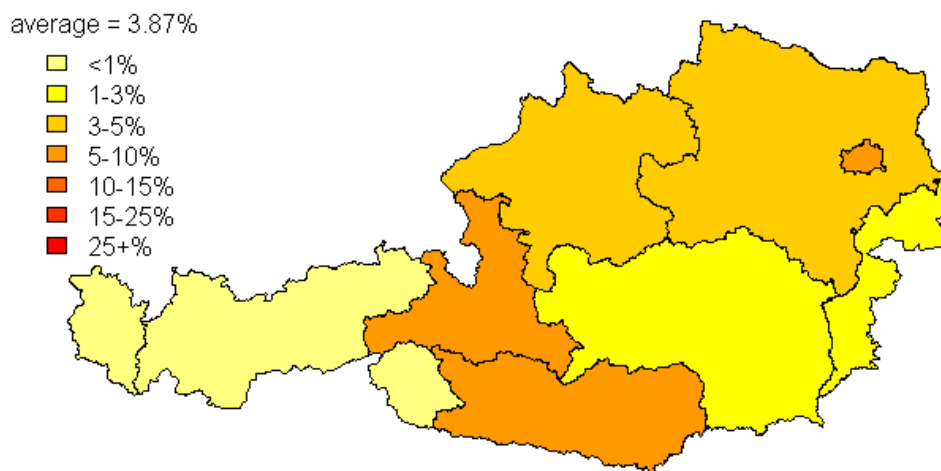


Abbildung 58: Prävalenz der HV-Personen mit chronischem Schmerz - Alter 30-44 nach Geschlecht auf Ebene Bundesland

Prevalence of Chronic Dolor Males 45-59 years (HVB/SA)



Prevalence of Chronic Dolor Females 45-59 years (HVB/SA)

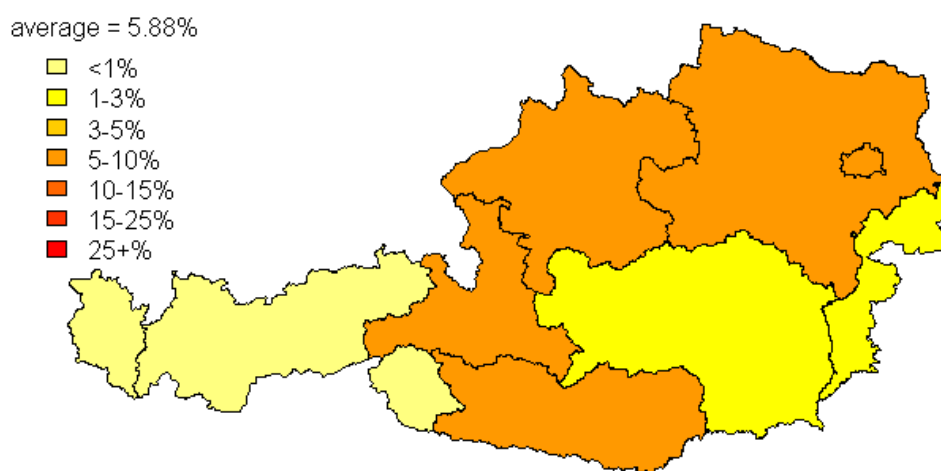
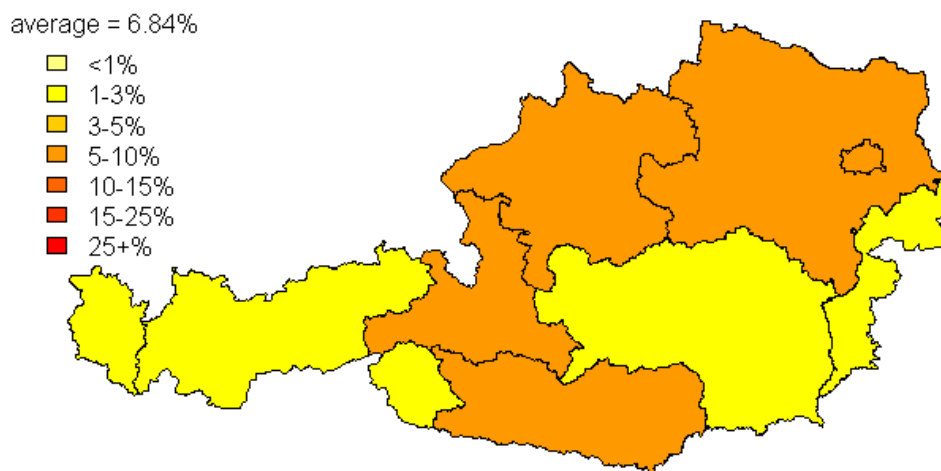


Abbildung 59: Prävalenz der HV-Personen mit chronischem Schmerz - Alter 45-59 nach Geschlecht auf Ebene Bundesland

Prevalence of Chronic Dolor Males 60-74 years (HVB/SA)



Prevalence of Chronic Dolor Females 60-74 years (HVB/SA)

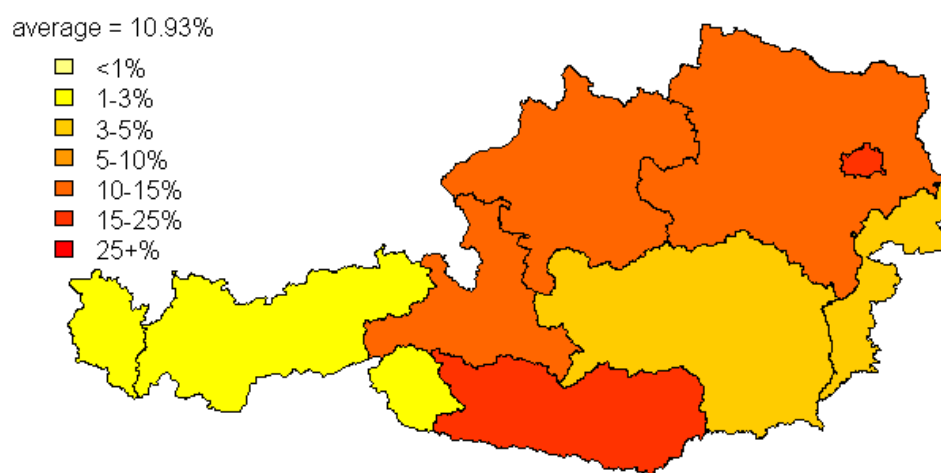
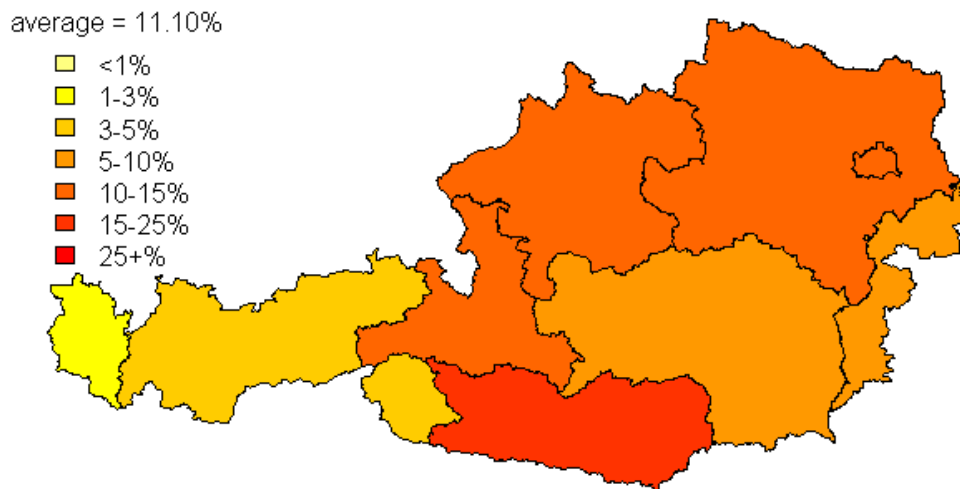


Abbildung 60: Prävalenz der HV-Personen mit chronischem Schmerz - Alter 60-74 nach Geschlecht auf Ebene Bundesland

Prevalence of Chronic Dolor Males 75+ years (HVB/SA)



Prevalence of Chronic Dolor Females 75+ years (HVB/SA)

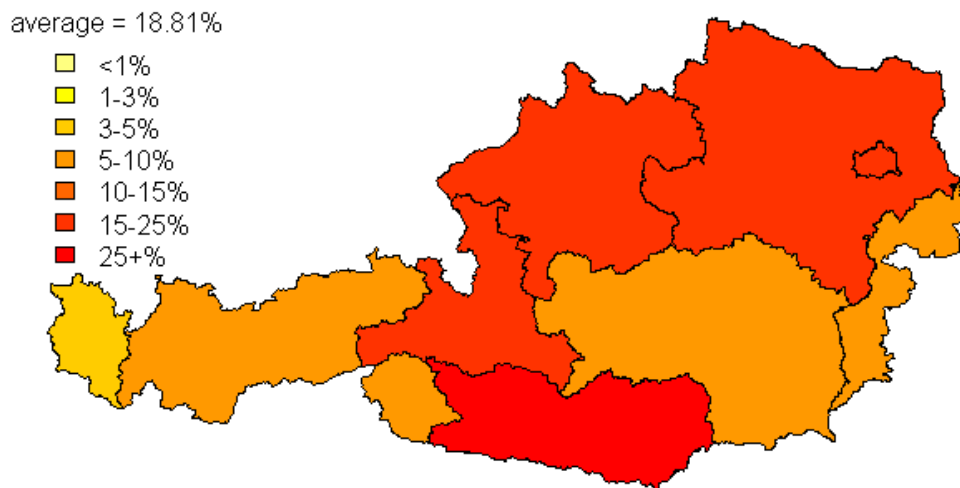


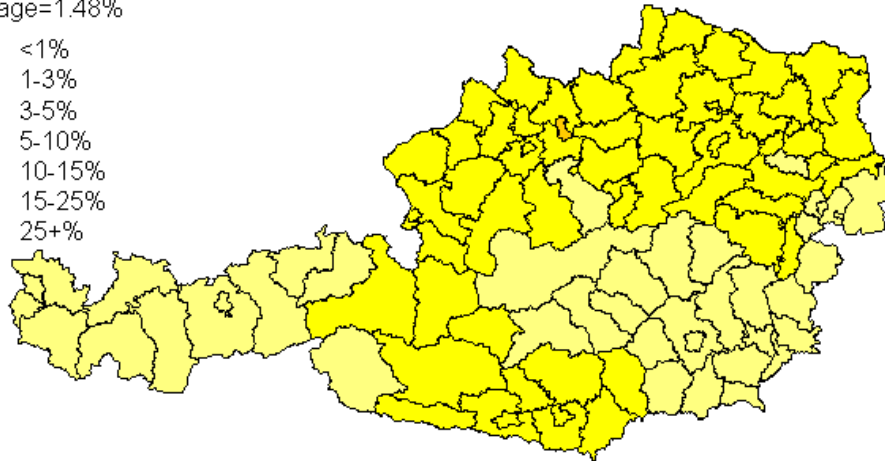
Abbildung 61: Prävalenz der HV-Personen mit chronischem Schmerz - Alter ab 75 nach Geschlecht auf Ebene Bundesland

6.3 Chronische Schmerzpatienten nach Altersgruppen und Geschlecht auf Bezirksebene (HV)

Prevalence of Chronic Dolor Males 30-44 years (HVB/SA)

average=1.48%

- <1%
- 1-3%
- 3-5%
- 5-10%
- 10-15%
- 15-25%
- 25+%



Prevalence of Chronic Dolor Females 30-44 years (HVB/SA)

average=2.36%

- <1%
- 1-3%
- 3-5%
- 5-10%
- 10-15%
- 15-25%
- 25+%

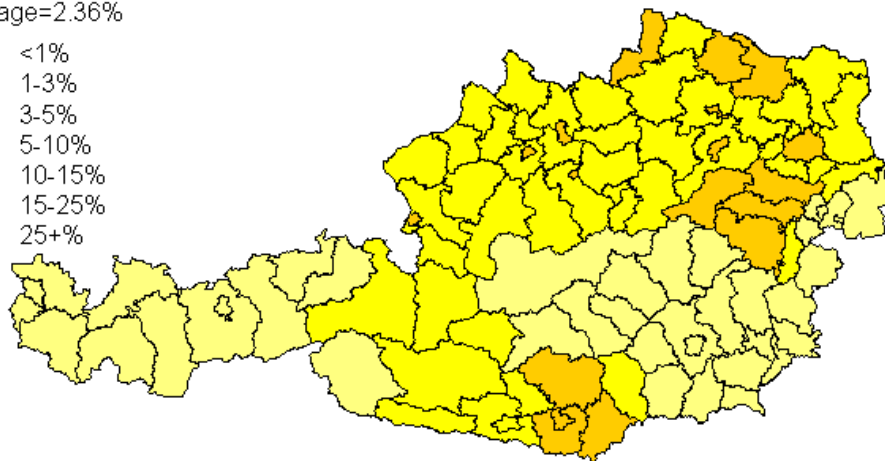
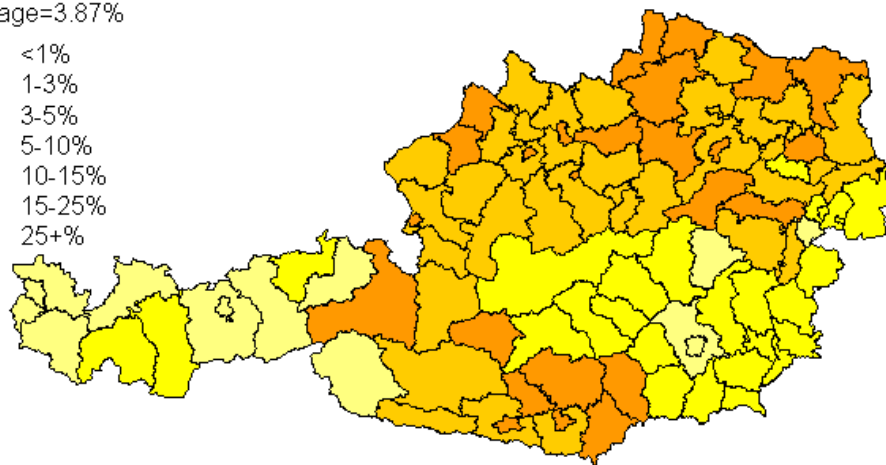


Abbildung 62: Prävalenz der HV-Personen mit chronischem Schmerz - Alter 30-44 nach Geschlecht auf Bezirksebene

Prevalence of Chronic Dolor Males 45-59 years (HVB/SA)

average=3.87%

- <1%
- 1-3%
- 3-5%
- 5-10%
- 10-15%
- 15-25%
- 25+%

**Prevalence of Chronic Dolor Females 45-59 years (HVB/SA)**

average=5.88%

- <1%
- 1-3%
- 3-5%
- 5-10%
- 10-15%
- 15-25%
- 25+%

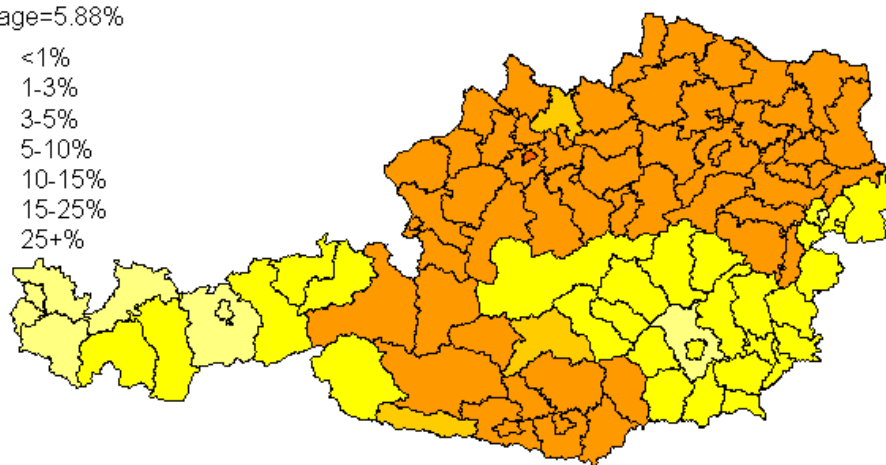
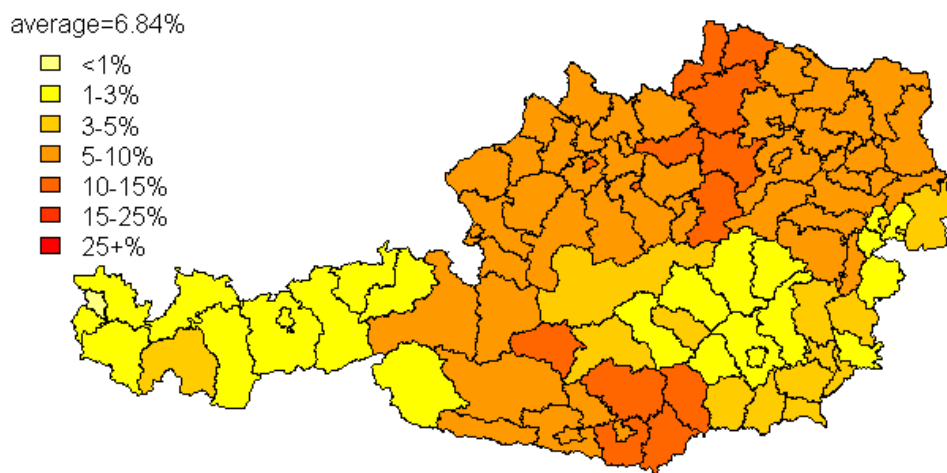


Abbildung 63: Prävalenz der HV-Personen mit chronischem Schmerz - Alter 45-59 nach Geschlecht auf Bezirksebene

Prevalence of Chronic Dolor Males 60-74 years (HVB/SA)



Prevalence of Chronic Dolor Females 60-74 years (HVB/SA)

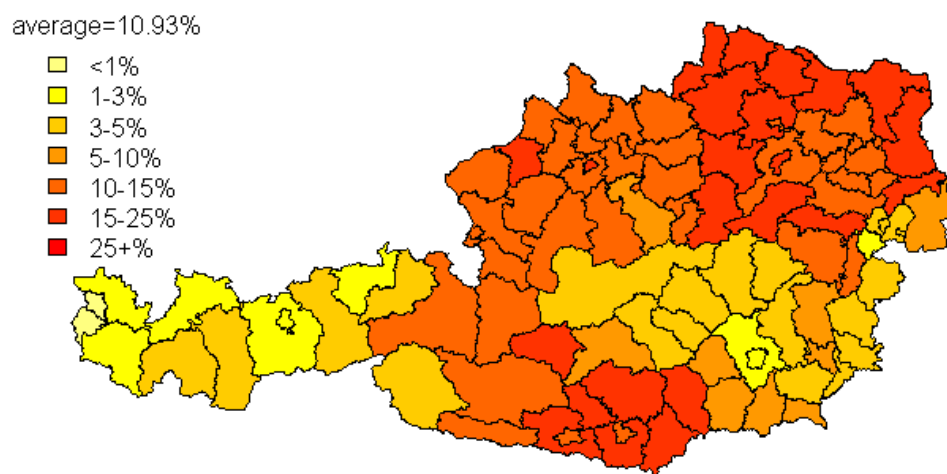
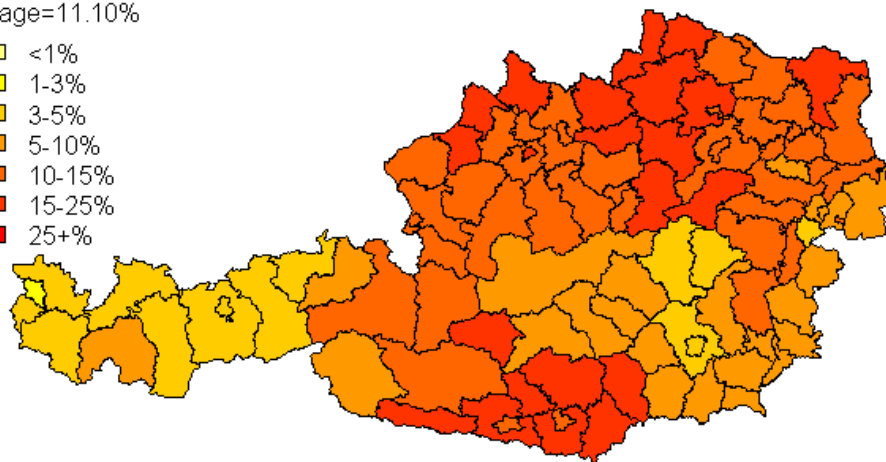


Abbildung 64: Prävalenz der HV-Personen mit chronischem Schmerz - Alter 60-74 nach Geschlecht auf Bezirksebene

Prevalence of Chronic Dolor Males 75+ years (HVB/SA)

average=11.10%

- <1%
- 1-3%
- 3-5%
- 5-10%
- 10-15%
- 15-25%
- 25+%

**Prevalence of Chronic Dolor Females 75+ years (HVB/SA)**

average=18.81%

- <1%
- 1-3%
- 3-5%
- 5-10%
- 10-15%
- 15-25%
- 25+%

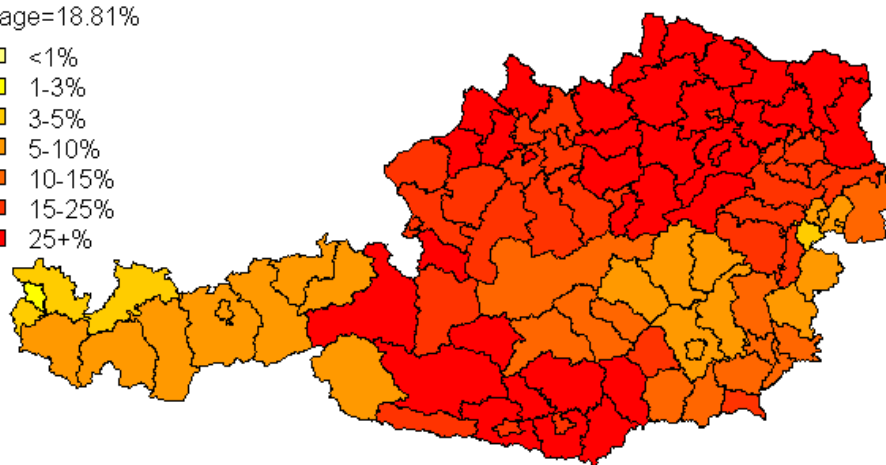


Abbildung 65: Prävalenz der HV-Personen mit chronischem Schmerz - Alter ab 75 nach Geschlecht auf Bezirksebene

Annex

Einige Anmerkungen

zum

GAP-DRG Datenmodell

Im Zuge der Umsetzung des Projekts OF-LÖG (siehe Projektplan vom 18. Feber 2010) war als einer der Aufgabenpunkte eine Sichtung des vom Hauptverband zur Verfügung gestellten Datenmodells GAP-DRG vorgesehen; als Ergebnis dieser Sichtung resultieren einige nachstehend angeführte Bemerkungen bzw., wenn man dies so verstehen will, Vorschläge zur Weiterentwicklung eines für die analytischen Ziele der EBM ausgelegten Datenmodells.

Die nachstehenden Überlegungen gehen – darin vielleicht von den Intentionen der Ersteller des GAP-DRG Datenmodells abweichend – von einer bewußt abstrahierenden Sicht auf die Daten im Sinne der EBM aus. Dies bedeutet, dass weniger die Strukturen der quellseitigen Datensysteme in Betracht gezogen werden, als vielmehr eine sachlogisch im Sinne der Nachvollziehbarkeit und Nachnutzbarkeit der Daten möglichst gut geeignete Modellierung angestrebt wird, die im Idealfall auch weitgehend *unabhängig* von Änderungen der Datenstrukturen der Quellsysteme oder der Bedingungen verschiedener Ladeprozesse, Integritätsprüfungen usw. gestaltet ist, um auf diese Weise eine stabile Schnittstelle zur statistischen Analytik zu ergeben.

Aus dieser Perspektive sollte klar sein, dass die angestellten Überlegungen (von einigen kleineren Aspekten einmal angesehen) dem Grunde nach nicht als eine *Kritik* des bestehenden GAP-DRG Datenmodells zu verstehen, sondern vielmehr von der Vorstellung getragen sind, unbeschadet vorhandener Datenquellen und -strukturen eine den Ambitionen der EBM angemessene Datenmodellierung zu erreichen. Zweifellos bilden die Datensammlungen des Hauptverbands der österreichischen Sozialversicherungsträger einen wesentlichen Beitrag zu einem solchen konsolidierten Datenpool; gleichzeitig ist aber offensichtlich, dass neben jetzt schon bestehenden vor allem in Zukunft absehbar weitere Datenquellen in Betracht zu ziehen sind bzw. sein werden – bspw., weil ein immer größerer Teil der medizinischen Versorgung außerhalb bestehender nationaler Abrechnungssysteme abgewickelt wird, sei es infolge der Inanspruchnahme außervertraglicher Versorgungsleistungen, sei es durch grenzüberschreitend erfolgende Medizindienstleistungen. Aus dem Blickwinkel der EBM kann letztlich nicht ein bestimmtes – wenngleich vielleicht traditionell privilegiertes oder konkurrenzloses – System der Organisation von Gesundheitsdaten unhinterfragt als Maßstab herangezogen werden, sofern dieses nicht alle erforderlichen oder jedenfalls vernünftig forderbaren Kriterien an *kanonischer* Repräsentation und Qualität hinsichtlich Modell und Inhalt schon vorab erfüllt.

Insoweit also die angestellten Überlegungen zu einem konsensuellen Modellentwurf hinführen, wäre in weiterer Folge daran zu denken, ggf. auch bereits bestehende Datenmodelle – wie natürlich auch GAP-DRG selbst – einer Revision zu unterziehen, bevor eine umfangreichere Datennutzung oder Datenzusammenführung unternommen wird. Zumindest hier wurde als Arbeitshypothese davon ausgegangen, dass der sich auf den Zeitraum 2006/07 bezogene Datensatz und seine Struktur als experimenteller Prototyp verstehen, der im Lichte sich ergebender Erfahrungen und Erkenntnisse schrittweise adaptiert werden kann und werden soll.

1 Konzeptionelle Einheiten und Konstrukte

Eine der wesentlichen Funktionen des (konkreten) Datenmodells besteht darin, um die zentralen Merkmalsträger herum verschiedene statistische Populationen von analytischem Interesse transparent formulieren und aus dem Datenbestand ableiten zu können. Die Analyse der vorfindlichen GAP-DRG Inhalte und Strukturen ergibt einen grundsätzlich dreischichtigen Aufbau, bestehend aus:

- **Entitäten** (im Sinne von primären Merkmalsträgern) und
- **Beziehungen** (zwischen Entitäten) als oberster Schicht,
- **Merkmalen** (als Schicht funktionaler Abbildungen der Merkmale von Entitäten und Beziehungen zu den diesen zugeordneten Wertebereichen), sowie den
- **Wertebereichen** selbst als dritter Schicht, die als sog. *Stammdaten* verwaltet wird.

Um Verwechslungen (vor allem im Kontext relationaler Datenbanksysteme und -modelle) zu vermeiden, wird nachfolgend anstelle von (inhaltlichen) Relationen stets von ‚Beziehungen‘ die Rede sein; technisch werden de facto Entitäten, Beziehungen und auch (in vielen Fällen) Wertebereiche relational – d.h. in Form von Tabellen – repräsentiert.

Im Kontext von GAP-DRG sind die relevanten Entitäten der Modellierung naturgemäß auf der einen Seite die **Patienten** („Versicherte“), denen auf der anderen Seite die **medizinischen Leistungsträger** (stationär, ambulant, niedergelassen) gegenüberstehen. Grundsätzlich wäre auch vorstellbar, eine Entität „Sozialversicherungsträger“ einzuführen; dies ist jedoch im Hinblick auf die analytischen Ziele der EBM wenig motiviert: vielmehr bietet es sich an, die Sozialversicherungsträger als spezifischen Wertebereich aufzufassen, der im Bereich der Stammdaten des Datenmodells repräsentiert werden kann.

Im Sinne einer prägnanten Darstellung wird nachfolgend von der Konvention Gebrauch gemacht, wesentliche Konstrukte durch Symbole zu kennzeichnen; konkret wird **(P)** für die Entität der Versicherten bzw. Patienten und **(M)** für die (verschiedenen Ausprägungen der) medizinischen Leistungsträger eingeführt, wobei die runde Klammerung auf den Typus ‚Entität‘ des Konstrukts verweist.

Das inhaltliche GAP-DRG Modell lagert nun im Grunde lediglich vier spezifische Beziehungen – entsprechend unterschiedlichen Formen von medizinischen Leistungsformen – an diese beiden Entitäten an, wie sich graphisch einfach in nachfolgender Abbildung veranschaulichen lässt (vgl. Abbildung 66, nächste Seite). Diese vier Beziehungen sind:

- Leistung im niedergelassen (bzw. ambulanten) Bereich,
- Leistung im stationären Bereich („Aufenthalt“),
- Verschreibung eines Heilmittels („Rezept“) und
- Bescheinigung einer temporären Arbeitsunfähigkeit („Krankmeldung“).

Ggf. kann ohne tieferegehende strukturelle Änderung des Modells eine Aufspaltung von niedergelassenem und (derzeit nicht berücksichtigtem) ambulantem Bereich vorgenommen werden. In Fortführung der symbolischen Kurzschreibweise werden die Kürzel **<A>** (für stationäre **A**ufenthalte), **<L>** (niedergelassene/ambulante **L**eistungen), **<H>** (für **H**eilmittelverschreibungen) und **<K>** (für **K**rankmeldungen) vorgeschlagen.

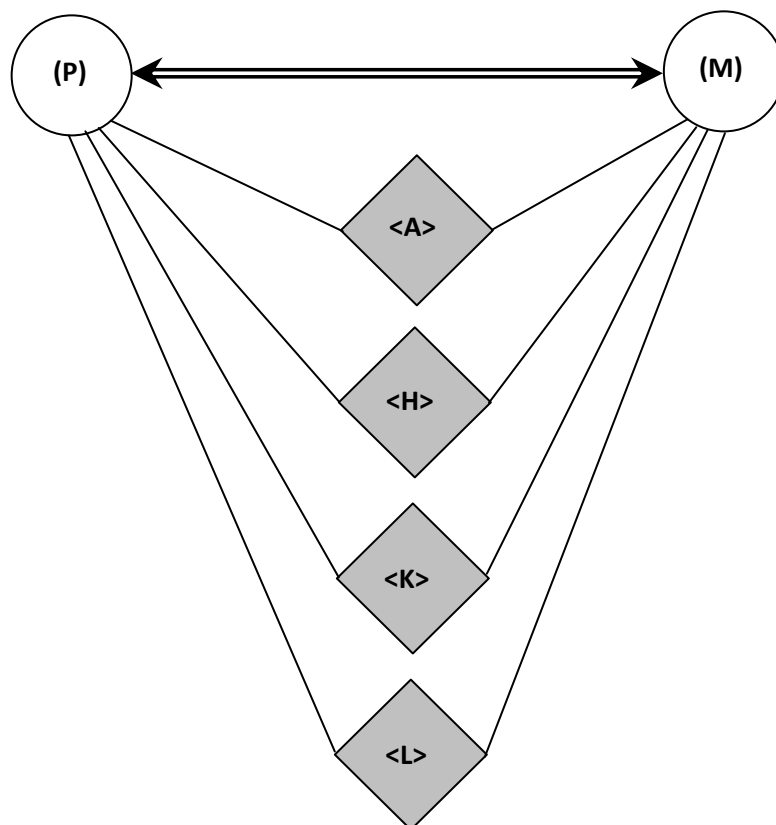


Abbildung 66: Elementares GAP-DRG ER-Modell

Dieses strukturelle Bild ist allerdings insofern noch nicht ganz korrekt, als die dargestellten Beziehungen im Grunde nicht direkt zwischen Patienten und Medizinleistungsträgern bestehen, sondern zwischen den diesen „vorgelagerten“ (Satelliten-) Entitäten (i) *Versicherungsverhältnis* (auf der Seite der Patienten) und (ii) *Vertragsverhältnis* (auf Seite der Leistungsträger). Unter Verwendung des Symbols **S** zur Kennzeichnung der **Sozialversicherungsträger** sind also die in der Abbildung eingetragenen Entitäten **(P)** bzw. **(M)** zu ersetzen durch **(<PS>)** – für die Versicherungsverhältnisse – bzw. **(<MS>)** für die „Leistungsträgerdynamik“. In beiden Fällen drückt die Symbolik (1:n)-Beziehungen aus, d.h. eine **P**-Instanz kann (zu jeder Zeit oder nacheinander) mit beliebig vielen **S**-Instanzen in Beziehung stehen, analog bzgl. **M**-Instanz.

Eine implizite *Konsistenzbedingung* ergibt sich für alle Instanzen der zwischen **(<PS>)**- und **(<MS>)**-Instanzen aufgespannten Beziehungen: die referenzierte **S**-Instanz muss jeweils *ident* sein, da eine versicherte Leistung seitens eines Patienten immer nur über eines „seiner“ Versicherungsverhältnisse in Anspruch genommen werden kann, aus dem heraus auch der Leistungsträger – wenn auch ggf. über einen zwischengeschalteten anderen („abrechnenden“) Versicherungsträger – refundiert wird.

Seitens der Entität **(P)** tritt als weitere Ergänzung des Datenmodells noch eine Beziehung **<PP>** der „Mitversicherten“ als Satellitenstruktur auf. Wie die Symbolik suggeriert, sind die Instanzen von **<PP>** Beziehungen *zwischen* Patienten zur Repräsentation indirekter Versicherungsverhältnisse, wie sie sich aus den Familienstrukturen (Status der Erwerbstätigkeit/Angehörige) ergeben. Allerdings unterliegt diese Beziehung einer Dynamik (Entstehen bzw. Ende von Mitversicherungsverhältnissen).

nissen) bzw. auch Multiplizität (mehrfache Mitversicherungsverhältnisse), die geeignet zu repräsentieren sind (siehe dazu später).

Auch bzgl. der Beziehungen ergeben sich teilweise *subsidiäre* Strukturen; dies betrifft einerseits die Beziehung **<A>** der (Anstalts-) Aufenthalte, zu denen Leistungen, Diagnosen und Verlegungen zugeordnet sind, andererseits die Beziehung **<K>** der Krankmeldungen, zu denen ebenfalls (jeweils mehrere) Diagnosen angeführt sein können. Entsprechend der eingeführten Symbolkonvention lassen sich diese Subsidiär-Beziehungen einfach durch **<AB>** (stationäre Ver- bzw. Belegungen), **<AD>** (Diagnosen zu stationären Aufenthalten), **<AL>** (abgerechnete stationäre Leistungen) bzw. **<KD>** (Krankmeldungsdiagnosen) kennzeichnen.

Alle übrigen Objekte des Gegenstandsbereichs fallen in die Kategorie von *Wertebereichen*, die unterschiedliche Strukturen aufweisen können, z.B. einfache Enumerationen (wie die Liste möglicher Staatsbürgerschaften) oder komplexe hierarchische Systeme von Wertausprägungen (wie die Systematik der pharmazeutischen Wirkstoffe ATC). Im Sinne der verwendeten Symbolik sind Wertebereiche durch eckige Klammern symbolisiert, z.B. **[nat]** für die Nationalität (Staatsbürgerschaft) oder **[atc]** für die Mehrsteller-Struktur der Wirkstoffe. Nachstehende Tabelle gibt einen Aufriss der in GAP-DRG genutzten Wertebereiche. Die Tabelle enthält auch einen Hinweis auf die *Varianten* einzelner Wertebereiche; diese werden meist zeitlicher Natur sein (d.h. eine spätere *Version* ersetzt eine vorangehend gültige), es sind aber auch gleichzeitig gültige Versionen durchaus vorstellbar. De facto sind im Zeitverlauf allerdings in den meisten Wertebereichen Anpassungen der Wertausprägungen zu erwarten, sodass in der Verwaltung dieser Stammdaten ein genereller Mechanismus zur Bestimmung von Gültigkeitszeiträumen anzustreben ist.

Überhaupt sind verschiedene Ebenen der Temporalität aus grundsätzlichen Erwägungen zu unterscheiden, nämlich:

- Dynamik der *Merkmalsträger* (Temporalität der Instanzen), bezogen auf die zeitliche Variation der Wertausprägungen einzelner Merkmale („Fieberkurve“);
- Dynamik der *Abbildungsfunktion* (Temporalität der Wertebereiche), wie sie sich in Varianten bzw. Versionen ausdrückt, sowie
- Dynamik der *Schemata* (Temporalität der Struktur der Repräsentation), z.B. durch „evolvierende“ Hinzunahme weiterer oder Entfall bisheriger Merkmale (bzw. gekoppelt in Form einer semantischen Redefinition einzelner Merkmale).

Der Dynamik von Abbildungsfunktionen wird meist durch Korrespondenzfunktionen Rechnung getragen; aber die Lebensdauer von Merkmalsträgern kann durchaus auch die Lebensdauer der Schemata überschreiten und so Anlass zu multiplen Repräsentationen (mit entsprechenden Übergangsfunktionen zur Wahrung der Objektidentität bzw. Rekonstruktion der Objektbiographien) geben. Die sich aus diesen Ebenen der Temporalität ergebenden Modellierungsfragen bzw. Aggregationslogiken werden in der einschlägigen Datenbank-Literatur (temporal data bases, spatio-temporal warehousing etc.) abgehandelt.

Eine andere Frage betrifft die Bildung von Wertbereich-Varianten (für verschiedene analytische Zwecke), z.B. durch Erweiterung der Grundwertebereiche um spezifische Wertausprägungen etwa zur Kennzeichnung von fehlenden Meldewerten, Kombinationen bestimmter Wertausprägungen, Gruppierungen und dgl. mehr. Bspw. könnte bzgl. des Gebietsstands – Wertebereich **[reg]** – eine Kombination der

Gliederung des österreichischen Gebietsstands (NUTS-Regionen, Gesundheitsregionen) mit dem anderer Staaten erfolgen.

Der in der Tabelle 52 angeführte „kombinierte Leistungskatalog“ **[lx]** ist vorerst lediglich als Platzhalter für eine Verschmelzung von **[hol]** und **[mel]** vorgesehen; generell sind Korrespondenzstrukturen vorzusehen, soweit (Mehrsteller-) Hierarchien oder andere Wertebereiche ineinander überzuführen sind (wie insb. im Fall von **[icd9]** <-> **[icd10]**).

Tabelle 52: Wesentliche Wertebereiche des GAP-DRG Strukturmodells

Kürzel	Bezeichnung	Typ	Variante*	Bedeutung
atc	Wirkstoff	Hierarchie	Nein	Systematik der pharmazeutischen Wirkstoffe
dgn	Diagnosen	Enumeration	ja ?	Systematik medizinischer Diagnosen (ICD)
dtyp	Diagnosetyp	Enumeration	nein	Typologie der Diagnosen (Haupt-/Nebendiagnose)
gen	Geschlecht (biolog.)	Enumeration	nein	Geschlecht der Personen
hm	Heilmittel	Hierarchie	nein	Liste der (jemals) zugelassenen Heilmittel
hol	Leistung gem. Honorarordnung	Hierarchie	nein	Generalliste aller Honorarordnungen
icd9	ICD-9	Hierarchie	ja	Systematik medizinischer Diagnosen (ICD-9) – Version, als Variante zu ICD-10
icd10	ICD-10	Hierarchie	ja	Systematik medizinischer Diagnosen (ICD-10) – Version, als Variante zu ICD-9
lx	Kombinierter Leistungskatalog	Hierarchie	nein	Integration von (meta-) HOL und MEL
mel	med. Einzelleistungen	Enumeration	nein	Katalog med. Einzelleistungen
mfg	med. Fachgebiete	Enumeration	nein	Liste der medizinischen Fachgebiete
mtyp	Typ des med. Leistungsträgers	Enumeration	nein	Verzeichnis der Typen med. Leistungsträger
nat	Staatsbürgerschaft	Enumeration	nein	Liste der Nationen
reg	Raumgliederung	Hierarchie	ja	Einteilung des Gebietsstands
svkat	Versicherungskategorien	Enumeration	nein	zeitabhängige Versicherungskategorie (im Versicherungsverhältnis)

* die Ausprägung ‚nein‘ bedeutet hier, dass eine gleichzeitige Nutzung verschiedener Versionen ausgeschlossen ist

Bzgl. der (Entität der) Sozialversicherungsträger (**S**) können diese aus praktischen Gesichtspunkten (wie oben dargelegt) ebenfalls technisch als Wertebereich **[svt]** und somit als Stammdaten konzipiert werden. Im Hinblick auf den Umstand, dass

es auch „vertragslose“ Zustände geben kann (einerseits privat versicherte oder sich selbst versichernde Patienten, andererseits Medizinleistungsträger ohne Vertrag mit einer Sozialversicherung), wäre der Wertebereich **[svt]** um eine Wertausprägung ‚vertragsfrei‘ zu erweitern, die in formaler Hinsicht wie die Wertausprägungen anderer registrierter Sozialversicherungsträger genutzt werden kann.

In jedem Fall ist der Ladevorgang der Quelldaten (der einzelnen Sozialversicherungsträger) in die Datenbank der geeignete Zeitpunkt, allfällige Inkohärenzen der Wertebereiche auszuräumen und auf GAP-DRG Standard-Terminologie umzustellen. Dabei lassen sich grundsätzlich drei Prüf-Ebenen der Datenübernahme unterscheiden, nämlich:

- Prüfung *individueller* Merkmale: auf dieser Ebene werden einzelne Merkmale für sich im Sinne des sog. *profiling* analysiert, um die Zulässigkeit von Wertausprägungen, Ausreißer, illegale Werte, fehlende Werte, Formatfehler und dgl. mehr zu erkennen und nach Möglichkeit automatisiert (unter Anlage eines Fehlerprotokolls) richtig zu stellen;
- Prüfung von Enitäten und Beziehungen auf *kollektiver* Ebene: vielfach lässt sich die Validität bzw. ggf. Korrektur von Quelldaten nicht anhand einer Prüfung einzelner Merkmale für sich feststellen, sondern nur aus Charakteristika einer gesamten Population ableiten, bspw. wenn die Plausibilität multivariater Strukturen (Stichwort: Ausreißer) oder Imputationen aus den gemeinsamen Verteilungseigenschaften mehrerer Merkmale gleichzeitig abgeleitet werden;
- Prüfung auf der Ebene der *Datenkonsolidierung*: schließlich treten bestimmte Konstellationen erst durch den Abgleich mehrerer (ganz oder teilweise, zeitlich und/oder sachlich) überlappender Datenkörper zutage, die ggf. entsprechende Korrekturmaßnahmen nach sich ziehen.

In jedem Fall ist natürlich wünschenswert, möglichst viele der Prüfvorgänge direkt und automatisiert in die Ladeprozesse (der Quelldaten in das GAP-DRG System) einzubeziehen, um den Aufwand manueller Interventionen gering zu halten und auf spezifische Maßnahmen der Qualitätssicherung zu beschränken.

2 Referenzierung von Instanzen (Schlüsselkonzept)

Alle Objekte (Instanzen) im GAP-DRG Datenmodell müssen naturgemäß über *eindeutige* Schlüssel referenzierbar sein. Die Schlüsselstruktur leitet sich unmittelbar aus dem entworfenen Datenmodell ab, muss aber praktisch aus den verfügbaren Attributen der Datenquellen ermittelt werden. In aller Regel empfiehlt es sich, wie das GAP-DRG bisher auch vorsieht, den Objekten (Instanzen) *intern* eindeutige Schlüssel zuzuweisen. Bei allen Stammdaten (Wertebereichen) ist dies relativ problemlos möglich; schwieriger liegt der Fall bei den Registern, also konkret bei den Entitäten **(P)** und **(M)**, da diese ihre Fremdschlüssel-Beziehungen aus den Quellsystemen an GAP-DRG „erben“. Generell ist aber zu empfehlen, im Zuge des Datenimports in GAP-DRG (Ladevorgang) die ohnehin unvermeidlichen Konsistenz- und Kohärenzprüfungen der Quelldaten (auch) zur Schlüsselkonversion zu nutzen, u.a. um auf diesem Weg Redundanzen zu vermeiden und von allfälligen Restrukturierungen der Quellsysteme unabhängig zu werden bzw. zu bleiben. Bspw. macht es vermutlich wenig Sinn, in GAP-DRG die Versicherten-Pseudonyme zu übernehmen, sondern diese besser in einer Hilfstabelle (Pseudonym-Register) für den Ladevorgang zu hinterlegen (damit wird bspw. erreicht, dass Personen, die aus we-

lichen Gründen auch immer ohne Pseudonyme ausgestattet sind, in den Datenbestand konsistent zu integrieren). In GAP-DRG selbst wäre vorteilhaft durchgängig mit einer intern vergebenen **P_ID** zu operieren. Diese Argumentation gilt analog für Entität (**M**). Selbstredend ist – nicht zuletzt in autodokumentativer Hinsicht – empfehlenswert, durchgängig und konsistent in der Implementierung von GAP-DRG die (wie auch immer) festgelegten Identifikatoren-Symbole zu nutzen. So wären bspw. die Instanzen der Primär-Beziehungen **<A>**, **<H>**, **<K>** und **<L>** einheitlich über die Fremdschlüssel **PS_ID** und **MS_ID** der Entitäten(-Relationen) (**<PS>**) bzw. (**<MS>**) identifizierbar, ggf. natürlich unter Hinzunahme weiterer Schlüsselkomponenten.

Die Tragweite einer solchen konsistenten Schlüsselorganisation wird insb. in der Herstellung der für die Analytik essentiellen Datenverknüpfungen (sog. „joins“ zur Ableitung multidimensionaler Tabellen) erkennbar. Typischer Weise werden bestimmte Patienten-Teilkollektive solchen Analysen zugrundeliegen (z.B. alle Versicherten mit einem diagnostizierten Diabetes innerhalb eines bestimmten Zeitraums); daraus ergibt sich, dass derartigen Datenselektionen zunächst einmal eine Verknüpfung zwischen den Entitäten (**P**) bzw. (**<PS>**) und den daran hängenden Beziehungen **<A>**, **<H>**, **<K>** sowie **<L>** (mitunter auch unter Einbeziehung subsidiärer Beziehungen) herzustellen ist (z.B., wenn die Auswahl von Diagnosen abhängt, die über die Subsidiär-Beziehungen **<AD>** bzw. **<KD>** repräsentiert sind). Infolge der jeweils 1:n-relationalen Verknüpfungen zwischen (**P**) und (**<PS>**) bzw. (**<PS>**) und den (primären) Beziehungen über die Fremdschlüssel **P_ID** bzw. **PS_ID** sind derartige Verkettungen grundsätzlich sehr einfach realisierbar. Analoges gilt spiegelbildlich für die Verknüpfung ausgehend von (**M**) bzw. (**<MS>**). Die Filterungen selbst erfolgen in aller Regel dann über bestimmte Wertausprägungen entsprechender Merkmale, z.B. Heilmittel- oder ICD-Codes in Heilmittelverschreibungen (**<H>**) oder Diagnosen (**<AD>** oder **<KD>**).

Besonderes Augenmerk ist in der Datenmodellierung der *zeitlichen* Dimension zu widmen, da viele – longitudinale – Analysen die temporale Struktur (Krankheitsverläufe, Inzidenzen, Prävalenzen, ...) als Ausgangspunkt verwenden. Um eine einheitliche Verarbeitungs-, Verknüpfungs- und Abfrage-logik zu erlauben, sind daher alle zeitbezogenen Merkmale konsistent zu gestalten und zu nutzen. I.A. wird als Granularität der Zeitachse das *Tagesdatum* heranzuziehen sein; in etlichen Fällen bezieht sich die Granularität aber auch nur auf Monate, Quartale oder Jahre. Semantisch ist darüber hinaus eine Differenzierung zwischen

- *Beginn(zeit)punkten* (bei Zustandseintritten, einsetzenden Vorgängen, ...);
- *End(zeit)punkten* (bei Zustandsrückführungen, endenden Vorgängen, ...), sowie
- *Ereignis(zeit)punkten* („punktuelle“ Zeitbezüge, z.B. Datum einer Diagnosestellung)

zu berücksichtigen. So sind etwa (*Zeit-*) *Dauern* nur sinnvoll zwischen Beginn- und Endpunkten bestimmbar, während ein Vergleich von Ereignispunkten lediglich zeitliche *Abstände* ergibt. Die Konsequenzen bzgl. der Aggregation von dauerhaften Vorgängen gegenüber punktuellen Ereignissen sind offensichtlich. In der Regel ist die temporale Semantik eines Merkmals festzulegen, da diese u.a. durch die gewählte Granularität mitbestimmt ist (so wird der „punktuelle“ Zeitbezug auf ein Monat nur wenige Tage andauernde Vorgänge als Ereignis subsumieren). Es em-

pfiehlt sich, die Semantik von Zeitmerkmalen am besten bereits in der Form der symbolischen Bezeichnung (z.B. einen Präfix) zum Ausdruck zu bringen.

In der praktischen Realisierung des Datenmodells von GAP-DRG wird in manchen Fällen eine stärkere Gliederung der Relationen als an sich logisch angezeigt erfolgen; z.B. wird die Entität **(P)** bislang aufgespalten in (i) Personen-Stammdatensätze (**gapdrg.personen**) und (ii) diesen zugeordnete Instanzen, die auf (monatlichen) Meldungen der einzelnen Versicherungsträger beruhen (**gapdrg.personendaten**). Grundsätzlich folgt diese Datenorganisation den Geboten der Zweckmäßigkeit, wobei allerdings dennoch im Sinne der Wahrung der Datenkonsistenz (die sich in der Kohärenz der darauf aufbauenden Analytik spiegelt) auf Konsolidierung Bedacht zu nehmen ist; siehe dazu gleich unten. Eine typische Quelle von Inkonsistenzen ergibt sich etwa in der Zusammenführung personenbezogener Stammdaten aus den Quellsystemen unterschiedlicher Versicherungsträger (etwa bzgl. der Merkmale Geschlecht, Todesdatum, oder auch Rezeptgebührenbefreiung). Vor allem im Hinblick auf Analysen der *Qualität* von Quelldaten mag es angezeigt sein, zunächst auch inkonsistente Daten aus den Quellsystemen in GAP-DRG zu übernehmen; insofern jedoch in die nachgelagerte Analytik unterschiedliche Varianten konsolidierter Werte einfließen, ist dies entsprechend explizit (am besten auf Merkmalsebene) zu dokumentieren. Nicht zweckmäßig ist hingegen die Praxis, auf *ad hoc* erfolgte Datenbereinigungen zurückzugreifen, da dies die Vergleichbarkeit analytischer Ergebnisse beeinträchtigen kann und ggf. die Nachvollziehbarkeit von Analysen überhaupt in Frage stellt. Werden bspw. im Zuge der Zusammenführung von Quelldaten verschiedener Sozialversicherungsträger voneinander abweichende Stammdaten einzelner Patienten (z.B. teilweise fehlendes Sterbedatum, variierende Geschlechtsangaben, unterschiedliche Wohnsitze zum gleichen Berichtszeitpunkt, wechselnder Ausweis des Status der Rezeptgebührenbefreiung, etc.) entdeckt und sind diese nicht im Sinne einer formalen Bereinigung während des Ladevorgangs eindeutig auflösbar, sollten jeder Analytik auf solchen Merkmalen entsprechende Konversionen der Meldedaten auf Ebene der Entitäten (**(P)** bzw. **<PS>**) – ggf. in *multiplen* Versionen – vorangehen, die dann als dokumentierte Varianten ein- und desselben Grundmerkmals in die Struktur aufgenommen werden (können). In Fortführung des Beispiels zum Merkmal der Rezeptgebührenbefreiung würde sich ev. anbieten, dieses Merkmal technisch aus der Relation zur Repräsentation der Entität **(P)** sogar ganz herauszulösen und den Verlauf der Rezeptgebührenbefreiung eines Versicherten separat abzubilden; in der Tat ist aber ein weitergehendes Konzept wie im nachfolgenden Abschnitt skizziert zu präferieren.

3 Praktische Überlegungen zur Modellierung des Versichertenregisters

Die folgenden Absätze konkretisieren verschiedene Konsequenzen der angestellten Überlegungen zur Datenmodellierung im Bereich der im gegenständlichen Projekt zentralen Patientendaten. Den Ausgangspunkt bilden dazu die beiden Tabellenstrukturen „Personen_gesamt“ bzw. „Personendaten“ aus dem Dokument „Semantische Datenbeschreibung Casemix-Daten laut technischem Datenmodell“ (Stand 03. Feber 2010). Diese Tabellen werden im Bericht weiter oben als **gapdrg.personen** bzw. **gapdrg.personendaten** referenziert. Anzumerken ist, dass gegenüber dem zitierten Berichtsstand inzwischen die Tabelle **gapdrg.personen** um ein Attribut ‚zpv‘

ergänzt wurde (das anzeigt, ob der Datensatz mit den Daten der ZPV, Zentrale Partnerverwaltung, abgeglichen und übereinstimmend ist).

Zur Tabelle **gapdrg.personen** liegt nachstehende Beschreibung vor:

Tabelle Personen_gesamt

Diese Tabelle beinhaltet Stamminformationen zu den Personen. In dieser Tabelle sind die Pers_ID unique.

PERS_ID	PK	Identifikationsnummer Person (numerischer Wert, aus vnump abgeleitet)
VNUMP		Pseudonym in Langform (in der ursprünglich von FOKO erhaltenen Form für die Person)
GEBJAHR		Geburtsjahr ¹
TODDATUM		Datum des Todes ²
Gebjahr_mehrfach		Nur mit J befüllt, wenn beim FOKO Import mehr als ein Geburtsjahr eingetragen war, ansonsten (null) ³

¹ Bei Personen, in denen *Gebjahr_mehrfach* (491323 Personen) mit J befüllt ist, kann nicht festgestellt werden, ob das Geburtsjahr valide ist und ist daher sinnvollerweise für eine Auswertung zu exkludieren

² Todesdatum ist häufig mit 04122006 befüllt, da durch eine Datenbankmigration bei denjenigen, die die Eigenschaft gestorben „ja“ vermerkt hatten, jedoch kein definitives Todesdatum gemeldet hatten, das Todesdatum mit dem Datumswert der Migration befüllt wurde

³ Alle zu einer Person aus FOKO importierten Geburtsjahre finden sich in der Tabelle Personendaten im Feld *Gebjahr* wieder

vnump= alle Anspruchsberechtigten, die eine Leistung/Verordnung/Krankenstand/Krankenhausaufenthalt im Zeitraum 2006/2007 hatten.

vnumpv= alle Versicherten, die entweder selbst eine Leistung/Verordnung/Krankenstand/Krankenhausaufenthalt im Zeitraum 2006/2007 hatten oder deren Angehörige eine Leistung/Verordnung/Krankenstand/Krankenhausaufenthalt im Zeitraum 2006/2007 hatten.

Analog liegt zur Tabelle **gapdrg.personendaten** nachstehende Beschreibung vor:

Tabelle Personendaten

Diese Tabelle beinhaltet weitere Stamminformationen zu Personen und ist mit Personen_gesamt über die pers_id, sowie Bezirk verknüpft. In dieser Tabelle können die Pers_IDs aufgrund verschiedener Einträge zu einer Person mehrfach vorkommen.

ID	PK	Identifikationsnummer (Fortlaufende Nummer innerhalb der Tabelle), Autowert
PERS_ID	FK	Identifikationsnummer Person (Feld pers_id in Tabelle personen_gesamt)
VTR		Meldender Versicherungsträger Jener Träger, der diesen Datensatz gemeldet hat. Es ist nicht klar, ob dies Leistungszuständiger oder Abrechnender Versicherungsträger ist. Ist fachlich mit dem Feld Herkunft ident.
VNUMV		Pseudonym des Versicherten in Langform, ist meistens (null) ¹
BEZIRK_ID	FK	Identifikationsnummer des Bezirks (Feld Bezirk_id in der Tabelle Bezirk)*
VERSKAT	FK	Versicherungskategorie (Beschreibung siehe Tabelle Verskat)*
REZGEB		Rezeptgebührenbefreiung (Ja/Nein)*
VON		Von Datum (Gültigkeit der in dem Datensatz befüllten Ausprägung)*
BIS		Bis Datum (Gültigkeit der in dem Datensatz befüllten Ausprägung)*
HERKUNFT	PK	Jener Träger, der diesen Datensatz gemeldet hat. Ist fachlich mit dem Feld VTR ident.
GESL		Geschlecht, auch wenn mehrfach vorhanden*
GEBJAHR		Geburtsjahr (wenn mehrfach vorhanden befüllt, ansonsten (null))*

*alle Ausprägungen gehören der pers_id an und nicht der vnumv

¹Um feststellen zu können, ob Patient Angehöriger ist muss man die Felder vnumv aus der Tabelle personen_gesamt und vnumv aus der Tabelle Personendaten vergleichen. Weichen die Werte ab, so drückt die pers_id einen Angehörigen aus.

Grundsätzlich ist anzumerken, dass unter den gegebenen Umständen eine Aufteilung der Attribute (im zitierten Dokument als ‚Felder‘ bezeichnet, eine gängige Konvention) auf mehrere untereinander verbundene Tabellen durchaus zweckmäßig ist; gleichzeitig ist die gewählte Form dieser Aufteilung aber zu hinterfragen. Die Entität (**P**), wie vorhin eingeführt, modelliert im Prinzip das *Register* aller Personen, die im Gesundheitssystem erfasst sind (der *semantische* Umfang dieser Population wird an dieser Stelle nicht weiter in Betracht gezogen). Abgesehen von einer allgemeinen Referenz, die ein Versicherter innerhalb des österreichischen Systems der gesetzlichen Sozialversicherung zugewiesen bekommt (Sozialversicherungsnummer, © 2011 MU Wien/Univ. Wien

die hier aus datenschutzrechtlichen Gründen allenfalls als Pseudonym in Erscheinung tritt) und die am besten durch einen internen GAP-DRG Schlüssel ersetzt wird, sind zu jeder Person einige Attribute zur Beschreibung naheliegend: Geburtsjahr, ggf. Sterbedatum (man beachte die Asymmetrie: der genaue *Todestag* hingegen scheint datenschutzrechtlich unproblematisch zu sein), sowie das (biologische) Geschlecht. Unbeschadet der prinzipiellen Konstanz dieser Attribute (vom Sonderfall des Geschlechts einmal abgesehen) liegt die Tücke diesfalls vor allem im Bereich der Datenkonsolidierung: die von den verschiedenen Sozialversicherungsträgern und über die Zeit einlaufenden Datenmeldungen sind nicht notwendig konsistent (wie sich empirisch ja auch zeigt). Daneben besteht immer auch noch die Möglichkeit, dass im Register Personen *ohne* Bezug zu einer (österreichischen) Sozialversicherung auftreten.

Die Instanzen der Entität **(P)** bilden also zuallererst einmal die Gesamtheit der erfassten Personen, denen alle stabilen und konsolidierten Merkmale direkt als Attribute zugeordnet werden können; dies ergibt i.W. ein Tabellen-Schema der Form von **gapdrg.personen**. Zu überlegen im Sinne eines Re-Designs wäre allerdings:

- Durch die Einführung eines GAP-DRG internen Schlüssels für die Entität, hier ‚PERS_ID‘ (im Sinne der vorgeschlagenen Symbolik wäre mnemotechnisch vermutlich **P_ID** prägnanter), wird die Mitspeicherung der Pseudonyme (soweit überhaupt vorhanden) redundant; die Beziehung zwischen internem Schlüssel und Pseudonym sollte besser auf die Ebene des Ladevorgangs in Form einer dort verwalteten Hilfstabelle verlagert werden; das redundante Feld ‚VNUMP‘ kann somit entfallen und dadurch das GAP-DRG Schema vereinfachen helfen; die in diesem Feld implizit codierte Information, ob die Instanzen im Zeitraum 2006/07 eine Versicherungsleistung in Anspruch genommen haben, ist definitionsgemäß nicht statisch und daher aus dem Schema der Tabelle sowieso zu exkludieren.
- Das Feld ‚GEBJAHR‘ als Zeitbezug mit Granularität ‚Jahr‘ sollte entweder eine valide Jahresangabe oder eine eindeutige Wertausprägung für eine fehlende Angabe (wie ‚0000‘) enthalten. Ob derartige Datensätze aus nachfolgenden Auswertungen auszuschließen sind oder nicht, hängt von den konkreten Umständen der Analytik ab und kann nicht ex ante entschieden werden.
- Das Feld ‚Gebjahr_mehrfach‘ ist eine wenig informative Hilfskonstruktion aus der Phase der Datenkonsolidierung. Eleganter wäre zweifellos stattdessen die Angabe eines *Status* (d.h., Meta- oder Paradata) dieses Attributs, der sich eines eigens dafür festgelegten Wertebereichs bedient, aus dem jedenfalls hervorgeht, ob GEBJAHR ein gemeldetes eindeutiges Datum enthält bzw. auf welche Weise die vermerkte Wertausprägung (im Ladevorgang bzw. der Datenvalidierungsphase) sonst zustande gekommen ist.
- Noch lebende Personen weisen zwangsläufig einen Fehlwert im Feld ‚TODDATUM‘ auf; analog zum Fall des Geburtsjahres sollte dementsprechend ein interpretierbarer Fehlwert (wie z.B. ‚39.19.0000‘) eingetragen sein (die Verwendung datenbank-spezifischer System-Nullwerte empfiehlt sich in solchen Fällen eher *nicht*, da es ev. auch andere – technische – Gründe geben kann, warum ein Eintrag fehlt). Keinesfalls darf jedoch ein sachlich willkürliches Sterbedatum angeführt werden, das gleichzeitig auch ein valides Sterbedatum von Versicherten sein könnte (wie dies bei ‚04.12.2006‘ zweifellos der Fall ist),

nur weil das Sterbedatum nicht bekannt ist: hier wäre eine weitere Wertausprägung zur Kennzeichnung eines unbekanntes Sterbedatums vorzusehen (z.B. ‚39.19.9999‘) oder, noch weitergehend, eine spezifische Wertausprägung zur Kennzeichnung eines unbekanntes Sterbedatums vor einem bestimmten Stichtag (z.B. ‚39.19.9990‘ für „unbekanntes Sterbedatum vor oder am 04.12.2006“). Analog zum Geburtsjahr sollte zudem auch hier ein Status-Attribut eingeführt werden, aus dem die Provenanz des Werts hervorgeht.

- Aus Gründen der Analogie wäre auch die Angabe des (biologischen) Geschlechts an dieser Stelle angebracht; die auftretenden Unsicherheiten infolge inkohärenter Meldungen der Versicherungsträger entsprechen jenen bei den Geburtsjahren und Sterbedaten und können ebenfalls über ein eigenes Status-Attribut, dessen Wertausprägung im Ladevorgang bzw. in der Phase der Datenvalidierung bestimmt wird, dokumentiert werden. Für nicht bestimmbare Wertausprägungen wäre ebenfalls ein eindeutiger Fehlwert-Kennzeichner in den Wertebereich **[gen]** aufzunehmen.

Zusammenfassend würde das resultierende Tabellenschema für **gapdrg.personen** also folgende Attribute (Felder) samt zugeordneten Wertebereichen in **(P)** umfassen:

- **P_ID** (Primärschlüssel): intern generierter Autowert
- GEBURT_JAHR: **temp @ jahr**
- **_?_GEBURT_JAHR**: **status**
- TOD_DATUM: **temp @ tag**
- **_?_TOD_DATUM**: **status**
- GESCHLECHT: **gen**
- **_?_GESCHLECHT**: **status**

Der (hierarchisch als 2-Steller organisierte) Wertebereich **[status]** könnte dazu Wertausprägungen etwa folgender Art zusammenfassen, wobei die erste Ebene der Hierarchie sich auf die *Meldequellen* (d.h. SV-Träger) und die zweite Ebene sich auf die *Zeitachse* bezieht. Die Bedeutung der Codes auf der 2. Ebene der Codierung hängt dabei vom Code der ersten Ebene ab. Die zeitliche Dimension geht „ontologisch“ von einer Unterscheidung in (nahtlos) aufeinanderfolgende Zeitabschnitte aus, die entweder (i) den Meldeperioden oder in den Meldungen genannten (ii) Referenzperioden (‚VON‘ – ‚BIS‘; siehe dazu gleich unten) entsprechen. Dabei bedeuten in Tabelle 53 (nächste Seite):

- *aktuell* die Meldewerte zum gegenwärtigen Zeitpunkt (Realzeit) oder der letzten verfügbaren Meldeperiode;
- *vorangehend* die Meldewerte der der letzten verfügbaren Meldeperiode vorangehenden Meldeperiode oder der der letzten Referenzperiode unmittelbar vorangehenden Referenzperiode;
- *früher* alle Meldewerte der der vorangehenden Meldeperiode oder Referenzperiode vorangehenden Zeiträume;
- *jüngst* die Menge der jeweils am wenigsten weit zurückliegenden Meldewerte der einzelnen Meldequellen, sowie
- *ältest* die Menge der jeweils am weitesten zurückliegenden Meldewerte der einzelnen Meldequellen.

So bedeutet etwa ein Statuswert **41**, dass die Wertausprägung, auf die der Status Bezug nimmt, als *der kleinste unter mehreren sonst gleichhäufig vorkommenden aktuellen Meldewerten aller Meldequellen* bestimmt wurde (was impliziert, dass das gegenständliche Merkmal einen geordneten Wertebereich zugewiesen hat).

Bezüglich *imputierter* Wertausprägungen ist im Sinne einer guten Dokumentation anzustreben, die zur Anwendungen kommenden Berechnungsverfahren (ebenefalls) auf der zweiten Hierarchie-Ebene der Codierung zu registrieren.

Tabelle 53: Modell der Wertebereiche für das Paramerkmal ‚status‘

Quelldimension		Zeitdimension	
Code	Bedeutung	Code	Bedeutung
0	Meldewert, eindeutig	0	(immer konstant)
1	<i>häufigster</i> Meldewert	0	aus <i>allen</i> Meldewerten
		1	aus den <i>aktuellen</i> Meldewerten
		2	aus den <i>vorangehenden</i> Meldewerten
		3	aus allen <i>früheren</i> Meldewerten
		4	aus den <i>jüngsten</i> Meldewerten
5	aus den <i>ältesten</i> Meldewerten		
2	<i>kleinster</i> Meldewert	wie vorhin	wie vorhin
3	<i>größter</i> Meldewert	wie vorhin	wie vorhin
4	<i>häufigst-kleinsten</i> Meldewert	wie vorhin	wie vorhin
5	<i>häufigst-größter</i> Meldewert	wie vorhin	wie vorhin
9	<i>berechneter</i> Wert (Imputation)	0	keine Verfahrensangabe (weitere Codes je nach registriertem Verfahren)

Neben den im Prinzip statischen Merkmalen der Entität (**P**), die – im Ladevorgang – mehr oder weniger eindeutig aus den Quelldaten ableitbar sind, ist eine Reihe dynamischer Merkmale zu berücksichtigen. Dies gibt Anlass, jeder **P**-Instanz eine *Historie* zuzuordnen. Diese Dynamik spiegelt im Übrigen der Meldeprozess wider, indem die verschiedenen Sozialversicherungsträger zu den einzelnen Versicherten (bzw. Mitversicherten) aktualisierte Datensätze übermitteln, mitunter in bis zu monatlicher Periodizität (Analoges dürfte für andere Datenquellen gelten, insofern sie Daten auf Personen-Ebene einspielen). Auch diese zeitbehafteten Daten sind in GAP-DRG zu konsolidieren. Der naheliegende Ort zur Darstellung dieser konsolidierten Datenbestände ist die Beziehung(-sentität) (<**PS**>) als „Verbindung“ von Personen und Sozialversicherungsträgern. Aus der Sicht der **P**-Instanzen entstehen somit „Personen-Historien“ (vulgo *Versicherungsverläufe*, die an sich für gesetzlich sozialversicherte Personen auch in der ZPV abgebildet sind), die sich aus zeitmärkierten (gültig ‚VON‘ – ‚BIS‘) und von den SV-Trägern gelieferten Datensätzen zusammensetzen.

Die nachfolgende Darstellung geht von der Annahme aus, dass zur Modellierung kein spezifisches Datenbanksystem zur Unterstützung temporaler Daten zum Einsatz kommt, sondern die temporale Semantik in der üblichen relationalen Struktur – unter Zuhilfenahme weiterer Einschränkungen (constraints) – ausgedrückt wird.

Merkmale von **P**-Instanzen mit zeitlicher Dynamik sind insb. der *Wohnsitz*, eine gewährte *Rezeptgebührenbefreiung* sowie die *Versicherungskategorie*, wobei letztere ggf. zwischen den einzelnen SV-Trägern durchaus konsistent auch zur *gleichen Zeit* variieren kann (unterschiedliche Versicherungskategorien bei gleichzeitiger Mehrfachversicherung bei verschiedenen SV-Trägern). Wie aus dem Tabellenschema **gapdrg.personendaten** ersichtlich, umfassen die gemeldeten Datensätze zusätzlich auch noch Felder wie ‚GESL‘ und ‚GEBJAHR‘ – obwohl dies im Grunde redundant ist und letztlich nur (noch) einer Prüfung der Datenqualität der Meldedaten dienen kann. Wesentlich sind die Datum-Felder ‚VON‘ und ‚BIS‘, durch die der Gültigkeitszeitraum der jeweils gemeldeten Merkmalskombination bestimmt ist. Das Feld ‚VNUMV‘ schließlich codiert implizit die Information, ob sich der Datensatz auf eine versicherte oder mitversicherte Person bzw. **P**-Instanz bezieht; auch dieser Umstand ist naturgemäß veränderlich. Hinsichtlich der Konsistenzbedingungen stellt die Mitversicherung allerdings einen Sonderfall dar, da zwar eine **P**-Instanz gleichzeitig bei auch mehreren anderen **P**-Instanzen mitversichert sein kann, alle diese Mitversicherungsverhältnisse aber *zeitgleich* beim Übergang in den Zustand eines „eigenen“ (Haupt-)Versicherungsverhältnisses enden müssen.

Die Konsolidierungslogik hängt stark vom Meldesystem der SV-Träger (bzw. allgemeiner: der einbezogenen Datenquellen) ab; im Weiteren wird hier diesbzgl. unterstellt, dass ein SV-Träger *nur dann* einen Datensatz einer bei ihm versicherten **P**-Instanz meldet, wenn eine *Veränderung* der Quell(-stamm-)daten erfolgt ist, d.h. also, die Änderung einer (jedenfalls quellseitigen) Wertausprägung in mindestens einem Meldeattribut. Infolge der unterschiedlichen Semantik scheint eine Aufspaltung der verschiedenen Merkmale hinsichtlich ihrer Abhängigkeit von einzelnen SV-Trägern angemessen.

Obwohl verschiedene SV-Träger (bzw. Datenquellen) Werte zu den Merkmalen Wohnsitz (BEZRIK_ID) und Rezeptgebührenbefreiung (REZGEB) melden, kann es zu jedem Zeitpunkt für eine **P**-Instanz jeweils nur eine einzige richtige Wertausprägung geben (auch wenn diese nicht immer eindeutig eruierbar sein mag). Dies legt nahe, dafür eine Satellitenentität zu (**P**) anzulegen, in der diese Merkmale – wiederum ergänzt um Status-Attribute zur Dokumentation der Provenanz aus dem Ladevorgang bzw. der Phase der Datenvalidierung analog zu Merkmalen wie ‚GESCHLECHT‘ (siehe vorhin) – zusammengefasst sind; sei diese Satellitenentität nachfolgend durch (**Pt**) symbolisiert. Den Primärschlüssel von (**Pt**) bildet **P_ID** gemeinsam mit dem Merkmal **GÜLTIG_ab** zur (einseitigen) Eingrenzung des Gültigkeitszeitraums – der *Referenzperiode* – der jeweiligen **Pt**-Instanz. Da es sich hier um *diachrone* Merkmale handelt (also solche, die jeweils nur eine einzige Wertausprägung zu einer Zeit annehmen können), kann die Angabe eines Merkmals ‚GÜLTIG_bis‘ entfallen; werden die Instanzen nach **GÜLTIG_ab** aufsteigend geordnet, ergibt sich das Ende einer Referenzperiode unmittelbar aus dem Beginn der darauffolgenden. Man beachte, dass bei Veränderung der Wertausprägung auch nur *eines* der diachronen Merkmale bereits eine neue (**Pt**)-Instanz resultiert. Zusammenfassend ergibt sich etwa folgendes (**Pt**)-Tabellenschema:

- **P_ID** (Fremdschlüssel zu **(P)**)
- **GÜLTIG_ab: temp @ tag**
- **WOHNBEZIRK: reg @ bezirk**
- **_?_WOHNBEZIRK: status**
- **RG_BEFREIUNG: 0/1**
- **_?_RG_BEFREIUNG: status**

Merkmale, die sowohl von der Zeit als auch vom SV-Träger abhängen, sind am besten *direkt* der (Beziehungs-) Entität (**<PS>**) – oder einfacher (**(PS)**), wenn unterstellt wird, das die SV-Träger tatsächlich wie vorgeschlagen als Wertebereich [**svt**] modelliert werden – zuzuordnen. Im gegenständlichen Fall dürfte das einzige in Frage kommende Merkmal die Versicherungskategorie ‚VERSKAT‘ sein, wenn einmal von der Möglichkeit abgesehen wird, auch *Mitversicherungsverhältnisse* hier mit zu berücksichtigen (und so die Beziehung **<PP>** gänzlich obsolet zu machen).

- **P_ID** (Fremdschlüssel zu **(P)**)
- **S_ID: svt**
- **GÜLTIG_von: temp @ tag**
- **GÜLTIG_bis: temp @ tag**
- **V_KATEGORIE: svkat**

P-Instanzen, zu denen keine korrespondierende (**(PS)**-Instanz zu einem bestimmten SV-Träger vorliegt, sind entsprechend auch nicht bei diesem SV-Träger versichert; während ein solches Versicherungsverhältnis aber besteht, muss dazu auch eine **V_KATEGORIE** gegeben sein (jedenfalls dann, wenn die Wertausprägung des Schlüsselmerkmals **S_ID** nicht den Sonderfall ‚vertragsfrei‘ betrifft). Das Merkmal ‚**V_KATEGORIE**‘ ist diesfalls *synchron* modelliert, d.h. überlappende Referenzperioden sind zulässig (sogar für idente Kombinationen von **P_ID** und **S_ID**); sollte eine diachrone Variante bevorzugt sein, wäre das Merkmal **GÜLTIG_bis** aus dem Schlüssel zu entfernen und nur zu nutzen, wenn ein Versicherungsverhältnis ohne Änderung der Versicherungskategorie endet. Insofern die Wertausprägung des (Schlüssel-) Merkmals **GÜLTIG_bis** nicht bekannt ist, wäre auch hier wiederum ein spezifischer Kennwert (z.B. ‚39.19.9999‘ in der Bedeutung von „zu diesem Zeitpunkt“) einzusetzen.

Wie bereits angedeutet, können Mitversicherungsverhältnisse in (**(PS)**) sehr einfach mitmodelliert werden, indem z.B. einfach ein weiteres Merkmal angefügt wird:

- **MITVERSICHERT_BEI (P_ID/Fremdschlüssel zu (P) oder NULL)**

Der SV-Träger des Mitversicherten kann sich aus Gründen der Konsistenz (im angeführten Gültigkeitszeitraum) allerdings nur auf ein Element aus der Menge jener SV-Träger beschränken, bei denen gleichzeitig ein Versicherungsverhältnis zum Hauptversicherten besteht. Auch der Fall einer Mehrfach-Mitversicherung bei einer mehrfach-versicherten **P**-Instanz – wohl ein theoretischer Sonderfall – kann in dieser Struktur abgebildet werden. Diese – somit recht elegante – Variante der Modellierung von Mitversicherungsverhältnissen ersetzt eine sonst eigens einzurichtende Beziehung **<PP>**.

Ehe bei zeitabhängigen Meldeattributen eine die Meldequellen übergreifende Konsolidierung der Wertausprägungen vorgenommen werden kann, ist als Vorbereitungsschritt eine *Angleichung* der Referenzperioden (d.h., der gemeldeten Gültigkeitszeiträume) herbeizuführen. Diese erfolgt über eine **Partitionierung** überlappender oder einander inkludierender Referenzperioden mit folgenden Fallunterscheidungen, wobei hier – ohne Beschränkung der Allgemeinheit – von lediglich zwei nicht-identen Referenzperioden (t'_0 t'_1) bzw. ($t^{}_0$ $t^{}_1$) (mit $t'_0 < t'_1$ bzw. $t^{}_0 \leq t^{}_1$ sowie $t'_0 \leq t^{}_0$) ausgegangen wird:

- *Überlappung*: $t'_0 < t^{}_0 < t'_1 < t^{}_1$, ergebend eine Partitionierung in (t'_0 $t^{}_0$), ($t^{}_0$ t'_1) und (t'_1 $t^{}_1$);
- *Inklusion*: $t'_0 < t^{}_1 \leq t'_1$, ergebend eine Partitionierung in die – bereits bestehende – Periode ($t^{}_0$ $t^{}_1$) sowie, falls $t^{}_1 < t'_1$, ($t^{}_1$ t'_1) und, falls $t'_0 < t^{}_0$, (t'_0 $t^{}_0$).

Die Wertausprägungen sind aus den ursprünglichen Datensätzen für die entstehenden Partitionierungen entsprechend zu replizieren; die Konsolidierung der Merkmale erfolgt sodann *innerhalb* der einzelnen Partitionen. Insofern sich herausstellt, dass unmittelbar aufeinander folgende Partitionen (bzw. Referenzperioden) – d.h., $t'_1 = t^{}_0$ – hinsichtlich aller Merkmale *ident*e Wertausprägungen aufweisen, kann schließlich eine *Verschmelzung* (t'_0 $t^{}_1$) vorgenommen werden (bzw., falls diese durchgehende Wertidentität bereits vorab festgestellt wird, kann die Partitionierung auch gleich unterbleiben).

Je nach Lage der Datenmeldungen durch die SV-Träger können sich auch Leerperioden ergeben (also Zeiträume ohne Vorliegen einer Datenmeldung). Dies ist inhaltlich äquivalent zu „leeren“ Meldeperioden, d.h. Datensätzen (Instanzen), in denen alle Merkmale außerhalb des Primärschlüssels mit systemseitigen NULL-Ausprägungen belegt sind.

4 Prozessorganisation

In der vorliegenden Form des technischen GAP-DRG Datenmodells stellt die Tabelle **gapdrg.personendaten** ihrem Wesen nach eine „Durchgangsstruktur“ dar, in der zunächst die einlangenden Meldedaten ohne weitere semantisch geleitete Aufbereitung angesammelt werden. Vor allem im Hinblick auf nicht oder kaum antizipierbare analytische Bedarfe macht eine Vorratsspeicherung dieser Art durchaus Sinn, um auch *ex post* noch unterschiedliche Imputationen oder Datenanpassungen („edits“) vornehmen zu können. Unbeschadet dieses subsidiären Nutzens ist aber eine Aufbereitung der Meldedaten nach den oben dargestellten Gesichtspunkten in die Strukturen **(P)**, **(Pt)** und **(PS)** (ggf. auch **<PP>**) vorzuziehen. Diese Argumentation lässt sich mutatis mutandis auch auf die anderen, hier nicht weiter betrachteten Komponenten von GAP-DRG übertragen.

Epilog: Zur Ontologie der Inanspruchnahme

Das oben skizzierte Datenmodell mit den Entitäten **<A>**, **<H>**, **<L>**, **<K>** entspricht weitgehend der bestehenden GAP-DRG Formulierung; ggf. sind aber auch in dieser Hinsicht radikalere Modellierungsvarianten zielführender. Ohne diesen Punkt hier tiefer auszuführen, sei lediglich angedeutet, dass sich eine verallgemeinerte „Inanspruchnahme“ als übergeordnete Modell-Entität zwischen Patienten (**<P>**) und

Medizin-Dienstleistern (<M>) anbietet, die einerseits die grundlegenden Typen *medizinischer Inanspruchnahme* repräsentiert, also etwa

- <In> ... niedergelassener Bereich
- <Ia> ... ambulanter Bereich
- <Is> ... stationärer Bereich

und andererseits die *inhaltliche Form* der Inanspruchnahme formal charakterisiert, also etwa durch eine Klassifikation in „Ereignisse“ (vgl. oben in Bezug auf Temporalität), darunter z.B.:

- Intervention
- Krank-/Gesundschreibung
- Verordnung eines Heilmittels
- Überweisung
- Aufnahme
- Verlegung
- Entlassung

Diese Ereignis-Typen können den Typen der Inanspruchnahme – matrixweise – gegenübergestellt werden (wobei vielleicht nicht alle Kombinationen logisch bzw. rechtlich sinnvoll oder möglich sind); jedenfalls wären die Ereignisse jeweils ausgestattet mit (i) einem „Anlass“ (Grund, Diagnose(n), ...) bzw. (ii) einer „Folge“, z.B. einem Rezept über ein Heilmittel oder einer Leistung aus einem entsprechenden Leistungskatalog. Insb. die Ereigniskategorie „Intervention“ würde wohl eine tiefergehende Systematisierung nach sich ziehen, z.B. unterscheidend zwischen Einzelintervention, Folgeintervention (einer Interventionsserie nach Therapieplan o.ä.), Vorbereitung/Erstgespräch, Nachsorge/Kontrolle und dgl. mehr. Ein Abgleich mit konzeptuellen Überlegungen medizinischer Aufzeichnungsstandards (wie EHR usw.) liegt diesbzgl. nahe.

Dieser Ansatz ist mit obigen Überlegungen (und auch GAP-DRG wie vorliegend) kompatibel, wenngleich natürlich bei der gegenwärtigen Datenversorgung viele Strukturelemente inhaltlich (noch) nicht befüllbar sind und teilweise die (zeitliche) Granularität der Datenbestände die semantische Ausdrucksstärke der Ereignisstruktur nicht ausspielen kann.

Dennoch wäre begründet davon auszugehen, dass eine derartige patienten- und gleichzeitig auch dienstleister-zentrierte Repräsentation der Inanspruchnahmen die angestrebte Multidimensionalität des epidemiologischen Auswertungsgeschehens gut unterstützt. Die Frage der Zusammenführung einzelner Ereignisse zu inhaltlich motivierten Ereignisketten („medizinischen Episoden“) – wie sie derzeit partiell im Bereich der stationären Aufenthalte implizit dargestellt werden – wäre damit in flexibler Weise bereits der Auswertungslogik zugeordnet. ♦