

Ergebnisbericht

(gemäß Nr. 14.1 ANBest-IF)

Konsortialführung:	Philipps-Universität Marburg ausführende Stelle: Prof. Dr. med. Norbert Donner-Banzhoff, Fachbereich Medizin, Abteilung für Allgemeinmedizin, Präventive und Rehabilitative Medizin
Förderkennzeichen:	01VSF17016
Akronym:	WirtMed
Projekttitel:	Die Verordnung von Arzneimitteln: Prüfung und Steuerung von Wirtschaftlichkeit und Qualität (WirtMed-Studie)
Autoren:	Prof. Norbert Donner-Banzhoff in Zusammenarbeit mit Dr. Ronja Flemming, Prof. Leonie Sundmacher (TP A), Dr. Dominika Franiel, Maïke Schulz (TP B, E), Nikoletta Lippert, Dr. Maria Sebastiao, Prof. Thomas Kühlein (TP C), Julia Gollnick (TP D) und Julia Muth
Förderzeitraum:	1. April 2018 – 31. Dezember 2021

Inhaltsverzeichnis

I.....	Abkürzungsverzeichnis.....	2
II.....	Abbildungsverzeichnis.....	3
III.....	Tabellenverzeichnis.....	4
1.....	Zusammenfassung.....	7
2.....	Beteiligte Projektpartner.....	9
3.....	Projektziele.....	11
4.....	Projektdurchführung.....	13
5.....	Methodik.....	15
6.....	Projektergebnisse.....	32
7.....	Diskussion der Projektergebnisse.....	66
8.....	Verwendung der Ergebnisse nach Ende der Förderung.....	70
9.....	Erfolgte bzw. geplante Veröffentlichungen.....	72
10.....	Literaturverzeichnis.....	76
11.....	Anhang.....	77
12.....	Anlagen.....	78

I. Abkürzungsverzeichnis

ACE	Angiotensine converting enzyme
ADHS	Attention deficit hyperactivity disorder
AM	Arzneimittel
AMNOG	Arzneimittelmarktneuordnungsgesetz
AOK	Allgemeine Ortskrankenkasse
AT1	Angiotensin1 receptor blocker
ATC	Anatomical Therapeutic Chemical Classification System
AVD	Arzneiverordnungs-Daten
BBSR	Bundesinstitut für Bau-, Stadt- und Raumforschung
BS	Betriebsstätte
BSNR	Betriebsstätten-Nummer
BTM	Betäubungsmittel-Gesetz
DDD	Defined daily doses
DOAK	Direkte orale Antikoagulanzen (Syn. NOAK)
EMA	European Medicines Agency
EPO	Erythropoetin
GAM	Generalized additive model
GKV	Gesetzliche Krankenversicherung
GLM	Generalized linear model
GnRH	Gonadotropine releasing hormone
HR	Hazard ratio
ICD	International classification of diseases
INKAR	Indikatoren und Karten zur Raum- und Stadtentwicklung
KBV	Kassenärztliche Bundesvereinigung
KHK	Koronare Herzkrankheit
LANR	Lebenslange Arzt-Nummer
LE	Lungenembolie
LMU	Ludwigs-Maximilians-Universität (München)
MS	Multiple Sklerose
NOAK	Neue orale Antikoagulanzen (Syn. DOAK)
nvVHF	Nicht-valvuläres Vorhofflimmern

O/E	Observed/expected
OR	Odds ratio
PPI	Protonenpumpen-Ihibitoren
SGB	Sozialgesetzbuch
TNF	Tumor-Nekrose-Faktor
SHAP	SHaple Additive exPlanation
S/V	Acubitril-Valsartan
TEP	Total-Endoprothese
TUM	Technische Universität München
TVT	Tiefe Beinvenen-Thrombose
VDA	Verordnungsdaten ambulant
WidO	Wissenschaftlicher Dienst der Ortskrankenkassen
WSV	Wirkstoff-Vereinbarung
ZE	Zielerreicher
Zi	Zentralinstitut für die Kassenärztliche Versorgung (Berlin)
ZV	Zielverfehler

II. Abbildungsverzeichnis

Abbildung 1: Schema der Analyseprozedur	32
Abbildung 2: Konzentrationskurve Verordnungsmenge Targin	35
Abbildung 3: Soziodemographische Angaben zu den Ärzt*innen der Einzelinterviews und Fokusgruppen.....	40
Abbildung 4: O/E-Indikator der durchschnittlichen Verordnungskosten im Quartal 4/2016	48
Abbildung 5: SHAP summary plot der durchschnittlichen Verordnungskosten im Quartal 4/2016	49
Abbildung 6: O/E-Indikator des Verordnungsanteils von Generika am Gesamtmarkt im Quartal 4/2016	51
Abbildung 7: SHAP summary plot des Verordnungsanteils von Generika am Gesamtmarkt im Quartal 4/2016	52
Abbildung 8: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 1/2016	54
Abbildung 9: SHAP summary plot der durchschnittlich verordneten PPI-Mengen	

im Quartal 2/2016	55
Abbildung 10: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 3/2016	56
Abbildung 11: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 4/2016	57
Abbildung 12: SHAP summary plot des Verordnungsanteils von DOAKs an allen Antikoagulantien im Quartal 1/2016	59
Abbildung 13: SHAP summary plot des Verordnungsanteils von DOAKs an allen Antikoagulantien im Quartal 2/2016	60
Abbildung 14: SHAP summary plot des Verordnungsanteils von DOAKs an allen Antikoagulantien im Quartal 3/2016	61
Abbildung 15: SHAP summary plot des Verordnungsanteils von DOAKs an allen Antikoagulantien im Quartal 4/2016	62
Abbildung 16: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 1/2016	64
Abbildung 17: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 2/2016	64
Abbildung 18: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 3/2016	65
Abbildung 19: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 4/2016	66

III. Tabellenverzeichnis

Tabelle 1: Monetäre Indikatoren	20
Tabelle 2: Qualitative Indikatoren	20
Tabelle 3: Indikatoren mit Bezug zur Wirkstoffvereinbarung (WSV-Ziele)	22
Tabelle 4: Indikatoren des Strukturbruchmodells	24
Tabelle 5: deskriptiv untersuchte Charakteristika	27
Tabelle 6: Untersuchte Aspekte der Verordnung	28
Tabelle 7: Indikatoren der Basisanalyse und der weiterführenden Analyse	29
Tabelle 8: Confounder der Basisanalyse	30
Tabelle 9: Confounder der weiterführenden Analysen	30
Tabelle 10: Vorliegen eines Strukturbruchs in der KV Bayerns und KV Hessen	38
Tabelle 11: Potentiell schwierige/nachteilige Systemeigenschaften aus der Sicht	

der befragten Zielverfehlenden	44
Tabelle 12: Aufteilung des Hausarzt-Betriebsstätten in der Kategorie der Auffälligkeiten für der Indikator durchschnittliche Kosten je Verordnung in der für das Quartal 4/2016	47
Tabelle 13: Haupteinflussfaktoren auf die durchschnittlichen Verordnungskosten im Quartal 4/2016	50
Tabelle 14: Aufteilung der Hausarzt-Betriebsstätten in die Kategorien der Auffälligkeiten für den Indikator Verordnungsanteil von Generika am Gesamtmarkt für das Quartal 4/2016	50
Tabelle 15: Haupteinflussfaktoren auf den Verordnungsanteil von Generika am Gesamtmarkt im Quartal 4/2016	53
Tabelle 16: Prozentuale Aufteilung der Betriebsstätten (Praxen) in jeweils eine der vier Kategorien für den Indikator verordnete PPI-Menge für die Quartal 1/2016 bis 4/2016 und das Gesamtjahr 2016	54
Tabelle 17: Haupteinflussfaktoren auf die jeweiligen durchschnittlich verordneten PPI-Mengen im Quartal 1/2016	55
Tabelle 18: Haupteinflussfaktoren auf die durchschnittlich verordneten PPI-Mengen im Quartal 2/2016	55
Tabelle 19: Haupteinflussfaktoren auf die durchschnittlich verordneten PPI-Mengen Quartal 3/2016	56
Tabelle 20: Haupteinflussfaktoren auf die durchschnittlich verordneten PPI-Mengen im Quartal 4/2016	57
Tabelle 21: Prozentuale Aufteilung der Betriebsstätten (Praxen) in jeweils eine den vier Kategorien für den Indikator des Verordnungsanteils von DOAKs an allen Antikoagulantien für die Quartale 1/2016 bis 4/2016 und das Gesamtjahr 2016	58
Tabelle 22: Haupteinflussfaktoren auf den Verordnungsanteil von DOAKs an allen Antikoagulantien im Quartal 1/2016.....	59
Tabelle 23: Haupteinflussfaktoren auf den Verordnungsanteil von DOAKs an allen Antikoagulantien im Quartal 2/2016	60
Tabelle 24: Haupteinflussfaktoren auf den Verordnungsanteil von DOAKs an allen Antikoagulantien im Quartal 3/2016	61
Tabelle 25: Haupteinflussfaktoren auf den Verordnungsanteil von DOAKs an allen Antikoagulantien im Quartal 4/2016	62
Tabelle 26: Prozentuale Aufteilung der Betriebsstätten (Praxen) in jeweils eine	

der vier Kategorien für den Indikator des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker für die Quartale 1/2016 bis 4/2016 und das Gesamtjahr 2016	63
Tabelle 27: Haupteinflussfaktoren auf den Verordnungsanteil von generischen DDDs im Bereich der Lipidsenker im Quartal 1/2016	64
Tabelle 28: Haupteinflussfaktoren auf den Verordnungsanteil von generischen DDDs im Bereich der Lipidsenker im Quartal 2/2016	65
Tabelle 29: Haupteinflussfaktoren auf den Verordnungsanteil von generischen DDDs im Bereich der Lipidsenker im Quartal 3/2016	65
Tabelle 30: Haupteinflussfaktoren auf den Verordnungsanteil von generischen DDDs im Bereich der Lipidsenker im Quartal 4/2016	66

1. Zusammenfassung

Hintergrund: Die Verordnung von Arzneimitteln unterliegt dem Wirtschaftlichkeitsgebot des §12 SGB V. Allerdings ist die herkömmliche Wirtschaftlichkeitsprüfung (WP) für den ambulanten Sektor (Richtgrößen-Überschreitung nach §106 SGB V) in die Kritik geraten. Der Gesetzgeber hat die Vertragspartner deshalb verpflichtet, neue Verfahren der Wirtschaftlichkeitsprüfung zu entwickeln (106b SGB V). Die bekannteste Alternative zur herkömmlichen Prüfung ist die bayrische Wirkstoffvereinbarung (WSV). - Unter einem Profiling verstehen wir die Identifizierung von Leistungserbringern, welche in nicht mehr akzeptablem Maße von einer Norm abweichen. Dafür sind international sehr differenzierte Methoden entwickelt worden. – Jegliche WP setzt bisher eine unabhängige Entscheidung der Verordner voraus. Es ist jedoch fraglich, ob dies für einzelne Verordner zutrifft, oder ob nicht eine Beeinflussung z.B. durch Krankenhäuser (Entlassmedikation) oder andere Praxen stattfindet. Im Folgenden sprechen wir von „beobachtungswürdigen Medikamenten“, deren Verordnung in Bezug auf Kosten oder Qualität kritisch ist bzw. sein kann.

Methodik: In 5 Teilprojekten (TP) wurden Forschungsfragen in Zusammenhang mit der Wirtschaftlichkeit der Arzneimittel-Verordnung sowie deren Prüfung und Steuerung untersucht. In TP A wurde die Verbreitung von neuen Medikamenten an Hand von Verordnungs- und Behandlungsdaten von AOK Bayern, AOK Hessen und AOK Plus erfasst. Verordnungsdaten des Zi dienten dem Vergleich der bayrischen WSV mit Hessen als einem Beispiel der Richtgrößenprüfung (TP B). Mit teilstrukturieren Interviews (tIV) und Fokusgruppen wurden die Einschätzungen und Bewertungen der WSV durch niedergelassene ÄrztInnen eingeholt (TP C). In TP D wurden Praxen untersucht, die dauerhaft die Verordnungsziele der WSV nicht erreichten (Verordnungsdaten; tIV). Schließlich konnten wir neue Methoden des Profiling erstmalig auf Verordnungsdaten aus Deutschland (Zi) anwenden (TP E).

Ergebnisse: Die Befunde aus TP A legen nahe, dass durch PatientInnen und andere Praxen die Verordnung neuer Medikamente beeinflusst wird. Nach einer stationären Behandlung tragen PatientInnen eine Verordnung in den ambulanten Sektor, was allerdings nur indirekt erfasst werden konnte. - Durch die Einführung der WSV mit ihrer Orientierung an Leitsubstanzen und Generikaquoten kam es nicht zu einer Steigerung der Kosten insgesamt. Für bewusst gesteuerten Bereiche lassen sich positive Effekte nachweisen. Allerdings sind für TP B methodische Einschränkungen zu bedenken. - Die WSV wird von den befragten Vertragsärzten mehrheitlich begrüßt (TP C). Zeitnahes Feedback bietet eine nützliche Orientierung, die Sorge vor einem Regress hat offenbar abgenommen. Allerdings wurde der Wunsch nach verbesserter Information und Handlungsleitfäden geäußert. Kritisiert wurde die bei rabattierten Arzneimitteln fehlende Transparenz für die Verordner. – Unter den Zielverfehlern waren vor allem Fachgruppen mit einem geringen Verordnungsspektrum vertreten (z.B. Orthopäden mit Osteoporose-Mitteln). Die Charakteristika „männlicher Arzt“, „städtische Praxislage“, „Fallzahl pro Quartal max. 600“ und „keine Zusatzbezeichnungen“ waren häufiger vertreten als in der Vergleichsgruppe der Zielerreicher (TP D). In TP E schließlich erwiesen sich moderne Methoden des Profiling geeignet, Ausreißer bei Medikamenten-Verordnungen zu identifizieren.

Diskussion: Die individuelle Verordnung wird durch ambulante und ambulant/stationäre Netzwerke beeinflusst. Die bayrische WSV hat offenbar steuernde Effekte, ohne dass die Gesamtkosten gestiegen wären. Auch aus der Perspektive der Vertragsärzte überwiegt eine positive Einschätzung. Die in diesem Projekt eingesetzten innovativen Methoden werden den verantwortlichen Körperschaften künftig in einer „Toolbox“ zugänglich gemacht (vorbehaltlich einer Entscheidung des Innovationsausschusses beim G-BA).

Akronym: WirtMed

Förderkennzeichen: 01VSF17016

Während die Verpflichtungen der KVen und der Krankenkassen zur Übermittlung von Verordnungs- und Leistungsdaten gesetzlich geregelt sind (§§ 296, 297, 300 SGB V), fehlt bisher eine Verpflichtung der Krankenhäuser, die bei der Entlassung nach stationärer Behandlung erforderliche Dauermedikation der Krankenkasse oder einer anderen Prüf-Einrichtung (z.B. dem MDK) mitzuteilen. Damit kann ein wesentlicher Einflussfaktor nicht erfasst werden; eine entsprechende gesetzliche Regelung erscheint uns dringend erforderlich.

2. Beteiligte Projektpartner

Philipps-Universität Marburg (PUM) Abteilung für Allgemeinmedizin, Präventive und Rehabilitative Medizin 35043 Marburg Konsortialführung Prof. Dr. med. Norbert Donner-Banzhoff, MHSc
Friedrich-Alexander-Universität Erlangen-Nürnberg UME) Universitätsklinikum Erlangen Allgemeinmedizinisches Institut 91054 Erlangen Prof. Dr. Thomas Kühlein
Technische Universität München (TUM) Lehrstuhl für Gesundheitsökonomie Fakultät für Sport-und Gesundheitswissenschaft 80992 München Prof. Dr. Leonie Sundmacher
Zentralinstitut für die kassenärztliche Versorgung in der Bundesrepublik Deutschland 10587 Berlin (Zi) Frau Maike Schulz
Politecnico di Milano (MIL) MOX - Modeling and Scientific Computing Dipartimento di Matematica "F. Brioschi" Mailand, Italien Prof. Anna-Maria Paganoni
Kassenärztliche Vereinigung Bayerns (KVB) 80687 München Dr. Wolfgang Krombholz, Dr. Franziska Hörbrand, Dr. Peter Killian
Kassenärztliche Vereinigung Hessen (KVH) 60486 Frankfurt Dr. Wolfgang LangHeinrich
AOK Bayern - Die Gesundheitskasse (AOK-B) 81739 München Herr Wolfgang Adolf
AOK – Die Gesundheitskasse in Hessen (AOK-H) 65760 Eschborn Frau Heike Plachetka
AOK Plus – Die Gesundheitskasse für Sachsen und Thüringen (AOK Plus) 01067 Dresden

Herr Andreas Mohr, Dr. Ulf Maywald

Aufgabe der TUM war die Auswertung der Verordnungsdaten unter dem Gesichtspunkt der raumzeitlichen Ausbreitung der Verordnungen. [TP A] Verordnungs- und Behandlungsdaten wurden von der AOK-B, der AOK-H und der AOK Plus zur Verfügung gestellt. Auswertung, Interpretation und Publikation wurden von PUM, Zi, KVB und KVH unterstützt.

Für den Vergleich der bayrischen WSV mit Hessen wurden Daten von KVB und AOK-H zur Verfügung gestellt [TP B] und vom Zi ausgewertet. Auswertung, Interpretation und Publikation wurden zudem von PUM, AOK Plus und KVH unterstützt.

Interviews zur Evaluation der bayrischen WSV aus der Sicht der VertragsärztInnen wurden von UME in enger Zusammenarbeit mit PUM geplant und durchgeführt. [TP C].

Die Untersuchung von zielverfehlenden Betriebsstätten oblag der PUM. Der qualitative Studienteil wurde mit enger Unterstützung von UME, der quantitative mit Unterstützung der KVB (u.a. Datenlieferung) durchgeführt [TP D].

Federführend bei der Entwicklung von Profiling-Methoden war MIL in enger Kooperation mit dem Zi (Datenlieferung, Auswertung, Interpretation, Publikation) sowie PUM [TP E].

Im Rahmen regelmäßiger Gesamtprojekt-Treffen wurden sämtliche Partner über den Stand der Teilprojekte informiert, diskutierten Studiendesign, -implementierung, -auswertung sowie die Interpretation und künftige Verwertung der Ergebnisse. Die Aktivitäten im Rahmen der Teilprojekte wurde von PUM koordiniert und unterstützt.

3. Projektziele

3.1 Einleitung und Begründung

Die Verordnung von Arzneimitteln unterliegt dem Wirtschaftlichkeitsgebot des §12 SGB V. Die gesetzliche Verpflichtung von Krankenkassen und Kassenärztlichen Vereinigungen zur Prüfung einer wirtschaftlichen Ordnungsweise nach 106b SGB V wurde bisher überwiegend in Form der sog. Richtgrößenprüfung umgesetzt. Danach wird auf der Basis bisheriger Verordnungen für jede Fachgruppe ein Richtwert in € pro behandeltem Patient festgelegt. Durch Multiplikation mit der Zahl der behandelten Patienten ergibt sich damit für die einzelne Betriebsstätte ein ideelles Gesamtbudget für die auf Kosten der GKV verordneten Medikamente, dessen Überschreitung zu einer Überprüfung auf Praxisbesonderheiten führt. Liegen diese nicht vor, sind Feedback, Information ggf. auch finanzielle Sanktionen (sog. Regress) vorgesehen.

An diesem Verfahren ist kritisiert worden, dass bei den über die gesamte Fachgruppe berechneten Richtgrößen Morbidität und Verordnungstiefe unzureichend berücksichtigt werden, etwa durch eine Benachteiligung von Praxen, die Patienten mit hohem Bedarf an Medikamenten behandeln und diese Verordnungen selbst ausstellen (anstatt dies anderen Praxen zu überlassen). Es besteht unter diesen Bedingungen ein Anreiz, Patienten allein für die Verordnung teurer Medikamente zu überweisen. Schließlich widerspricht die Annahme, Verordner seien autonom, der Tatsache, dass oft mehrere niedergelassene Ärzte gemeinsam Patienten versorgen und einmal initiierte Verordnungen anderer Ärzte (Facharzt, Krankenhaus) weiterführen. Die Preise von Medikamenten, für welche die jeweilige Krankenkasse mit einem Hersteller einen Rabattvertrag abgeschlossen hat, sind den Verordnern nicht bekannt; damit kann das Risiko für eine Richtgrößenüberschreitung nicht mehr abgeschätzt werden. Schließlich wird vermutet, ohne dass sich dies verlässlich nachweisen ließe, dass wegen der Drohung eines Regress Patienten wirksame Behandlungen vorenthalten werden. Vor diesem Hintergrund hat der Gesetzgeber die Vertragspartner verpflichtet, neue Verfahren der Wirtschaftlichkeitsprüfung zu entwickeln (106b SGB V).

Die bekannteste Alternative zur herkömmlichen Prüfung ist die bayrische Wirkstoffvereinbarung (WSV), welche sich nicht mehr an Verordnungskosten, sondern vielmehr an Leitsubstanzen und Generikaquoten orientiert. Wenn eine Verordnung den vorgegebenen Leitsubstanzen entspricht, entfällt die Sorge um die Überschreitung des Richtgrößen-Budgets. Außerdem entfällt die Wirtschaftlichkeit als ein Grund, an andere Praxen zu überweisen. Mit der Orientierung an Leitsubstanzen wird der Gesichtspunkt der Verordnungs-Qualität gegenüber einer rein monetären Orientierung gestärkt. Durch zeitnahes Feedback soll eine an Qualität und Wirtschaftlichkeit orientierte Steuerung erfolgen, im Gegensatz zur bisherigen Prüfung mit der Drohung von Sanktionen.

International sind Instrumente zum sog. Profiling in der medizinischen Versorgung entwickelt worden. (Spiegelhalter et al. 2012) Darunter versteht man die Identifizierung von Einheiten (Ärzten, Krankenhäusern, Praxen usw.), welche in nicht mehr akzeptablem Maße von einer Norm abweichen; dies kann sich auf Prozesse (Verhalten) und auf patientenbezogene Outcomes beziehen. Auch die herkömmliche Wirtschaftlichkeitsprüfung ist also ein „Profiling“ im Sinne dieser Definition.

Ergebnisse der Diffusions- und Implementierungs-Forschung legen nahe, dass sich Innovationen nicht gleichmäßig in einer Gruppe oder einem Netz von Akteuren ausbreiten. (Rogers 2003) Vielmehr sind Zentren, welche ihrerseits ihre Umgebung beeinflussen, von großer Be-

deutung. Bei der Verordnung von Medikamenten sind dies vor allem Krankenhäuser (Entlassmedikation), Spezial- und Schwerpunkt-Praxen. Mit den Methoden der Diffusionsforschung lassen sich diese Prozesse beschreiben. Damit wird die Annahme der bisherigen Wirtschaftlichkeitsprüfung überprüft, dass die betrachteten Einheiten (Betriebsstätten) unabhängig voneinander seien [Teilprojekt A].

Anhand von Verwaltungsdaten haben wir die Auswirkungen der bayerischen WSV untersucht. Hier ist zum einen von Interesse, ob eine Orientierung an Leitsubstanzen und Generikaquoten tatsächlich eine entsprechende Verbesserung erreicht. Zum anderen stellt sich die Frage, ob als unerwünschter Effekt die Verordnungskosten Steigerungen erfahren [Teilprojekt B]. Da eine Steuerung des Verwaltungs-Geschehens auf die Kooperation der Beteiligten, vor allem die der Vertragsärzte, angewiesen ist, sind Einschätzungen, Akzeptanz und Umgangsweisen der Ärzteschaft zu untersuchen [Teilprojekte C].

Jede systematische Steuerung und Prüfung produziert „Zielverfeher“. Diese Gruppe verdient besonderes Interesse, um die Treffsicherheit (Validität) des Steuerungs- und Prüfungssystems zu reflektieren, aber auch um evtl. Maßnahmen zu entwickeln, welche den auffälligen Verordnern eine Verbesserung ermöglichen [Teilprojekt D].

Eine Zusammenstellung von Erkenntnissen und Ressourcen zur Steuerung und Prüfung von Wirtschaftlichkeit und Qualität ist wünschenswert, um den zuständigen Institutionen (KVen, Prüfstellen, Krankenkassen) aktuelle Methoden und Verfahren zugänglich zu machen.

3.2 Projektziele

Untersuchung der räumzeitlichen Ausbreitung der Verordnung von ausgewählten Medikamenten, der gegenseitigen Beeinflussung von Praxis (Betriebsstätten) und des stationären Sektors [Teilprojekt A].

Vergleich der bayrischen WSV mit einem Bundesland mit herkömmlicher Richtgrößenprüfung (Hessen) in Bezug auf Verordnungskosten und -qualität bzw. mit den übrigen Bundesländern. Als empirische Grundlage dienen Verwaltungsdaten [Teilprojekt B].

Evaluation der bayrischen WSV aus der Sicht der VertragsärztInnen. Als empirische Grundlagen dienen teilstrukturierte Interviews und Fokusgruppen [Teilprojekt C]

Charakterisierung von in Bezug auf die WSV zielverfehlenden Betriebsstätten bzw. ÄrztInnen im Vergleich zu Zielerreichern [Teilprojekt D].

Anpassung des internationalen Wissensstandes in Bezug auf das Profiling an die Prüfung von Wirtschaftlichkeit und/oder Qualität der Medikamenten-Verordnung angepasst und entsprechende Erprobung [Teilprojekt E].

4. Projektdurchführung

4.1 Teilprojekt A

Nach Rücksprache mit Konsortialführung und -partnern wurden die zu untersuchenden Medikamente, Fachgruppen, Regionen usw. festgelegt; auf dieser Grundlage erstellten wir eine ausführlichen Datensatzbeschreibung. Entsprechend wurden von AOK Hessen, AOK PLUS, AOK Bayern die erforderlichen Daten an die LMU geliefert. In einem Fall war aufgrund eines fehlerhaften Datenabzugs eine Neulieferung erforderlich.

Mit netzwerkanalytischen Methoden untersuchten wir die Verbreitung des Medikaments Sacubitril/Valsartan. Außerdem demonstrierten wir Patientenpfade, durch welche ein Einfluss auf die Verordnung des Wirkstoffs ausgeübt wird. Diese wurden auch für die DOAKs Rivaroxaban, Apixaban, Edoxaban und Dabigatran untersucht. Durch die Erfassung regionaler Versorgungsunterschiede konnten wir den Einfluss von stationären Behandlungen aufzeigen.

Unsere Analysen wurden so aufbereitet, dass das Instrumentarium für zukünftige Anwender in der Toolbox verfügbar sein wird.

4.2 Teilprojekt B

In Rücksprache mit Konsortialführung und den -partnern wurde ein detaillierter Analyseplan erstellt, der sowohl monetäre als auch qualitative Indikatoren (bezieht sich auf die medizinische Versorgungsqualität) umfasst. Die Einheit der Auswertung (Betriebsstätten) und die untersuchten Zeiträume wurden festgelegt. Damit ergaben sich zwei Kohorten (Hessen, Bayern) von Betriebsstätten / Verordnungen von 2010 bis 2016. Die entsprechenden Daten lagen beim Zi vor und wurden vor ihrer Auswertung umfassend auf ihre Plausibilität geprüft.

Die statistischen Auswertungen wurden im Längs- und Querschnittsvergleich vorgenommen. Hier wurde mit dem Strukturbruchmodell nach Chow (Chow 1960) Neuland betreten.

Das Vorgehen wurde so dargestellt und erläutert, dass es künftig in der Toolbox Nutzern zur Verfügung steht.

4.3 Teilprojekt C

Das auf eine qualitative Methodik ausgerichtete Projekt wurde mit einem ausführlichen Interview-Leitfaden vorbereitet, der mit den Partnern des TP D abgestimmt wurde. Nach Einholen eines Ethikvotums wurden niedergelassene Ärztinnen aus dem weiteren Umland des Universitätsklinikums Erlangen rekrutiert. Mit 20 von ihnen konnten wir Einzelinterviews führen, außerdem fanden 7 Fokusgruppen statt (z.T. online). Die audio-aufgezeichneten Interviews wurden transkribiert und ausgewertet, ein systematischer Code-Baum wurde iterativ entwickelt. Die EDV-gestützte Auswertung fand in Zusammenarbeit mit den Partnern des TP D statt.

Das methodische Instrumentarium dieser Befragung im Rahmen von teilstrukturierten Einzelinterviews und Fokusgruppen wurde für die Toolbox zur künftigen Nutzung dargestellt und erläutert.

4.4 Teilprojekt D

Nach einem positiven Ethikvotum wurde in Absprache mit Konsortialführung und -partnern eine Definition der „Zielverfehlung“ festgelegt und für die Auswertung bzw. Rekrutierung ope-

Akronym: WirtMed

Förderkennzeichen: 01VSF17016

rationalisiert. Mit einem Anschreiben der KV Bayerns wurde diese um ihre Einwilligung an teilstrukturierten Interviews gebeten, 26 Interviews konnten wir durchführen. Die audio-aufgezeichneten Interviews wurden transkribiert und ausgewertet, ein systematischer Code-Baum wurde iterativ entwickelt. Die EDV-gestützte Auswertung fand in Zusammenarbeit mit den Partnern des TP C statt.

Für die deskriptive Charakterisierung der Zielverfeher wurde ein von der KV Bayerns erstellter Datensatz quantitativ ausgewertet.

Das methodische Instrumentarium (qualitativ, quantitativ) wurde für die Toolbox zur künftigen Nutzung dargestellt und erläutert.

4.5 Teilprojekt E

Dieses Teilprojekt wurde als externer Auftrag an das Politecnico di Milano (Italien) vergeben.

5. Methodik

5.1 Teilprojekt A

5.1.1 Auswahl der beobachtungswürdigen Medikamente

Im Teilprojekt A wurde zunächst eine Auswahl von beobachtungswürdigen Medikamenten als Grundlage für die Analyse der Diffusionsmuster erarbeitet. Als Kriterien für ein beobachtungswürdiges Medikament wurden die folgenden definiert:

- kein oder nur ein geringer Zusatznutzen während des Untersuchungszeitraums;
- umstrittene Evidenz der Wirksamkeit oder existierende Kontraindikationen und gravierende Nebenwirkungen;
- hohe gesellschaftliche Kosten durch einen hohen Preis;
- Arzneimittel als positives Beispiel aus der gleichen Wirkstoffgruppe eines beobachtungswürdigen Arzneimittels.

Der Auswahlprozess erfolgte im Rahmen eines mehrstufigen Verfahrens innerhalb eines Expertenpanels, das sich aus Mitgliedern des Konsortiums zusammensetzte. Das Panel bestand aus PharmazeutInnen, MedizinerInnen, ExpertInnen des Erstattungs geschehens der Krankenkassen sowie GesundheitsökonomInnen. In mehreren Telefonkonferenzen und der Möglichkeit von schriftlichem Feedback wurde eine Liste von insgesamt 34 Wirkstoffen erarbeitet, denen entsprechend ihrer Relevanz für die zu untersuchende Fragestellung eine Priorität 1 bis 3, bzw. die Kategorie „*Positivwirkstoff*“ durch das Expertenpanel zugewiesen wurde (siehe Anhang 11.1.1 Wirkstoffliste, enthält Wirkstoffe der Priorität 1, 2 und Positivwirkstoffe).

In den im Teilprojekt A durchgeführten Analysen wurden insgesamt sieben Wirkstoffe untersucht, welche alle die höchste Priorität des Expertenpanels zugesprochen bekamen, bzw. Positivwirkstoffe einer zur Analyse ausgewählten Wirkstoffgruppe darstellten. Die untersuchten Wirkstoffe umfassten Sacubitril/Valsartan (Entresto®) (ATC C09DX04), Pregabalin (Lyrica®) (ATC N03AX16), Oxycodon & Naloxon (Targin®) (ATC N02AA55), sowie die vier sogenannten neuen oralen Antikoagulanzen (DOAKs) Apixapan (ATC B01AF02), Edoxaban (ATC B01AF03), Rivaroxaban (ATC B01AF01) und Dabigatran (ATC B01AE07).

5.1.2 Datensatz und Datenlinkage

Grundlage für die Analysen im Teilprojekt A bildeten Routinedaten der drei beteiligten regionalen Krankenkassen AOK Bayern, AOK Hessen und AOK Plus, bei denen Menschen aus den Bundesländern Bayern, Hessen, Thüringen und Sachsen gesetzlich versichert sind. Die Variablen der lebenslangen Arztnummern, der Praxisidentifikationsnummern sowie der Krankenhausidentifikationsnummern wurden in allen drei Krankenkassen gleichermaßen pseudonymisiert, sodass die Daten über diese Variablen verknüpfbar wurden. Die Verknüpfung und Qualitätssicherung der drei Datensätze erfolgte in der Vertrauensstelle der LMU.

Dieser Routinedatensatz umfasste Abrechnungsdaten der eingeschlossenen PatientInnen aus dem ambulanten und stationären Sektor, Patientenstammdaten, Diagnose-, Operations- und Prozedurendaten sowie Apothekendaten innerhalb des Beobachtungszeitraums von Januar 2013 bis einschließlich Dezember 2017. Das Kriterium für den Einschluss der PatientInnen in den Datensatz war mindestens eine Verordnung und Abrechnung von mindestens einem der ausgewählten 34 Wirkstoffe während des Beobachtungszeitraums. Ausgeschlossen wurden

PatientInnen, die weniger als 350 Tage pro Jahr bei der Krankenkasse versichert waren. Verstorbene wurden von diesem Ausschlusskriterium ausgenommen. Der dadurch übermittelte Analysedatensatz umfasste die Abrechnungsdaten von ca. 1,25 Millionen Versicherten der drei Krankenkassen.

An diesen Datensatz wurden weitere Daten herangespielt:

- Die **wido Stammtabelle** enthält detaillierte Informationen über die Arzneimittel (wie z. B. Datum der Marktzulassung, Packungsgröße und Preis) und wurde über die Pharmazentralnummer mit den Routinedaten der Krankenkassen verknüpft.
- Für die Berücksichtigung von regionalen Einflussfaktoren auf die Verbreitung von Arzneimitteln wurden weiterhin **öffentlich zugängliche Daten zu regionalen Kennzeichen**, wie z. B. Krankenhausbetten pro Bevölkerung und Kreis, Urbanität oder Ärztedichte, von dem Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR – INKAR Datenbank) herangezogen, die über die Kreiskennziffer mit dem Analysedatensatz verknüpft wurden.

Für die durchgeführten Analysen wurde der Analysedatensatz weiterhin jeweils reduziert auf die relevanten Zeiträume und verordnenden Leistungserbringer bzw. PatientInnen mit relevanter Verordnung. Eine Übersicht der zugrundeliegenden Datensätze für die vier durchgeführten Analysen ist im Anhang 11.1.2 zusammengefasst.

5.1.3 Auswertungsstrategien

*Sacubitril/Valsartan*¹

Ziel dieser Analyse war es, die Verbreitung des Wirkstoffs Sacubitril/Valsartan (S/V) im ambulanten Sektor ab Markteinführung im Januar 2016 bis Dezember 2017 zu analysieren und hierbei insbesondere den Einfluss der PatientInnen und ihrer Pfade zu untersuchen.

S/V ist für die Behandlung von PatientInnen mit Herzinsuffizienz (und ohne Diabetes Mellitus) zugelassen und durchlief das AMNOG-Verfahren. Dabei wurde zwar ein beträchtlicher Zusatznutzen im Vergleich zur Standardtherapie festgestellt, allerdings gab es für diese Wirksamkeit nur einen Anhaltspunkt (niedrigste Evidenzkategorie) und die Jahrestherapiekosten waren 2016 etwa 50 Mal höher als die Kosten für die Vergleichstherapie. (Gemeinsamer Bundesausschuss 2016)

Um die Verbreitung des Wirkstoffs zu analysieren wurde erhoben, welche ÄrztInnen zu welchem Zeitpunkt das erste Mal das Arzneimittel verordnet haben (sogenannte Erstverordnungen) und es wurde in Ereigniszeitanalysen untersucht, wie verschiedene Einflussfaktoren diesen Zeitpunkt beeinflusst haben. Dabei wurden zwei Arten von Erstverordnungen unterschieden: Die Erstverordnung einer/s ÄrztIn konnte einerseits eine Verordnung sein, die auch für den/die PatientIn eine Erstverordnung darstellte und andererseits eine Verordnung, die für den/die PatientIn eine Folgeverordnung darstellte.

¹ Diese Analyse wurde im Paper „*Diffusion of a new drug among ambulatory physicians – The impact of patient pathways*“ zur Veröffentlichung eingereicht. Weitere Details zur Methodik können der Veröffentlichung entnommen werden.

Es wurden insgesamt vier verschiedene Modelle erhoben und für diese jeweils eine Ereigniszeitanalyse durchgeführt:

In *Modell 1* wurden alle ÄrztInnen mit einer Erstverordnung berücksichtigt (sogenannt *Neuverordner*), während wir uns in *Modell 2* auf die ÄrztInnen konzentrierten, deren Erstverordnung auch eine Erstverordnung für die PatientInnen darstellten. In den *Modellen 3* und *4* unterteilten wir die ÄrztInnen des Modells 2 danach, ob diese Erstverordnungen der PatientInnen den Empfehlungen der nationalen Versorgungsleitlinie entsprachen (*Modell 3*) oder ob es sich bei der Verordnung um eine potentiell nicht empfohlene Therapieoption handelte (*Modell 4*).

Die Auswahl der genauen Methode bei Ereigniszeitanalysen hängt unter anderem von der abhängigen und den erklärenden Variablen ab sowie von der Frage, ob der Zeitpunkt des Ereignisses intervallzensiert oder kontinuierlich vorliegt. Der Zeitpunkt der Einführung von S/V je ÄrztIn wurde monatlich ermittelt, auch wenn der tatsächliche Zeitpunkt der Einführung an verschiedenen Tagen innerhalb dieses Intervalls liegen kann. Folglich handelte es sich um eine intervallzensierte Datenstruktur, bei der die Erstverordnung je ÄrztIn pro Monat aggregiert erhoben wurde.

Es sollten außerdem zeitvariable und zeitkonstante erklärende Variablen einbezogen werden. Wir gingen weiterhin davon aus, bzw. haben durch erste Analysen festgestellt, dass einige Effekte der erklärenden Variablen im Zeitverlauf keinen konstanten Effekt auf die abhängige Variable hatten. Links- und Rechtszensierung der Daten konnte durch die vorliegende Datenstruktur ausgeschlossen werden, sodass ein generalized linear model (GLM) mit log-log Verknüpfungsfunktion angewendet werden konnte. Bei diesem diente die monatlich ermittelte binäre Variable, ob ein/e ÄrztIn S/V verordnet hat oder nicht, als Ergebnisvariable.

Da erste Analysen ergeben haben, dass es möglicherweise einen nichtlinearen Zeittrend bei der Verbreitung des Arzneimittels gibt, wurden weiterhin sogenannte smooth functions verwendet, die für die Zeit, gemessen in Monaten seit der Marktzulassung, kontrollieren. Der Zusammenhang zwischen abhängiger und erklärender Variablen kann in generalisierten additiven Modellen (GAM) geschätzt werden, ohne dass eine lineare Beziehung angenommen wird. (Hastie and Tibshirani 1986)

Direkte orale Antikoagulanzen (DOAKs)

Ziel dieser Analyse war es, die Verbreitung, bzw. Nutzung der direkten oralen Antikoagulanzen (Apixaban, Dabigatran, Edoxaban und Rivaroxaban) zu untersuchen. Seit 2015 sind vier dieser Arzneimittel auf dem deutschen Markt verfügbar, wobei die Evidenz ihrer Wirksamkeit unterschiedlich und teilweise umstritten ist. Zudem sind die Arzneimittel wesentlich teurer als z. B. die Vergleichstherapie mit Vitamin-K-Antagonisten. Zwei der vier Wirkstoffe (Apixaban und Edoxaban) sind als sogenannte *Positivwirkstoffe* in die Liste der beobachtungswürdigen Medikamente der Studie aufgenommen worden.

In der Analyse sollte untersucht werden, ob es Regionen in Deutschland gibt, in denen einer der vier Wirkstoffe vermehrt verordnet wurde als in den anderen und welche Faktoren auf Patienten-, Arzt- und Regionsebene die Verordnung des einen oder anderen Arzneimittels beeinflusst haben. Um diese Faktoren untersuchen zu können, wurden Mehrebenenmodelle mit binärer bzw. multinomialer Ergebnisvariable errechnet. PatientInnen stellten dabei die unterste Ebene des Modells dar, die ÄrztInnen die zweite und die Regionen (Kreise) die dritte Ebene. Die drei untersuchten Ebenen sind hierarchisch angeordnet, sodass mehrere Patien-

tInnen der/m gleichen ÄrztIn zugeordnet werden können, wobei immer nur der/die erstverordnende ÄrztIn betrachtet wurde. Die ÄrztInnen konnten durch ihre Praxiszugehörigkeit wiederum einem Kreis zugeordnet werden, sodass die Kreise die dritte und höchste Ebene der Analyse darstellten.

Pregabalin

Ziel dieser Analyse war es, das Ordnungsverhalten von ambulanten tätigen ÄrztInnen in empirischen Ärztenetzwerken zu untersuchen. Das Medikament wurde ausgewählt, da die Wirksamkeit umstritten ist und gleichzeitig ein erhöhtes Abhängigkeitspotential erkannt wurde.

In dieser Analyse wurde untersucht, ob die verordnenden ÄrztInnen von PatientInnen, die eine auffällig hohe Menge an Pregabalin verordnet bekommen haben, loser miteinander in empirischen Ärztenetzwerken (sogenannten Ordnernetzen) verbunden waren als die verordnenden ÄrztInnen von PatientInnen mit durchschnittlichen Ordnungsmengen. Die sich daraus ergebenden zwei Gruppen an PatientInnen wurden durch univariate statistische Tests miteinander verglichen, um so Unterschiede hinsichtlich der Patientencharakteristika und Ordnernetze ermitteln zu können.

In einem ersten Schritt wurden die PatientInnen dieses Analysedatensatzes daher in zwei Gruppen unterteilt: Die eine Gruppe bildeten die PatientInnen, die eine größere Menge als empfohlen im Zeitraum eines Jahres verordnet bekommen haben und somit das Medikament potentiell missbräuchlich verwendet haben können. Zur Klassifikation wurde die Empfehlung der European Medicine Agency (EMA) herangezogen, die in einem Bericht über Pregabalin eine maximale Dosis von 600mg pro Tag empfiehlt. (European Medicines Agency 2010) Diese Menge entspricht zwei definierten Tagesdosen (Defined Daily Doses [DDD]) des Medikaments. Die andere Gruppe bestand aus den PatientInnen, die durchschnittlich eine geringere Menge während eines Jahres verordnet bekommen haben.

Um die beiden Patientengruppen abschließend miteinander vergleichen zu können, wurden univariate statistische Tests durchgeführt. Für kategoriale Variablen wurde jeweils ein Chi²-Test durchgeführt, es sei denn eine Variable enthielt weniger als fünf Beobachtungen. Bei solch kleineren Gruppen wurde stattdessen der exakte Fisher-Test durchgeführt (Heumann et al. 2016). Für kontinuierliche Variablen wurde der nichtparametrische Wilcoxon-Mann-Whitney U-Test durchgeführt, um zu untersuchen, ob Werte der einen Gruppe signifikant größer (oder kleiner) als die der anderen Gruppe waren. Zuletzt wurde noch für multiples Testen korrigiert, da wiederholt statistische Tests auf dem gleichen Datensatz angewendet wurden. Dies könnte zu einem erhöhten Risiko für einen Fehler erster Art (falsch positiv) führen, andererseits können durch eine Bonferroni-Korrektur die sich ergebenden p-Werte korrigiert werden. (Bretz et al. 2016)

Oxycodon & Naloxon (Targin™)

Diese Analyse wurde methodisch an die Analyse von Pregabalin angelehnt und unterscheidet sich lediglich durch die durchgeführte Gruppenbildung sowie die ermittelten Patientenmerkmale, da andere medizinische Indikationen zu einer Verschreibung des Arzneimittels führen. Es wurden insgesamt drei verschiedene Patientengruppen gebildet, für die im Anschluss untersucht wurde, ob sich die PatientInnen der Gruppen systematisch voneinander unterscheiden und ob sich die verordnenden ÄrztInnen voneinander unterscheiden. Die Gruppen wurden mithilfe rein deskriptiver Statistiken miteinander verglichen.

In einem ersten Schritt wurden die PatientInnen dieses Analysedatensatzes daher in drei Gruppen unterteilt: Die eine Gruppe bildeten die PatientInnen, die eine sehr große Menge (365 oder mehr definierte Tagesdosen (DDD)) des Arzneimittels im Zeitraum eines Jahres verordnet bekommen haben. Da die DDD einen täglichen durchschnittlichen Verbrauch des Medikaments definieren, kann bei einem jährlichen Verordnungsvolumen von mehr als 365 Tagesdosen angenommen werden, dass das Medikament potentiell missbräuchlich verwendet wurde. Die andere Gruppe bestand aus den PatientInnen, die im Zeitraum eines Jahres nur eine Verordnung Targin erhalten haben. Diese Gruppe von PatientInnen stellen insofern eine Besonderheit dar, als dass sie das Medikament nicht regelmäßig verordnet bekommen. Alle anderen PatientInnen wurden einer dritten Gruppe zugeordnet und können somit als durchschnittliche Verbraucher betrachtet werden. In den durchgeführten Vergleichen wurden sowohl Merkmale der PatientInnen zwischen den beiden Gruppen verglichen als auch Merkmale der Verordnernetzwerke.

In den durchgeführten Vergleichen wurden sowohl Merkmale der PatientInnen zwischen den beiden Gruppen verglichen als auch Merkmale der Verordnernetzwerke.

5.2 Teilprojekt B

Im Teilprojekt B wurde der Fokus auf die Bayerische Wirkstoffvereinbarung (WSV) gelegt, die am 01.12.2014 in Kraft trat. Die WSV legt ihr Hauptaugenmerk auf die Steuerung des Verordnungsgeschehens durch eine zeitnahe Rückmeldung zum Ordnungsverhalten des jeweiligen Arztes. Dies steht im Gegensatz zu einer Prüfsystematik, die, im Nachgang und oft um Jahre versetzt, eine wirtschaftliche Ordnungsweise von Arzneimitteln prüft. Dazu werden Quoten für Generika-, Leitsubstanz- und Mengenziele festgelegt, die zudem nach ärztlichen Vergleichsgruppen aufgefächert bzw. differenziert sind. Mit Hilfe eines Längsschnitt- und eines Querschnittvergleichs wurden die Auswirkungen des innovativen Steuerungsansatzes analysiert. Für die Analysen wurde zum einen die Ordnungssituation in Bayern vor und nach Einführung der Wirkstoffvereinbarung betrachtet (Längsschnittvergleich) und zum anderen ein Vergleich der KV-Gebiete Bayern und Hessen durchgeführt (Querschnittvergleich).

5.2.1 Ausgewählte Indikatoren

Im Zuge der Analyse wurden sowohl monetäre als auch qualitative (bzgl. der Ordnungsqualität) Wirtschaftlichkeitsindikatoren ausgewertet, um beide Aspekte der Wirtschaftlichkeit abbilden zu können. Zudem wurden einige bayernspezifische Indikatoren (ausgewählte Ziele der WSV) analysiert. Folgende Indikatoren wurden ausgewählt:

Tabelle 1: Monetäre Indikatoren

durchschnittliche Kosten je Verordnung
durchschnittliche Kosten je Patient
durchschnittliche Kosten pro DDD bei den folgenden ATC-Gruppen: <ol style="list-style-type: none">1. ATC-Gruppe: Thrombozytenaggregationshemmer, exkl. Heparin (B01AC)2. ATC-Gruppe: Mittel mit Wirkung auf das Renin-Angiotensin-Aldosteron-System (C09)3. ATC-Gruppe: Mittel, die den Lipidstoffwechsel beeinflussen (C10)4. ATC-Gruppe: Mittel bei obstruktiven Atemwegserkrankungen (R03)
Verordnungsanteil von Generika am Gesamtmarkt (insb. auch im Bereich Antidiabetika)
Verordnungsanteil von Medikamenten, die als Me-Too-Präparate eingeordnet sind (laut Me-Too Liste der KV Nordrhein ² (Stand 2018), nur Präparate ohne Rabattvertrag)
Anteil der Rabattarzneimittel am Gesamtmarkt
Aut-idem-Quote

Tabelle 2: Qualitative Indikatoren

PPI-Mengen (DDD je Arzneimittelpatient)
Ezetimib-Mengen (gemessen in DDD je Patient, als Wirkstoff in Kombi- und Monopräparaten)
Patientenanteil mit DOAK-Verordnung an allen Patienten mit Antikoagulantienverordnung
Verordnungstiefenvergleich vor und nach Einführung der Wirkstoffvereinbarung
das verordnungsrelevante Überweisungsverhalten von Hausärzten zu Fachärzten bzw. Fachärzten zu Hausärzten (am Beispiel von Entresto)
Erreichung der allgemeinen Leitsubstanzquote, d.h. Verordnungsanteil aller Leitsubstanzen (innerhalb der KV-eigenen Leitsubstanzgruppen) über alle Leitsubstanzwirkstoffgruppen hinweg
Anteil neuer Arzneimittel (d.h. AMNOG-bewerteter Arzneimittel)

² aktuelle Me-Too Liste aufrufbar auf https://www.kvno.de/fileadmin/shared/pdf/online/beratung/regressvermeidung/me-too-liste_2017.pdf

Anteil Medikamente mit einem Zusatznutzen in mindestens einer Subgruppe an allen AM-NOG-bewerteten Präparaten
Anteil von Biosimilars am Beispiel der TNF- α -Inhibitoren
Übereinstimmung mit dem Medikationskatalog der Kassenärztlichen Bundesvereinigung (Medikationskatalog: Stand 2016 (V4.0) – inkl. ATC-Historie 2014-2016)
Anteil Patienten mit Polypharmazie
Anteil der Asthmapatienten und Patienten mit chronischer obstruktiver Lungenerkrankung (COPD) mit inhalativer Medikation
Anteil der Hypertoniker mit medikamentöser Therapie
Anteil der Patienten mit koronarer Herzkrankheit und Thrombozytenaggregationshemmer
Anteil der betreuten Patienten nach Myokardinfarkt mit Beta-Blockerbehandlung
Anteil der Patienten mit koronarer Herzkrankheit, die Medikamente mit Zusatznutzen erhalten (diese existiert nur für KHK-Patienten mit akutem Koronarsyndrom)
Anteil der Patienten mit Herzinsuffizienz und mit einem ACE-Hemmer oder AT1-Blocker
Anteil der Patienten mit Herzinsuffizienz und mit einem Beta-Rezeptorenblocker
Anteil der Patienten mit Herzinsuffizienz, die bei Vorhofflimmern mit DOAKs oder Vitamin-K-Antagonisten behandelt werden
Anteil der Herzinsuffizienzpatienten, die Medikamente mit Zusatznutzen erhalten (diese existiert nur für Patienten mit symptomatischer chron. Herzinsuffizienz)
Versorgung von Osteoporosepatienten mit einer medikamentösen Therapie

Tabelle 3: Indikatoren mit Bezug zur Wirkstoffvereinbarung (WSV-Ziele)

Anteil generischer DDDs im Bereich Analgetika außer BTM-pflichtige Opioide
Anteil generischer DDDs im Bereich der Mittel bei obstruktiven Atemwegserkrankungen
Anteil generischer DDDs im Bereich der Mittel mit Wirkung auf das Renin-Angiotensin-System
Anteil generischer DDDs im Bereich Thrombozytenaggregationshemmer (ohne Ticagrelor)
Anteil der bayerischen Leitsubstanz-DDDs in der Gruppe der MS-Therapeutika
Anteil der bayerischen Leitsubstanz-DDDs in der Gruppe der TNF- α -Inhibitoren

Die untersuchten Indikatoren wurden aus verschiedenen Quellen zusammengetragen, mit dem Ziel, eine möglichst breitgefächerte Messung von Kosten und Qualität zu erzielen. Dabei wurde zum einen auf Indikatoren aus den bestehenden Prüfvereinbarungen der beteiligten KVen zurückgegriffen und zum anderen passende Indikatoren aus dem Qualitätsindikatoren-system für die ambulante Versorgung (QiSA) des AQUA-Instituts ausgewählt, welche auf den geltenden Medikationsleitlinien beruhen.

5.2.2 Grundgesamtheit und Studienpopulation

Der Analysedatensatz basiert auf den Arzneiverordnungsdaten nach §300 (2) SGB V und den ambulanten Versorgungsdaten, welche ein Extrakt der vertragsärztlichen Abrechnungsdaten darstellen (§295 SGB V). Die Analysen wurden im Auftrag der KVen gemäß §80 SGB X durchgeführt. Die Grundgesamtheit enthält alle Patienten, die in den Untersuchungsjahren 2010-2016 mindestens einen Arztkontakt hatten und alle Betriebsstätten, die in diesen Untersuchungsjahren Arzneimittel durchgängig verordnet haben. In die Grundgesamtheit eingegangen sind Betriebsstätten jedoch nur, wenn sie in jedem der betrachteten Untersuchungsjahre mindestens 20 verschiedenen Patienten ein Arzneimittel verordnet haben. Damit sollen nicht verordnende Betriebsstätten, wie z. B. Psychotherapiepraxen, unberücksichtigt bleiben und zudem Datenartefakte, die z.B. durch Fehler beim Drucken oder Einlesen der Rezepte entstehen können, minimiert werden. Da jeder Fehler bzgl. der Betriebsstätte ein neues Pseudonym in den Daten generiert, welchem in der Regel nur einzelne Verordnungen zugeordnet sind, ist eine Mindestanzahl an Verordnungen ein probates Mittel, nur die echten, verordnenden Betriebsstätten zu filtern. Mit diesem Vorgehen bleiben 99% aller Verordnungen und ca. 80% der pseudonymisierten BSNRs erhalten. So bleibt eine Anzahl an Betriebsstätten erhalten, die plausibel erscheint.

Der erste Auswertungsblock besteht aus zwei Teilen: dem Längsschnittvergleich für Bayern und dem Querschnittsvergleich zwischen Bayern und Hessen jeweils für die Jahre 2010, 2013 und 2016. Für diese Auswertungen wurden nur die Daten der drei Beobachtungsjahre aller

Betriebsstätten der oben beschriebenen Studienpopulation aufgenommen, d.h. für die monetären Indikatoren alle Muster-16-Arzneimittelverordnungen und für die qualitativen Indikatoren zudem die Patienten mit den entsprechenden Diagnosen in den jeweiligen Beobachtungsjahren.

Dabei flossen in die Auswertung nur Patienten mit gesicherten Diagnosen ein, die mindestens für zwei Quartale im jeweils betrachteten Kalenderjahr vorzufinden waren (M2Q-Kriterium).

Im zweiten Auswertungsblock wurden alle Betriebsstätten aus der oben beschriebenen Studienpopulation für den Gesamtzeitraum 2010-2016 für eine Gesamtzeitraumanalyse ausgewählt.

5.2.3 Auswertung

Für den Längsschnittvergleich wurden für die Jahre 2010, 2013 und 2016 Vergleichsanalysen mithilfe des gepaarten t-Tests durchgeführt – jeweils für jeden gelisteten Indikator für die KV Bayerns, die KV Hessen und den Bund. Zudem wurden die Anstiege zwischen den Beobachtungsjahren jedes Indikators gebildet und ebenfalls mithilfe des gepaarten t-Tests auf statistisch signifikante Unterschiede der Mittelwerte getestet, um die Entwicklung über die Zeit und auch die Entwicklungsveränderung abzubilden.

Für die Querschnittsanalysen wurden Vergleichsanalysen mithilfe des ungepaarten t-Tests durchgeführt. Auf diese Weise wurden die mittleren Anstiege zwischen den Beobachtungzeitpunkten 2013 und 2016 analysiert und verglichen, um auf signifikante Unterschiede zwischen Bayern und Hessen zu testen.

In einem zweiten Auswertungsblock wurden für Bayern und Hessen zusätzliche Längsschnittanalysen mithilfe des Strukturbruchmodells nach Chow (Chow 1960) durchgeführt, welches auf eine geeignete Auswahl an monetären und qualitativen Indikatoren angewandt wurde. Dieses Modell ermittelt, ob zu einem gegebenen Interventionszeitpunkt (hier: Einführung der Wirkstoffvereinbarung) ein messbarer Bruch innerhalb des aufgestellten Modells vorliegt. Für diese Analyse wurden einige geeignete, schon mittels gepaarten und ungepaarten t-Tests untersuchte Indikatoren gewählt.

Diese Auswahl sollte zum einen beide Aspekte der Wirtschaftlichkeit abdecken und zum anderen einige schon im Vergleich der verbundenen Stichproben gefundene Ergebnisse überprüfen und ggf. validieren. Es wurden die folgenden Indikatoren mithilfe des Strukturbruchmodells analysiert:

Tabelle 4: Indikatoren des Strukturbruchmodells

durchschnittliche Kosten je Verordnung
durchschnittliche Kosten je Patient
Verordnungsanteil von Generika am Gesamtmarkt
Patientenanteil mit DOAK-Verordnung an allen Patienten mit Antikoagulantienverordnung
Erreichung von Leitsubstanzquoten
Bayernspezifische Generika-DDD
Leitsubstanz-DDD-Ziel in der Gruppe der MS-Therapeutika

Des Weiteren wurden die oben aufgelisteten Indikatoren noch für die folgenden Größen adjustiert bzw. differenziert:

- Alter des Patienten,
- Geschlecht des Patienten,
- Multimorbidität des Patienten,
- Polypharmazie des Patienten (d.h. mehr als fünf verordnete Wirkstoffe),
- Größe der Betriebsstätte und
- Verordnungstiefe der Betriebsstätte.

5.3 Teilprojekt C

5.3.1 Studiendesign und Leitfadententwicklung

Es wurde eine qualitative Studie mittels teilstrukturierter Interviews und Fokusgruppen durchgeführt. Im Rahmen der Studie wurde im Teilprojekt D ein teilstrukturierter Leitfaden entwickelt und für die Durchführung der Interviews für Teilprojekt C angepasst. Der Leitfaden umfasst vier große Abschnitte nach der mündlichen Studienaufklärung: 1. Allgemeine Informationen zu Person und Praxiserfahrung (Fachrichtung und typische Patient*innenfälle), 2. Allgemeines zum Ordnungsverhalten (Kriterien zur Auswahl von Arzneimittel-Therapien), Rolle der Patient*in, Umgang mit neuen Medikamenten, Umgang mit initiierten Therapien von anderen (Fach-) Ärzt*innen), 3. Gesamtsituation und Bayerische Wirkstoffvereinbarung (Veränderung von Ordnungsverhalten, Vor- und Nachteile, Wünsche und Verbesserungsmöglichkeiten, Zufriedenheit, Vereinbarkeit von Wirtschaftlichkeit und Qualität) und 4. Abschluss und Fazit (Motivation, weitere Aspekte, abschließendes Fazit zur WSV). Unter Punkt 2 wurden zudem Fallbeispiele für die jeweilige Fachrichtung vorbereitet. Für die Durchführung der Fokusgruppen wurde der Interviewleitfaden adaptiert (gekürzt, z.B. Wegfall Fallbeispiel). Im Zuge der Nachrekrutierung im Juli 2020 wurden für die weiteren Fokusgruppen leichte Anpassungen vorgenommen, um die aktuellen Entwicklungen hin zur WSV 3.0 abzubilden.

Alle Befragten füllten zusätzlich einen soziodemographischen Fragebogen mit Angaben zu ihrer Person und Praxis aus.

5.3.2 Rekrutierung

Für die Studie wurden niedergelassene Ärzt*innen rekrutiert mit folgenden Einschlusskriterien:

- Niedergelassene Ärzt*innen im Raum Bayern
- Fachrichtungen: Hausärzt*innen, Innere Medizin (Kardiologie, Diabetologen, Pneumologen, Hämato-Onkologen), Neurologen und Psychiater
- WSV-Zielerreichung von mind. 90 % in den Quartalen 2/ 2017 bis 1/2018.

Die Rekrutierung fand über Anschreiben durch die Kassenärztliche Vereinigung Bayerns und über die Ärztenetze des Allgemeinmedizinischen Institut Erlangens statt.

Da sich in der ersten Rekrutierungswelle nur Hausärzt*innen meldeten, wurden in der zweiten Rekrutierungswelle (Juli 2020) nur Ärzt*innen anderer Fachrichtungen angeschrieben. Über andere Rekrutierungswege (Newsletter des Allgemeinmedizinischen Instituts, persönliche Ansprache, Anschreiben von Lehrarztpraxen) konnte eine weitere Person gewonnen werden. Eine Person sprang aufgrund der Online-Durchführung wieder ab.

5.3.3 Datenaufbereitung und Auswertung

Die Audiodateien wurden nach von Dresing und Pehl vollständig transkribiert. (Dresing and Pehl 2017) Die Auswertung erfolgte in Anlehnung an die qualitative Inhaltsanalyse nach Mayring mittels eines deduktiv-induktiven Verfahrens, mit Hilfe des Softwareprogramms MAXQDA Plus 2020. Das Projektteam in Erlangen diskutierte die Kategorien kontinuierlich (konsensuelle Validierung). Mit dem endgültigen Kategoriensystem wurde schließlich das gesamte Datenmaterial kodiert. Eine kommunikative Validierung der Ergebnisse fand mit den dem Institut angeschlossenen Lehr- und Weiterbildungsärzten im Rahmen des "Tages der Allgemeinmedizin/Weiterbildung/Lehre" am 10.10.2020 statt. Das Kategoriensystem wurde auch bei den Fokusgruppen angewandt und um weitere Kategorien ergänzt wie beispielsweise der Vergleich zwischen WSV 2.0 und 3.0.

5.4 Teilprojekt D

5.4.1 Definition

Im Rahmen der Bayerischen Wirkstoffvereinbarung schaffen es einige Betriebsstätten und darin arbeitende Ärzte nicht, langfristig die vorgegebenen Arzneimittelquoten zu erreichen. Diese so genannten Zielverfeher sind wie folgt definiert:

a) Betriebsstätten, die in mindestens zwei aufeinander folgenden Quartalen innerhalb des festgelegten Beobachtungszeitraumes (Quartal 2/2017 bis Quartal 1/2018) weniger als 85% des Gesamtziels im Rahmen der Bayerischen Wirkstoffvereinbarung erreicht haben

und

b) sich in einem auf das nächste Quartal um weniger als 3% bei der Gesamtzielerreichung verbessert haben.

Zielerreicher sind im Gegensatz dazu Betriebsstätten, die

a) mindestens 85% ihres Gesamtziels in der Bayerischen Wirkstoffvereinbarung erreicht haben,

oder

b) sich – im Falle einer Gesamtzielerreichung von weniger als 85% – hinsichtlich ihrer Gesamtzielerreichung in zwei aufeinanderfolgenden Quartalen um mehr als 3% verbessert haben.

Die Definition der Zielerreicher und Zielverfehlen galt für das ganze TP-D und wurde sowohl für die Identifikation von Zielverfehlen-Betriebsstätten als Grundlage für die Rekrutierung der Interviewpartner als auch für die deskriptiven Analysen der Routinedaten angewendet.

Welche Hintergründe und Motive zur Auffälligkeit beitragen könnten wurde anhand von teilstrukturierten, leitfadengestützten Interviews exploriert. Hierfür wurden Ärzte aus Betriebsstätten befragt, die die Zielverfehlen-Kriterien erfüllen und im Wirkungsbereich der Bayerischen Wirkstoffvereinbarung (ambulanter Sektor) liegen. Weiterhin wurden anhand von deskriptiven Analysen eines Routinedatensatzes der Kassenärztlichen Vereinigung Bayerns Charakteristika von Ärzten und Betriebsstätten – jeweils in der Gruppe der Zielverfehlen und Zielerreicher - sowie Verordnungen untersucht.

5.4.2 Qualitativer Arm – teilstrukturierte, leitfadengestützte Einzelinterviews

Um den aktuellen Datenschutzbestimmungen gerecht zu werden und die Anonymität der Ärzt*innen zu wahren, wurde ein umfangreiches Rekrutierungsverfahren entwickelt (siehe Anhang 11.4.1).

Ein von den Universitäten entwickeltes Anschreiben zur Einladung zu den Interviews wurde an alle n=401 identifizierten Betriebsstätten angeschrieben. Bei Studieninteresse konnten sich die entsprechenden Betriebsstätten mit einem beiliegenden Schreiben zur Interessensbekundung bei den universitären Einrichtungen melden. Diese versendeten daraufhin Studienunterlagen mit Informationsschreiben, Datenschutzerklärung und Einwilligung (siehe Anlage TP-D 1: Studienunterlagen). Es wurde ein Interviewtermin ausgemacht, für den die beiden Interviewerinnen in die Praxis kamen.

Für die Interviews wurde ein Interviewleitfaden entwickelt, der folgende Themenblöcke abdeckte und sich durch den gewählten Aufbau über ein Kennenlernen des Arztes und zunehmendes Vertrauen hin zu brisanteren Themen aufbaute: Allgemeine Informationen zu Person und Praxis, Rolle als Arzt, Allgemeines zum Ordnungsverhalten, Patientenbezogenes Ordnungsverhalten, Gesamtsituation, Bayerische Wirkstoffvereinbarung, Up-to-date bleiben, Abschluss.. Die Interviews wurden aufgezeichnet und extern transkribiert, bevor sie codiert und ausgewertet wurden. Die Entwicklung des Codebaumes erfolgte gemischt induktiv-deduktiv, die Auswertung folgte der qualitativen Inhaltsanalyse nach Mayring.

5.4.3 Quantitativer Arm – deskriptive Analyse von Routinedaten

Für die deskriptiven Analysen wurde ein Analyseleitfaden entwickelt. Sowohl in der Gruppe der Zielerreicher als auch der Zielverfehlen wurden folgende Charakteristika untersucht:

Tabelle 5: deskriptiv untersuchte Charakteristika

Fachgruppe und der jeweilige Anteil an Zielverfehlen-Betriebsstätten
Zusatzbezeichnungen
Alter
Geschlecht
Größe der Betriebsstätte
i. Quartalsmäßige Fallzahl
ii. Anzahl Arztsitze pro Betriebsstätte
Lage der Betriebsstätte
a) Städtisch/ländlich
b) Vier verschiedene Stufen an siedlungsstrukturellen Kreistypen

Hierbei wurden die Häufigkeiten verschiedener Merkmalsausprägungen untersucht. Bei potentiellen Unterschieden zwischen der Gruppe der Zielerreicher und Zielverfehlen wurde der Chi-Quadrat-Test durchgeführt, um auf einen Zusammenhang von Unterschied und Gruppenzugehörigkeit zu testen. Eine ursprünglich geplante Analyse auf Fachgruppenebene wurde nicht weiter ausgewertet, da auf Grund der zum Teil sehr geringen Fachgruppengröße von n=1 Betriebsstätten keine aussagekräftigen Ergebnisse erzielt werden konnten.

Auswertungen fanden mit Hilfe des Statistikprogramms SPSS statt.

Weiterhin wurde ebenfalls an Daten der Kassenärztlichen Vereinigung Bayerns eine Analyse von Verordnungen durchgeführt. Hierbei wurde für 16 von 25 Fachgruppen die Anzahl relevanter Einzelziele identifiziert und verglichen. Als relevante Einzelziele sind hier Ziele mit einem hohen Verordnungsvolumen im Praxisalltag gemeint, die das Hauptverordnungsspektrum der jeweiligen Fachgruppe abbilden. Für weitere Analysen wurden acht Fachgruppen ausgewählt, deren Verordnungsziele weiter untersucht wurden.

Siehe Anhang 11.4.3: betrachtete Fachgruppen für weitere Verordnungsanalysen

Innerhalb der Fachgruppen wurden folgende Aspekte der Arzneimittelverordnung im Rahmen der Bayerischen Wirkstoffvereinbarung untersucht.

Tabelle 6: Untersuchte Aspekte der Verordnungen

Welche Einzelziele werden von der Fachgruppe nicht erreicht?

Welche Ziele sind die verordnungsmengenmäßig umfangreichsten Ziele?

Welche jeweils TOP 3 Arzneimittel stehen einer Zielerreichung jeweils entgegen?

Wieviele relevante Ziele haben die jeweiligen Fachgruppen?

Um welche Art von Zielen handelt es sich jeweils?

Welche Auffälligkeiten ergeben sich jeweils?

Die Verordnungsdaten enthalten für die verschiedenen Einzelziele auf Fachgruppenebene die insgesamt innerhalb dieses Einzelziels verordnete DDD-Menge (defined-daily-doses) sowie die prozentuale Zielerreichung. Wurden die vorgegebenen Quoten der Bayerischen Wirkstoffvereinbarung erreicht, ist die Zielerreichung auf 100% gesetzt, wurden die Quoten übererfüllt, ist die Zielerreichung bei über 100% und bei einer Zielverfehlung weniger als 100%.

Auf Fachgruppenebene wurde hierbei betrachtet, welche Einzelziele von den acht Fachgruppen jeweils mit weniger als 90 Prozent der Zielerreichung verfehlt wurden. Ebenso wurden die jeweils drei Einzelziele mit dem in DDD gemessenen höchsten Verordnungsvolumen identifiziert.

Für die identifizierten Einzelziele wurde weiter untersucht, um welche Art der Zielvorgaben es sich handelt (Generikaziele, Leitsubstanzziele, Mengenziel), welche drei konkreten Arzneimittel einer Zielerreichung am meisten entgegenstehen (gemessen am verordneten DDD-Volumen) und ob sich weitere Auffälligkeiten wie fachgruppenübergreifend häufig verfehlt Einzelziele oder besonders häufig verordnete Arzneimittel zeigen.

Die Auswertungen fanden anhand von Excel-Tabellen statt.

5.5 Teilprojekt E

Zielsetzung und Indikatoren

Das primäre Ziel der Analyse ist die Identifikation von auffälligen Betriebsstätten, wobei sich die Auffälligkeit auf die Wirtschaftlichkeit des jeweiligen Ordnungsverhaltens bezieht. Für die Abbildung der Wirtschaftlichkeit des Ordnungsverhaltens werden aussagekräftige Indikatoren aus dem TP B genutzt.

Für die Identifikation der auffälligen Betriebsstätten wird der Ansatz der Profiling-Analyse genutzt, der in diesem Kontext (d.h. des wirtschaftlichen Ordnens) noch nicht zum Einsatz kam. Bisher wurde diese Analyseart für die Identifikation von auffälligen Einheiten - zumeist Krankenhäuser - gewählt, wo Auffälligkeiten, wie z.B. eine erhöhte Sterberate der Patienten oder eine erhöhte Rückkehrquote von bereits entlassenen Patienten untersucht wurden. Die Profiling-Analyse bietet den Vorteil, Unterschiede zwischen den zu betrachtenden Einheiten (hier: Betriebsstätten) zu berücksichtigen und in das Modell als Adjustierungsvariablen miteinfließen zu lassen. Somit können verschiedenen Faktoren wie Praxisgröße, Anzahl der Ärzte je

Betriebsstätte, Patientenstruktur (z.B. Altersverteilung, Morbidität etc.) mitberücksichtigt werden und durch die Analyse echte Auffälligkeiten gefunden werden, die nicht nur durch diese Merkmale bedingt sind. Somit ist ein wichtiger Schritt der Miteinbezug des jeweiligen Case Mix und eine geeignete Adjustierung dafür.

Ziel ist die Entwicklung und Anwendung einer Analyseprozedur, die sowohl interpretierbar als auch robust ist und künftigen Anwendern in der Entscheidungsfindung bzgl. „echt auffälligen“ Betriebsstätten hilft. Diese zu entwickelnde Analyseprozedur soll aus einer Kombination eines multivariaten Ansatzes der Ausreißerbestimmung, Regressionsverfahren und der Bayesschen Inferenz bestehen.

Für die Entwicklung des Modells und der Analyseprozedur wurden fünf Indikatoren im Sinne einer Machbarkeitsstudie gewählt (s. Tabelle 7), welche verschiedene Aspekte einer wirtschaftlichen Verordnungsweise widerspiegeln sollen.

In einem ersten Block wurde - als Basisanalyse - für zwei Indikatoren am Beispiel einer KV für das Quartal 4/2016 die Analyseprozedur entwickelt und erprobt (s. Tabelle 8). Hierbei wurden verschiedene Methoden in einer neuen und innovativen Weise zusammengeführt, um eine Ermittlung von auffälligen Einheiten unabhängig von den jeweiligen Prüfsystemen der KVen durchzuführen – nur basierend auf statistischen Methoden kombiniert mit dem Machine-Learning-Ansatz.

In einem zweiten Block wurde die Analyseprozedur für drei weitere Indikatoren ebenfalls am Beispiel einer KV über die Quartale der Jahres 2016 erprobt und feinadjustiert (s. Tabelle 9).

Tabelle 7: Indikatoren der Basisanalyse und der weiterführenden Analyse

Basisanalyse
durchschnittliche Kosten je Verordnung
Verordnungsanteil von Generika am Gesamtmarkt
Weiterführende Analyse
verordnete PPI-Mengen (gemessen in DDD je Arzneimittelpatient)
Anteil verordneter DOAK-DDDs an allen verordneten Antikoagulantien-DDDs
Anteil von generischen DDDs im Bereich der Lipidsenker

Für die Analysen wurden die folgenden Confounder gewählt (s. Tabellen 8 und 9):

Tabelle 8: Confounder der Basisanalyse

Verordnungstiefe einer Betriebsstätte (gemessen in DDDs)
Anteil der Patienten mit einer Diagnoseanzahl über dem 90%-Perzentil (aller Patientendiagnosen)
Anteil der Polypharmaziepatienten
Patientenanzahl je Arzt in einer Betriebsstätte
Anteil von männlichen Patienten an allen Patienten
Durchschnittsalter der Patienten mit mindestens einer Arzneimittelverordnung
Anteil der Patienten, die in einem ländlichen Gebiet wohnen
Anteil der Patienten, die in einem städtischen Gebiet wohnen
Neurologisch-psychiatrische Praxis-Kennung (nimmt den Wert 1 an - ansonsten 0), falls in der jeweiligen Hausarztpraxis noch zusätzlich neurologisch-psychiatrische Fachgruppen vertreten sind
Kardiologische Praxis-Kennung (nimmt den Wert 1 an - ansonsten 0), falls in der jeweiligen Hausarztpraxis noch zusätzlich die Fachgruppe der Kardiologen vertreten ist

Tabelle 9: Confounder der weiterführenden Analysen

Verordnungstiefe einer Betriebsstätte (gemessen in DDDs)
Anteil der Patienten mit einer Diagnoseanzahl über dem 90%-Perzentil (aller Patientendiagnosen)
Anteil der Polypharmaziepatienten
Patientenanzahl je Arzt in einer Betriebsstätte
Anteil von weiblichen Patienten an allen Patienten
Durchschnittsalter der Patienten mit mindestens einer Arzneimittelverordnung
Anteil der Patienten, die in einem ländlichen Gebiet wohnen
Anteil der Patienten, die in einem städtischen Gebiet wohnen
Betriebsstättenkennung (nimmt den Wert 1 an - ansonsten 0), falls in der jeweiligen Hausarztpraxis noch zusätzliche Nichthausarzt-Fachgruppen vertreten sind.

Im Teilprojekt E wurden nur Daten aus dem Jahr 2016 analysiert, weil erst hier ein Abgleich der Indikatoren mit der Wirkstoffvereinbarung möglich war. Im Zuge dessen zeigte sich, dass es mit den vorab gewählten Kontrollgrenzen zu einer Unterschätzung der Anzahl der auffälligen Einheiten kam. Für die Wahl von passenden Kontrollgrenzen, die das Modell adäquat adjustieren, wäre ein Abgleich mit der Menge von echt auffälligen Betriebsstätten nötig gewesen, welcher nicht möglich war aufgrund von unterschiedlichen Pseudonymisierungen der Betriebsstätten in den KVen und im Zentralinstitut.

Dennoch wurde innerhalb des Konsortiums zusätzlich eine historische Betrachtung analog zu TP B diskutiert. Hiervon wurde jedoch Abstand genommen, da für die Erprobung der Methode durch die Einbeziehung weiterer Jahre kein zusätzlicher Erkenntnisgewinn erzielt werden konnte und auch inhaltlich wäre die Nutzung von Indikatoren, die in den jeweiligen Jahren nicht Gegenstand der Prüfvereinbarung waren, hinsichtlich der Ergebnisinterpretation fragwürdig.

5.5.1 Grundgesamtheit und Studienpopulation

Der Analysedatensatz basiert auf den Arzneiverordnungsdaten nach §300 (2) SGB V und den ambulanten Versorgungsdaten, welche ein Extrakt der vertragsärztlichen Abrechnungsdaten darstellen (§295 SGB V). Die Analysen wurden im Auftrag der KVen gemäß §80 SGB X durchgeführt. Die Grundgesamtheit enthält alle Patienten, die im Untersuchungsjahr 2016 mindestens einen Arztkontakt hatten und alle Betriebsstätten, die in diesen Untersuchungsjahren Arzneimittel verordnet haben. In die Grundgesamtheit eingegangen sind Betriebsstätten jedoch nur, wenn sie in jedem der betrachteten Untersuchungsjahre mindestens 20 verschiedenen Patienten ein Arzneimittel verordnet haben.

Für diese Auswertungen wurden nur die Daten des Beobachtungsjahres 2016 aller Betriebsstätten der oben beschriebenen Studienpopulation aufgenommen, d.h. für die Indikatoren alle Muster-16-Arzneimittelverordnungen und für die Confounder zudem die Patienten mit den entsprechenden Diagnosen im Beobachtungsjahr 2016. Dabei flossen in die Auswertung nur Patienten mit gesicherten Diagnosen ein, die mindestens für zwei Quartale im jeweils betrachteten Kalenderjahr vorzufinden waren (M2Q-Kriterium).

5.5.2 Analyseplan

Die Analyseprozedur bestand aus folgenden Schritten:

1. Bayessche Inferenz
2. Bestimmung von auffälligen Betriebsstätten mithilfe der „natürlichen“ Verteilung (Bayessche Inferenz)
3. Regression Forest
4. Bestimmung der auffälligen Betriebsstätten auf Basis des Regression Forests mithilfe eines Observed/Expected-Indikators
5. Interpretation der Ergebnisse mithilfe von Shapley Values

In einem ersten Schritt wird für jeden betrachteten Wirtschaftlichkeitsindikator mithilfe der „natürlichen“ Verteilung, der Bayesschen Inferenz, des Indikators ein Wahrscheinlichkeitsmodell aufgestellt. Mithilfe dieser „natürlichen“ Verteilung wird ein Benchmark definiert und im direkten Vergleich dazu können Betriebsstätten ermittelt werden, die auffällig sind.

In einem weiteren Schritt wird die jeweilige Indikatorvariable mithilfe des Ansatzes des Regression Forests als endogene Variable modelliert, wobei sie von mehreren ausgewählten,

exogenen Variablen abhängt. Die so mittels Regression-Forest-Ansatz berechneten Variablenwerte werden genutzt, um einen Observed/Expected-Indikator - hier als Differenz der beobachteten und der vom Modell vorhergesagten Werte - zu bestimmen. Auch mit diesem Vorgehen können auffällige Betriebsstätten ermittelt werden, indem extreme Residuenwerte des Observed/Expected-Indikators identifiziert werden.

Im Weiteren werden diese beiden Ansätze der Identifikation von (möglichen) auffälligen Einheiten kombiniert, indem der mittels Random Forest Regression geschätzte Indikator in das Modell der Bayesschen Wahrscheinlichkeit eingebettet wird. Somit erfolgt eine kombinierte Analyse und Entscheidungsfindung, die die Vorteile beider Ansätze vereint: zum einen ein Wahrscheinlichkeitsmodell, dass die zugrundeliegenden Unsicherheiten vollständig abbildet und vorab bekannte Informationen und Annahmen beinhaltet und zum anderen die Möglichkeit exogene Faktoren zu berücksichtigen, die die zu analysierende Größe maßgeblich beeinflussen.

Um die Analyseergebnisse final interpretieren zu können, was bei der Anwendung von Machine Learning Modellen nicht trivial ist, wird die Methode der Shapley Values aus der Spieltheorie genutzt. Diese hilft die Einflussnahme der exogenen Variablen richtig einzuordnen.

Die so durchgeführte Identifikation von (möglichen) auffälligen Einheiten - hier Betriebsstätten - soll zum einen Betriebsstätten ermitteln, die offensichtlich auffällig sind, und zum anderen Betriebsstätten, die auffällig erscheinen, wo aber im Nachgang noch eine individuelle Überprüfung der Verordnungsweise anzuraten ist. Mit diesem Vorgehen könnten u.a. potentiell auffällige Praxen bestimmt werden, die in einer genaueren Nachuntersuchung auf Auffälligkeit überprüft werden können. Zudem kann gezielt untersucht werden, ob u.a. Praxisbesonderheiten vorliegen. Dies könnte helfen echt auffällige von scheinbar auffälligen Betriebsstätten gezielt zu trennen.

Der gesamte Prozess wird in folgender Abbildung (Abb. 1) visualisiert:

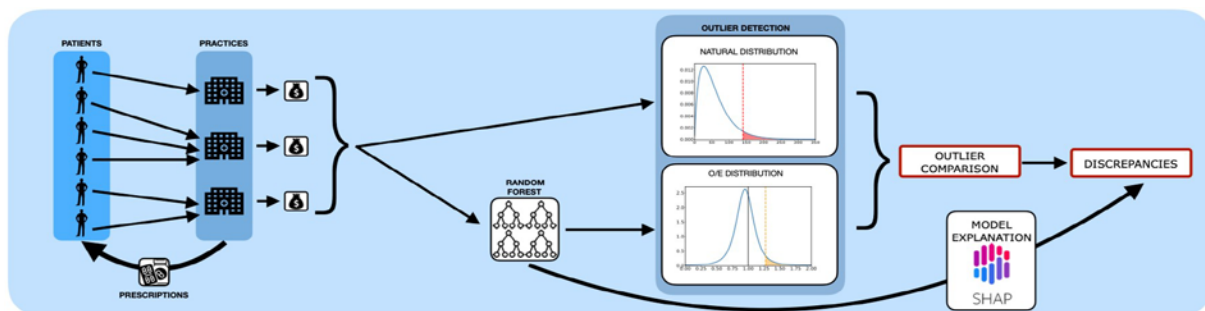


Abbildung 1: Schema der Analyseprozedur

6. Projektergebnisse

6.1 Teilprojekt A

Für die Verbreitung und Verwendung von verschreibungspflichtigen Arzneimitteln haben wir mehrere Einflussfaktoren demonstrieren können.

6.1.1 Sacubitril/Valsartan

Aus dieser Analyse lässt sich ableiten, dass PatientInnen durch ihre Kontakte zu verschiedenen niedergelassenen ÄrztInnen zu einer Verbreitung von Sacubitril/Valsartan beitragen. Die Regressionsergebnisse dieser Analyse werden in Form von Hazard Ratios (HRs) dargestellt. Diese geben das Verhältnis der Hazard Rates S/V zu verordnen für zwei ÄrztInnen an, die sich nur in Bezug auf die untersuchte erklärende Variable unterscheiden, wobei alle anderen Variablen konstant bleiben. Es zeigt sich, dass ÄrztInnen eine höhere Hazard Rate aufweisen, S/V zum ersten Mal zu verordnen, wenn sie mit vielen anderen ÄrztInnen im empirischen Ärztenetz verbunden sind, die das Medikament im Vormonat bereits verordnet haben. Je mehr bereits verordnende ÄrztInnen verbunden sind, umso höher ist die Wahrscheinlichkeit S/V zum ersten Mal zu verordnen (HR 1,593 für Expositionswert $\leq 25\%$ bis HR 2,755 für Expositionswert $> 75\%$ in Modell 1). Dieses Ergebnis ist für alle vier geschätzten Modelle gültig. In Bezug auf die Patientenpfade ergeben die Regressionsmodelle, dass ÄrztInnen eine höhere Hazard Rate aufweisen, wenn sie PatientInnen behandeln, die zuvor von anderen ÄrztInnen das Medikament bereits verordnet bekommen haben (HR 3,351 in Modell 1). Dieser Effekt ist für alle vier Modelle positiv und signifikant (5 % - Signifikanzniveau), jedoch am stärksten für das Modell 1. Im Modell 1 ist auch ein positiver Effekt zu verzeichnen, wenn PatientInnen eine Vormedikation mit S/V von einer/m ÄrztIn einer anderen Praxis verordnet bekommen haben (HR 2,760). Dieser Effekt ist nicht zu erkennen, wenn in den Modellen 2 bis 4 nur die ÄrztInnen betrachtet werden, die für PatientInnen eine Medikation mit S/V initiieren.

6.1.2 Direkte orale Antikoagulanzen

Hier wurde die räumliche Verbreitung der vier direkten oralen Antikoagulanzen Apixaban, Dabigatran, Edoxaban und Rivaroxaban untersucht. Diese sind in Deutschland seit unterschiedlich langer Zeit zum Beispiel zur Prophylaxe von Thrombosen und Schlaganfall sowie zur Behandlung von tiefen Venenthrombosen (TVT) oder Lungenembolien (LE) zugelassen. Die direkten oralen Antikoagulanzen sind insbesondere beobachtungswürdig, weil sie im Vergleich zur Standardtherapie mit z. B. Vitamin-K-Antagonisten wesentlich teurer sind und zudem ihre Wirksamkeit je Indikation unterschiedlich und teilweise umstritten ist. Die Tabelle im Anhang 11.1.1 bietet eine Übersicht über die untersuchten Wirkstoffe. Edoxaban und Apixaban wurden in bzw. nach 2011 zugelassen, sodass beide in einer Nutzenbewertung im AMNOG Verfahren hinsichtlich ihrer Wirksamkeit geprüft wurden. Für beide Wirkstoffe wurde dabei für einzelne Anwendungsgebiete ein geringer Zusatznutzen im Vergleich zur jeweiligen Vergleichstherapie festgestellt. Diese werden entsprechend und auch wegen des geringeren Preises im Vergleich zu Dabigatran und Rivaroxaban als Positivwirkstoff in der Gruppe der DOAKs gewertet.

In diesem Bericht wird insbesondere auf die logistische Regression eingegangen, bei der zwischen einer Verordnung der Positivwirkstoffe (Apixaban & Edoxaban) sowie den beiden anderen DOAKs (Dabigatran & Rivaroxaban) unterschieden wurde. Die ermittelten Odds Ratios (ORs) geben an, wie sich eine Veränderung der betrachteten erklärenden Variable auf die Wahrscheinlichkeit einen Positivwirkstoff verordnet zu bekommen, auswirkt, vorausgesetzt alle anderen erklärenden Variablen bleiben konstant. Die Ergebnisse zeigen, dass auf den drei

untersuchten Ebenen (Patient, Arzt, Region), verschiedene Faktoren die Entscheidung für einen der beiden Positivwirkstoffe signifikant beeinflussen. Für Patient*innen mit einem höheren Alter und Patient*innen mit nichtvalvulärem Vorhofflimmern ist es wahrscheinlicher, einen Positivwirkstoff verordnet zu bekommen als für jüngere bzw. PatientInnen ohne nichtvalvuläres Vorhofflimmern (OR Alter: 1,02; OR nichtvalvuläres Vorhofflimmern: 1,628). Ein Krankenhausaufenthalt vor der Erstverordnung mit einem DOAK steigerte die Wahrscheinlichkeit einen Positivwirkstoff verordnet zu bekommen um den Faktor 1,3. Für Ärzt*innen zeigt sich, dass es wahrscheinlicher für KardiologInnen war einen Positivwirkstoff zu verordnen als für Allgemeinmediziner*innen (OR 1,243), wobei die anderen Fachärzt*innen eine geringere Wahrscheinlichkeit als Allgemeinmediziner*innen hatten. Auf der Ebene der Regionen zeigt sich, dass in Kreisen, in der eine Wirtschaftlichkeitsverordnung zum Zeitpunkt der Erstverordnung den Umgang mit DOAKs regelte, die Wahrscheinlichkeit höher war, dass ein Positivwirkstoff verordnet wurde als in Regionen, in denen eine solche Regelung nicht aktiv war (OR 1,505)³.

³ Die Ursache dieses Effekts kann auch durch andere Faktoren beeinflusst werden, wie zum Beispiel durch einen zeitlichen Effekt.

6.1.3 Pregabalin

Die deskriptiven Ergebnisse⁴ dieser Auswertung zeigen, dass die Gruppe der Patient*innen, die als „normale Pregabalin-Nutzer“ eingestuft wurden, für alle dargestellten Merkmale ähnliche Werte wie die Gesamtpopulation aufwies. Die meisten Merkmale der Gruppe der „normalen Nutzer“ und die der Patient*innen, die Pregabalin potentiell missbräuchlich nutzen, unterscheiden sich jedoch häufig systematisch. Die Ergebnisse zeigen, dass insbesondere männliche und mittelalte Patienten eine auffällig hohe Menge an Pregabalin verordnet bekommen haben und eine diagnostizierte Suchterkrankung aufweisen. Auf Seiten der Verordner zeigt sich, dass die meisten Patient*innen ihre erste Verordnung für Pregabalin von einer/m Allgemeinmediziner*in (62 %), Neurochirurg*innen oder Neurolog*innen erhielten. Dieses Merkmal unterschied sich nur geringfügig und nicht signifikant zwischen den beiden Patientengruppen. Weiterhin erhielten die Pregabalin-Nutzer im Laufe eines Jahres durchschnittlich sechs Verschreibungen von zwei verschiedenen Ärzt*innen. Im Gegensatz dazu erhielten Patient*innen, die Pregabalin potentiell missbräuchlich verwendeten, im Durchschnitt etwa 13 Rezepte von drei verschiedenen Ärzt*innen. Ihre Verschreibungsnetzwerke waren also deutlich größer als die von „normalen NutzerInnen“. Schließlich teilten sich die ÄrztInnen, die einem „normalen Nutzer“ Pregabalin verschrieben, im Durchschnitt 48 Patient*innen. Dieser Wert war bei den Verordnern der anderen Gruppe deutlich geringer (33 PatientInnen). Die Verschreibungsnetze von Patient*innen, die Pregabalin potentiell missbräuchlich verwendeten waren also in Bezug auf die Zahl der gemeinsam behandelten Patient*innen weniger eng miteinander verbunden. Um weitere Einblicke in die Verschreibungsnetzwerke zu erhalten, wurde die maximale geografische Entfernung zwischen den Verordnern berechnet und zwischen den beiden Gruppen verglichen. Es zeigt sich, dass bei den Patient*innen, die Pregabalin potentiell missbräuchlich nutzten, die maximale Entfernung im Durchschnitt etwa 16 Kilometer betrug, während die Verordner von „normalen Nutzern“ weniger als halb so weit voneinander entfernt waren.

6.1.4 Oxycodon/Naloxon

Für die Analyse des Ordnungsverhalten der Ärzt*innen in Bezug auf das Kombinationspräparat von Oxycodon & Naloxon (Targin™) unterschieden wir drei Patientengruppen. Die erste Gruppe bestand aus Patient*innen, die mehr als 365 Tagesdosen innerhalb eines Jahres verordnet bekommen haben, die Gruppe 3 aus Patient*innen mit nur einer Verordnung für Targin und die Gruppe 2 aus Patient*innen mit durchschnittlichen Verordnungsmengen. Es haben insgesamt ca. 38.000 PatientInnen mindestens eine Erstverordnung des Arzneimittels erhalten. Die Verteilung der insgesamt ausgestellten Verordnungsmenge gemessen in DDD der im Datensatz enthaltenen Leistungserbringer ist in Form einer Konzentrationskurve in Abbildung

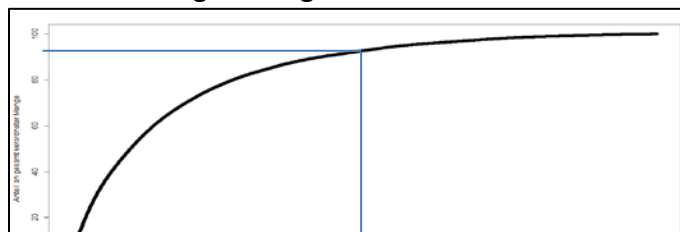


Abbildung 2: Konzentrationskurve Verordnungsmenge Targin

⁴ Die deskriptiven Ergebnisse sowie die Regressionsergebnisse sind in Form eines Papers veröffentlicht: Fleming R. Patterns of pregabalin prescribing in four German federal states: analysis of routine data to investigate potential misuse of pregabalin. *BMJ Open* 2022;12:e060104. doi: 10.1136/bmjopen-2021-060104.

abzulesen. Hier zeigt sich, dass etwa 50 % der Patient*innen mit den größten Verordnungsmengen des Arzneimittels über 90 % der insgesamt verordneten Menge Oxycodon & Naloxon ausmachen.

In Bezug auf die drei gebildeten Gruppen zeigt sich, dass nur ein sehr kleiner Anteil an Patient*innen (< 1 %) eine deutlich erhöhte Menge an Targin verordnet bekommen hat. Die Gruppe 3 mit den PatientInnen mit nur einer Verordnung innerhalb eines Jahres macht etwa 43 % der PatientInnen aus und stellte somit für die insgesamt verordnete Menge an Targin nur einen sehr geringen Teil dar (siehe Abb. 2).

Die Tabelle 11.1.4 im Anhang fasst die ermittelten Merkmale der drei Gruppen zusammen und ermöglicht sowohl den Vergleich zwischen den drei Gruppen als auch einen Vergleich zwischen den einzelnen Gruppen und der Grundgesamtheit, um so spezifische Charakteristika der einzelnen Gruppen herausarbeiten zu können.

Es zeigt sich zum Beispiel, dass auch bei dieser Analyse männliche Patienten in der Gruppe 1 überrepräsentiert sind im Vergleich zu den anderen Gruppen (45 % vs. 34 bis 36 %). Auch die Altersverteilung ist insbesondere bei der Gruppe 1 auffällig und abweichend von den anderen Gruppen mit vergleichsweise vielen Patient*innen, die zwischen 50 und 59 Jahre alt sind. Es scheint außerdem insbesondere in städtischen Regionen vermehrt zu großen Verordnungsmengen von Targin zu kommen, so wohnen etwa 46 % der Patient*innen aus Gruppe 1 in städtischen Regionen, wohingegen es in den anderen Gruppen nur etwa 40 % der Patient*innen sind.

Bei den medizinischen Ursachen, die in den Routinedaten identifiziert werden konnten, waren insbesondere Schmerzen die häufigste Indikation. Dieses Ergebnis ist in allen Gruppen zu beobachten. Auffällig ist bei den Indikationen, dass insbesondere in Gruppe 1 ein größerer Anteil (22 %) der Patient*innen mehrere der untersuchten Erkrankungen aufwies im Vergleich zu den restlichen Gruppen.

Alle Patient*innen der Gruppe 1 erhielten vor der Erstverordnung mit Targin mindestens eine Verordnung für Opioide, von den Patient*innen der Gruppe 3 erhielten das dagegen nur knapp 40 %. Eine Obstipation als Nebenwirkung von Oxycodon als Monopräparat war auch insbesondere häufig bei den Patient*innen der Gruppe 1 aufgetreten und in den Routinedaten abbildbar. Es traten auch trotz der Gabe von Targin Obstipationen im Nachhinein auf. Hier war die Rate in der Gruppe 1 mit knapp 15 % am höchsten und in der Gruppe 3 mit 6 % am niedrigsten.

Hinsichtlich der verordnenden Ärzt*innen wird ersichtlich, dass in der Gruppe 1 die Anästhesiologie nach der Allgemeinmedizin am häufigsten verantwortlich für die Erstverordnung mit Targin war. Die Anzahl der verordnenden Ärzt*innen und Praxen für die drei Gruppen ist recht gering und übersteigt im Mittelwert nicht mehr als 2,2 Ärzt*innen.

6.2 Teilprojekt B

6.2.1 Mittelwertvergleiche der Indikatoren für Wirtschaftlichkeit mithilfe von gepaarten und ungepaarten t-Tests

Aus den Analysen des Vergleichs mithilfe des gepaarten t-Tests für die Beobachtungsjahre 2010, 2013 und 2016 und den Analysen mithilfe des ungepaarten t-Tests für den Anstieg zwischen den Jahren 2013 zu 2016 in der KV Bayerns und der KV Hessen ergaben sich die folgenden Hauptergebnisse:

- Bei der Kostenentwicklung sind in Bayern, Hessen und im Bund ähnliche Entwicklungen zu verzeichnen. Es zeigt sich ein Anstieg der Kosten um 29% in Bayern, 27% in Hessen und 30% im Bund zwischen den Jahren 2010 und 2016.
- Bei Indikatoren, die sich konkret auf in Bayern gesteuerte Bereiche beziehen (Generikaquote, Erreichung der allgemeinen Leitsubstanzquote, Leitsubstanzquote im Bereich der Antikoagulantien (DOAKs)), zeigt sich, dass durch die Einführung der Wirkstoffvereinbarung in diesen Bereichen ein positiver Effekt im Sinne der Wirtschaftlichkeit in der KV Bayerns messbar ist.
-

6.2.2 Hauptergebnisse des Strukturbruchmodells nach Chow:

Die Analyse des Strukturbruchmodells nach Chow ergab die folgenden Ergebnisse (siehe auch Anlage 12.1):

Tabelle 10: Vorliegen eines Strukturbruchs in der KV Bayerns und KV Hessen

Indikator	Strukturbruch in Bayern	Strukturbruch in Hessen
Kosten je Verordnung	nein	nein
Kosten je Patient	nein	nein
Verordnungsanteil von Generika am Gesamtmarkt	ja	nein
Patientenanteil mit DOAK-Verordnung an allen Patienten mit Antikoagulantienverordnung	ja	nein
Erreichung der allgemeinen Leitsubstanzquote	ja	ja
Anteil generischer DDDs im Bereich Analgetika außer BTM-pflichtige Opioide	nein	nein
Anteil generischer DDDs im Bereich der Mittel bei obstruktiven Atemwegserkrankungen	ja	ja
Anteil generischer DDDs im Bereich der Mittel mit Wirkung auf das Renin-Angiotensin-System	ja	ja
Anteil generischer DDDs im Bereich Thrombozytenaggregationshemmer (ohne Ticagrelor)	nein	nein
Anteil der bayerischen Leitsubstanz-DDDs in der Gruppe der MS-Therapeutika	ja	ja

Die Ergebnisse des Strukturbruchmodells nach Chow der Indikatoren liefern die folgenden Schlussfolgerungen:

- Die Tendenzen der Kostenentwicklung sind in Bayern und Hessen ähnlich und es liegt kein Strukturbruch zum Einführungszeitpunkt der Wirkstoffvereinbarung vor. Das bedeutet auch, dass trotz der Abkehr von der Steuerung der Kosten hin zu Generika- und Leitsubstanzquoten in Bayern die Gesamtkosten (gesteuerter und ungesteuerter Markt) nicht überproportional gestiegen sind.
- Die Steuerung wirkt: in den allgemeinen gesteuerten Bereichen Generika, Antikoagulantien (DOAKs) und Erreichung der allgemeinen Leitsubstanzquote ist global ein Strukturbruch in Bayern messbar. Allerdings ist das aufgestellte Modell für die Erreichung der allgemeinen Leitsubstanzquote kaum aussagefähig, da die dort auftretende schlechte Modellqualität zu berücksichtigen ist. Dies ist vermutlich auf die heterogenen Entwicklungen in den einzelnen Substanzklassen zurückzuführen.
- Bei faktisch allen Indikatoren mit Bezug zur Wirkstoffvereinbarung ist bei Kombination der Ergebnisse des Strukturbruchmodells und der Ergebnisse der Mittelwertvergleiche eine positive Entwicklung innerhalb der KV Bayerns für die bayerischen Ziele zu verzeichnen. Auch hier wirkt die Steuerung.
- In Hessen ist hingegen weder beim Indikator Antikoagulantien (DOAKs) (nicht gesteuert) noch Generika ein Strukturbruch identifizierbar. Im Bereich Erreichung der allgemeinen Leitsubstanzquote ist in Hessen global ein Strukturbruch messbar. Die Veränderung der Leitsubstanzquote ist jedoch (im Vergleich zu Bayern) trotz Strukturbruch eher gering.

6.3 Teilprojekt C

6.3.1 Studienpopulation

Es wurden Einzelinterviews mit 20 Ärzt*innen und sieben Fokusgruppen mit insgesamt 36 Ärzt*innen zwischen 10/2019 und 10/2020 durchgeführt. Die Ärzt*innen waren 41 bis über 60 Jahre alt. Verschiedene Fachrichtungen waren vertreten: Hausärztliche Medizin, fachärztlich Innere Medizin (Pneumologie, Kardiologie, (Hämato)Onkologie, Rheumatologie, Gastroenterologie, Nephrologie), Neurologie/Psychiatrie und Chirurgie. Verschiedene Praxistypen, wie Einzel- oder Gemeinschaftspraxen, und Standorte (ländlich, klein/mittelgroß, Großstadt) waren vertreten.

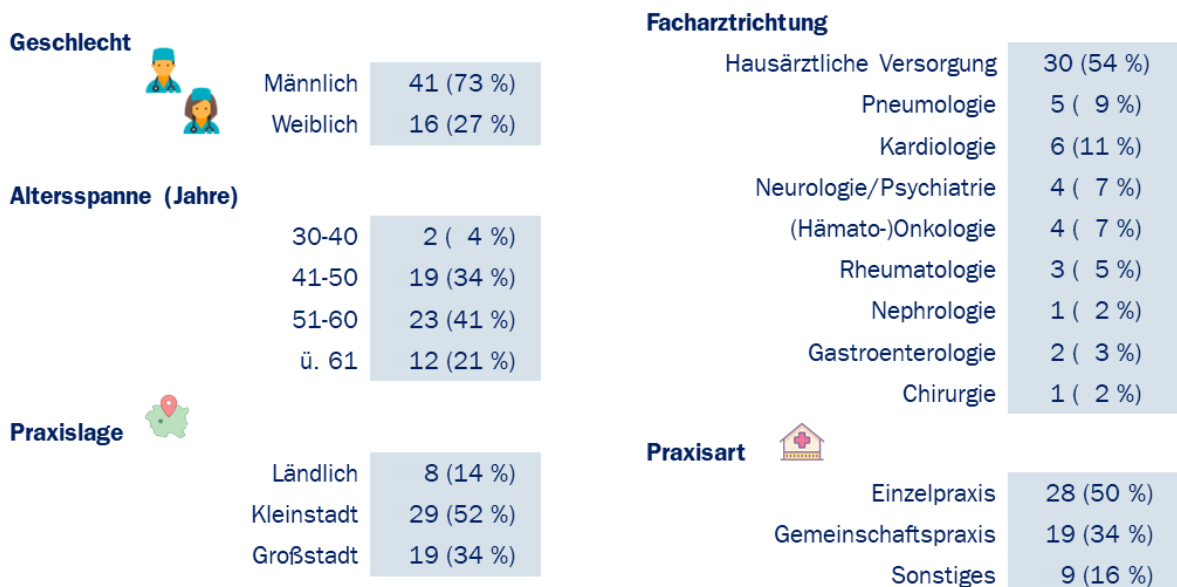


Abbildung 3: Soziodemographische Angaben zu den Ärzt*innen der Einzelinterviews und Fokusgruppen

6.3.2 Wahrnehmung der WSV

Es ging hier darum, wie das Programm der KVB von den niedergelassenen Vertragsärzt*innen wahrgenommen wird; zu welchen Reaktionen das Feedback führt; welche Schwierigkeiten, Barrieren und Nachteile die niedergelassenen Ärzt*innen sehen.

Grundsätzlich sehen die Ärzt*innen wirtschaftliches Handeln als notwendig an. Dabei scheinen einige von Ihnen das Wirtschaftlichkeitsprinzip verinnerlicht (TN-03) zu haben. Dennoch sehen sich auch manche Ärzt*innen „nicht verantwortlich“ für das Gesamtsystem (FG-03). Durch die Vorgabe der Wirtschaftlichkeit befinden sie sich dauerhaft in einem Spannungsfeld: „Wir sind in diesem Spannungsfeld [...] zwischen hohen Anforderungen, anspruchsvollen Patienten, anspruchsvollen Leitlinien auf der einen Seite. Und auf der anderen Seite den harten wirtschaftlichen Zwängen. Ich fühle mich da täglich in der Zwickmühle.“ (FG-06).

Die Bewertung der Ärzt*innen zur WSV 2.0 fällt insgesamt positiv aus.

Folgende positiven Aspekte der quartalsweisen Trendmeldungen nennen die Ärzt*innen:

- Zeitnahes Feedback durch die Trendmeldungen
- Leicht verständlich
- Zeitökonomisch
- Anregend zur Selbstreflexion
- Mehr Kontrollmöglichkeiten
- Mehr Gerechtigkeit
- Weniger Regresssorgen
- Einführung der WSV bietet nur Vorteile im Vergleich zur Richtgrößenprüfung

Die Ärzt*innen üben auch Kritik. Dabei fällt auf, dass Kritik insgesamt weniger von Hausärzt*innen als von anderen Facharzttrichtungen berichtet wird.

Der größte Kritikpunkt fast aller befragten Ärzt*innen sind die Rabattverträge. Hier herrscht großes Unverständnis mit vielerlei Begründungen. Die Rabattverträge sind/werden:

- schwierig für die Patient*innen, da sie infolgedessen häufig neue Präparate bekommen (Verwechslungen von Medikamenten, doppelte Einnahme oder gar keine).
- schwierig für die Ärzt*in durch Mehraufwand (Suche im System nach Rabattarzneimittel; zusätzlich abhängig von Kasse). Laut Selbstbericht ist nicht immer eine einfache Entscheidungsabgabe an die Apotheke möglich.
- undurchsichtig und sind laut Ärzt*innen nicht sinnvoll (Infrage stellen von Kosten-Nutzen; auch für das Gesamtsystem).
- als Ungleichbehandlung wahrgenommen („Wir als Ärzteschaft sollen alles Offenlegen und die Kassen dürfen Ihre eigenen Sachen abwickeln“).
- verbunden mit der Forderung, sie alle abzuschaffen oder für alle Kassen zu vereinheitlichen (FG-06).

Weitere Kritikpunkte sind:

- Wenig(er) Therapiefreiheit
- Fehlende Transparenz zu Kriterien und deren Entstehung
- (weiterhin bestehende) Regressorgen
- Keine Leitlinienorientierung
- Hemmung innovativer Medikamente
- Fehlende Praxisnähe
- Keine Mehraufwandsvergütung

Wenngleich die Ärzt*innen Kritik äußern, so überwiegen die positiven Aspekte für sie, allerdings nur bei der WSV 2.0. Bereits bestehende Kritik an der WSV 2.0 hat sich laut den befragten Ärzt*innen unter 3.0 verschärft.

Manchen Ärzt*innen stehen der Einführung der WSV aus folgenden Gründen gleichgültig gegenüber:

- Allgemeine Orientierung an Wirtschaftlichkeitsgebot -> WSV hat keine Auswirkung auf ihre Arbeit
- Gleichgültigkeit aufgrund höheren Alters -> führt zu:
 - weniger finanziellen Sorgen und/oder Akzeptanz gegenüber regelmäßigen Vorgabenänderungen
 - fehlendem Interesse an und Wissen zur WSV

6.3.3 Welche Anregungen lassen sich aus den Erfahrungen für zukünftige Regelungen ableiten?

Die Verbesserungsvorschläge, sowohl von den Ärzt*innen selbst genannt, als auch Ableitungen aus dem Datenmaterial lassen sich in drei Bereiche gliedern:

- a) Verbesserungen der Kommunikation seitens der Kassenärztlichen Vereinigung (Inhalt und Form)
- b) Verbesserungen in der Arzneimittelvereinbarung im Allgemeinen und
- c) in der Trendberichterstattung im Besonderen

Zu a): Bei der inhaltlichen Kommunikation wünschen sich die Ärzt*innen mehr Erläuterungen zu den Neuerungen in der Trendberichterstattung, zu den arzneimittelvertragskonformen Leitlinien sowie eine pharmaindustrieunabhängige Arzneimittelbewertung. Hinsichtlich des

Formats bewerten sie die Beratungen mit den KV-Pharmazeuten positiv, ebenso wie die Einführungsschulung für neue Ärzt*innen. Sie wünschen sich einen anderen Informationstransfer als E-Mails. Die offizielle Website sollte intuitiver gestaltet sein, damit wichtige Informationen zu Medikamenten schnell abgerufen werden können.

Zu b): Die Ärzt*innen wünschen sich statt einer Anweisung lieber einen Handlungsleitfaden. Die Ärzte möchten an der Entwicklung der Kriterien beteiligt werden. Sie fordern mehr Transparenz bei der Kriterienentwicklung. Einige Ärzt*innen wünschen sich mehr Flexibilität und Ausnahmen (z.B. Schweregradeinstufung). Einige wünschen sich eine dynamischere Anpassung an neue Gegebenheiten, andere wiederum mehr Stabilität. Die Rabattverträge sollten hier gesondert betrachtet werden. Die Vorschläge reichten von der Abschaffung bis zur Vereinheitlichung von Rabattverträgen (kassenübergreifend) und für alle Packungsgrößen geltend.

Zu c): Die Ampel in der Trendberichterstattung sollte beibehalten werden, mit der Möglichkeit der Fehlermeldung an die KV Bayerns. Weitere Aufschlüsselungen sollten möglich sein, wie z.B. ein Vergleich von Praxen oder der prozentuale Einfluss eines einzelnen Ziels auf das Gesamtergebnis.

6.3.4 Welche Auswirkungen auf Verordnungsverhalten, Kosten, Überweisungen (Ausweichstrategien), Versorgungstiefe haben die bayerische Wirkstoffvereinbarung (einschl. Feedback) sowie von anderen KVen implementierte Verfahren?

Insgesamt zeigt die qualitative Erhebung, dass die Einführung der WSV bei manchen Ärzt*innen zu einer Veränderung des Verordnungsverhaltens führt. Im Folgenden werden die unterschiedlichen Verhaltensweisen getrennt nach „Veränderung“ und „keine Veränderung im Verordnungsverhalten“ dargestellt. Es werden Umstellungsschwierigkeiten sowie mögliche Ausweichstrategien dargestellt.

Veränderung im Verordnungsverhalten: Neu- und Umstellung

Einige Ärzt*innen reagieren auf das Feedback bzw. die Rückmeldungen mit Neu- oder Umstellung der Medikation. Dabei kommt es häufig zu einer Umstellung auf Generika sowie auf Leitsubstanzen oder Rabattvertragspräparate (TN-06). Außerdem verordnen sie vermehrt N3 statt N1-Packungsgrößen, beispielsweise bei generischen Blutdruckmitteln, oder stellen Kassen- statt Privatrezepte, wie z.B. bei Ibu-600 aus. Kombinationspräparate sowie Aut-idem-Kreuze meiden viele Ärzt*innen zunehmend (TN-09).

Keine Veränderung im Verordnungsverhalten

Diesen stehen auch einige Ärzt*innen, die ihr Verordnungsverhalten nicht ändern. Diese Ärzt*innen führen als Begründung an, dass keine Alternativen vorhanden seien, was die Bio-Äquivalenz von Generika oder die Wirkstoffzusammensetzung von beispielsweise Schilddrüsenmedikamenten betrifft (FG-07). Außerdem wird die Beibehaltung der Medikation durch patient*innenbezogene Faktoren begründet (FG-07, TN-06). Vereinzelt geben Ärzt*innen an, den Austausch von Medikamenten über die Verwendung von Aut-idem zu vermeiden (TN-06). Einige Ärzt*innen stützen sich eher auf Leitlinien und die Indikation sowie auf Selbst- und Fremdkompetenz (TN-04).

Regresssorgen unter der WSV

Die Regresssorgen der Hausärzt*innen (HÄ) haben nach Einführung der WSV mehrheitlich abgenommen. Dabei thematisieren die Ärzt*innen die Regresssorgen oftmals von alleine. Es zei-

gen sich dennoch drei Verordnungstypen: HÄ mit geringen Regresssorgen, mit größeren Regresssorgen sowie Resignierte. Andere Facharzttrichtungen finden sich vermehrt in den Gruppen größere Regresssorgen und Resignierte wieder. Diesen Typen lassen sich Tendenzen im Verordnungsverhalten zuordnen. Ärzt*innen mit geringen Regresssorgen haben Vertrauen in die WSV. Auch vereinzelt rote Punkte machen ihnen keine Sorgen (TN-10:81). Sie sichern sich bei kostenintensiven Weiterverordnungen vorsorglich über eine gute Dokumentation vor Prüfverfahren ab (TN-10:79). Ärzt*innen mit größeren Regresssorgen suchen zu vielen teuren Medikamenten gleichwertige Alternativen. Das Nicht-(Weiter)verordnen führt laut HÄ in den meisten Fällen zu grünen Punkten. Größere Regresssorgen gehen mit größerem Fokus auf Wirtschaftlichkeit einher (TN-07:79). Eine chronische Hilflosigkeit zeigt sich bei der dritten Gruppe: „*Da war ich eine Zeit lang grün, jetzt bin ich langsam gelb. [...] Letztendlich, bei der Prüfung, muss man sagen, das ist nicht umsetzbar. [...] Das ist einfach völlig wirklichkeitsfern.*“ (TN-20:51). Dementsprechend resignieren manche und verordnen – sich im Teufelskreis befindend – auch teure Medikamente weiter (TN-19:30).

Umstellungsschwierigkeiten

Die Ärzt*innen äußern auch, dass die Umstellung des Verordnungsverhaltens nach dem Feedback schwierig sein kann. Diese Schwierigkeiten sind zum einen auf Schnittstellenprobleme zurückzuführen. Als Problem wird hierbei die Umsetzung bzw. Umstellung von Krankenhaus- bzw. Facharzt*innen-Therapien angeführt. Auch an der Schnittstelle mit der Apotheke treten manchmal Probleme auf. Es befindet sich eine Publikation im Review-Verfahren, welche die Schnittstelle stationär/ambulant für die Hausarzt*innen erläutert. Allgemeine Ein- und Umstellungsschwierigkeiten begründen Ärzt*innen durch den enormen Aufwand, welcher durch die Neuverordnungen entsteht. Dabei wird sowohl ein hoher logistischer Aufwand, als auch ein hoher zeitlicher Aufwand für die Aufklärung bzw. Überzeugung der Patient*innen berichtet.

Strategien

Die Ärzt*innen berichten nur selten davon, Ausweichstrategien zu verwenden. Überweisungen werden beispielsweise aufgrund der Gesamtzieelerreichung durchgeführt, was zu einem „Verschiebebahnhof“ für teure Medikamente wie z.B. Biologika führt (TN-17). In den Daten finden sich vereinzelt Aussagen dazu, dass die Verordnung von Medikamenten zwischen Facharzt*in und Hausarzt*in geteilt wird, z.B. bei Inhalativa. Daneben werden auch Absicherungsstrategien angewendet (Dokumentation, insbesondere die richtige Indikation) (FG-05). Außerdem berichten viele Ärzt*innen, die KV-Beratung regelmäßig in Anspruch zu nehmen. Seltener dagegen werden Beratungen durch Pharma(-vertreter*innen) oder Jurist*innen wahrgenommen (z.B. TN-01). Als sehr selten vorkommende Strategie berichten Ärzt*innen von sog. Verdünnerscheinen.

6.4 Teilprojekt D

6.4.1 Qualitativer Studienarm

Es wurden 26 Ärzt*innen aus verschiedenen Fachrichtungen, Praxisarten und unterschiedlichen Alters befragt. Die 5 weiblichen und 21 männlichen Ärzte arbeiteten sowohl in Einzel- als auch Gemeinschaftspraxen, aber auch in MVZ und anderen Praxisformen. Die jeweiligen Betriebsstätten lagen in verschiedenen geographischen Regionen, die sowohl ländliche, als auch mittel- und großstädtische Regionen umfassten. Eine detaillierte Stichprobenbeschreibung ist im Anhang 11.4.4 zu finden.

Eine detailliertere Beschreibung der Stichprobe ist der Anlage TP-D 4: Stichprobenbeschreibung zu entnehmen.

Die Auswertung und Abstraktion des Interviewmaterials ergibt drei Ebenen, den verschiedene Einflussfaktoren für Auffälligkeit zugeordnet werden können: Arzt – System der bayerischen Wirkstoffvereinbarung – Gesundheitssystem.

Arzt

Auf der Ebene des Arztes sind vor allem die persönlichen Wertevorstellungen und Ansprüche, mit den Ärzt*innen in den Praxisalltag gehen, relevant. Für die interviewten Ärzt*innen ist das patientenorientierte, verantwortungsbewusste Ordnungsverhalten und das Arbeiten auf dem aktuellen Stand der Wissenschaft die am häufigsten genannten Ansprüche an die eigene Arbeit. Weiterhin werden Eigenschaften wie „Hohe Motivation zur Arbeit“ und das „Eiserne Nervenkostüm“, Eigenschaften, die für persönliche Stärke stehen, wichtig. Auch wenn Wirtschaftlichkeit für die Ärzt*Innen durchaus relevant ist, so gehört sie dennoch nicht zu den höchsten Leitmotiven für die ärztliche Arbeit.

Für eine genaue Auflistung der von den Ärzten angegeben wichtigen Eigenschaften für Ärzt*innen und deren Häufigkeiten siehe Anlage zu TP-D: Rolle als Ärzt*in, wichtige Ärzt*innen-Eigenschaften.

Einen weiteren Einflussfaktor auf der Arztebene stellt der Informationsstatus dar. Sowohl hinsichtlich der Regularien der Bayerischen Wirkstoffvereinbarung als auch bezüglich der Studienlage zu verschiedenen Arzneimitteln gibt es keinen einheitlichen Informationsstand unter den Ärzt*innen, einigen fehlen detaillierte Informationen zu den zu erfüllen Quoten und das Verständnis für das Zustandekommen der Ziele. Zum Teil stimmen die Informationen, die die Ärzt*innen haben, nicht mit den gültigen Regularien über ein – beispielhaft sei hier die Annahme erwähnt, dass der Fachgruppenausgleich weggefallen sei, der aber noch gilt. Teilweise werden die Vorgaben der zuvor gültigen Richtgrößenprüfung mit den Vorgaben zur der aktuellen bayerischen Wirkstoffvereinbarung vermischt. Problematisch sei hier häufig, dass die Informationen zwar in der Praxis angekommen, aber in dem Praxisalltag und der großen Menge an Post untergehen.

Neben den Informationen rund um das Steuerungssystem, sind auch die Informationen zu konkreten Arzneimitteln nicht einheitlich.

Auch bisherige Regresserfahrungen beeinflussen die Ärzt*innen in ihrem Ordnungsalltag und werden als teils ständig begleitende Bedrohung empfunden.

Die Zufriedenheit der Ärzt*innen mit ihrer Ordnungssituation wird stark davon beeinflusst wie gut Ärzt*innen ihre Patient*innen versorgt wissen, andererseits aber auch von der Auffälligkeitssituation. Dies führt zu einer überwiegend mittelmäßigen Zufriedenheit mit der aktuellen Ordnungssituation (siehe Anlage TP-D: Zufriedenheit mit der eigene Ordnungssituation).

Bayerische Wirkstoffvereinbarung

In der Wahrnehmung der Ärzt*innen führt ein regulierendes System häufig zu schlechter Zusammenarbeit zwischen Ärzt*innen und zu entwickelten Ausweichstrategien. In ihren Augen fangen nur wenige Ärzt*innen viele der unwirtschaftlichen Verordnungen auf, was zu deren Auffälligkeit führt.

Auch hier hat der Informationsstatus wieder einen Einfluss. Manche Ärzt*innen sind von ihren Informationen zu verschiedenen Medikamenten überzeugt und verordnen nach diesem Informationsstatus. Differenziert ihre Meinung zu Arzneimitteln von den Informationen, die für die Entwicklung der Bayerischen Wirkstoffvereinbarung zu Grunde liegen, führt dies zu Verordnungen, die nicht im Sinne der Bayerischen Wirkstoffvereinbarung sind.

Auch gewisse Eigenschaften der Bayerischen Wirkstoffvereinbarung führen aus der Sicht der befragten Ärzt*innen zu Schwierigkeiten bei der Zielerreichung. Problematisch sind vor allem folgende Aspekte:

Tabelle 11: Potentiell schwierige/nachteilige Systemeigenschaften aus der Sicht der befragten Zielverfehlenden

Dokumentationsaufwand immens und viel zu hoch

Gut eingestellte Langzeittherapien müssen teils umgestellt werden, um Vorgaben annähernd zu erreichen

Ausweichstrategien werden entwickelt

Schon wenige Patient*innen/Verordnungen haben große Auswirkungen, wodurch die Ampelpunkt für die Ärzt*innen an Aussagekraft verlieren

System ist zu komplex und zu kompliziert, wodurch es undurchsichtig wirkt

Rabattverträge und Änderungen kommen als Information bei den Ärzt*innen häufig schlecht an

Informationen zum System und seinen Inhalten und Vorgaben kommen bei Ärzten häufig schlecht an

Ziele werden zu schnell und zu stark verändert, die Umstellung und Anpassung von Therapien benötigte teils mehr Zeit

Ungerechtigkeit, dass manche Fachgruppen viele und andere wenige Wirkstoffmöglichkeiten und damit kaum Ausgleichsmöglichkeiten haben

Steuerungsmechanismus ist zu undifferenziert und nur auf Standardpatienten ausgelegt

Abhängigkeit von sich ständig ändernden Rabattverträgen sei zu hoch

Vergleichsgruppe erscheint einigen Ärzt*innen zu inhomogen

Die Ärzt*innen gaben an, für Therapien gerade stehen zu müssen, die sie eigentlich nicht ansetzen würden. Durch den Wechsel auf die Arzneimittel im Sinne der Vorgaben der Bayerischen Wirkstoffvereinbarung, müsse man den Patienten den Wechsel immer wieder erklären und damit sich seine ärztliche Kompetenz in Frage stellen. Auch mit der Auswahl der Arzneimittel sind die interviewten Ärzt*innen nicht immer zufrieden bzw. einverstanden gewesen.

Gesundheitssystem

Auch auf das gesamte Gesundheitssystem bezogen, gibt es in den Augen der Ärzt*innen Aspekte, die die Verordnungssituation verbessern könnten. Einerseits ändert sich die Marktlage immer wieder, Arzneimittel werden vom Markt genommen, manche Arzneimittel gibt es bestimmten Stärken und Packungsgrößen nur von einer einzigen Firma oder Lieferprobleme beeinträchtigen die Verfügbarkeit. Dieser Aspekt wird in den meisten Steuer- und Regelsystem

nicht abgebildet und kann die Verordnungssituation erschweren. Andererseits wünschen sich die Ärzt*innen mehr Vertrauen und weniger Reglementierung, eine nicht immer direkt anklagende Kommunikation und eine Praxissoftware, die direkt für den Patienten und die Indikation schon das richtige Arzneimittel im Sinne einer wirtschaftlichen Verordnung anzeigt – ggf. nach Vorgabe des Wirkstoffes oder aber auch direkt nur auf die Indikation bezogen. Diesen Vorschlag könne man dann annehmen oder sich bewusst und begründet für ein anderes Arzneimittel entscheiden.

Zusätzlich wird gefordert, dass die Krankenkassen mehr Öffentlichkeitsarbeit im Sinne einer Aufklärung über Arzneimittelverordnung leisten. Die derzeitigen Informationen, mit den Patienten in die Arztpraxen kommen, würden besagen, dass ein/e Ärzt*in in einem jedes Arzneimittel verordnen könne und dürfe. Dies führt häufig zu Spannungen in den Arzt-Patienten-Beziehungen. Eine Aufklärungsarbeit, dass Patient*innen informiert werden, dass Ärzt*innen gewisse Vorgaben bei der Arzneimittelverordnung zu erfüllen haben, könne hier Zeit für Erklärungen in der Praxis sparen und das Vertrauen und damit die Compliance der Patient*Innen fördern. Jede nicht ausgestellte Überweisung für Verordnungen und auch jedes nicht durchgeführt Arzt-Shopping würde Folgekosten wie neu ausgelöste Fallpauschalen und Medikationsfehler und –missverständnisse sparen.

6.4.2 Quantitativer Studienarm

Im Rahmen der deskriptiven Analysen von Routinedatensätzen, wurde untersucht, wie hoch der Anteil der jeweiligen Zielverfeher in den einzelnen Fachgruppen ist (siehe Anlage TP-D: Fachgruppenverteilung).

Das Ergebnis zeigt, dass vor allem Fachgruppen mit einem geringen Verordnungsspektrum und – umfang die oberen Plätze der Fachgruppen mit den höchsten Zielverfeher-Anteilen bilden.

Weiterhin erscheinen in der Gruppe der Zielverfeher die Charakteristika „männlicher Arzt“, „städtische Praxislage“, „Fallzahl pro Quartal max. 600“ und „keine Zusatzbezeichnungen“ mit signifikant höherer Häufigkeit als in der Vergleichsgruppe der Zielerreicher (siehe Anhang 11.4.8: Vergleich ZE-ZV hinsichtlich Charakteristika mit potenziellem Einfluss auf Auffälligkeit).

Die Charakteristika „Anzahl Arztsitze pro Betriebsstätte“ und „Alter“ zeigen keine erhöhte Auffälligkeit in einer der beiden Gruppen.

Zusätzlich zu den deskriptiven Analysen von Arzt- und Praxischarakteristika wurden auch Verordnungen betrachtet.

In Bayern gibt es 25 verschiedene Fachgruppen, die Arzneimittel im Rahmen der verschiedenen Einzelziele verordnen. Betrachtet man die Anzahl relevanter Einzelziele, Ziele mit hohem Verordnungsumfang, so stellt man fest, dass es Fachgruppen gibt, für die viele Einzelziele relevant sind. In der Gruppe der Hausärzte werden 23 verschiedenen Einzelziele bedient. Dem gegenüber stehen Fachgruppen, die nur wenige Einzelziele bedienen, wie beispielsweise die Augenärzte, die nur zwei verschiedene Einzelziele bedienen. Insofern bestehen für die verschiedenen Fachgruppen unterschiedliche Ausgleichsmöglichkeiten von verfehlten und erreichten Einzelzielen bzw. verfehlte Einzelziele fallen unterschiedlich stark ins Gewicht.

Für acht ausgewählte Fachgruppen zeigt sich weiterhin, dass die verfehlten Ziele meist Ziele mit geringem Verordnungsumfang sind, während hingegen die je drei umfangreichsten Ziele pro Fachgruppe in der Regel erreicht werden. Die umfangreichsten Ziele sind vorwiegend Generikaziele, bei den verfehlten Zielen hingegen ist der Anteil an Leitsubstanzziele erhöht. In

der Fachgruppe der Kinder- und Jugendmedizin werden auffällig viele Ziele nicht erreicht. Das Ziel der Knochenerkrankungen, in das die Osteoporose fällt, zeigt als TOP 3 Arzneimittel, die der Zielerreichung entgegenstehen, nur parenterale Darreichungsformen; die Wahl einer parenterale Therapie, z.B. zur Umgehung der bekannten oralen Nebenwirkungen, führt also zwangsläufig zur Einstufung als unwirtschaftlichen Arzneimittelverordnung.

6.5 Teilprojekt E

6.5.1 Basisanalysen

Durchschnittliche Kosten je Verordnung

Für den Indikator der durchschnittlichen Kosten je Verordnung wurde die Analyse nach der beschriebenen Analyseprozedur durchgeführt, wobei die in Tabelle 12 aufgelisteten Confounder der Basisanalyse berücksichtigt wurden.

Mit einer gewählten Kontrollgrenze (control limit) von 5% wurden bei 2667 betrachteten Betriebsstätten insgesamt die folgenden Ergebnisse gefunden (siehe Tabelle 12 und Abb. 4).

Tabelle 12: Aufteilung der Hausarzt-Betriebsstätten in die Kategorien der Auffälligkeiten für den Indikator durchschnittliche Kosten je Verordnung in der für das Quartal 4/2016

Kategorie der Auffälligkeit	Anzahl auffälliger Betriebsstätten	prozentualer Anteil
„natürlicher“ Outlier	134	5%
„O/E“-Outlier ⁵	129	4,8%
sowohl „natürlicher“ als auch „O/E“-Outlier ⁶	72	2,7%
auffällig nach der „natürlichen“ Verteilung aber nicht nach dem Machine-Learning-Modell	62	2,3%
nicht auffällig nach der „natürlichen“ Verteilung aber auffällig nach dem Machine-Learning-Modell	57	2,1%

⁵ Outlier gemäß des Observed/Expected-Indikators

⁶ Outlier gemäß der „natürlichen“ Verteilung als auch des Observed/Expected-Indikators

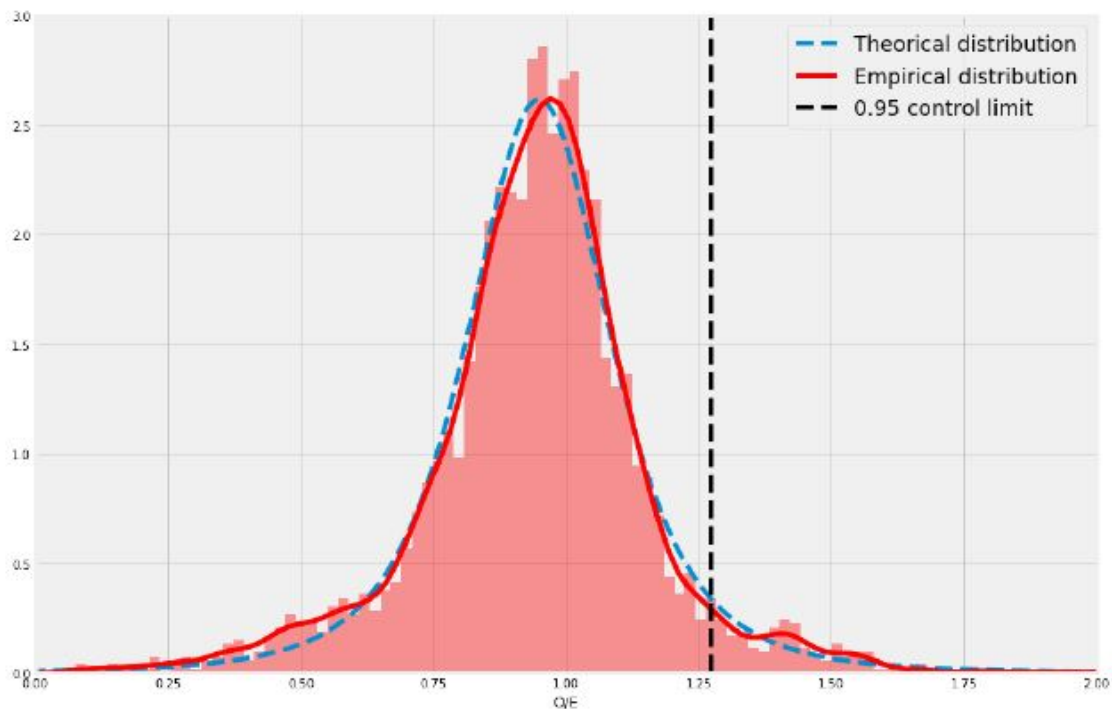


Abbildung 4: O/E-Indikator der durchschnittlichen Verordnungskosten im Quartal 4/2016. Rot abgetragen ist die empirische Verteilung (empirical distribution), die mittels Kerndichteschätzung (kernel density estimation) geschätzt wurde – blau abgetragen ist die theoretische Verteilung (theoretical distribution). Zudem ist schwarz abgetragen die 95%-Kontrollgrenze (control limit), die mit der Bayesschen Inferenz erzeugt wurde.

In Abb. 4 ist die Menge aller betrachteten Betriebsstätten (roter Bereich) unter dem Gesichtspunkt ihrer potentiellen Auffälligkeit bzgl. ihrer durchschnittlichen Kosten je Verordnung dargestellt. Betriebsstätten, sind Outlier gemäß des Observed/Expected-Indikators, wenn sie sich in denjenigen Bereichen der Abbildung befinden, in denen die rote Verteilungsfunktion über der blauen liegt. Dort fallen die beobachteten Werte (Observed-Werte) der Betriebsstätten höher aus, als vom Modell vorhergesagt (Expected-Werte). Die Outlier gemäß der „natürlichen“ Verteilung befinden sich rechts von der Kontrollgrenze. Betriebsstätten, die sowohl gemäß des Observed/Expected-Indikators als auch der „natürlichen“ Verteilung auffällig sind, befinden sich rechts von der Kontrollgrenzen in dem Bereich, wo die rote Verteilung über der blauen liegt. Betriebsstätten in diesem Bereich sind eindeutig auffällig. Erfüllt eine Betriebsstätte keine der hier genannten Auffälligkeitskriterien, so ist sie als eindeutig unauffällig einzustufen.

Global wurden für diesen Indikator die folgenden Haupteinflussgrößen mittels der SHAP Values bestimmt (siehe Abb. 5 und Tabelle 13):

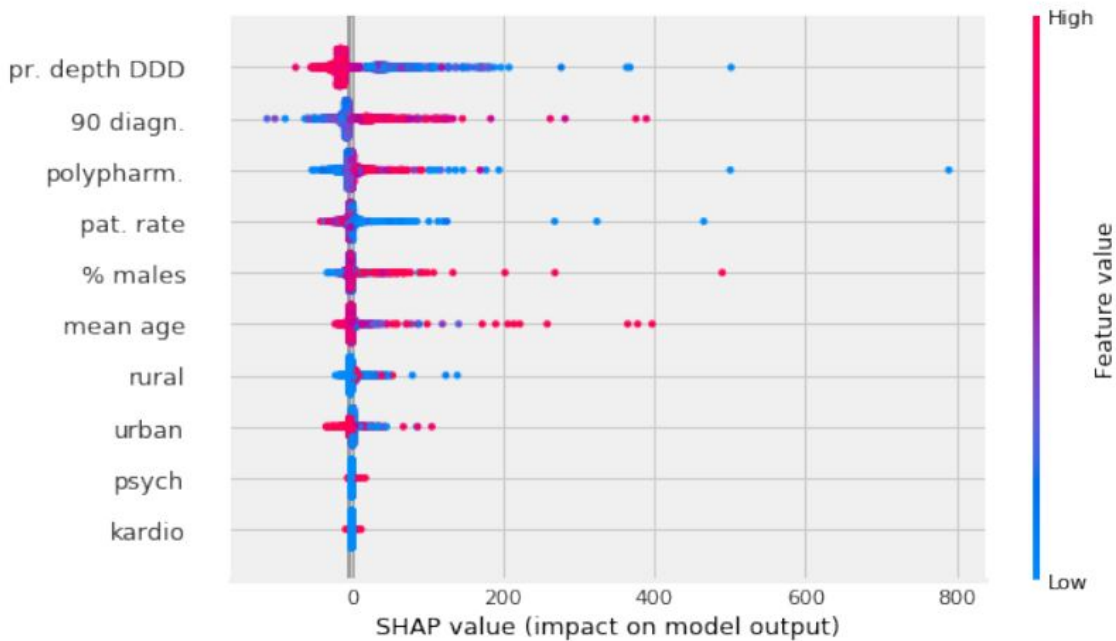


Abbildung 5: SHAP summary plot der durchschnittlichen Verordnungskosten im Quartal 4/2016 – absteigend gelistet sind die Haupteinflussfaktoren auf die durchschnittlichen Verordnungskosten über alle betrachteter Betriebsstätten hinweg. Mit der farblichen Abgrenzung (roter Impact= hoher Impact und blauer Impact= geringer Impact) wird die jeweilige Impactrichtung abgebildet.

In Abb. 5 werden die Haupteinflussvariablen absteigend nach ihrem Impact sortiert dargestellt. Die Einflussvariablen sind dabei mit ihrem Impact (roter Impact= hoher Impact und blauer Impact= geringer Impact) und die jeweilige Impactrichtung abgebildet. So bildet jede Betriebsstätte in jeder Variablenzeile jeweils einen eigenen Punkt mit einer eigenen Farbe und einer eigenen Impactposition (stark positiv/negativ oder wenig positiv/negativ).

Bei der ersten Variable unseres Rankings, der Verordnungstiefe, ist eine starke Häufung von roten Punkten im negativen Bereich festzustellen, was bedeutet, dass eine hohe Verordnungstiefe die durchschnittlichen Verordnungskosten sinken lässt. Demgegenüber lassen hohe Werte beim Anteil der Patient*innen mit Diagnoseanzahl über dem 90%-Perzentil, beim Anteil von Polypharmaziepatienten und beim Anteil von männlichen Patienten die durchschnittlichen Verordnungskosten eher steigen. Die übrigen Variablen sind mit diesem Verfahren allerdings schlechter zu interpretieren.

Tabelle 13: Haupteinflussfaktoren auf die durchschnittlichen Verordnungskosten im Quartal 4/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable	Einflussart (Impactrichtung)
Verordnungstiefe	Negativ
Anteil der Patient*innen mit Diagnoseanzahl über dem 90%-Perzentil	Positiv
Anteil von Polypharmaziepatienten	Positiv
Anteil von männlichen Patienten	Positiv

Verordnungsanteil von Generika am Gesamtmarkt

Für den Indikator des Verordnungsanteils von Generika am Gesamtmarkt wurde die Analyse ebenfalls mit den in Tabelle 12 aufgelisteten Confoundern der Basisanalyse durchgeführt.

Mit einer gewählten Kontrollgrenze (control limit) von 5% wurden bei 2664 betrachteten Betriebsstätten insgesamt die folgenden Ergebnisse gefunden (siehe Tabelle 14 und Abb 6).

Tabelle 14: Aufteilung der Hausarzt-Betriebsstätten in die Kategorien der Auffälligkeiten für den Indikator Verordnungsanteil von Generika am Gesamtmarkt für das Quartal 4/2016

Kategorie der Auffälligkeit	Anzahl auffälliger Betriebsstätten	prozentualer Anteil
„natürlicher“ Outlier	234	8,7%
„O/E“-Outlier ⁷	192	7,2%
sowohl „natürlicher“ als auch „O/E“-Outlier ⁸	163	6,1%
auffällig nach der „natürlichen“ Verteilung aber nicht nach dem Machine-Learning-Modell	71	2,7%
nicht auffällig nach der „natürlichen“ Verteilung aber auffällig nach dem Machine-Learning-Modell	29	1,1%

⁷ Outlier gemäß des Observed/Expected-Indikators

⁸ Outlier gemäß der „natürlichen“ Verteilung als auch des Observed/Expected-Indikators

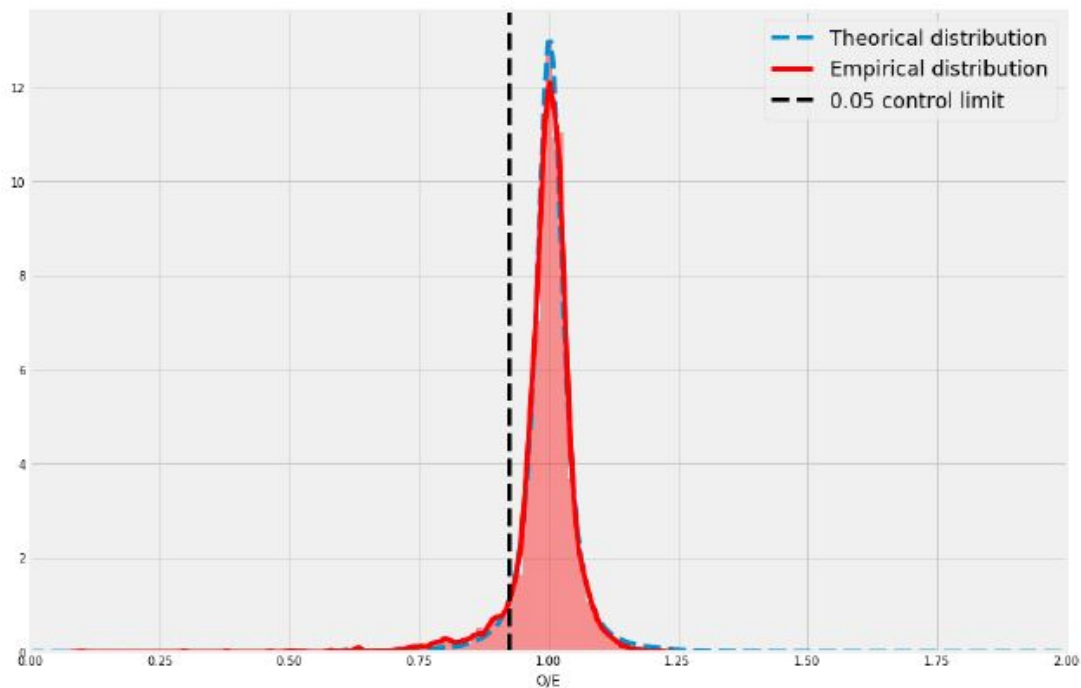


Abbildung 6: O/E-Indikator des Ordnungsanteils von Generika am Gesamtmarkt im Quartal 4/2016. Rot abgetragen ist die empirische Verteilung (empirical distribution), die mittels Kerndichteschätzung (kernel density estimation) geschätzt wurde – blau abgetragen ist die theoretische Verteilung (theoretical distribution). Zudem ist schwarz abgetragen die 95%-Kontrollgrenze (control limit), die mit der Bayesschen Inferenz erzeugt wurde.

In Abb. 6 ist die Menge aller betrachteten Betriebsstätten (roter Bereich) unter dem Gesichtspunkt ihrer potentiellen Auffälligkeit bzgl. ihres Ordnungsanteils von Generika am Gesamtmarkt dargestellt. Outlier gemäß des Observed/Expected-Indikators befinden sich hier in den Bereichen, in denen die rote Verteilungsfunktion unter der blauen liegt. Dort fallen die beobachteten Werte (Observed-Werte) niedriger aus, als vom Modell vorhergesagt (Expected-Werte). Die Outlier gemäß der „natürlichen“ Verteilung befinden sich links von der Kontrollgrenze. Betriebsstätten, die sowohl gemäß des Observed/Expected-Indikators als auch der „natürlichen“ Verteilung auffällig sind, befinden sich links von der Kontrollgrenzen in dem Bereich, wo die rote Verteilung unter der blauen liegt. Betriebsstätten in diesem Bereich sind eindeutig auffällig. Erfüllt eine Betriebsstätte keine der hier genannten Auffälligkeitskriterien, so ist sie als eindeutig unauffällig einzustufen.

Global wurden für diesen Indikator die folgenden Haupteinflussgrößen mittels der SHAP Values bestimmt (siehe Abb. 7 und Tabelle 15):

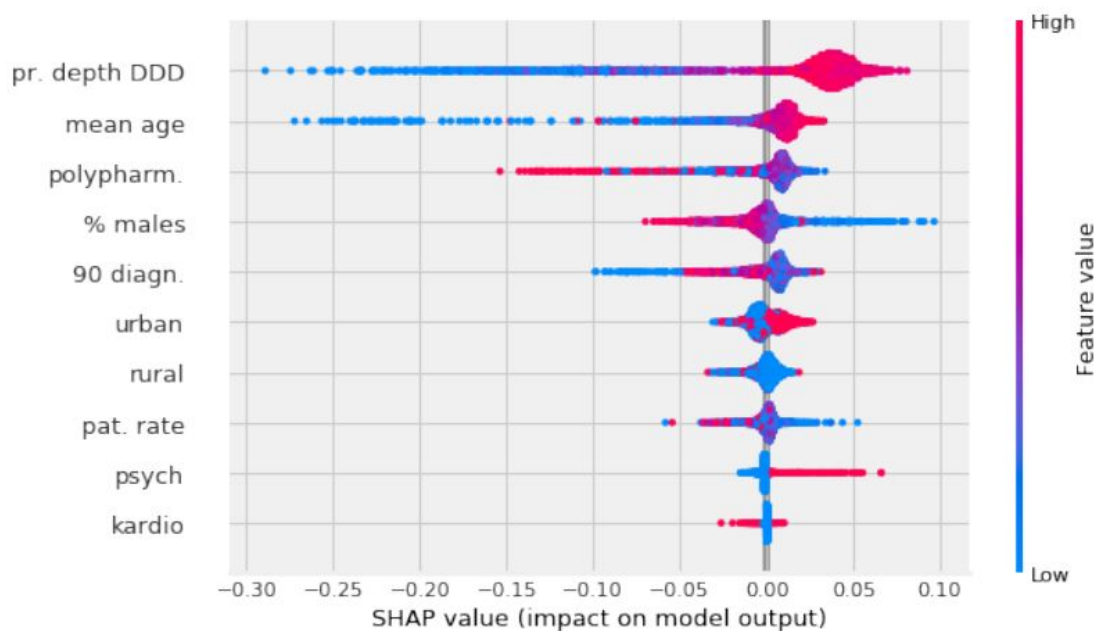


Abbildung 7: SHAP summary plot des Verordnungsanteils von Generika am Gesamtmarkt im Quartal 4/2016 – absteigend gelistet sind die Haupteinflussfaktoren auf den Verordnungsanteil von Generika über alle betrachteten Betriebsstätten hinweg. Mit der farblichen Abgrenzung (roter Impact= hoher Impact und blauer Impact= geringer Impact) wird die jeweilige Impactrichtung abgebildet.

In Abbildung 7 sind die einflussgebenden Variablen mit ihrem Impact (roter Impact= hoher Impact und blauer Impact= geringer Impact) und die jeweilige Impactrichtung auf die Indikatorvariable Verordnungsanteil von Generika am Gesamtmarkt abgebildet. Jede Betriebsstätte bildet in jeder Variablenzeile jeweils einen eigenen Punkt mit einer eigenen Farbe und einer eigenen Impactposition (stark positiv/negativ oder wenig positiv/negativ).

Insgesamt ist zu erkennen, dass die Verordnungstiefe mit Abstand den stärksten Einfluss auf den Verordnungsanteil von Generika hat, wobei der Verordnungsanteil steigt bzw. fällt, wenn auch der Wert der Verordnungstiefe steigt oder fällt. Eine ähnliche Interpretation kann bzgl. der Einflussvariable Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung getroffen werden: es handelt sich um die zweitstärkste Einflussvariable, deren steigender bzw. fallender Wert den Verordnungsanteil von Generika ebenfalls steigen oder fallen lässt. Bei den Variablen Anteil von männlichen Patienten und Anteil von Polypharmaziepatienten ergibt sich eine entgegengesetzte Interpretation. Hier steigt der Verordnungsanteil von Generika, wenn die Einflussvariable sinkt und umgekehrt. Die anderen Variablen weisen insgesamt nur einen geringen Einfluss auf. Zudem ist eine eindeutige Interpretation schwierig, da nicht klar ist, in welcher Weise sie den Entscheidungsprozess des Modells beeinflussen.

Tabelle 15: Haupteinflussfaktoren auf den Verordnungsanteil von Generika am Gesamtmarkt im Quartal 4/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable	Einflussart (Impactrichtung)
Verordnungstiefe	Positiv
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Positiv
Anteil von männlichen Patienten	Negativ
Anteil von Polypharmaziepatienten	Negativ

Anwendung und Adjustierung der Analyseprozedur für die Quartale 1/2016 bis 4/2016 sowie das Gesamtjahr 2016

Weiterhin wurden beispielhaft für eine KV für die Quartale 1/2016 bis 4/2016 und das Gesamtjahr 2016 die in Tabelle 15 aufgeführten Indikatoren der weiterführenden Analyse analysiert. Dabei wurden in allen Analysen die in Tabelle 16 aufgelisteten Confounder der weiterführenden Analysen berücksichtigt und eine Kontrollgrenze von 5% gewählt.

Im Zuge der Analyse wurden die betrachteten Betriebsstätten einer der vier folgenden Kategorien zugeordnet:

- **Kategorie 1:** Eine Betriebsstätte ist nicht auffällig, d.h. sie ist weder nach der natürlichen Verteilung noch nach dem Machine-Learning-Ansatz ein Outlier.
- **Kategorie 2:** Eine Betriebsstätte ist mittels beider Ansätze auffällig, d.h. sie ist sowohl ein „natürlicher“ als auch „O/E“-Outlier.
- **Kategorie 3:** Eine Betriebsstätte ist auffällig nach der „natürlichen“ Verteilung aber nicht nach dem Machine-Learning-Modell.
- **Kategorie 4:** Eine Betriebsstätte ist nicht auffällig nach der „natürlichen“ Verteilung aber auffällig nach dem Machine-Learning-Modell sind.

Im Folgenden werden die Analyseergebnisse der einzelnen Indikatoren präsentiert.

6.5.2 Praxisanwendungen der Analysen

Verordnete PPI-Mengen (DDDs je Arzneimittelpatient)

Es wurden 6205 PPI-verordnende Betriebsstätten analysiert und jeweils in eine der vier Kategorien eingeordnet. Insgesamt wurden hierbei die folgenden Ergebnisse für die einzelnen Quartale und auch das Gesamtjahr 2016 ermittelt (s. Tabelle 16):

Tabelle 16: Prozentuale Aufteilung der Betriebsstätten (Praxen) in jeweils eine der vier Kategorien für den Indikator verordnete PPI-Mengen für die Quartale 1/2016 bis 4/2016 und das Gesamtjahr 2016

Zeitraum	Anzahl Praxen gesamt	Anteil Kategorie 1	Anteil Kategorie 2	Anteil Kategorie 3	Anteil Kategorie 4
1/2016	6205	93,1%	3,0%	1,1%	2,8%
2/2016	6205	92,7%	3,1%	1,1%	3,0%
3/2016	6205	93,4%	2,9%	1,2%	2,5%
4/2016	6205	93,3%	3,0%	1,1%	2,6%
2016 gesamt	6205	88,3% (5478)	1,2% (73)	0,2% (10)	0,4% (22)

Während sich in den einzelnen Quartalen die Anteile der Praxen innerhalb der Kategorien 2, 3 und 4 bei ca. 3%, 1,1% und zwischen 2,5% und 3% bewegen, so liegen die über das ganze Jahr befindlichen Betriebsstätten dieser Kategorien lediglich bei 1,2%, 0,2% und 0,4%.

Um den Einfluss der einzelnen exogenen Variablen zu bestimmen, werden die SHAP summary plots der einzelnen Quartale (s. Abb. 8 – 11) genutzt, die die Haupteinflussfaktoren bestimmen (s. Tabellen 17– 20).

Ganz konsistent, ist überall (s. Abb. 8 – 11) die gleiche Einflussverteilung zu sehen: der Anteil an Polypharmaziepatienten hat in allen Quartalen mit Abstand den stärksten Einfluss auf die PPI-Verordnungen, gefolgt von der Verordnungstiefe. Ein Anstieg in einer dieser beiden Variablen resultiert auch in einem Anstieg der durchschnittlich verordneten PPI-Menge. Nicht ganz so relevant, aber immer noch einflussgebend sind die Variablen ist das Durchschnittsalter der Patienten und der Anteil der Patienten, die in einem städtischen Gebiet wohnen. Auch diese Variablen führen mit steigenden Werten zu höheren PPI-Verordnungsrate.

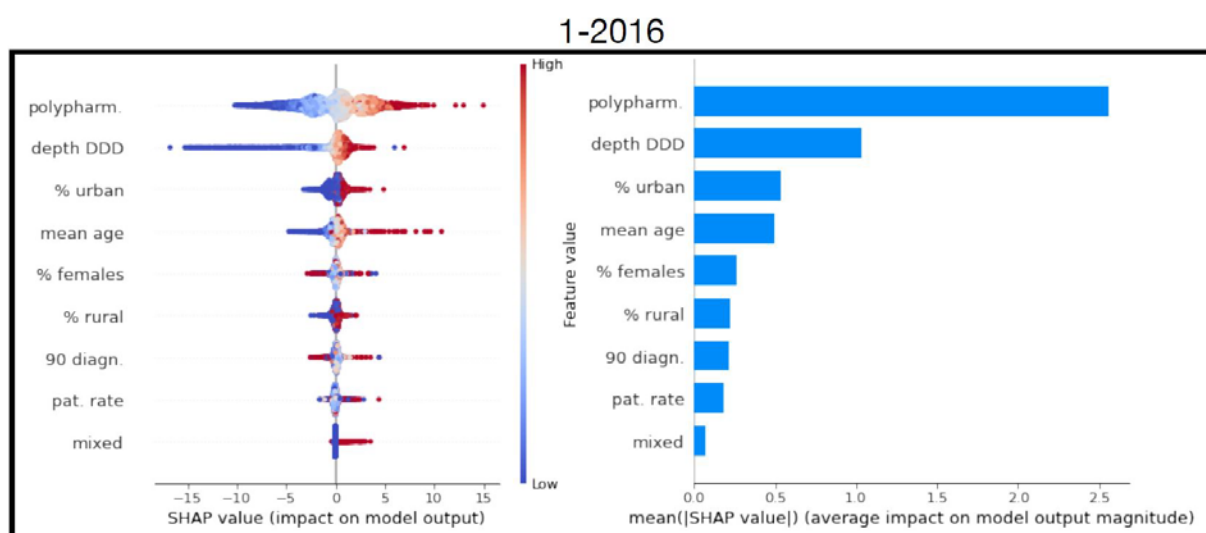


Abbildung 8: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 1/2016

Tabelle 17: Haupteinflussfaktoren auf die durchschnittlich verordneten PPI-Mengen im Quartal 1/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 1/2021)	Einflussart (Impactrichtung)
Anteil von Polypharmaziepatient*innen	Positiv
Verordnungstiefe	Positiv
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Positiv
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Positiv

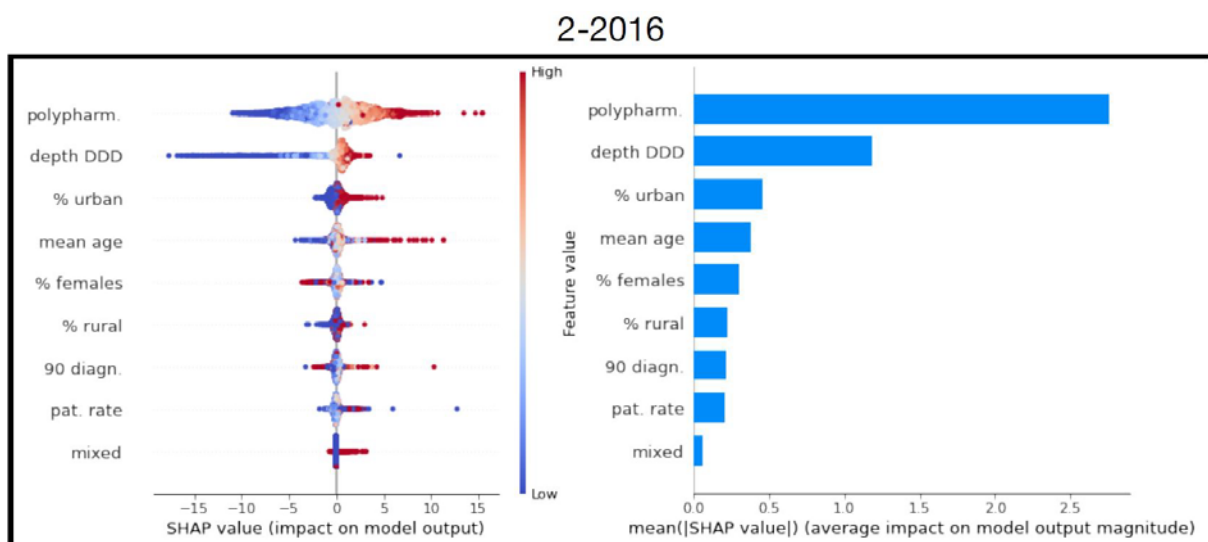


Abbildung 9: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 2/2016

Tabelle 18: Haupteinflussfaktoren auf die durchschnittlich verordneten PPI-Mengen im Quartal 2/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 2/2021)	Einflussart (Impactrichtung)
Anteil von Polypharmaziepatient*innen	Positiv
Verordnungstiefe	Positiv
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Positiv
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Positiv

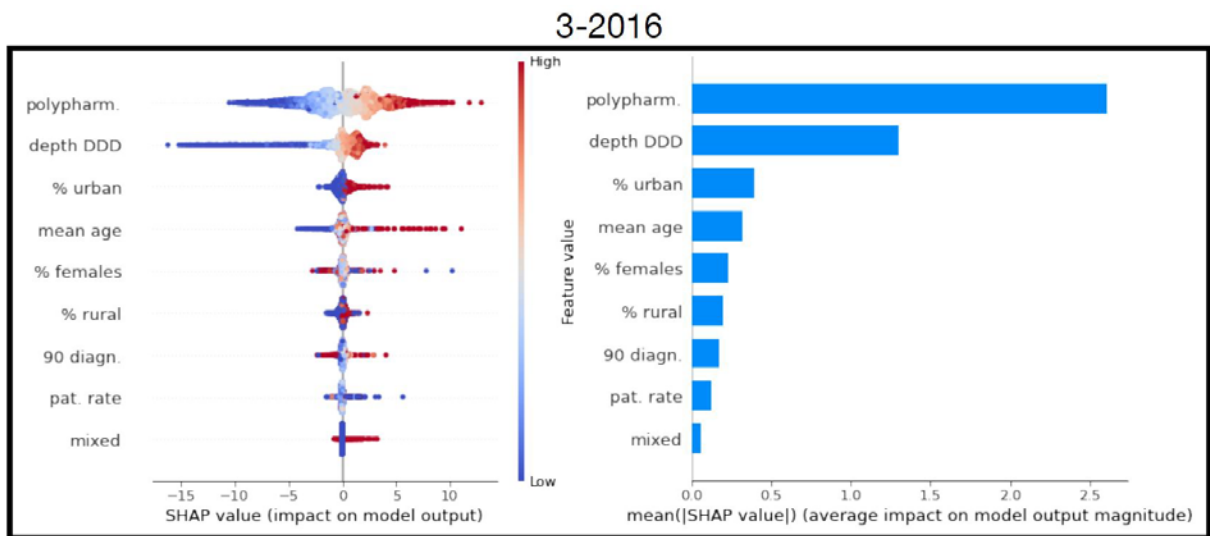


Abbildung 10: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 3/2016

Tabelle 19: Haupteinflussfaktoren auf die durchschnittlich verordneten PPI-Mengen im Quartal 3/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 1/2021)	Einflussart (Impactrichtung)
Anteil von Polypharmaziepatient*innen	Positiv
Verordnungstiefe	Positiv
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Positiv
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Positiv

4-2016

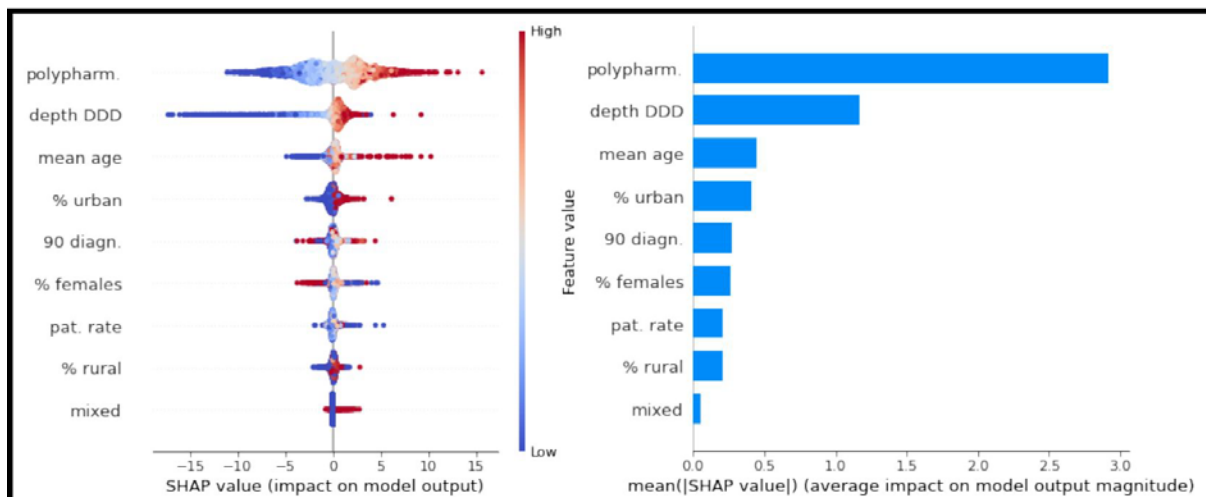


Abbildung 11: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 4/2016

Tabelle 20: Haupteinflussfaktoren auf die durchschnittlich verordneten PPI-Mengen im Quartal 4/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 1/2021)	Einflussart (Impactrichtung)
Anteil von Polypharmaziepatient*innen	Positiv
Verordnungstiefe	Positiv
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Positiv
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Positiv

Dies lässt sich folgendermaßen interpretieren: Bei Polypharmaziepatienten kann es durchaus angebracht sein, verstärkt PPIs zu verordnen, um den Magen des Patienten zu schützen bzw. durch Medikamente verursachte Beschwerden zu behandeln. Daher werden PPIs oft als Begleitmedikation eingesetzt. Weist die jeweilige Betriebsstätte eine ausgeprägte Verordnungstiefe auf, so werden den dortigen Patienten die meisten Verordnungen von dieser Praxis ausgestellt und somit auch die der PPIs. Da ältere Menschen oft multimorbide und zudem Polypharmaziepatienten sind, werden Ihnen zum Schutz im Schnitt häufiger PPIs verordnet als jungen, nicht-multimorbiden Patienten. Zudem dürften ältere Patienten PPIs in manchen Fällen auch als Ulkus-Prophylaxe erhalten haben.

Verordnungsanteil von DOAKs an allen Antikoagulantien (DDDs)

Es wurden 5878 Antikoagulantien-verordnende Betriebsstätten analysiert und jeweils in eine der vier Kategorien eingeordnet. Insgesamt wurden hierbei die folgenden Ergebnisse für die einzelnen Quartale und auch das Gesamtjahr 2016 ermittelt (s. Tabelle 21):

Tabelle 21: Prozentuale Aufteilung der Betriebsstätten (Praxen) in jeweils eine der vier Kategorien für den Indikator des Verordnungsanteils von DOAKs an allen Antikoagulantien für die Quartale 1/2016 bis 4/2016 und das Gesamtjahr 2016

Zeitraum	Anzahl Praxen gesamt	Anteil Kategorie 1	Anteil Kategorie 2	Anteil Kategorie 3	Anteil Kategorie 4
1/2016	5878	94,7%	2,9%	0,5%	1,9%
2/2016	5878	94,7%	2,6%	0,6%	2,1%
3/2016	5878	94,7%	2,0%	0,8%	2,4%
4/2016	5878	94,6%	2,2%	0,8%	2,4%
2016 gesamt	5878	88,7% (5214)	0,4% (23)	0,1% (4)	0,1% (4)

Während sich in den einzelnen Quartalen die Anteile der Praxen innerhalb der Kategorien 2, 3 und 4 bei 2%-2,9%, 0,5%-0,8% und 1,9%-2,4% bewegen, so liegen die Anteile der über das ganze Jahr in diesen Kategorien befindlichen Betriebsstätten lediglich bei 0,4%, 0,1% und 0,1%.

Um den Einfluss der einzelnen exogenen Variablen zu bestimmen, werden die SHAP summary plots der einzelnen Quartale (s. Abb. 12 – 15) genutzt, die die Haupteinflussfaktoren bestimmen (s. Tabellen 22 – 25).

Zu sehen ist die folgende Einflussverteilung (über die vier Quartale hinweg) der exogenen Variablen: die Haupteinflussfaktoren sind der Anteil der Patient*innen, die in einem städtischen Gebiet wohnen, gefolgt vom Anteil der Patient*innen, die in einem ländlichen Gebiet wohnen. Dabei bekommen Patient*innen aus städtischen Gebieten DOAKs häufiger verschrieben als in ländlichen Gebieten. Zudem hat das durchschnittliche Patientenalter den drittgrößten Einfluss auf den Anteil der DOAK-Verordnungen: ein Anstieg im durchschnittliche Patientenalter bewirkt ein Absinken im Verordnungsanteil, d.h. ältere Patient*innen bekommen eher andere Antikoagulantien als die DOAKs verordnet. Nicht ganz so einflussreich, aber immer noch relevant sind die exogenen Variablen Anteil an weiblichen Patientinnen und Anteil an Polypharmaziepatienten, deren Anstieg auch in beiden Fällen einen Anstieg der DOAK-Verordnungsraten und somit -anteile nach sich zieht.

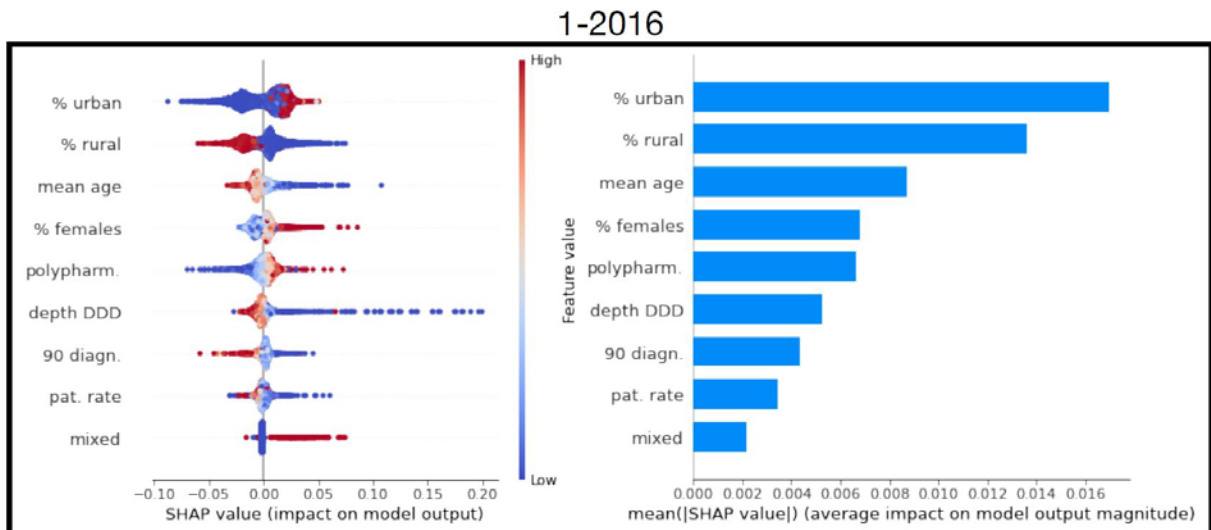


Abbildung 12: SHAP summary plot des Verordnungsanteils von DOAKs an allen Antikoagulantien im Quartal 1/2016

Tabelle 22: Haupteinflussfaktoren auf den Verordnungsanteil von DOAKs an allen Antikoagulantien im Quartal 1/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 1/2021)	Einflussart (Impactrichtung)
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Positiv
Anteil der Patient*innen, die in einem ländlichen Gebiet wohnen	Negativ
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Negativ
Anteil von weiblichen Patientinnen	Positiv
Anteil von Polypharmaziepatient*innen	Positiv

2-2016

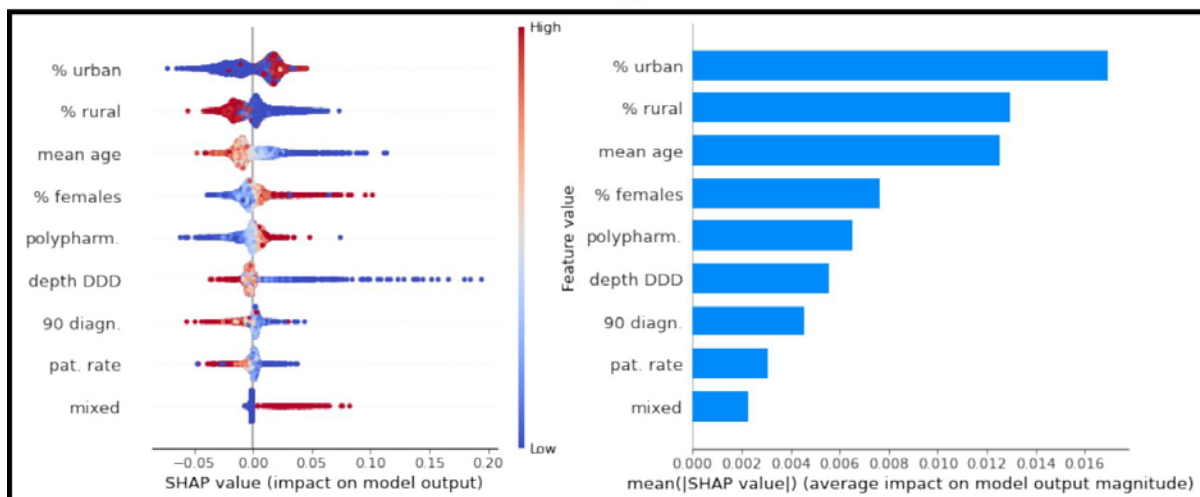


Abbildung 13: SHAP summary plot des Verordnungsanteils von DOAKs an allen Antikoagulantien im Quartal 2/2016

Tabelle 23: Haupteinflussfaktoren auf den Verordnungsanteil von DOAKs an allen Antikoagulantien im Quartal 2/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 2/2021)	Einflussart (Impactrichtung)
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Positiv
Anteil der Patient*innen, die in einem ländlichen Gebiet wohnen	Negativ
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Negativ
Anteil von weiblichen Patientinnen	Positiv
Anteil von Polypharmaziepatient*innen	Positiv

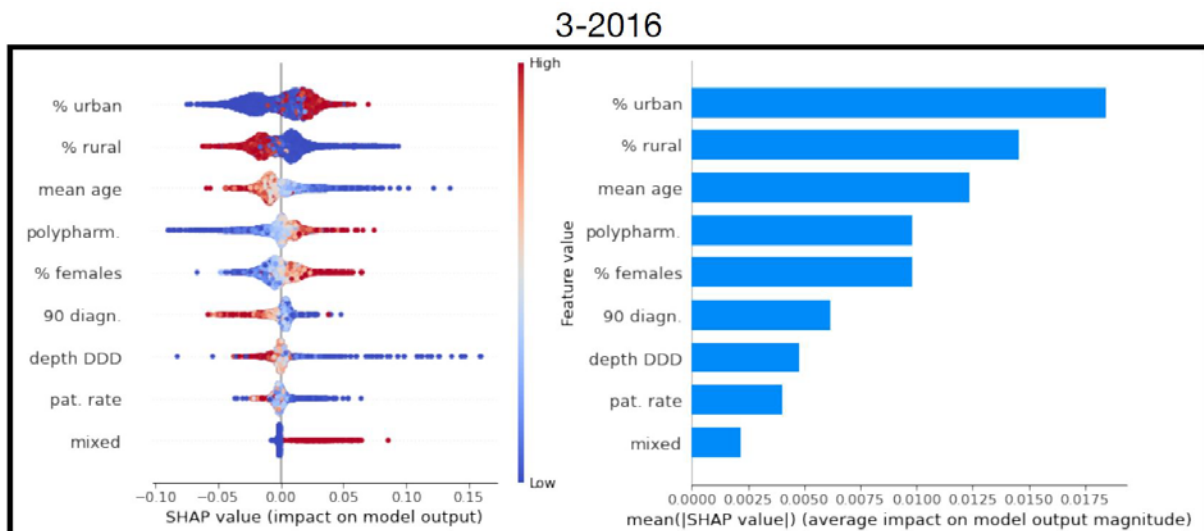


Abbildung 14: SHAP summary plot des Verordnungsanteils von DOAKs an allen Antikoagulantien im Quartal 3/2016

Tabelle 24: Haupteinflussfaktoren auf den Verordnungsanteil von DOAKs an allen Antikoagulantien im Quartal 3/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 3/2021)	Einflussart (Impactrichtung)
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Positiv
Anteil der Patient*innen, die in einem ländlichen Gebiet wohnen	Negativ
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Negativ
Anteil von weiblichen Patientinnen	Positiv
Anteil von Polypharmaziepatient*innen	Positiv

4-2016

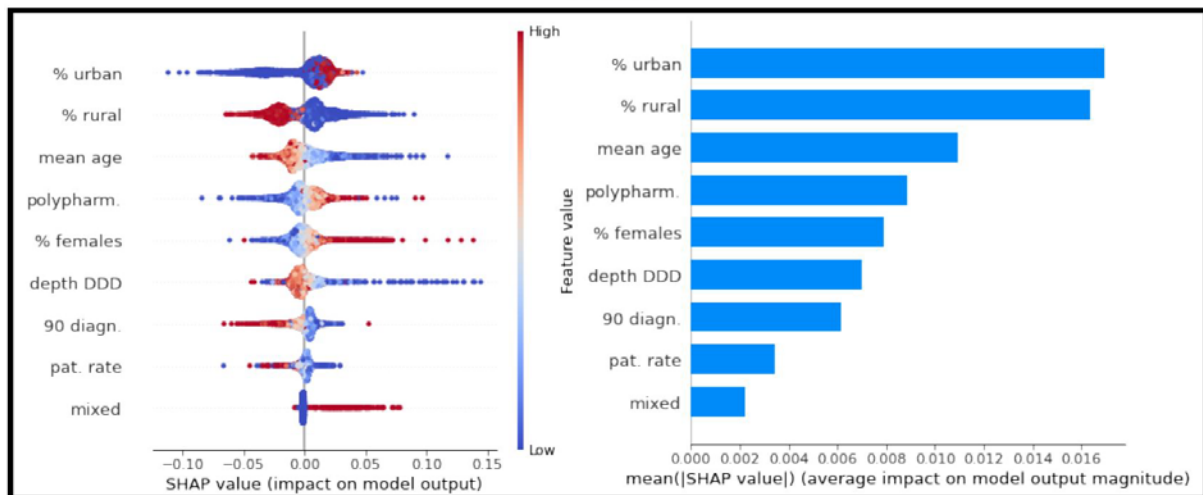


Abbildung 15: SHAP summary plot des Verordnungsanteils von DOAKs an allen Antikoagulantien im Quartal 4/2016

Tabelle 25: Haupteinflussfaktoren auf den Verordnungsanteil von DOAKs an allen Antikoagulantien im Quartal 4/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 4/2021)	Einflussart (Impactrichtung)
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Positiv
Anteil der Patient*innen, die in einem ländlichen Gebiet wohnen	Negativ
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Negativ
Anteil von weiblichen Patientinnen	Positiv
Anteil von Polypharmaziepatient*innen	Positiv

Dies lässt sich folgendermaßen interpretieren:

Das häufigere Verordnen von DOAKs in städtischen Gebieten könnte darauf zurückzuführen zu sein, dass verschiedene Disseminations-Mechanismen bei städtischen Praxen intensiver bzw. früher wirken (z.B. Besuche von Pharmareferenten, Fortbildungsveranstaltungen mit dieser Zielsetzung). Zudem ist in Städten sowohl die Facharzt-dichte als auch die Krankenhaus-dichte höher. Erhalten Patient*innen ihre Antikoagulationserstverordnung aus einer Facharztpraxis oder eine Antikoagulation im Rahmen einer Entlassmedikation aus dem Krankenhaus, so handelt es sich zumeist um ein DOAK. Im Anschluss erwarten diese Patient*innen oft diesen Wirkstoff auch von/m eigener/n Hausärzt*in verordnet zu bekommen.

Bei zeitlich neueren Ersteinstellungen handelt es sich vermehrt um jüngere Patient*innen, die in der Regel auf DOAKs eingestellt werden und zudem eher im städtischen Bereich wohnen.

In ländlichen Gebieten wohnen dahingegen eher ältere Patient*innen, die ihre Antikoagulation überwiegend schon vor längerer Zeit eingestellt bekommen haben (auf Vitamin-K-Antagonisten). Hat sich diese Einstellung bewährt, wird oft nicht mehr zu DOAKs gewechselt.

Anteil von generischen DDDs im Bereich der Lipidsenker

Es wurden 6139 Lipidsenker-verordnende Betriebsstätten analysiert und jeweils in eine der vier Kategorien eingeordnet. Insgesamt wurden hierbei die folgenden Ergebnisse für die einzelnen Quartale und auch das Gesamtjahr 2016 ermittelt (s. Tabelle 26):

Tabelle 26: Prozentuale Aufteilung der Betriebsstätten (Praxen) in jeweils eine der vier Kategorien für den Indikator des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker für die Quartale 1/2016 bis 4/2016 und das Gesamtjahr 2016

Zeitraum	Anzahl Praxen gesamt	Anteil Kategorie 1	Anteil Kategorie 2	Anteil Kategorie 3	Anteil Kategorie 4
1/2016	6139	92,0%	1,7%	0,0%	6,3%
2/2016	6139	92,0%	1,4%	0,0%	6,5%
3/2016	6139	92,3%	1,4%	0,0%	6,3%
4/2016	6139	92,3%	1,5%	0,1%	6,1%
2016 gesamt	6139	85,4% (5240)	0,5% (30)	0,0% (0)	1,7% (79)

Während sich in den einzelnen Quartalen die Anteile der Praxen innerhalb der Kategorien 2, 3 und 4 bei ca. 3%, 1,1% und zwischen 2,5% und 3% bewegen, so liegen die über das ganze Jahr befindlichen Betriebsstätten dieser Kategorien lediglich bei 1,2%, 0,2% und 0,4%.

Um den Einfluss der einzelnen exogenen Variablen zu bestimmen, werden die SHAP summary plots der einzelnen Quartale (s. Abb. 16 – 19) genutzt, die die Haupteinflussfaktoren bestimmen (s. Tabellen 27– 30).

Die Einflussverteilung der exogenen Variablen ist die folgende: In den Quartalen 1/2016 bis 3/2016 ist der Haupteinflussfaktor der Anteil der Patient*innen, die in einem städtischen Gebiet wohnen, gefolgt vom durchschnittlichen Patientenalter. Dabei sinkt die Generikarate mit steigendem Anteil von Patient*innen aus städtischen Gebieten und steigt mit dem durchschnittlichen Patientenalter. Der drittgrößte Einflussfaktor ist der Anteil an weiblichen Patienten, wobei mit steigendem Anteil der Anteil an generischen DDDs sinkt. Die übrigen Variablen verhalten sich in den Quartalen ganz unterschiedlich und erlauben keine Schlussfolgerung bzgl. eines möglichen Einflusses auf den Anteil von generischen DDDs im Bereich der Lipidsenker.

Im Quartal 4/2016 wird das durchschnittliche Patient*innenalter zum Haupteinflussfaktor, dicht gefolgt vom Anteil von Patienten aus städtischen Gebieten. Der Anteil an weiblichen Patientinnen verschwindet sogar ganz aus der Liste der einflussnehmenden Variablen.

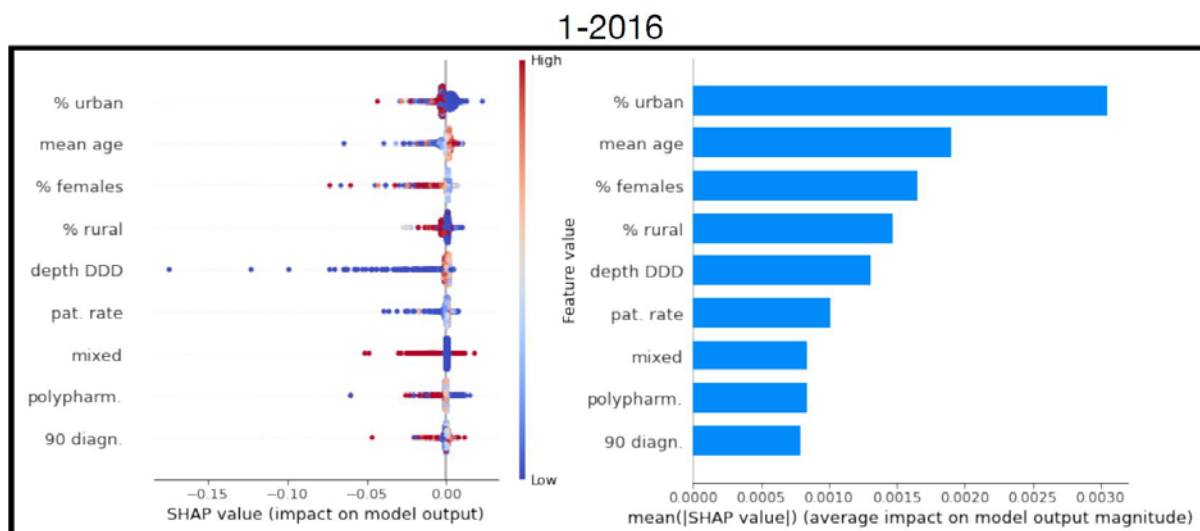


Abbildung 16: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 1/2016

Tabelle 27: Haupteinflussfaktoren auf den Verordnungsanteil von generischen DDDs im Bereich der Lipidsenker im Quartal 1/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 1/2021)	Einflussart (Impactrichtung)
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Negativ
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Positiv
Anteil von weiblichen Patientinnen	Negativ

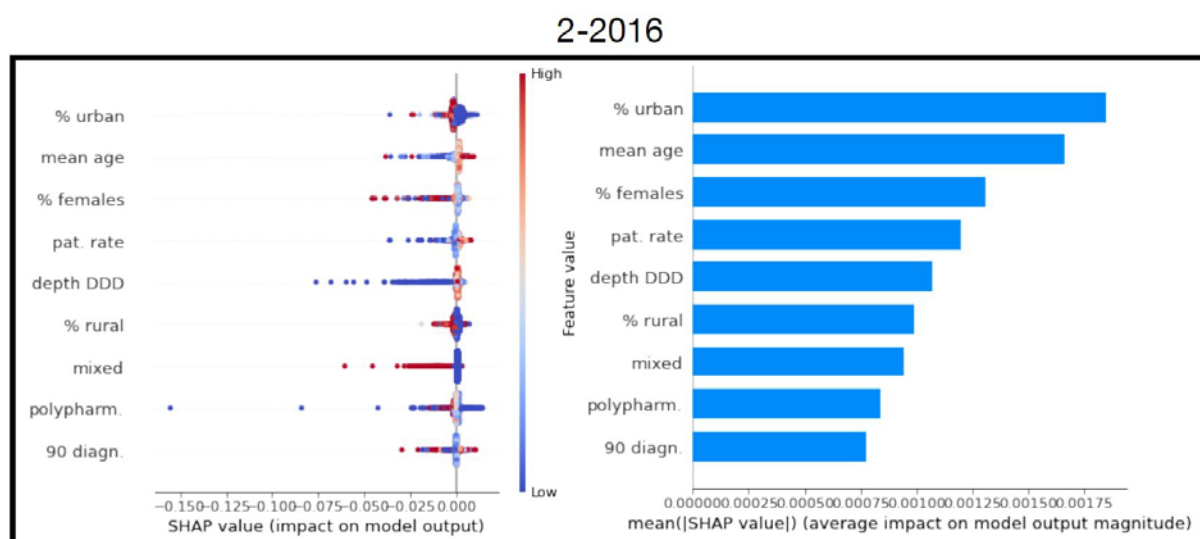


Abbildung 17: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 2/2016

Tabelle 28: Haupteinflussfaktoren auf den Verordnungsanteil von generischen DDDs im Bereich der Lipidsenker im Quartal 2/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 2/2021)	Einflussart (Impactrichtung)
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Negativ
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Positiv
Anteil von weiblichen Patientinnen	Negativ

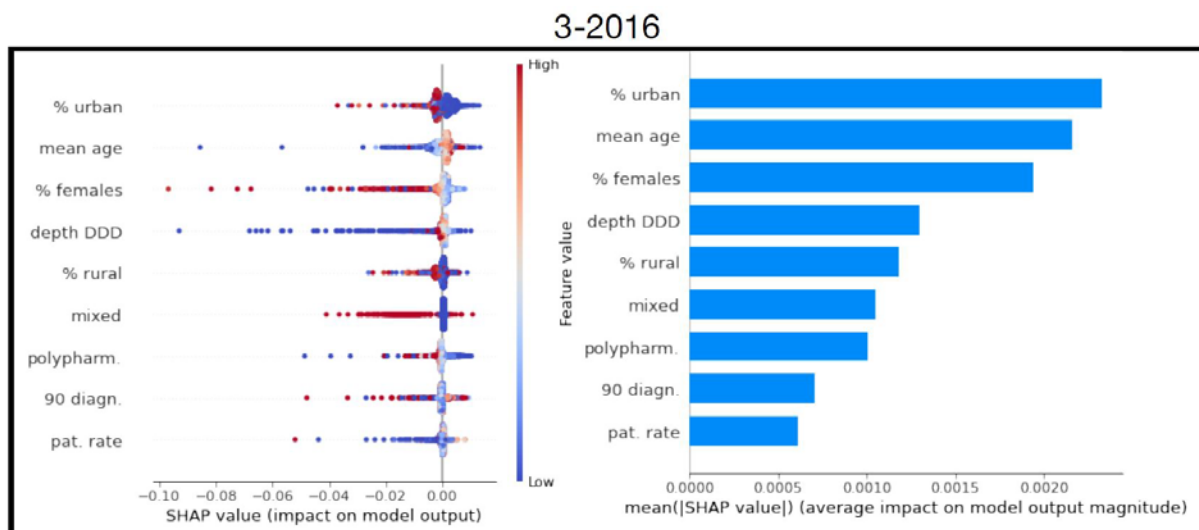


Abbildung 18: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 3/2016

Tabelle 29: Haupteinflussfaktoren auf den Verordnungsanteil von generischen DDDs im Bereich der Lipidsenker im Quartal 3/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 3/2021)	Einflussart (Impactrichtung)
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Negativ
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Positiv
Anteil von weiblichen Patientinnen	Negativ

4-2016

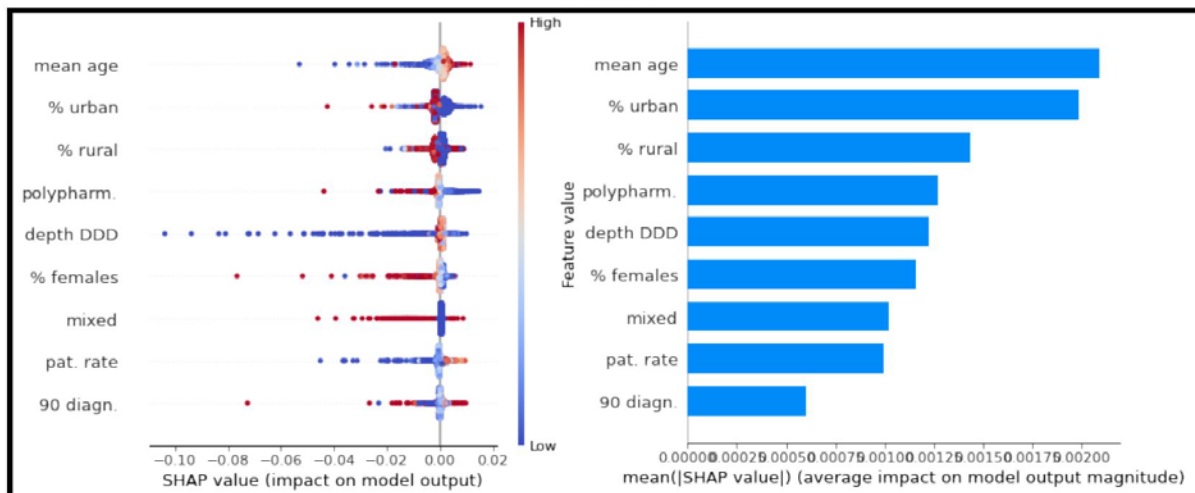


Abbildung 19: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 4/2016

Tabelle 30: Haupteinflussfaktoren auf den Verordnungsanteil von generischen DDDs im Bereich der Lipidsenker im Quartal 4/2016 – absteigend nach dem Betrag des jeweiligen Impacts sortiert

Einflussvariable (Quartal 4/2021)	Einflussart (Impactrichtung)
Durchschnittsalter der Patient*innen mit mindestens einer Arzneimittelverordnung	Positiv
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	Negativ

Die für diesen Indikator vorliegende Einflussverteilung der exogenen Variablen ist schwerer zu interpretieren als bei den anderen Indikatoren:

Da zumeist ältere Patient*innen Lipidsenker erhalten, übt sowohl die Variable des Durchschnittsalters der Patient*innen als auch die Variable städtischer Wohnort einen großen (aber auch entgegengesetzten) Einfluss auf das Verordnungsgeschehen von Lipidsenkern aus. Je mehr ältere Patient*innen eine Praxis versorgt, desto höher sind die die Verordnungsraten der Lipidsenker. Zudem verordnen hausärztliche Praxen bei einem mittleren bis großen Verordnungsvolumen an Lipidsenkern überwiegend generisch. So ist bei Praxen mit vielen Lipidsenkerverordnungen ein großer Anteil an generischen Verordnungen zu verzeichnen. Insgesamt steigt somit bei einer Erhöhung des Verordnungsvolumens an Lipidsenkern auch der Anteil an generischen Lipidsenker. Da viele jüngere Patient*innen eher in städtischen Gebieten wohnen, werden dort insgesamt weniger Lipidsenker verordnet. Im Quartal 4 des Jahres 2016 ist eine Zunahme von Verordnungen an nicht-generischen Lipidsenkern (sowohl Fibrate als auch Ezetimib) zu verzeichnen, die oft weiblichen Patientinnen verschrieben werden. Dies führt dazu, dass die exogene Variable „Anteil an weiblichen Patientinnen“ an Einfluss auf den Generikaanteil bei den Lipidsenkern verliert, da für weibliche Patientinnen im 4. Quartal 2016 mehr Originalpräparate verordnet wurden. Mit insgesamt fallenden generischen DDDs nimmt auch der Anteil der weiblichen generischen DDDs ab.

7. Diskussion der Projektergebnisse

7.1 Teilprojekt A

Die Ergebnisse der im Rahmen von Teilprojekt A durchgeführten Analysen zeigen auf, dass der Wirtschaftlichkeitsaspekt bei der Verordnung von Arzneimitteln durch mehrere Faktoren beeinflusst wird. So konnte durch die 1. Analyse (siehe 6.1.1 Sacubitril/Valsartan) dargestellt werden, dass bei der Verbreitung eines neuen Medikaments neben dem Einfluss der ärztlichen KollegInnen auch der Einfluss der Patient*innen eine wichtige Rolle zu spielen scheint, da eine Medikation mit diesem Medikament zu weiteren Verordnungen durch andere ÄrztInnen führen kann, wodurch eine unmittelbare Verbreitung des Medikaments erfolgt. In dieser Analyse zeigt sich weiterhin, dass auch die Verbindung zwischen ambulanten Ärzt*innen einen Einfluss auf das Ordnungsverhalten nimmt und auch hier eine gewisse Beeinflussung bei der Entscheidung für ein neues Medikament ableitbar ist. Die 2. Analyse (siehe 6.1.2 Direkte orale Antikoagulanzen) erweitert den Blick auf die Makroebene und untersucht unter anderem regionale Einflussfaktoren auf die Verbreitung/Entscheidung für oder gegen Medikamente. In dieser Analyse zeigt sich unter anderem auch die Verflechtung von stationärem und ambulanten Sektor, in dem Sinne, dass PatientInnen offenbar durch einen stationären Aufenthalt Medikationen in den ambulanten Sektor tragen. Die 3. und 4. Analyse (siehe 6.1.3 Pregabalin und 6.1.4 Oxycodon/Naloxon) verdeutlichen abschließend noch einmal die Notwendigkeit der Kooperation im ambulanten Sektor, wenn es um die Ausstellung von Verordnungen geht, und zeigen auf, wie Patient*innen Probleme und Brüche der ambulanten Versorgung möglicherweise zu ihren Gunsten ausnutzen können. Insbesondere bei der Verordnung beobachtungswürdiger Medikamente (sei es aus Gründen der umstrittenen Evidenz der Wirksamkeit, potentieller Nebenwirkungen oder der Wirtschaftlichkeit) ist somit eine enge Kooperation der ambulanten Leistungserbringer notwendig.

Bei der Interpretation der Ergebnisse ist zu berücksichtigen, dass die untersuchten Routinedaten kein perfektes Abbild des Ordnungs-Geschehens darstellen. Der schwerste Mangel besteht darin, dass Entlassmedikationen nach stationärer Entlassung und die Induktion von Weiterverordnungen nicht (sicher) erfasst werden können. Auch können die Mikroprozesse von Kommunikation, Abwägung und Beeinflussung nicht erfasst werden.

7.2 Teilprojekt B

Im Vergleich der Kostenentwicklung von Bayern und Hessen konnte gezeigt werden, dass die Umstellung auf die WSV zu keiner Expansion auf der Seite der Verordnungskosten geführt hat. Positiv zeigt sich eine Wirkung der WSV in den Bereichen Generikaanteil, Antikoagulantien (DOAKs) und Erreichung der allgemeinen Leitsubstanzquote. Bei den meisten Indikatoren mit Bezug zur Wirkstoffvereinbarung ist eine positive Entwicklung innerhalb der KV Bayerns für die bayerischen Ziele zu verzeichnen.

Die Verordnung von Medikamenten in der ambulanten Versorgung ist ein komplexes Geschehen. Vorgaben und Steuerungen in Bezug auf Wirtschaftlichkeit und Qualität sind nur ein Faktor unter vielen. Hinzukommt, dass komplexe Interventionen wie die WSV erst über einen längeren Zeitraum nachhaltig wirken, da hier langfristige Routinen und Einstellungen geändert werden müssen. Teilweise wird dies erst über einen Generationswechsel möglich sein. Deshalb sind künftig Analysen mit noch längeren Untersuchungszeiträumen sinnvoll.

7.3 Teilprojekt C

In qualitativen Fokusgruppen zeigte sich eine insgesamt positive Einschätzung der WSV; Zeitnahes Feedback, Verständlichkeit, mehr Kontrollmöglichkeiten und Gerechtigkeit wurden hervorgehoben, die Regressorgen hätten abgenommen.

Kritisch wurde die Intransparenz für die Verordner bei Rabattverträgen gesehen, Einschränkung der Therapie-Freiheit durch die WSV und Widersprüche zu Leitlinien-Empfehlungen gesehen. Teilweise wurde von Schwierigkeiten berichtet, bei chronisch kranken Patient*innen auf Leitsubstanzen umzustellen.

Nicht alle Ärzt*innen hatten verstanden, dass Überweisungen wegen der Verordnungen „teurer“ Medikamente unter den Regelungen WSV kaum noch plausibel sind.

Bei Schlussfolgerungen aus qualitativen Fokusgruppen sind verschiedene Möglichkeiten des Bias zu bedenken. Ein Selektions-Bias ergibt sich dadurch, dass die Ärzt*innen, die sich zur Teilnahme bereit erklären, von den übrigen sich unterscheiden können. Die Dynamik einer Fokusgruppe schätzen wir als produktiv ein, um Erkenntnisse zu erhalten, die mit Einzelinterviews oder Fragebögen nicht möglich wären. Durch die Dominanz Einzelner kann sich jedoch, auch bei kompetenter Moderation, einer Beeinflussung anderer Teilnehmer*innen ergeben.

Die von den Teilnehmer*innen erfassten Äußerungen können auch dazu genutzt werden, die Inhalts-Validität künftiger Befragungs-Instrumente zu verbessern (z.B. repräsentative Surveys von Vertragsärzt*innen).

7.4 Teilprojekt D

Unter den Betriebsstätten, die als Zielverfeher eingestuft wurde, waren kleine Praxen, männliche Praxisinhaber und solche ohne Zusatzbezeichnungen etwas häufiger vertreten.

Neben diesen eher objektiven Kriterien ergeben die qualitativen Interviews Aufschlüsse über Einstellungen und individuelle Erfahrungen. Diese Befragungen können wegen der geringen Zahl sowie einem offensichtlichen Ungleichgewicht der Fachgruppen-Zugehörigkeit keine unmittelbare Repräsentativität beanspruchen. Auch wenn die Befragungen durch Angehörige neutraler akademischer Institutionen durchgeführt wurden, kann ein Bias der sozialen Erwünschtheit nicht ausgeschlossen werden.

Arbeiten auf dem aktuellen Stand der Wissenschaft, Verantwortungsbewusstsein und eine Einbeziehung von Patient*innen in Entscheidungen wurden als erste Priorität gesehen. Die Wirtschaftlichkeit der Verordnungen spielte eine geringere Rolle, auch die Unabhängigkeit von Industrieinteressen wurde als weniger wichtig empfunden. Entsprechen hoch war der Stellenwert von Pharmavertretern als Informationsquelle.

Informationslücken bestanden in Bezug auf das System der Wirkstoffvereinbarung.

Zur Zielverfehlung führten vor allem Einzelziele, die eher einen Randbereich des typischen Verordnungsspektrums ausmachen und somit nur einen geringen Verordnungsumfang aufweisen.

Konfliktträchtig sind vor allem orale Antikoagulanzen und die Umstellung auf Biosimilars.

Setzt man alle bisherigen Ergebnisse in Beziehung, so stellt man fest, dass es mehrere Ebenen gibt, die das Verordnungsgeschehen beeinflussen und zur Auffälligkeit beitragen könnten (siehe Anlage TP-D: Ebenen mit potenziellen Einflussfaktoren auf die Auffälligkeit).

Gibt es ungünstige Voraussetzungen an einem dieser Bereiche, kann eine Auffälligkeit möglicherweise resultieren. Je weiter unten in dem Modell der Konflikt auftritt oder je mehr Ebenen betroffen sind, desto größer ist das Risiko einer Auffälligkeit.

Möchte man Ansatzpunkte für die Auffälligkeit herausarbeiten, so stellt man fest, dass drei dieser vier Ebenen Potential hierfür aufzeigen:

1. **Ärzt*innen:** Wertevorstellungen, Überzeugungen und Erfahrungen sind schwer zu beeinflussen. Eine gute Informationsbasis zu Vor- und Nachteilen von Arzneimitteln und auch den aktuell geltenden Regularien kann hier jedoch eine Möglichkeit darstellen, Überzeugungen und Erfahrungen zu beeinflussen
2. **Praxis und Patientenklientel:** Diese Faktoren werden in dem aktuellen Gesundheitssystem als gegeben gesehen und stellen keinen Ansatzpunkt für eine Verbesserung dar
3. **Vorgaben der bayerischen Wirkstoffvereinbarung:** Die von den Ärzt*innen genannten Nachteile und Probleme, sowie häufige problematische Verordnungen (z.B. Antikoagulantienziel) bieten möglicherweise einen Ansatzpunkt, um das System zu verbessern
4. **Gesundheitssystem:** Da sich die Patientenversorgung und das Gesundheitssystem sehr facettenreich darstellt und viele Bereiche auf Grund der expandierenden Kosten reguliert sind, ist es sinnvoll bei der Erstellung von Regularien die Verknüpfung mit anderen Bereichen im Blick zu haben. Eine Singularisierung des Problems und Projektion auf die einzelne Arztperson bzw. Betriebsstätte scheint hier Potential für Schwierigkeiten mit sich zu bringen. Arzneimittelkosten hängen somit nicht nur von dem einzelner/n Ärzt*in, Patient*innen und Erkrankungsstadium, sondern auch von dem Zusammenspiel des Gesundheitssektors ab.

Auch wenn die WSV in einzelnen Aspekten von den Zielverfehlern durchaus unterschiedlich bewertet wurde, so wurde weder die ihr zugrundeliegende Absicht, noch das Prozedere selbst grundsätzlich in Frage gestellt. Die Kombination von qualitativen und quantitativen Methoden erlaubte neue Einblicke in die Hintergründe des Phänomens „Zielverfehlung“, die über den Geltungsbereich der WSV hinaus relevant sein dürften.

7.5 Teilprojekt E

Ziel dieses Teilprojekts war es, ein allgemeines Analyseverfahren zur Auffindung von auffälligen Betriebsstätten zu entwickeln, das nicht von einzelnen Prüfvereinbarungen der Wirtschaftlichkeitsprüfung abhängt. Dies wurde mit dem hier präsentierten Analyseverfahren umgesetzt.

Hervorzuheben ist, dass sich dieses neue Verfahren aus der Bayesschen Inferenz und dem Machine-Learning-Ansatz zusammensetzt. Diese innovative Zusammensetzung bietet zum einen die Möglichkeit, auf der wirklich angenommenen Verteilung der zu bewertenden Indikatorvariable zu arbeiten, und zum anderen trotz Vorliegen einer Nicht-Normalverteilung ein Regressionsmodell aufzustellen, das nicht von der Normalverteilungsannahme abhängt. Somit wird weder die vorliegende Verteilung durch zu treffenden Annahmen verzerrt, noch die Möglichkeit der Ermittlung der einflussgebenden Variablen eingeschränkt. Zusammen mit dem spieltheoretischen Ansatz der Shapley Values kann nicht nur bestimmt werden, welche exogenen Variablen einen Einfluss auf die jeweilige Zielvariable ausüben, sondern auch einordnen, wie groß diese Einflüsse jeweils sind.

Mithilfe dieses Analyseverfahrens, kann mithilfe einer finalen Feinjustierung der Kontrollgrenzen die Menge aller echt auffälligen Einheiten (hier: Betriebsstätten) bestimmt werden. Diese Feinjustierung hängt vom jeweils vorliegenden Setting innerhalb einer KV ab (z.B. Kriterien, ab welchem Schwellenwert Betriebsstätten als auffällig gelten sollen) und müsste individuell durchgeführt werden.

Unabhängig von dieser Feinjustierung erhält man durch dieses Analyseverfahren diejenigen Variablen, die die zu betrachtenden Größe der Wirtschaftlichkeitsmessung maßgeblich beeinflussen. Dadurch kann besser unterschieden werden, welche Einheiten tatsächlich auffällig sein könnten und welche es nur zunächst erscheinen, aber nicht sind. Zudem kann aus der Menge der betrachteten Betriebsstätten die Untergruppe der Grenzfälle ermittelt werden, für welche sich eine genauere Betrachtung und Analyse des jeweiligen Ordnungsverhaltens lohnt, um festzustellen, ob diese Praxis schon auffällig oder die Verordnungen noch begründbar sind im Sinne des Wirtschaftlichkeitsgebots. Allerdings fehlte bei dieser Untersuchung der „Referenzstandard“, der sich z.B. durch eine individuelle „händische“ Prüfung ergeben würde. Anzustreben sind deshalb Untersuchungen, welche die Resultate eines Profiling mit der Prüfung des Ordnungsverhaltens einzelner Betriebsstätten vergleichen. Dem könnten Stichproben von auffälligen und nicht-auffälligen Betriebsstätten zugrunde liegen.

Dieses hier präsentierte Analyseverfahren führt zu gut interpretierbaren Ergebnissen und ist dabei allgemein genug, um auf beliebige Indikatoren und die jeweils gewünschten Adjustierungen anwendbar zu sein. Dabei ist es nicht eingeschränkt auf die klassischen Regressionsannahmen, wie z.B. der Normalverteilungsannahme. Damit kann die Wirklichkeit gut modelliert und analysiert werden ohne durch nötige Annahmen verzerrt zu werden. Insgesamt ist dieses Verfahren in der Umsetzung zwar komplex, aber wegen der damit erzeugten Ergebnisse lohnenswert.

8. Verwendung der Ergebnisse nach Ende der Förderung

8.1 Die Wirtschaftlichkeits-Vereinbarung der KVB

Die Wirkstoffvereinbarung (WSV) der Kassenärztlichen Vereinigung Bayerns wird unabhängig von der Förderung weitergeführt. Die Befragungsergebnisse machen deutlich, dass die niedergelassenen Ärzte die Zielsetzung der WSV unterstützen. Auch die Implementierung wird überwiegend als positiv eingeschätzt. Allerdings werden die Erkenntnisse des WirtMed-Projekts genutzt, um die WSV kontinuierlich weiterzuentwickeln.

8.2 Dauerhafte Nutzung von Erkenntnissen und Instrumenten: die Toolbox

Das WirtMed hat Erkenntnisse über die Verordnung von Medikamenten in der ambulanten Versorgung erbracht, relevante Faktoren identifiziert und die Wirksamkeit von Maßnahmen untersucht, die Qualität bzw. die Wirtschaftlichkeit der Verordnungen positiv zu beeinflussen. Wesentlich sind jedoch die Entwicklung und Erprobung von Methoden gewesen, die Wirtschaftlichkeit und Qualität der Verordnungen zu erfassen. Konkret haben diese einen doppelten Fokus:

- 1) Die Identifizierung von Einheiten (Betriebsstätten), welche in Bezug auf die Wirtschaftlichkeit oder die Qualität ihrer Verordnungen von einem Standard bzw. der Vergleichsgruppe in nicht angemessener Weise abweichen – „individuelles Profiling“ (Teilprojekte A, D, E)
- 2) Die Evaluation von Verfahren der Prüfung bzw. Steuerung von Qualität bzw. Wirtschaftlichkeit (Teilprojekte B, C, D)

Die methodischen Erkenntnisse aus dem WirtMed-Projekt haben wir in einer „Toolbox“ zusammengestellt. Diese soll für Kassenärztliche Vereinigungen und andere Körperschaften mit Verantwortung für die Wirtschaftlichkeit der Medikamenten-Verordnungen in der ambulanten Versorgung zugänglich sein und konkrete Hilfen für die genannten zwei Aufgaben bieten.

Im Einzelnen enthält die Toolbox folgende Komponenten (html-Format mit verlinkten Dokumenten):

- Einleitung und Überblick
- Methodische Einführung: Auswahl von Indikatoren einschließlich deren Begründung, Vor- und Nachteilen; Berücksichtigung von Störfaktoren (Confoundern)
- Individuelles Profiling: Identifizierung von „Outliern“ in Bezug auf einschlägige Indikatoren
- Diffusionsanalysen: Stellen die Zusammenhänge zwischen den einzelnen Betriebsstätten dar (Diffusion, Verordner-Netze, Patientenpfade)
- Evaluation von Strategien: Quantitative Analyse von Verordnungsdaten
- Evaluation von Strategien: qualitative Erhebungen
- Ziel-Verfehlen: Untersuchung von auffälligen Betriebsstätten auf Praxisbesonderheiten (d.h. evtl. Einstufung als „falsch positive“) sowie auf Ursachen und Motive für nicht optimale Medikamenten-Verordnungen sowie Barrieren gegen eine optimale Verordnungs-Praxis.

Diese Toolbox wird nur für Angehörige der GKV-Selbstverwaltung (v.a. KVen, Krankenkassen, Prüfstellen) sowie von wissenschaftlichen Einrichtungen zugänglich.

Das Zentralinstitut für die kassenärztliche Versorgung (Zi) in Berlin hat sich bereiterklärt, diese Toolbox zu hosten. Diese Toolbox soll im Licht wissenschaftlicher Ergebnisse und praktischer

Erfahrungen genutzt werden. Dies schließt die Option der Weiterentwicklung durch involvierte Akteure ein. Soweit dies der Zielsetzung dient, können auch Hinweise auf externe Ressourcen (methodische Publikationen, Beispielsprojekte, Online-Fortbildungen externer Anbieter usw.) erfolgen.

Als wichtigste Handlungs- und Transferempfehlung ist in diesem Projekt die Toolbox zu bewerten, deren umfangreichen Methoden für die Regelversorgung angewendet werden können.

8.3 Der „blinde Fleck“ der Wirtschaftlichkeitsprüfung

Krankenhäuser üben einen wesentlichen Einfluss auf die ambulante Verordnung von Medikamenten aus. Neben intraprofessionellem Austausch (informell, Fortbildungs-Veranstaltungen u.ä.) ist hier vor allem eine während der stationären Behandlung begonnene Dauertherapie von Bedeutung. § 115 Abs. 1 SGB V sieht vor, dass die Krankenhäuser dem weiterbehandelnden Vertragsarzt die dazu erforderlichen Wirkstoffe mitteilen. Abs. 2 fordert ausdrücklich, dass auch hier nur zweckmäßige und wirtschaftliche Wirkstoffe zum Einsatz kommen sollen. Entsprechendes sieht § 8 der Arzneimittelrichtlinie vor.

In §§ 296, 297 und 300 SGB V sind die Verpflichtungen der KVen und der Krankenkassen zur Übermittlung von Verordnungs- und Leistungsdaten geregelt.

Bisher fehlt eine Verpflichtung der Krankenhäuser, die bei der Entlassung nach stationärer Behandlung initiierte Dauermedikation der Krankenkasse oder einer anderen Prüf-Einrichtung (z.B. dem MDK) mitzuteilen.

Dies führt dazu, dass die im ambulanten Bereich durchgeführten Verfahren der Prüfung und Steuerung einen wesentlichen Einflussfaktor nicht erfassen können, nämlich eine während einer stationären Behandlung begonnene oder geänderte Dauermedikation. Die in diesem Fall erstmalig durch einen Vertragsarzt ausgestellte Verordnung wird ihm in der gegenwärtigen rechtlichen Situation zugerechnet, obwohl es sich de facto um eine Folgeverordnung handelt. Wie weit die während einer stationären Behandlung initiierten Verordnungen den Kriterien von Zweckmäßigkeit und Wirtschaftlichkeit (§ 115c Abs. 2 SGB V) genügen, ist nicht bekannt, da eine objektive Datenbasis für entsprechende Untersuchungen fehlt. Allerdings legen Rückmeldungen niedergelassener Ärztinnen, die im Rahmen des WirtMed-Projektes erfasst wurden, durchaus Defizite nahe. Da keine belastbare Datenbasis vorliegt, fehlt auch die Möglichkeit, auf die Wirtschaftlichkeit und Qualität dieser Verordnung Einfluss zu nehmen, sei dies durch Feedback, Benchmarking, spezifische medizinische Information oder Sanktionen.

Wir schlagen vor, dass eine bundeseinheitliche Regelung getroffen wird, auch die von Krankenhäusern bei Entlassung verordneten Medikamente in die Prüfung der Wirtschaftlichkeit nach § 12 SGB V einzubeziehen.

9. Erfolgte bzw. geplante Veröffentlichungen

9.1 Teilprojekt A

9.1.1 Kongressbeiträge

Flemming, R. für das WirtMed Konsortium (2021). Diffusion of a new drug among ambulatory physicians – The influence of patients' pathways. Vortrag im Rahmen der Vortragsitzung "Arzneimittel: Bepreisung, Zugang, Nebenwirkungen" bei der 13. dggö Jahrestagung 2021, 09.03.2021 [online].

Sundmacher, L. Flemming, R. für das WirtMed Konsortium (2020). Diffusion of a new drug among ambulatory physicians in Germany. Vortrag bei der EUHEA student supervisor conference, 04.09.2020 [online]

Sundmacher, L. Flemming, R. für das WirtMed Konsortium (2020). Diffusion of a new drug among ambulatory physicians in Germany. Vortrag bei dem Deutschen Kongress für Versorgungsforschung, 30.09.2020 [online]

9.1.2 Wissenschaftliche Artikel in Zeitschriften

Flemming R. Patterns of pregabalin prescribing in four German federal states: analysis of routine data to investigate potential misuse of pregabalin. *BMJ Open* 2022;12:e060104. doi: 10.1136/bmjopen-2021-060104.

Flemming, R., Frölich, F., Donner-Banzhoff, N., Sundmacher, L. Diffusion of a new drug among ambulatory physicians – The impact of patient pathways. *Health Economics*. 2023; 1– 13. doi: 10.1002/hec.4650
Flemming, R., Donner-Banzhoff, N., Sundmacher, L. (in Planung). Untersuchung der regionalen Nutzung neuer oraler Antikoagulanzen in Deutschland [Arbeitstitel]. (in Planung).

9.2 Teilprojekt B

9.2.1 Kongressbeiträge

Franiel, D., Schulz M., von Stillfried, D., Donner-Banzhoff, N., Muth, J., Gollnick, J., Hörbrand, F., Langheinrich, W. (2020). Die Bayerische Wirkstoffvereinbarung - ein Vorher-Nachher- bzw. Parallel-Vergleich. Postervortrag auf dem 19. *Deutschen Kongress für Versorgungsforschung*, 30. September - 01. Oktober 2020, Digital.

9.2.2 Wissenschaftliche Artikel in Zeitschriften

Franiel, D., Schulz, M., von Stillfried, D., Hörbrand, F., Killian, P. (in Arbeit). The effects of implementing an active substance agreement – a methodological before and after comparison. [Arbeitstitel] (in Arbeit).

Franiel, D., Schulz, M., von Stillfried, D., Hörbrand, F., Killian, P., (in Planung). Die WirtMed-Studie – Bewertung alternativer Methoden zur Steuerung von Arzneimittelverordnungen. [vorläufiger Titel] (in Planung).

9.3 Teilprojekt C

9.3.1 Kongressbeiträge

Lippert, N., Gollnick, J., Muth, J., Hörbrand, F., Donner-Banzhoff, N., Kühlein, T., & Sebastiao, M. (2021). Ärztliche Erfahrungen mit der bayerischen Wirkstoffvereinbarung als Systematik zur Steuerung einer wirtschaftlichen Arzneimittelverordnung – Qualitative Ergebnisse aus der

WirtMed-Studie. Vortrag auf dem 55. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Lübeck, 16. – 18. September 2021. <https://dx.doi.org/10.3205/21degam133>

Muth, J., Gollnick, J., Schulz, M., Franiel, D., Sundmacher, L., Flemming, R., Kühlein, T., Lippert, N., Hörbrand, F., Killian, P., Langheinrich, W., Mohr, A., Adolf, W., Plachteka, H., Paganoni, A.M., Vedove, D.D., Hirsch, O., & Donner-Banzhoff, N. (2021). WirtMed-Studie – Entwicklung alternativer Methoden zur Steuerung von Arzneimittelverordnungen. Vortrag auf dem 55. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Lübeck, 16. – 18. September 2021. <https://dx.doi.org/10.3205/21degam021>

Gollnick, J., Lippert, N., Muth, J., Hörbrand, F., Kühlein, T., & Donner-Banzhoff, N. (2021). Unwirtschaftliche Arzneimittelverordnungen – realistische Auswege aus der Unwirtschaftlichkeit? Vortrag auf dem 55. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Lübeck, 16. – 18. September 2021. <https://dx.doi.org/10.3205/21degam020>

Gollnick, J., Lippert, N., Muth, J., Hörbrand, F., Kühlein, T., & Donner-Banzhoff, N. (2021). Wirtschaftlichkeit bei Arzneimittelverordnungen - lediglich durch das Ordnungsverhalten von Ärzt*innen beeinflusst oder gibt es auch andere Einflussfaktoren? Vortrag auf dem 55. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Lübeck, 16. – 18. September 2021. <https://dx.doi.org/10.3205/21degam019>

Lippert, N., Gollnick, J., Muth, J., Fischaleck, J., Donner-Banzhoff, N., Sebastiao, M., & Kühlein, T. (2020). Evaluation der bayerischen Wirkstoffvereinbarung aus Sicht niedergelassener ÄrztInnen – Die WirtMed-Studie zur Wirtschaftlichkeit und Qualität von Medikamenten-Verordnungen (Teilprojekt C). Vortrag auf dem 54. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Erlangen, 16. – 19. Dezember 2020.

Gollnick, J., Lippert, N., Muth, J., Fischaleck, J., Kühlein, T., Donner-Banzhoff, N. (2020). 1 Slide 5 Minutes Vortrag: Regulating drug prescription in Bavaria: why are physicians fail to meet their efficiency targets? Wonca Congress, digital.

Lippert, N., Gollnick, J., Hermann, M., Muth, J., Hörbrand, F., Donner-Banzhoff, N., & Kühlein, T. (2020). Regresse und Wirtschaftlichkeitsprüfverfahren: Quo vadis? - Erfahrungen aus der WirtMed-Studie. Vortrag am Tag der Allgemeinmedizin, Tag der Weiterbildung und Tag der Lehre, 10. Oktober 2020, Digital.

Lippert, N., Gollnick, J., Gögsen, A., Hermann, M., Muth, J., Hörbrand, F., Donner-Banzhoff, N., Kühlein, T., & Sebastiao, M. (2020). "Ich übernehme nicht alles unkritisch." Arzneimittelverordnungen nach Krankenhausbehandlungen als Herausforderung für Hausärzt*innen. Eine qualitative Befragung im Rahmen der WirtMed-Studie. Vortrag auf dem 19. Deutschen Kongress für Versorgungsforschung, 30. September - 01. Oktober 2020, Digital.

Gollnick, J., Lippert, N., Muth, J., Hörbrand, F., Kühlein, T., & Donner-Banzhoff, N. (2020). WirtMed-Studie zur Wirtschaftlichkeit der Arzneimittel-Verordnung – immer diese rote Ampel! Vortrag auf dem 19. Deutschen Kongress für Versorgungsforschung, 30. September - 01. Oktober 2020, Digital.

Lippert, N., Gollnick, J., Muth, J., Fischaleck, J., Donner-Banzhoff, N., & Kühlein, T. (2019). Symposiumsbeitrag: Kritische Reflexion zur Rekrutierung und Durchführung qualitativer Einzelinterviews in der Versorgungsforschung. Symposiumsvortrag „Womit haben wir es eigentlich zu tun?“ – Reflexion verschiedener Studiendesigns in der allgemeinmedizinischen Versorgungsforschung auf dem 53. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Erlangen, 12. – 14. September 2019.

Lippert, N., Gollnick, J., Muth, J., Fischaleck, J., Donner-Banzhoff, N., van der Keylen, P., & Kühlein, T. (2019). Evaluation der bayerischen Wirkstoffvereinbarung aus Sicht der niedergelassenen ÄrztInnen – Die WirtMed-Studie zur Wirtschaftlichkeit und Qualität von Medikamenten-Verordnungen. 1 Slide 5 Minutes Vortrag auf dem 53. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Erlangen, 12. – 14. September 2019.

9.3.2 Wissenschaftliche Artikel in Zeitschriften

Zeschick, N., Gollnick, J., Muth, J., Hörbrand, F., Killian, P., Krombholz, W., Donner-Banzhoff, N., Kühlein, T., & Sebastião, M. (under review). Ordnungsverhalten von bayerischen Hausärzt*innen an der stationär-ambulanten Schnittstelle vor dem Hintergrund der Bayerischen Wirkstoffvereinbarung – Qualitative Ergebnisse der WirtMed-Studie. *Bundesgesundheitsblatt* (under review).

Zeschick, N., Gollnick, J., Muth, J., Hörbrand, F., Killian, P., Donner-Banzhoff, N., Kühlein, T., & Sebastião, M. (eingereicht). Physicians assessment of the Bavarian drug-expenditure control system: A qualitative study. *The European Journal of Health Economics* (eingereicht).

9.4 Teilprojekt D

9.4.1 Kongressbeiträge:

Muth, J., Gollnick, J., Schulz, M., Franiel, D., Sundmacher, L., Flemming, R., Kühlein, T., Lippert, N., Hörbrand, F., Killian, P., Langheinrich, W., Mohr, A., Adolf, W., Plachteka, H., Paganoni, A.M., Vedove, D.D., Hirsch, O., & Donner-Banzhoff, N. (2021). WirtMed-Studie – Entwicklung alternativer Methoden zur Steuerung von Arzneimittelverordnungen. Vortrag auf dem 55. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Lübeck, 16. – 18. September 2021.

Gollnick, J., Lippert, N., Muth, J., Hörbrand, F., Kühlein, T., Donner-Banzhoff, N. (2021). Unwirtschaftliche Arzneimittelverordnungen – realistische Auswege aus der Unwirtschaftlichkeit? Vortrag auf dem 55. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Lübeck, 16. – 18. September 2021.

Gollnick, J., Lippert, N., Muth, J., Hörbrand, F., Kühlein, T., Donner-Banzhoff, N. (2021). Wirtschaftlichkeit bei Arzneimittelverordnungen - lediglich durch das Ordnungsverhalten von Ärzt*innen beeinflusst oder gibt es auch andere Einflussfaktoren? Vortrag auf dem 55. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Lübeck, 16. – 18. September 2021.

Gollnick, J., Lippert, N., Muth, J., Fischaleck, J., Kühlein, T., Donner-Banzhoff, N. (2020). 1 Slide 5 Minutes Vortrag: Regulating drug prescription in Bavaria: why are physicians fail to meet their efficiency targets? Wonca Congress, digital, 16.-19. Dezember 2020.

Gollnick, J., Lippert, N., Muth, J., Hörbrand, F., Kühlein, T., Donner-Banzhoff, N. (2020). WirtMed-Studie zur Wirtschaftlichkeit der Arzneimittel-Verordnung – immer diese rote Ampel! Vortrag auf dem 19. Deutschen Kongress für Versorgungsforschung, 30. September - 01. Oktober 2020, Digital.

Gollnick, J., Lippert, N., Muth, J., Fischaleck, J., Kühlein, T., Donner-Banzhoff, N. (2019) Ich sehe nur noch rot! Die WirtMed-Studie zur Wirtschaftlichkeit der Medikamenten-Verordnung. Vortrag auf dem 53. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Erlangen, 12.-14. September 2019

Lippert, N., Gollnick, J., Muth, J., Fischaleck, J., Donner-Banzhoff, N., Kühlein, T. (2019). Symposium: Kritische Reflexion zur Rekrutierung und Durchführung qualitativer Einzelinterviews

in der Versorgungsforschung. „Womit haben wir es eigentlich zu tun?“ – Reflexion verschiedener Studiendesigns in der allgemeinmedizinischen Versorgungsforschung'. Symposiumsbeitrag auf dem 53. Kongress für Allgemeinmedizin und Familienmedizin (DEGAM), Erlangen, 12. – 14. September 2019.

9.4.2 Wissenschaftliche Artikel in Zeitschriften:

J. Gollnick, F. Hörbrand, N. Lippert, J. Muth, T. Kühlein, N. Donner-Banzhoff. Drug prescribing – The characterization of drug groups and their impact of missing the efficiency targets in the outpatient sector (*Journal of Health Economics*) (in Arbeit)

J. Gollnick, F. Hörbrand, N. Lippert, J. Muth, T. Kühlein, N. Donner-Banzhoff. Economically inefficient drug-prescribing – not only caused by reluctant prescribers: Reasons, influencing factors and potential solutions (*Healthcare*) (in Arbeit)

9.5 Teilprojekt E

9.5.1 Kongressbeiträge

Delle Vedove, D., Paganoni, A., Ieva, F., Spreafico, M., Franiel, D., Schulz, M., Hirsch, O., Muth, J., Gollnick, J., Donner-Banzhoff, N. (2020). WirtMed study: Development and testing of evaluation strategies for a suitable efficiency evaluation. Postervortrag auf dem *19. Deutschen Kongress für Versorgungsforschung*, 30. September - 01. Oktober 2020, Digital.

9.5.2 Wissenschaftliche Artikel in Zeitschriften

Delle Vedove, D., Paganoni, A., Franiel, D., Schulz, M., Donner-Banzhoff, N. (in Arbeit). A novel methodology for profiling medical prescribing. [Arbeitstitel] (in Arbeit).

10. Literaturverzeichnis

- Bretz F, Hothorn T, Westfall P, editors. Multiple Comparisons Using R. Chapman and Hall/CRC; 2016.
- Chow GC. Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*. 1960;28(3):591. doi:10.2307/1910133.
- Dresing T, Pehl T. Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende. 7th ed. Marburg: Eigenverlag; 2017.
- European Medicines Agency. Lyrica. 2010. <https://www.ema.europa.eu/en/medicines/human/EPAR/lyrica>. Accessed 20 Oct 2022.
- Gemeinsamer Bundesausschuss. Beschlüsse über die Nutzenbewertung von Arzneimitteln mit neuen Wirkstoffen nach § 35a SGB V – Sacubitril/Valsartan. 2016. https://www.g-ba.de/downloads/40-268-3823/2016-06-16_AM-RL-XII_Sacubitril_Valsartan_D-207_TrG.pdf. Accessed 20 Oct 2011.
- Hastie T, Tibshirani R. Generalized Additive Models. *Statist. Sci.* 1986. doi:10.1214/ss/1177013604.
- Heumann C, Schomaker M, Shalabh. Introduction to statistics and data analysis: With Exercises, solutions and applications in R. Cham: Springer; 2016.
- Rogers EM. Diffusion of innovations. New York, London, Toronto, Sydney: Free Press; 2003.
- Spiegelhalter D, Sherlaw-Johnson C, Bardsley M, Blunt I, Wood C, Grigg O. Statistical methods for healthcare regulation: rating, screening and surveillance. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2012;175(1):1–47. doi:10.1111/j.1467-985X.2011.01010.x.

Akronym: WirtMed
Förderkennzeichen: 01VSF17016

11. Anhang

Anhang 1: Teilprojekt A

Anhang 2: Teilprojekt B

Anhang 3: Teilprojekt C

Anhang 4: Teilprojekt D

Anhang 5: Teilprohekt E

Akronym: WirtMed
Förderkennzeichen: 01VSF17016

12. Anlagen

Anlage 1: Teilprojekt B – Ergebnistabellen der Hauptergebnisse

Anlage 2: Toolbox

Anhang 1: Teilprojekt A

1.1 Wirkstoffliste

Wirkstoff	Handelsname	ATC	Priorität
Insulin glargin	Toujeo	A10AE04	2
Sitagliptin	Januvia	A10BH01	1
Saxagliptin	Onglyza	A10BH03	1
Liraglutid	Victoza	A10BJ02	2
Dulaglutid	Trulicity	A10BJ05	1
Dapagliflozin	Forxiga	A10BK01	1
Canagliflozin	Invokana	A10BK02	1
Empagliflozin	Jardiance	A10BK03	<i>Positivwirkstoff</i>
Prasugrel	Efient	B01AC22	2
Ticragelor	Brilique	B01AC24	1
Dabigatran	Pradaxa	B01AE07	1
Rivaroxaban	Xarelto	B01AF01	1
Apixaban	Eliquis	B01AF02	<i>Positivwirkstoff</i>
Edoxaban	Lixiana	B01AF03	<i>Positivwirkstoff</i>
Dronedaron	Multaq	C01BD07	2
Ivabradin	Procoralan	C01EB17	1
Eplerenon	Inspra	C03DA04	1
Sacubitril/Valsartan	Entresto	C09DX04	1
Ezetimib/Simvastatin	Inegy	C10BA02	1
Solifenacin	Vesikur	G04BD08	2
Etanercept	Enbrel	L04AB01	<i>Positivwirkstoff</i>
Febuxostat	Adenuric	M04AA03	1
Denosumab	Prolia	M05BX04	1
Oxycodon/Naloxon	Targin	N02AA55	1
Tapentadol	Palexia	N02AX06	1
Pregabalin	Lyrica	N03AX16	1
Duloxetine	Cymbalta	N06AX21	1
Agomelatin	Valdoxan	N06AX22	1
Atomoxetin	Strattera	N06BA09	2
Lisdexamfetamin	Elvanse	N06BA12	1
Dimethylfumarat	Tecfidera	N07XX09	2
Indacaterol/Glycopyrronium	Ultibro Breezhaler	R03AC18	2
Umeclidinium/Vilanterol	Anoro	R03AL03	2
Tiotropium/Olodaterol	Spiolto	R03AL06	2

1.2 Beschreibung der Datensätze je Analyse

Analyse	Beschreibung des Analysedatensatzes
<p>1. Analyse Sacubitril/Valsartan (ATC C09DX04)</p>	<p><u>Aufgreifkriterien</u></p> <ul style="list-style-type: none"> – Ärzt*innen, die im Zeitraum von 2016 bis 2017 mindestens einmal den Wirkstoff Sacubitril/Valsartan verordnet haben und alle Ärzt*innen, die im Beobachtungszeitraum mindestens sechsmal und im Jahr 2016 mindestens einmal ein Medikament aus der Gruppe der Wirkstoffe, die auf das Renin-Angiotensin-System wirken (ATC C09), verordnet haben. – Patient*innen des Analysedatensatzes, die bei den eingeschlossenen ÄrztInnen zwischen 2015 und 2017 in Behandlung waren. <p><u>Datenbasis</u> Vollständige Abrechnungsdaten der eingeschlossenen Patient*innen des Zeitraums von 2015 bis 2017. Analyseeinheit sind Ärzt*innen. → <u>Anzahl ÄrztInnen: 7.533</u></p>
<p>2. Analyse NOAKs Apixapan (ATC B01AF02), Edoxaban (ATC B01AF03), Rivaroxaban (ATC B01AF01), Dabigatran (ATC B01AE07)</p>	<p><u>Aufgreifkriterien</u></p> <ul style="list-style-type: none"> – Patient*innen, die ab August 2015 bis März 2017 eine initiale Verordnung* für eines der neuen oralen Antikoagulanzen (ATC B01AF02, B01AF03, B01AF01, B01AE07) und keinen der anderen Wirkstoffe zuvor verordnet bekommen haben. – Ausschluss von den im Beobachtungszeitraum verstorbenen Patient*innen. <p><u>Datenbasis</u> Vollständige Abrechnungsdaten der eingeschlossenen Patient*innen des Zeitraums von August 2014 bis 2017. Kleinste Analyseeinheit sind Patient*innen. → Anzahl Patient*innen: 106.061</p>
<p>3. Analyse Pregabalin (ATC N03AX16)</p>	<p><u>Aufgreifkriterien</u></p> <ul style="list-style-type: none"> – Patient*innen mit einer initialen Verordnung* des Wirkstoffs Pregabalin (ATC N03AX16) im Zeitraum von 2014 bis 2017. – Ausschluss von Patient*innen jünger als 12 Jahre, Patient*innen mit unvollständigen Stammdaten, Patient*innen, die im Beobachtungszeitraum verstorben sind oder Patient*innen mit weniger als drei Verordnungen des Wirkstoffs im Beobachtungszeitraum von einem Jahr. – Alle im Beobachtungszeitraum behandelten Patient*innen der Ärzt*innen, die Pregabalin an die betrachtete Kohorte der – Patient*innen verordnet haben.

	<p><u>Datenbasis</u> Vollständige Abrechnungsdaten der eingeschlossenen Patient*innen. Der Beobachtungszeitraum je Patient*in war jeweils ein Jahr vor und ein Jahr nach Initialverordnung. Kleinste Analyseeinheit sind Patient*innen. ➔ Anzahl Patient*innen: 53.049</p>
<p>4. Analyse Oxycodon & Naloxon (ATC N02AA55)</p>	<p><u>Aufgreifkriterien</u></p> <ul style="list-style-type: none"> – Patient*innen mit einer initialen Verordnung* des Arzneimittels Oxycodon & Naloxon (ATC N02AA55) im Zeitraum von 2014 bis 2017. – Ausschluss von Patient*innen jünger als 18 Jahre, Patient*innen mit unvollständigen Stammdaten, Patient*innen, die im Beobachtungszeitraum von einem Jahr verstorben sind. <p><u>Datenbasis</u> Vollständige Abrechnungsdaten der eingeschlossenen Patient*innen. Der Beobachtungszeitraum je Patient*in war jeweils ein Jahr vor und ein Jahr nach Initialverordnung. Kleinste Analyseeinheit sind Patient*innen. ➔ Anzahl Patient*innen: 38.053</p>

Hinweis: * Die initiale Verordnung eines Medikaments wird so definiert, dass es keine Verordnung während des Vorjahres gab.

1.3 Übersicht der vier NOAKs in der 2. Analyse

	Zul.	AMNOG	Vergleichstherapie	Anwendungsgebiet
Dabigatran	2008	Nein	Heparine ➔ vergleichbare Wirksamkeit aber teurer	Zur Thromboseprophylaxe bei Knie- oder Hüftgelenkersatz (TEP); Schlaganfallprophylaxe bei Vorhofflimmern; Behandlung von TVT und LE
Rivaroxaban	2008	Nein	Heparine ➔ überlegene Wirksamkeit auf die Thromboembolie	Zur Thromboseprophylaxe bei Knie- oder Hüft-TEP; Schlaganfallprophylaxe bei Vorhofflimmern; Behandlung von TVT und LE ; zusätzlich in Kombination zur Sekundärprophylaxe nach Akuten Koronarsyndrom

Apixaban, Positivwirkstoff	2011	Geringer/ Kein Zusatznut- zen	2011 Heparine (Enoxaparin) → kein Zusatznutzen für Knie TEP, aber Hinweis für geringen Zusatz- nutzen bei Hüft TEP 2013 Vitamin-K-Antagonisten → Hinweis für geringen Zusatznut- zen 2015 Heparine & Warfarin → Hinweis für geringen Zusatznut- zen, bzw. keinen Zusatznutzen	Zur Thromboseprophylaxe bei Knie- oder Hüft-TEP; Schlaganfallprophylaxe bei Vorhofflimmern; Behandlung von TVT und LE
Edoxaban Positivwirk-	2015	Geringer/ Kein Zusatz- nutzen	Vitamin-K-Antagonisten (Warfa- rin) → Hinweis für geringen Zu- satznutzen Vitamin-K-Antagonisten (Warfa- rin) → Kein Zusatznutzen	Schlaganfallprophylaxe bei Vorhofflimmern; Behandlung von TVT und LE

1.4 Deskriptive Ergebnisse der 4. Analyse (Targin)

	Gesamt		Gruppe 1		Gruppe 2		Gruppe 3	
	N/Ø	%	N/Ø	%	N/Ø	%	N/Ø	%
Anzahl Patient*innen	38.053	100,0	249	0,7	21,422	56,3	16,382	43,1
Männer	13.521	35,5	112	45,0	7,484	34,9	5,925	36,2
Frauen	24.532	64,5	137	55,0	13,938	65,1	10,457	63,8
Alter 18-29 Jahre	441	1,2	5	2,0	177	0,8	259	1,6
30-39 Jahre	891	2,3	9	3,6	394	1,8	488	3,0
40-49 Jahre	2.497	6,6	34	13,7	1,204	5,6	1,259	7,7
50-59 Jahre	5.389	14,2	60	24,1	2,841	13,3	2,488	15,2
60-69 Jahre	6.442	16,9	47	18,9	3,472	16,2	2,923	17,8
> 69 Jahre	22.393	58,8	94	37,8	13,334	62,2	8,965	54,7
Wohnort PatientIn- nen								
Städtisch	15.213	40,0	15	46,2	8,630	40,3	6,468	39,5
Ländlich	22.790	59,9	134	53,8	12,772	59,6	9,884	60,3
Diagnostizierte Indikation								
Schmerzen	20.890	54,9	128	51,4	11,533	53,8	9,229	56,3
Neubildungen	2.076	5,5	15	6,0	1,198	5,6	863	5,3
Restless Legs	574	1,5	2	0,8	376	1,8	196	1,2
Mehrere	7.151	18,8	56	22,5	4,362	20,4	2,733	16,7
Keine	7.362	19,3	48	19,3	3,953	18,5	3,361	20,5
Vormedikation mit Oxycodon	2.543	6,7	49	19,7	1,687	7,9	807	4,9
Obstipation	3.476	9,1	41	16,5	2,230	10,4	1,205	7,4
Oxycodon & Obstipation	306	0,8	15	6,0	197	0,9	94	0,6
Krankenhausaufenthalt	13.827	36,3	52	20,9	7,772	36,3	6,003	36,6
Keine Indikation	3.730	9,8	31	12,4	1,985	9,3	1,714	10,5
Obstipation (Nebenwirkung)	3.266	8,6	37	14,9	2,215	10,3	1,014	6,2

Opioide	25.696	67,5	249	100,0	19,075	89,0	6,372	38,9
Fachgebiet der Erstverordnung								
<i>Allgemeinmedizin</i>	31.348	82,4	177	71,1	17,725	82,7	13,446	82,1
<i>Orthopädie</i>	1.583	4,2	6	2,4	744	3,5	833	5,1
<i>Anästhesiologie</i>	953	2,5	37	14,9	597	2,8	319	1,9
<i>Neurologie</i>	715	1,9	5	2,0	429	2,0	281	1,7
<i>Neurochirurgie</i>	513	1,3	2	0,8	224	1,0	287	1,8
<i>Sonstige</i>	2.941	7,7	22	8,8	1,703	7,9	1,216	7,4
Anzahl Verordnungen	3,9		12,7		6,0		1	
Anzahl verordnender ÄrztInnen	1,4		2,2		1,7		1	
Anzahl verordnender Praxen	1,2		1,7		1,4		1	

1.5 Deskriptive Ergebnisse der 1. Analyse (S/V)

	Modell 1 Durchschnitt [Min; Max] Absolut (%)	Modell 2 Durchschnitt [Min; Max] Absolut (%)	Modell 3 Durchschnitt [Min; Max] Absolut (%)	Modell 4 Durchschnitt [Min; Max] Absolut (%)
Anzahl Ärzt*innen	n = 7.198	n = 4.689	n = 2.718	n = 1.980
Monat der Erstver- ordnung	12,72 [1; 24]	12,26 [1; 24]	12,24 [1; 24]	12,27 [1; 24]
Praxistyp				
Einzelpraxis	1.823 (25,33%)	1.349 (28,71%)	760 (27,96%)	589 (29,75%)
Gemeinschaftspraxis	5.315 (73,84%)	3.313 (70,52%)	1.944 (71,52%)	1.369 (69,14%)
Sonstige	60 (0,83%)	36 (0,77%)	14 (0,52%)	22 (1,11%)
Facharztgruppe				
Allgemeinmedizin	6.217 (86,37%)	3.989 (84,91%)	2.260 (83,15%)	1.729 (87,32%)
Internistische Medizin	197 (2,74%)	136 (2,89%)	77 (2,83%)	59 (2,98%)
Kardiologie	613 (8,52%)	485 (10,32%)	330 (12,14%)	155 (7,83%)
Sonstige	171 (2,38%)	88 (1,87%)	51 (1,88%)	37 (1,87%)
Städtische Region	3.163 (43,94%)	1.947 (41,44%)	1.119 (41,17%)	828 (41,82%)
Krankenhausbetten (pro Kreis und 100.000 Einw.)	627,7 [0; 2.986]	627,6 [0; 2.986]	627,6 [0; 2.986]	627,6 [0; 2.986]

1.6 Regressionsergebnisse der 1. Analyse (S/V)

	Modell 1 HR	Modell 2 HR	Modell 3 HR	Modell 4 HR	p-value
Expositionswert (ref. 0%)					
≤ 25%	1,593	1,298	1,190	1,452	0,000
≤ 50%	2,043	1,343	1,263	1,450	0,000
≤ 75%	2,736	1,436	1,335	1,619	0,000
≤ 100%	2,755	1,261	1,224	1,325	0,111
Anzahl verbundener Ärzt*innen im empirischen Ärztenetzwerk	1,002	1,003	1,003	1,003	0,069
Patienten mit einer Vormedikation mit S/V (ref. keine)					
ausgestellt von einer/m Ärzt*in der gleichen Praxis	3,351	1,511	1,438	1,436	0,079

ausgestellt von einer/m Ärzt*in ei- ner anderen Praxis	2,76 0	0,000	1,127	0,238	1,108	0,410	1,102	0,578
Praxistyp (ref. Einzelpraxis)								
Gemeinschaftspraxis	0,88 1	0,001	0,875	0,003	0,874	0,022	0,878	0,060
Sonstiges	0,90 6	0,547	1,131	0,504	0,841	0,603	1,373	0,162
Verordnung von einer/m Kolleg*in[†]	1,17 6	0,000	1,068	0,100	1,013	0,809	1,135	0,039
Facharztgruppe (ref. Allgemeinme- dizin)								
Internistische Medizin	1,17 8	0,049	1,172	0,101	1,192	0,171	1,103	0,515
Kardiologie	1,64 1	0,000	1,468	0,000	1,530	0,000	1,354	0,002
andere	0,90 1	0,306	0,982	0,888	0,989	0,948	0,931	0,728
Ländliche Region (ref. städtisch)	1,08 1	0,002	1,090	0,005	1,086	0,043	1,100	0,046
Anzahl Krankenhausbetten	1,00 0	0,662	1,000	0,533	1,000	0,632	1,000	0,767
Intercept	0,03 2	0,000	0,053	0,000	0,056	0,000	0,047	0,000
Smooth-Terme								
Zeit (in Monaten)	Ja	Ja	Ja	Ja	Ja	Ja	Ja	Ja
Zeit x Expositionswert	Teilweise	Nein	Nein	Nein	Nein	Teilweise	Teilweise	Teilweise
Zeit x PatientIn mit Vormedikation	Teilweise	Nein	Nein	Nein	Nein	Nein	Nein	Nein
Zeit x Praxistyp	Teilweise	Teilweise	Teilweise	Nein	Nein	Nein	Nein	Nein
Zeit x Facharztgruppe	Teilweise	Teilweise	Teilweise	Teilweise	Teilweise	Teilweise	Nein	Nein

Hinweise: Die Hinweise zur statistischen Signifikanz in dieser Tabelle (Ja, teilweise, nein) beziehen sich auf ein Signifikanzniveau von 10 %. [†] Kolleg*in meint hier ein/e Ärzt*in der gleichen Praxis.

1.7 Deskriptive Ergebnisse der 2. Analyse (NOAKs)

	Apixaban n = 39.297		Dabigatran n = 8.307		Edoxaban n = 12.053		Rivaroxaban n = 46.404		Gesamt n = 106.061		
Level 1 (Patienten)											
Geschlecht (weiblich)	22.130	56,3%	4.404	53,0%	6.398	53,1%	25.511	55,0%	58.443	55,1%	
Alter	75.0		73.5		74.7		69.7		72,6		
TEP	1.495	3,8%	355	4,3%	101	0,8%	5.627	12,1%	7.578	7,1%	
TVT	4.841	12,3%	838	10,1%	1.480	12,3%	9.803	21,1%	16.962	16,0%	
LE	3.233	8,2%	571	6,9%	842	7,0%	5.076	10,9%	9.722	9,2%	
nvVHF	25.140	64,0%	5.586	67,2%	8.706	72,2%	19.577	42,2%	59.009	55,6%	
Keine Indikation	7.605	19,4%	1.540	18,5%	1.794	14,9%	10.370	22,3%	21.309	20,1%	
Vormedikation mit VKA	8.218	20,9%	2.006	24,1%	3.112	25,8%	6.596	14,2%	19.932	18,8%	
<i>davon Warfarin</i>	83	1,0%	18	0,9%	34	1,1%	61	0,9%	196	1,0%	
<i>davon Phenprocoumon</i>	8.144	99,1%	1.991	99,3%	3.084	99,1%	6.543	99,2%	19.762	99,1%	
Vormedikation mit Heparinen	7.564	19,2%	1.671	20,1%	2.792	23,2%	9.369	20,2%	21.396	20,2%	
<i>davon Enoxaparin</i>	5.223	69,1%	1.120	67,0%	1.863	66,7%	6.425	68,6%	14.631	68,4%	
Keine Vormedikation	26.961	68,6%	5.496	66,2%	7.465	61,9%	33.382	71,9%	73.304	69,1%	
Keine Indikation & keine Vormedikation	5.755	14,6%	1.131	13,6%	1.224	10,2%	7.885	17,0%	15.995	15,1%	
Krankenhausaufenthalt	15.323	39,0%	3.035	36,5%	2.254	18,7%	14.567	31,4%	35.179	33,2%	
Level 2 (Ärzte, LE_ID)											
	n = 13.290		n = 4.817		n = 5.356		n = 14.538		n = 18.276		
Hausärztliche Versorgung	12.022	90,5%	4.378	90,9%	4.813	89,9%	12.327	84,8%	15.377	84,1%	
Kardiologie	537	4,0%	235	4,9%	335	6,3%	523	3,6%	710	3,9%	
Internistische Medizin	266	2,0%	96	2,0%	103	1,9%	331	2,3%	444	2,4%	
Angiologie	61	0,5%	12	0,2%	22	0,4%	86	0,6%	93	0,5%	
Orthopädie	54	0,4%	32	0,7%	4	0,1%	459	3,2%	514	2,8%	
Sonstige	350	2,6%	64	1,3%	79	1,5%	812	5,6%	1.138	6,2%	
Level 3 (Region, Kreise)											
	n = 158										

1.8 Regressionsergebnisse der 2. Analyse (NOAKs)

Patient*innen mit einer Neuverordnung für NOAKs			95% Conf.
N = 106.061		Odds Ratio (SE)	Int.
Alter		1,02*** (0,0006)	[1.018; 1.021]
Geschlecht (ref. männlich)		1 (0,0148)	[0.971; 1.029]
Pflegestufe (ref. keine)		1,261*** (0,0259)	[1.212; 1.313]
Nichtvalvuläres Vorhofflimmern		1,628*** (0,0285)	[1.574; 1.685]
Totalendoprothese		0,372*** (0,0132)	[0.347; 0.399]
Tiefe Venenthrombose		0,799*** (0,0179)	[0.764; 0.835]
Lungenembolie		0,865*** (0,0227)	[0.822; 0.911]
Krankenhausaufenthalt		1,288*** (0,0233)	[1.243; 1.335]
Vormedikation mit Warfarin		1,363 (0,2276)	[0.982; 1.891]
Vormedikation mit Phenprocoumon		1,185*** (0,0235)	[1.14; 1.232]
Vormedikation mit Enoxaparin		0,987 (0,0216)	[0.946; 1.03]
Level 2 (LE_ID)			
FG (ref. HA)	Orthopädie	0,137*** (0,017)	[0.107; 0.175]
		0,758*** (0,0542)	[0.659; 0.872]
	Internistische Medizin		
		0,965 (0,1205)	[0.755; 1.232]
	Angiologie		
		1,243*** (0,0615)	[1.128; 1.369]
	Kardiologie		
		0,461*** (0,0291)	[0.408; 0.522]
	Sonstige		
			[0.973; 1.081]
Praxistyp (ref. Einzelpraxis)	Gemeinschaftspraxis	1,025 (0,0276)	

sonstiges	0,964 (0,0979)	[0.79; 1.177]
Level 3 (Region)		
KH-Betten	1,018 (0,0143)	[0.99; 1.047]
Arztdichte	0,981 (0,0123)	[0.957; 1.005]
WSV	1,505*** (0,0432)	[1.423; 1.592]
Region Typ (ref. städtisch)	0,812* (0,0674)	[0.69; 0.956]
Intercept	0,152*** (0,0243)	[0.111; 0.208]
Var(Region)	0.223 (0.028)	[0.174; 0.285]
Var(LE_ID)	0.676 (0.02)	[0.638; 0.716]

Hinweise: *** p-value < 0,001; ** p-value < 0,01; * p-value < 0,05

Anhang 2: Teilprojekt B

Im Folgenden werden die deskriptiven Befunde und die Testergebnisse für die jeweils gepaarten und ungepaarten t-Tests (Mittelwertvergleich) gezeigt.

2.1 Monetäre Indikatoren:

Durchschnittliche Kosten je Verordnung (gemessen in Euro):

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	52,39	102,61	1,00	
2013	57,34	115,13	1,12	
2016	65,22	166,31	1,62	
Anstieg von 2013 zu 2016	7,88	98,27	0,96	0,000
Anstieg von 2010 zu 2013	4,95	42,95	0,42	0,000
Änderung der Anstiegs- raten	2,93	96,04	0,94	0,002

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	50,47	123,42	1,79	
2013	54,92	114,73	1,67	
2016	61,66	145,32	2,11	
Anstieg von 2013 zu 2016	6,74	68,97	1,00	0,000
Anstieg von 2010 zu 2013	4,46	82,39	1,20	0,000
Änderung der Anstiegs- raten	2,28	75,37	1,09	0,037

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	56,48	128,39	0,50	
2013	61,40	153,99	0,60	
2016	69,04	188,18	0,74	

Anstieg von 2013 zu 2016	7,64	96,61	0,38	0,000
Anstieg von 2010 zu 2013	4,92	96,04	0,38	0,000
Änderung der Anstiegsraten	2,72	131,78	0,52	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	1,14	0,468

Durchschnittliche Kosten je Patient*in (gemessen in Euro):

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	259,25	591,07	5,77	
2013	304,16	829,87	8,10	
2016	345,47	1000,66	9,76	
Anstieg von 2013 zu 2016	41,32	418,18	4,08	0,000
Anstieg von 2010 zu 2013	44,90	385,07	3,76	0,000
Änderung der Anstiegsraten	-3,59	517,43	5,05	0,477

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	268,69	639,30	9,28	
2013	288,10	715,67	10,39	
2016	327,87	867,80	12,60	
Anstieg von 2013 zu 2016	39,77	290,84	4,22	0,000
Anstieg von 2010 zu 2013	19,40	368,07	5,35	0,000
Änderung der Anstiegsraten	20,37	431,18	6,26	0,001

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	308,06	749,39	2,94	
2013	349,18	1010,63	3,96	
2016	392,75	1209,95	4,75	
Anstieg von 2013 zu 2016	43,57	527,00	2,07	0,000
Anstieg von 2010 zu 2013	41,12	559,16	2,19	0,000
Änderung der Anstiegs- raten	2,45	756,24	2,97	0,409

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	1,54	0,818

Durchschnittliche Kosten pro DDD bei bestimmten ATC-Gruppen (gemessen in Euro):

ATC-Gruppe 1: B01AC (Thrombozytenaggregations-Hemmer exkl. Heparin)

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,79	1,70	0,007	
2013	0,55	0,42	0,005	
2016	0,37	3,27	0,007	
Anstieg von 2013 zu 2016	-0,21	0,55	0,007	0,000
Anstieg von 2010 zu 2013	-0,22	0,44	0,006	0,000
Änderung der Anstiegs- raten	0,01	0,73	0,010	0,407

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,69	0,55	0,011	
2013	0,53	2,91	0,020	
2016	0,37	2,34	0,007	
Anstieg von 2013 zu 2016	-0,20	2,90	0,057	0,000
Anstieg von 2010 zu 2013	-0,20	0,97	0,019	0,000
Änderung der Anstiegs- raten	0,05	1,84	0,037	0,202

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,70	5,20	0,019	
2013	0,51	4,90	0,017	
2016	0,34	6,35	0,029	
Anstieg von 2013 zu 2016	-0,16	4,66	0,024	0,000
Anstieg von 2010 zu 2013	-0,19	2,86	0,015	0,000
Änderung der Anstiegs- raten	0,03	6,80	0,035	0,476

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,007	0,849

ATC-Gruppe 2: C09 (Mittel mit Wirkung auf das Renin-Angiotensin-Aldosteron-System)

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,31	0,17	0,002	
2013	0,26	0,16	0,002	

2016	0,22	0,15	0,002	
Anstieg von 2013 zu 2016	-0,05	0,16	0,002	0,000
Anstieg von 2010 zu 2013	-0,05	0,16	0,002	0,000
Änderung der Anstiegsraten	0,01	0,25	0,003	0,086

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,31	0,19	0,003	
2013	0,26	0,16	0,003	
2016	0,22	0,16	0,003	
Anstieg von 2013 zu 2016	-0,05	0,17	0,003	0,000
Anstieg von 2010 zu 2013	-0,05	0,16	0,003	0,000
Änderung der Anstiegsraten	0,003	0,25	0,005	0,567

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,33	0,22	0,001	
2013	0,27	0,30	0,001	
2016	0,22	0,27	0,001	
Anstieg von 2013 zu 2016	-0,05	0,20	0,001	0,000
Anstieg von 2010 zu 2013	-0,06	0,20	0,001	0,000
Änderung der Anstiegsraten	0,01	0,28	0,001	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,005	0,168

ATC-Gruppe 3: C10 (Mittel, die den Lipidstoffwechsel beeinflussen)

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,42	0,25	0,003	
2013	0,35	0,27	0,003	
2016	0,34	0,35	0,004	
Anstieg von 2013 zu 2016	-0,01	0,27	0,004	0,000
Anstieg von 2010 zu 2013	-0,08	0,18	0,003	0,000
Änderung der Anstiegs- raten	0,06	0,35	0,005	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,48	2,75	0,056	
2013	0,34	0,25	0,005	
2016	0,36	0,36	0,006	
Anstieg von 2013 zu 2016	-0,002	0,24	0,005	0,661
Anstieg von 2010 zu 2013	-0,14	2,77	0,056	0,014
Änderung der Anstiegs- raten	0,14	2,78	0,056	0,015

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,51	2,50	0,006	
2013	0,37	0,35	0,001	
2016	0,36	0,54	0,002	
Anstieg von 2013 zu 2016	-0,01	0,31	0,002	0,000

Anstieg von 2010 zu 2013	-0,08	1,25	0,006	0,000
Änderung der Anstiegsraten	0,06	0,95	0,005	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,013	0,043

ATC-Gruppe 4: R03 (Mittel bei obstruktiven Atemwegserkrankungen)

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	1,42	0,80	0,007	
2013	1,37	0,99	0,010	
2016	1,51	2,19	0,024	
Anstieg von 2013 zu 2016	0,12	2,17	0,025	0,000
Anstieg von 2010 zu 2013	-0,04	0,74	0,008	0,000
Änderung der Anstiegsraten	0,18	2,29	0,027	0,000

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	1,42	1,00	0,017	
2013	1,37	0,92	0,015	
2016	1,41	1,95	0,029	
Anstieg von 2013 zu 2016	0,02	1,80	0,031	0,520
Anstieg von 2010 zu 2013	-0,04	1,05	0,018	0,016

Änderung der Anstiegsraten	0,07	2,25	0,039	0,059
----------------------------	------	------	-------	-------

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	1,42	0,99	0,004	
2013	1,39	2,46	0,004	
2016	1,52	2,45	0,011	
Anstieg von 2013 zu 2016	0,14	2,35	0,011	0,000
Anstieg von 2010 zu 2013	-0,05	1,02	0,005	0,000
Änderung der Anstiegsraten	0,20	2,51	0,012	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,103	0,015

Verordnungsanteil⁹ von Generika am Gesamtmarkt (allgemein):

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,634	0,152	0,001	
2013	0,683	0,154	0,002	
2016	0,717	0,155	0,002	
Anstieg von 2013 zu 2016	0,034	0,078	0,001	0,000
Anstieg von 2010 zu 2013	0,049	0,073	0,001	0,000

⁹ Der Verordnungsanteil bezieht sich auf die Grundmenge aller in Apotheken eingelösten Verordnungen.

Änderung der Anstiegsraten	-0,016	0,111	0,001	0,000
----------------------------	--------	-------	-------	-------

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,641	0,159	0,002	
2013	0,693	0,159	0,002	
2016	0,703	0,165	0,002	
Anstieg von 2013 zu 2016	0,010	0,068	0,001	0,000
Anstieg von 2010 zu 2013	0,053	0,076	0,001	0,000
Änderung der Anstiegsraten	-0,042	0,105	0,002	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,631	0,160	0,001	
2013	0,684	0,162	0,001	
2016	0,698	0,168	0,001	
Anstieg von 2013 zu 2016	0,014	0,071	0,0003	0,000
Anstieg von 2010 zu 2013	0,053	0,228	0,001	0,000
Änderung der Anstiegsraten	-0,038	0,244	0,001	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,023	0,000

Verordnungsanteil¹⁰ von Generika im Bereich Antidiabetika:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,783	0,136	0,002	
2013	0,751	0,145	0,002	
2016	0,702	0,162	0,002	
Anstieg von 2013 zu 2016	-0,061	0,109	0,001	0,000
Anstieg von 2010 zu 2013	-0,034	0,117	0,002	0,000
Änderung der Anstiegs- raten	-0,027	0,172	0,002	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,833	0,134	0,003	
2013	0,791	0,138	0,003	
2016	0,711	0,175	0,003	
Anstieg von 2013 zu 2016	-0,095	0,107	0,002	0,000
Anstieg von 2010 zu 2013	-0,046	0,112	0,002	0,000
Änderung der Anstiegs- raten	-0,049	0,162	0,003	0,000

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,825	0,141	0,001	
2013	0,793	0,150	0,001	
2016	0,717	0,172	0,001	

¹⁰ Der Verordnungsanteil bezieht sich auf die Grundmenge aller in Apotheken eingelösten Verordnungen im Bereich Antidiabetika.

Anstieg von 2013 zu 2016	-0,080	0,107	0,001	0,000
Anstieg von 2010 zu 2013	-0,034	0,117	0,001	0,000
Änderung der Anstiegsraten	-0,046	0,162	0,001	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,033	0,000

Verordnungsanteil¹¹ von Me-Too-Präparaten:

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,005	0,008	0,0001	
2013	0,004	0,010	0,0001	
2016	0,003	0,006	0,0001	
Anstieg von 2013 zu 2016	-0,0010	0,009	0,0001	0,000
Anstieg von 2010 zu 2013	-0,0004	0,008	0,0001	0,000
Änderung der Anstiegsraten	-0,0007	0,0 16	0,0002	0,000

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,005	0,018	0,0003	
2013	0,005	0,014	0,0002	
2016	0,004	0,010	0,0002	
Anstieg von 2013 zu 2016	-0,0007	0,009	0,0002	0,000

¹¹ Der Verordnungsanteil bezieht sich auf die Grundmenge aller in Apotheken eingelösten Verordnungen.

Anstieg von 2010 zu 2013	0,00006	0,014	0,0003	0,807
Änderung der Anstiegsraten	-0,0007	0,020	0,0004	0,044

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,005	0,011	0,0005	
2013	0,005	0,011	0,0005	
2016	0,004	0,009	0,0004	
Anstieg von 2013 zu 2016	-0,0007	0,008	0,0004	0,000
Anstieg von 2010 zu 2013	-0,0006	0,009	0,0004	0,000
Änderung der Anstiegsraten	-0,0008	0,014	0,0006	0,208

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,0004	0,021

Anteil der Rabattarzneimittel am Gesamtmarkt

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,409	0,178	0,002	
2013	0,444	0,198	0,002	
2016	0,494	0,191	0,002	
Anstieg von 2013 zu 2016	0,050	0,099	0,001	0,000
Anstieg von 2010 zu 2013	0,036	0,094	0,001	0,000
Änderung der Anstiegsraten	0,014	0,145	0,001	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,420	0,189	0,003	
2013	0,456	0,198	0,003	
2016	0,502	0,194	0,003	
Anstieg von 2013 zu 2016	0,046	0,094	0,001	0,000
Anstieg von 2010 zu 2013	0,037	0,099	0,001	0,000
Änderung der Anstiegs- raten	0,009	0,138	0,002	0,000

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,387	0,182	0,0007	
2013	0,435	0,199	0,0008	
2016	0,484	0,197	0,0008	
Anstieg von 2013 zu 2016	0,049	0,097	0,0004	0,000
Anstieg von 2010 zu 2013	0,048	0,103	0,0004	0,000
Änderung der Anstiegs- raten	0,0007	0,145	0,0006	0,240

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,004	0,027

Aut-idem-Quote:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,070	0,123	0,001	
2013	0,073	0,107	0,001	

2016	0,068	0,098	0,001	
Anstieg von 2013 zu 2016	-0,005	0,071	0,001	0,000
Anstieg von 2010 zu 2013	0,003	0,097	0,001	0,000
Änderung der Anstiegsraten	-0,009	0,127	0,001	0,000

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,085	0,144	0,002	
2013	0,060	0,105	0,002	
2016	0,057	0,092	0,001	
Anstieg von 2013 zu 2016	-0,002	0,062	0,001	0,008
Anstieg von 2010 zu 2013	-0,025	0,116	0,002	0,000
Änderung der Anstiegsraten	0,023	0,138	0,002	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,126	0,186	0,001	
2013	0,108	0,151	0,001	
2016	0,093	0,131	0,001	
Anstieg von 2013 zu 2016	-0,015	0,092	0,0004	0,000
Anstieg von 2010 zu 2013	-0,018	0,139	0,001	0,000
Änderung der Anstiegsraten	0,003	0,172	0,001	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,003	0,011

Anteil von Biosimilars am Beispiel der TNF- α -Inhibitoren:

Nur Fertigarzneimittel:

Bayern	Mittelwert	Standard- abweichung
2016	0,050	0,174

Hessen	Mittelwert	Standard- abweichung
2016	0,061	0,193

Bund	Mittelwert	Standard- abweichung
2016	0,040	0,149

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
2016	-0,011	0,285

Fertigarzneimittel und individuelle Zubereitungen:

Bayern	Mittelwert	Standard- abweichung
2016	0,055	0,180

Hessen	Mittelwert	Standard- abweichung
2016	0,071	0,204

Bund	Mittelwert	Standard- abweichung
2016	0,044	0,154

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
2016	-0,016	0,170

2.2 Qualitative Indikatoren:

PPI-Mengen (gemessen in DDDs je AM-Patient):

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	19,70	18,90	0,204	
2013	26,29	24,73	0,266	
2016	31,87	29,79	0,317	
Anstieg von 2013 zu 2016	6,26	10,46	0,112	0,000
Anstieg von 2010 zu 2013	7,45	9,48	0,102	0,000
Änderung der Anstiegs- raten	-1,13	12,74	0,14	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	19,48	21,42	0,353	
2013	26,58	26,94	0,441	
2016	33,16	34,16	0,554	
Anstieg von 2013 zu 2016	7,52	11,92	0,193	0,000
Anstieg von 2010 zu 2013	8,02	9,93	0,163	0,000
Änderung der Anstiegs- raten	-0,352	12,88	0,21	0,099

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	20,68	21,63	0,092	
2013	27,97	28,02	0,119	
2016	33,40	33,52	0,140	
Anstieg von 2013 zu 2016	6,09	11,09	0,046	0,000
Anstieg von 2010 zu 2013	8,35	11,33	0,048	0,000
Änderung der Anstiegs- raten	-2,17	14,06	0,060	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-1,26	0,000

Ezetimib-Mengen (gemessen in DDDs je AM-Patient*in):

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	1,67	1,91	0,027	
2013	1,39	1,78	0,024	
2016	1,65	2,75	0,029	
Anstieg von 2013 zu 2016	0,32	1,34	0,020	0,000
Anstieg von 2010 zu 2013	-0,31	1,38	0,021	0,000
Änderung der Anstiegs- raten	0,639	1,82	0,027	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	1,53	1,92	0,046	
2013	1,27	2,10	0,051	

2016	1,37	2,11	0,050	
Anstieg von 2013 zu 2016	0,36	1,00	0,023	0,000
Anstieg von 2010 zu 2013	-0,10	0,88	0,021	0,000
Änderung der Anstiegsraten	0,463	1,39	0,032	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	1,80	3,00	0,012	
2013	1,50	2,40	0,012	
2016	1,76	2,82	0,013	
Anstieg von 2013 zu 2016	0,32	1,49	0,009	0,000
Anstieg von 2010 zu 2013	-0,32	1,69	0,010	0,000
Änderung der Anstiegsraten	0,637	2,32	0,014	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,04	0,161

Patient*innenanteil mit DOAK-Verordnung an allen Patient*innen mit Antikoagulanzenverordnung:

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,132	0,282	0,016	
2013	0,349	0,231	0,003	
2016	0,562	0,225	0,003	
Anstieg von 2013 zu 2016	0,190	0,147	0,002	0,000

Anstieg von 2010 zu 2013	0,252	0,169	0,009	0,000
Änderung der Anstiegsraten	-0,065	0,217	0,012	0,000

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,118	0,267	0,017	
2013	0,352	0,238	0,005	
2016	0,638	0,212	0,004	
Anstieg von 2013 zu 2016	0,262	0,144	0,003	0,000
Anstieg von 2010 zu 2013	0,257	0,156	0,010	0,000
Änderung der Anstiegsraten	-0,006	0,209	0,013	0,659

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,162	0,324	0,006	
2013	0,371	0,270	0,001	
2016	0,603	0,223	0,001	
Anstieg von 2013 zu 2016	0,235	0,143	0,001	0,000
Anstieg von 2010 zu 2013	0,249	0,175	0,003	0,000
Änderung der Anstiegsraten	-0,027	0,216	0,004	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,066	0,000

Verordnungstiefenvergleich vor und nach Einführung der Wirkstoffvereinbarung (Fachärzt*innen):

Anteil der Verordnungen:

Bayern	Mittelwert	Standard-abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,205	0,112	0,0016	
2013	0,205	0,109	0,0016	
2016	0,198	0,107	0,0016	
Anstieg von 2013 zu 2016	-0,007	0,035	0,0005	0,000
Anstieg von 2010 zu 2013	-0,0002	0,037	0,0005	0,700
Änderung der Anstiegsraten	-0,0068	0,053	0,0008	0,000

Hessen	Mittelwert	Standard-abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,205	0,118	0,0024	
2013	0,205	0,118	0,0024	
2016	0,198	0,115	0,0024	
Anstieg von 2013 zu 2016	-0,0086	0,038	0,0008	0,000
Anstieg von 2010 zu 2013	0,0005	0,038	0,0008	0,511
Änderung der Anstiegsraten	-0,0091	0,057	0,0012	0,000

Bund	Mittelwert	Standard-abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,201	0,113	0,0006	
2013	0,198	0,108	0,0006	
2016	0,193	0,107	0,0006	
Anstieg von 2013 zu 2016	-0,005	0,036	0,0002	0,000

Anstieg von 2010 zu 2013	-0,003	0,039	0,0002	0,000
Änderung der Anstiegsraten	-0,0014	0,057	0,0003	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,0016	0,070

b.) Anteil der insgesamt verordneten DDDs:

Bayern	Mittelwert	Standard-abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,183	0,144	0,0021	
2013	0,180	0,142	0,0021	
2016	0,172	0,136	0,0020	
Anstieg von 2013 zu 2016	-0,008	0,038	0,0005	0,000
Anstieg von 2010 zu 2013	-0,003	0,043	0,0006	0,000
Änderung der Anstiegsraten	-0,005	0,060	0,0009	0,000

Hessen	Mittelwert	Standard-abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,164	0,144	0,0030	
2013	0,162	0,143	0,0029	
2016	0,152	0,135	0,0028	
Anstieg von 2013 zu 2016	-0,009	0,041	0,0008	0,000
Anstieg von 2010 zu 2013	-0,002	0,041	0,0008	0,008

Änderung der Anstiegsraten	-0,007	0,060	0,0012	0,000
----------------------------	--------	-------	--------	-------

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,176	0,139	0,0008	
2013	0,170	0,135	0,0008	
2016	0,163	0,131	0,0007	
Anstieg von 2013 zu 2016	-0,006	0,039	0,0002	0,000
Anstieg von 2010 zu 2013	-0,007	0,043	0,0002	0,000
Änderung der Anstiegsraten	0,0002	0,061	0,0003	0,484

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,0012	0,205

Verordnungstiefenvergleich vor und nach Einführung der Wirkstoffvereinbarung (Hausärzt*innen):

Anteil der Verordnungen:

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,589	0,116	0,0016	
2013	0,583	0,111	0,0015	
2016	0,570	0,120	0,0016	
Anstieg von 2013 zu 2016	-0,013	0,071	0,0010	0,000
Anstieg von 2010 zu 2013	-0,006	0,058	0,0008	0,000

Änderung der Anstiegsraten	-0,007	0,098	0,0013	0,000
----------------------------	--------	-------	--------	-------

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,596	0,106	0,0022	
2013	0,604	0,096	0,0021	
2016	0,585	0,117	0,0025	
Anstieg von 2013 zu 2016	-0,019	0,081	0,0013	0,000
Anstieg von 2010 zu 2013	0,008	0,063	0,0017	0,000
Änderung der Anstiegsraten	-0,027	0,106	0,0022	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,579	0,112	0,0006	
2013	0,575	0,108	0,0006	
2016	0,567	0,118	0,0006	
Anstieg von 2013 zu 2016	-0,008	0,073	0,0004	0,000
Anstieg von 2010 zu 2013	-0,003	0,061	0,0003	0,000
Änderung der Anstiegsraten	-0,005	0,100	0,0005	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,006	0,001

Anteil der insgesamt verordneten DDDs:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,642	0,133	0,0018	
2013	0,632	0,125	0,0017	
2016	0,617	0,135	0,0019	
Anstieg von 2013 zu 2016	-0,014	0,081	0,0011	0,000
Anstieg von 2010 zu 2013	-0,010	0,068	0,0009	0,000
Änderung der Anstiegs- raten	-0,004	0,112	0,002	0,007

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,596	0,123	0,0026	
2013	0,604	0,113	0,0024	
2016	0,585	0,133	0,0028	
Anstieg von 2013 zu 2016	-0,019	0,091	0,0019	0,000
Anstieg von 2010 zu 2013	0,008	0,073	0,0016	0,750
Änderung der Anstiegs- raten	-0,021	0,121	0,0026	0,000

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,632	0,128	0,0007	
2013	0,624	0,123	0,0007	
2016	0,615	0,133	0,0007	
Anstieg von 2013 zu 2016	-0,009	0,082	0,0005	0,000

Anstieg von 2010 zu 2013	-0,008	0,071	0,0004	0,000
Änderung der Anstiegsraten	-0,0003	0,115	0,0006	0,673

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,008	0,000

Verordnungsrelevanter Arztkontaktwechsel von Hausärzt*innen zu Fachärzt*innen (am Beispiel von Entresto):

Anteil der Patient*in mit einem Hausärzt*in -Fachärzt*in -Verordnungswechsel innerhalb der ersten drei Verordnungen

Bayern	Mittelwert	Standardabweichung
2016	0,782	0,305

Hessen	Mittelwert	Standardabweichung
2016	0,858	0,278

Bund	Mittelwert	Standardabweichung
2016	0,806	0,290

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
2016	-0,077	0,174

Anteil der Patient*innen mit zweimaligem Hausarzt*in-Facharzt*in-Verordnungswechsel innerhalb der ersten drei Verordnungen:

Bayern	Mittelwert	Standard-abweichung	Standardfehler des Mittelwertes
2016	0,774	0,295	0,051

Hessen	Mittelwert	Standard-abweichung	Standardfehler des Mittelwertes
2016	0,775	0,362	0,128

Bund	Mittelwert	Standard-abweichung	Standardfehler des Mittelwertes
2016	0,740	0,325	0,026

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
2016	-0,001	0,992

Verordnungsrelevanter Arztkontaktwechsel von Facharzt*innen zu Hausarzt*innen (am Beispiel von Entresto):

Anteil der Patient*innen mit einem Facharzt*innen-Hausarzt*innen-Verordnungswechsel innerhalb der ersten drei Verordnungen:

Bayern	Mittelwert	Standard-abweichung	Standardfehler des Mittelwertes
2016	0,678	0,296	0,025

Hessen	Mittelwert	Standard-abweichung	Standardfehler des Mittelwertes
2016	0,660	0,308	0,044

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes
2016	0,609	0,316	0,012

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
2016	0,018	0,718

Anteil der Patient*innen mit zweimaligem Fachärzt*innen-Hausärzt*innen-Verordnungswechsel innerhalb der ersten drei Verordnungen:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes
2016	0,410	0,337	0,053

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes
2016	0,311	0,235	0,061

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes
2016	0,284	0,267	0,017

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
2016	0,098	0,231

Verordnungsanteil von Leitsubstanzen:

Verordnungsanteil von bayerischen Leitsubstanzen:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,733	0,405	0,004	

2013	0,467	0,315	0,003	
2016	0,407	0,268	0,003	
Anstieg von 2013 zu 2016	-0,060	0,246	0,003	0,000
Anstieg von 2010 zu 2013	-0,267	0,340	0,004	0,000
Änderung der Anstiegsraten	0,206	0,463	0,005	0,000

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,727	0,417	0,007	
2013	0,451	0,321	0,005	
2016	0,373	0,269	0,004	
Anstieg von 2013 zu 2016	-0,077	0,244	0,004	0,000
Anstieg von 2010 zu 2013	-0,276	0,347	0,006	0,000
Änderung der Anstiegsraten	0,199	0,461	0,008	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,719	0,418	0,002	
2013	0,459	0,321	0,001	
2016	0,375	0,270	0,001	
Anstieg von 2013 zu 2016	-0,084	0,242	0,001	0,000
Anstieg von 2010 zu 2013	-0,259	0,345	0,001	0,000
Änderung der Anstiegsraten	0,175	0,459	0,002	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,017	0,000

Verordnungsanteil von hessischen Leitsubstanzen:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,740	0,304	0,003	
2013	0,747	0,305	0,003	
2016	0,727	0,306	0,003	
Anstieg von 2013 zu 2016	-0,020	0,295	0,003	0,000
Anstieg von 2010 zu 2013	0,007	0,301	0,003	0,028
Änderung der Anstiegs- raten	-0,027	0,501	0,005	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,762	0,298	0,005	
2013	0,766	0,302	0,005	
2016	0,746	0,299	0,005	
Anstieg von 2013 zu 2016	-0,020	0,310	0,005	0,000
Anstieg von 2010 zu 2013	0,004	0,310	0,005	0,356
Änderung der Anstiegs- raten	-0,024	0,520	0,008	0,002

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,741	0,304	0,001	
2013	0,755	0,300	0,001	

2016	0,740	0,300	0,001	
Anstieg von 2013 zu 2016	-0,016	0,305	0,001	0,000
Anstieg von 2010 zu 2013	0,014	0,309	0,001	0,000
Änderung der Anstiegsraten	-0,029	0,517	0,002	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,004	0,946

Anteil neuer (AMNOG-bewerteter) Arzneien:

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2013	0,002	0,005	0,00004	
2016	0,013	0,019	0,00018	
Anstieg von 2013 zu 2016	0,011	0,017	0,00017	0,000

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2013	0,002	0,006	0,00008	
2016	0,013	0,021	0,00031	
Anstieg von 2013 zu 2016	0,011	0,018	0,00027	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2013	0,002	0,006	0,00002	
2016	0,014	0,022	0,00008	
Anstieg von 2013 zu 2016	0,011	0,019	0,00007	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,00002	0,952

Anteil Medikamentenverordnung mit einem Zusatznutzen in mindestens einer Subgruppe an allen AMNOG-bewerteten Präparaten:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2013	0,239	0,278	0,0034	
2016	0,237	0,223	0,0025	
Anstieg von 2013 zu 2016	0,013	0,222	0,0025	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2013	0,323	0,302	0,0058	
2016	0,280	0,248	0,0042	
Anstieg von 2013 zu 2016	-0,010	0,224	0,0038	0,012

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2013	0,285	0,302	0,0015	
2016	0,249	0,233	0,0010	
Anstieg von 2013 zu 2016	-0,008	0,231	0,0010	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,022	0,000

Übereinstimmung mit dem Medikationskatalog der KBV:

Anteil der Verordnungen der Prioritäten 1, 2 oder 3:

Bayern	Mittelwert	Standard- abweichung
2016	0,320	0,235

Hessen	Mittelwert	Standard- abweichung
2016	0,313	0,243

Bund	Mittelwert	Standard- abweichung
2016	0,326	0,249

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
2016	0,0073	0,085

Anteil der Verordnungen der Prioritäten 1 oder 2:

Bayern	Mittelwert	Standard- abweichung
2016	0,264	0,201

Hessen	Mittelwert	Standard- abweichung
2016	0,258	0,210

Bund	Mittelwert	Standard- abweichung
2016	0,268	0,214

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
2016	0,0060	0,102

Anteil Patient*innen mit Polypharmazie:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,102	0,062	0,0006	
2013	0,111	0,069	0,0007	
2016	0,119	0,073	0,0007	
Anstieg von 2013 zu 2016	0,008	0,028	0,0003	0,000
Anstieg von 2010 zu 2013	0,010	0,027	0,0003	0,000
Änderung der Anstiegs- raten	-0,002	0,042	0,0004	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,105	0,073	0,0011	
2013	0,112	0,077	0,0011	
2016	0,120	0,083	0,0012	
Anstieg von 2013 zu 2016	0,008	0,031	0,0005	0,000
Anstieg von 2010 zu 2013	0,007	0,028	0,0004	0,000
Änderung der Anstiegs- raten	0,001	0,047	0,0007	0,178

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,120	0,075	0,0003	
2013	0,133	0,083	0,0003	
2016	0,137	0,085	0,0003	
Anstieg von 2013 zu 2016	0,004	0,031	0,0001	0,000

Anstieg von 2010 zu 2013	0,013	0,031	0,0001	0,000
Änderung der Anstiegsraten	-0,009	0,048	0,0002	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,0005	0,302

Nach Indikationsgebiet differenzierte qualitative Indikatoren:

Asthma und COPD

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,524	0,132	0,0013	
2013	0,516	0,123	0,0012	
2016	0,506	0,121	0,0012	
Anstieg von 2013 zu 2016	-0,010	0,117	0,0012	0,000
Anstieg von 2010 zu 2013	-0,008	0,127	0,0011	0,000
Änderung der Anstiegsraten	-0,002	0,202	0,0020	0,245

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,491	0,132	0,0019	
2013	0,500	0,123	0,0018	
2016	0,492	0,128	0,0019	
Anstieg von 2013 zu 2016	-0,007	0,137	0,0020	0,001
Anstieg von 2010 zu 2013	0,007	0,137	0,0020	0,000
Änderung der Anstiegsraten	-0,014	0,233	0,0034	0,000

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,514	0,126	0,0005	
2013	0,520	0,119	0,0005	
2016	0,505	0,121	0,0005	
Anstieg von 2013 zu 2016	-0,015	0,114	0,0004	0,000
Anstieg von 2010 zu 2013	0,006	0,120	0,0005	0,000
Änderung der Anstiegs- raten	-0,020	0,194	0,0007	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,003	0,140

Bluthochdruck

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,807	0,167	0,0016	
2013	0,823	0,149	0,0015	
2016	0,819	0,149	0,0015	
Anstieg von 2013 zu 2016	-0,004	0,134	0,0013	0,001
Anstieg von 2010 zu 2013	0,016	0,139	0,0014	0,000
Änderung der Anstiegs- raten	-0,021	0,230	0,0023	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,825	0,172	0,0025	

2013	0,842	0,151	0,0022	
2016	0,832	0,157	0,0023	
Anstieg von 2013 zu 2016	-0,010	0,146	0,0021	0,000
Anstieg von 2010 zu 2013	0,017	0,150	0,0022	0,000
Änderung der Anstiegsraten	-0,027	0,284	0,0036	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,826	0,164	0,0006	
2013	0,840	0,148	0,0006	
2016	0,831	0,152	0,0006	
Anstieg von 2013 zu 2016	-0,009	0,135	0,0005	0,000
Anstieg von 2010 zu 2013	0,014	0,140	0,0005	0,000
Änderung der Anstiegsraten	-0,024	0,231	0,0009	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,005	0,033

Koronare Herzkrankheit

Anteil der KHK-Patient*innen, die Thrombozytenaggregationshemmer erhalten

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,298	0,156	0,0015	
2013	0,297	0,150	0,0015	
2016	0,315	0,151	0,0015	

Anstieg von 2013 zu 2016	0,019	0,158	0,0016	0,000
Anstieg von 2010 zu 2013	-0,001	0,158	0,0016	0,426
Änderung der Anstiegsraten	0,020	0,266	0,003	0,000

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,307	0,164	0,0025	
2013	0,323	0,161	0,0024	
2016	0,320	0,161	0,0024	
Anstieg von 2013 zu 2016	-0,003	0,175	0,0026	0,265
Anstieg von 2010 zu 2013	0,016	0,178	0,0027	0,000
Änderung der Anstiegsraten	-0,019	0,298	0,004	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,327	0,156	0,0006	
2013	0,343	0,154	0,0006	
2016	0,345	0,153	0,0006	
Anstieg von 2013 zu 2016	0,002	0,151	0,0006	0,000
Anstieg von 2010 zu 2013	0,016	0,155	0,0006	0,000
Änderung der Anstiegsraten	-0,014	0,259	0,0010	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,022	0,000

Anteil der betreuten Patient*innen nach Myokardinfarkt, die eine Therapie mit Betablockern erhalten

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,596	0,402	0,0042	
2013	0,684	0,343	0,0036	
2016	0,689	0,318	0,0033	
Anstieg von 2013 zu 2016	0,005	0,425	0,0044	0,286
Anstieg von 2010 zu 2013	0,089	0,473	0,0049	0,000
Änderung der Anstiegs- raten	-0,084	0,758	0,0079	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,649	0,361	0,006	
2013	0,694	0,333	0,005	
2016	0,683	0,322	0,005	
Anstieg von 2013 zu 2016	-0,011	0,407	0,006	0,096
Anstieg von 2010 zu 2013	0,045	0,427	0,007	0,000
Änderung der Anstiegs- raten	-0,055	0,708	0,011	0,000

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,667	0,346	0,0014	
2013	0,715	0,311	0,0013	
2016	0,706	0,301	0,0012	
Anstieg von 2013 zu 2016	-0,008	0,381	0,0016	0,000

Anstieg von 2010 zu 2013	0,047	0,409	0,0017	0,000
Änderung der Anstiegsraten	-0,055	0,669	0,0027	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,015	0,048

Anteil der KHK-Patient*innen, die Medikamente mit Zusatznutzen erhalten (Patient*innen mit instabiler Angina pectoris oder Myokardinfarkt ohne ST-Strecken-Hebung (NSTEMI) mit Ticagrelor)

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2013	0,028	0,126	0,0014	
2016	0,076	0,186	0,0021	
Anstieg von 2013 zu 2016	0,048	0,221	0,0025	0,000

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2013	0,038	0,148	0,0026	
2016	0,113	0,241	0,0042	
Anstieg von 2013 zu 2016	0,076	0,279	0,0049	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2013	0,038	0,140	0,0006	
2016	0,084	0,192	0,0008	
Anstieg von 2013 zu 2016	0,046	0,235	0,0010	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,028	0,000

Herzinsuffizienz (nur Linksherzinsuffizienz mit reduzierter Pumpfunktion):

Anteil der Patient*innen mit Herzinsuffizienz, die mit einem ACE-Hemmer oder AT1-Blocker behandelt werden:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,602	0,362	0,004	
2013	0,659	0,293	0,003	
2016	0,671	0,259	0,003	
Anstieg von 2013 zu 2016	0,012	0,341	0,003	0,000
Anstieg von 2010 zu 2013	0,057	0,409	0,004	0,000
Änderung der Anstiegs- raten	-0,045	0,631	0,006	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,657	0,341	0,005	
2013	0,677	0,311	0,005	
2016	0,668	0,302	0,005	
Anstieg von 2013 zu 2016	-0,009	0,383	0,006	0,133
Anstieg von 2010 zu 2013	0,020	0,405	0,006	0,001
Änderung der Anstiegs- raten	-0,029	0,673	0,010	0,005

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
-------------	------------	-------------------------	------------------------------------	-------------

2010	0,658	0,332	0,001	
2013	0,698	0,274	0,001	
2016	0,697	0,247	0,001	
Anstieg von 2013 zu 2016	-0,001	0,320	0,001	0,475
Anstieg von 2010 zu 2013	0,039	0,373	0,001	0,000
Änderung der Anstiegsraten	-0,040	0,582	0,002	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,021	0,002

Anteil der Patient*innen mit Herzinsuffizienz, die mit einem Beta-Rezeptorenblocker behandelt werden:

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,499	0,350	0,004	
2013	0,570	0,292	0,003	
2016	0,592	0,262	0,003	
Anstieg von 2013 zu 2016	0,022	0,337	0,003	0,000
Anstieg von 2010 zu 2013	0,071	0,400	0,004	0,000
Änderung der Anstiegsraten	-0,049	0,622	0,006	0,000

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,561	0,335	0,005	
2013	0,606	0,309	0,005	
2016	0,613	0,300	0,005	

Anstieg von 2013 zu 2016	0,007	0,382	0,006	0,247
Anstieg von 2010 zu 2013	0,044	0,401	0,006	0,000
Änderung der Anstiegsraten	-0,038	0,667	0,010	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,553	0,325	0,001	
2013	0,613	0,276	0,001	
2016	0,625	0,250	0,001	
Anstieg von 2013 zu 2016	0,012	0,318	0,001	0,000
Anstieg von 2010 zu 2013	0,060	0,370	0,001	0,000
Änderung der Anstiegsraten	-0,048	0,582	0,002	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,015	0,026

Anteil der Patient*innen mit Herzinsuffizienz, die bei Vorhofflimmern mit DOAKS oder Vitamin-K-Antagonisten behandelt werden:

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,432	0,403	0,004	
2013	0,570	0,369	0,004	
2016	0,683	0,333	0,004	
Anstieg von 2013 zu 2016	0,113	0,444	0,005	0,000
Anstieg von 2010 zu 2013	0,138	0,476	0,005	0,000

Änderung der Anstiegsraten	-0,025	0,777	0,008	0,003
----------------------------	--------	-------	-------	-------

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,242	0,427	0,035	
2013	0,293	0,452	0,037	
2016	0,408	0,487	0,040	
Anstieg von 2013 zu 2016	0,115	0,659	0,055	0,037
Anstieg von 2010 zu 2013	0,051	0,606	0,050	0,307
Änderung der Anstiegsraten	0,063	1,054	0,087	0,469

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,405	0,410	0,004	
2013	0,541	0,391	0,004	
2016	0,641	0,371	0,004	
Anstieg von 2013 zu 2016	0,100	0,485	0,005	0,000
Anstieg von 2010 zu 2013	0,136	0,498	0,005	0,000
Änderung der Anstiegsraten	-0,036	0,834	0,008	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,0013	0,981

Anteil der Patient*innen mit Herzinsuffizienz, die Medikamente mit Zusatznutzen erhalten (nur bei symptomatischer chron. Herzinsuffizienz):

Bayern	Mittelwert	Standard- abweichung
2016	0,0058	0,0229

Hessen	Mittelwert	Standard- abweichung
2016	0,0043	0,0148

Bund	Mittelwert	Standard- abweichung
2016	0,0051	0,0160

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
2016	0,0015	0,000

Osteoporose:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,416	0,180	0,002	
2013	0,384	1,169	0,002	
2016	0,360	0,164	0,002	
Anstieg von 2013 zu 2016	-0,024	0,168	0,002	0,000
Anstieg von 2010 zu 2013	-0,031	0,176	0,002	0,000
Änderung der Anstiegs- raten	0,008	0,292	0,003	0,008

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,392	0,190	0,003	
2013	0,383	0,185	0,003	
2016	0,370	0,184	0,003	
Anstieg von 2013 zu 2016	-0,012	0,194	0,003	0,000
Anstieg von 2010 zu 2013	-0,009	0,187	0,003	0,001
Änderung der Anstiegs- raten	-0,003	0,323	0,005	0,512

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,404	0,178	0,001	
2013	0,391	0,172	0,001	
2016	0,371	0,169	0,001	
Anstieg von 2013 zu 2016	-0,020	0,163	0,001	0,000
Anstieg von 2010 zu 2013	-0,013	0,167	0,001	0,000
Änderung der Anstiegs- raten	-0,007	0,278	0,001	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,011	0,001

2.3 Bayernspezifische Indikatoren für Wirtschaftlichkeit:

Zusätzlich zu den allgemeinen Indikatoren wurden insgesamt sechs bayernspezifische Indikatoren analysiert – sowohl für die KV Bayerns, die KV Hessen als auch den Bund, um einen adäquaten Vergleich zu haben und evtl. Effekte in der KV Bayerns besser einordnen zu können. Bei den sechs Indikatoren handelt es sich um sechs Generika- und Leitsubstanzziele aus der Bayerischen Wirkstoffvereinbarung, die in Bayern 2010 und 2013 noch nicht, dafür aber 2016 gesteuert wurden. In der KV Hessen und im Bund werden diese Bereiche nicht gesteuert.

Bayernspezifische Indikatoren – Generikaziele

Anteil der Generikaverordnungen im Bereich Analgetika (außer BTM-pflichtige Opiode):

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,752	0,244	0,002	
2013	0,764	0,245	0,002	
2016	0,889	0,234	0,002	
Anstieg von 2013 zu 2016	0,125	0,273	0,003	0,000
Anstieg von 2010 zu 2013	0,012	0,278	0,003	0,000
Änderung der Anstiegs- raten	0,114	0,466	0,005	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,787	0,243	0,004	
2013	0,787	0,244	0,004	
2016	0,876	0,248	0,004	
Anstieg von 2013 zu 2016	0,089	0,280	0,004	0,000
Anstieg von 2010 zu 2013	0,0004	0,279	0,004	0,915
Änderung der Anstiegs- raten	0,089	0,472	0,007	0,000

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,766	0,238	0,001	
2013	0,784	0,232	0,001	
2016	0,891	0,225	0,001	
Anstieg von 2013 zu 2016	0,107	0,260	0,001	0,000

Anstieg von 2010 zu 2013	0,018	0,266	0,001	0,000
Änderung der Anstiegsraten	0,089	0,444	0,002	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,036	0,000

Anteil der Generikaverordnungen im Bereich der Mittel bei obstruktiven Atemwegserkrankungen:

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,359	0,258	0,003	
2013	0,370	0,265	0,003	
2016	0,387	0,270	0,003	
Anstieg von 2013 zu 2016	0,016	0,302	0,003	0,000
Anstieg von 2010 zu 2013	0,011	0,306	0,003	0,000
Änderung der Anstiegsraten	0,005	0,515	0,005	0,364

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,401	0,272	0,004	
2013	0,420	0,279	0,004	
2016	0,418	0,281	0,004	
Anstieg von 2013 zu 2016	-0,002	0,312	0,005	0,688
Anstieg von 2010 zu 2013	0,020	0,316	0,005	0,000

Änderung der Anstiegsraten	-0,021	0,533	0,008	0,009
----------------------------	--------	-------	-------	-------

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,382	0,271	0,001	
2013	0,400	0,279	0,001	
2016	0,401	0,281	0,001	
Anstieg von 2013 zu 2016	0,001	0,322	0,001	0,395
Anstieg von 2010 zu 2013	0,019	0,323	0,001	0,000
Änderung der Anstiegsraten	-0,017	0,549	0,002	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,018	0,002

Anteil der Generikaverordnungen im Bereich der Mittel mit Wirkung auf das Renin-Angiotensin-System:

Bayern	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,612	0,321	0,004	
2013	0,750	0,326	0,004	
2016	0,833	0,319	0,003	
Anstieg von 2013 zu 2016	0,083	0,378	0,004	0,000
Anstieg von 2010 zu 2013	0,139	0,374	0,004	0,000
Änderung der Anstiegsraten	-0,056	0,641	0,007	0,000

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,619	0,330	0,005	
2013	0,759	0,328	0,005	
2016	0,844	0,312	0,005	
Anstieg von 2013 zu 2016	0,085	0,373	0,006	0,000
Anstieg von 2010 zu 2013	0,141	0,378	0,006	0,000
Änderung der Anstiegs- raten	-0,056	0,636	0,010	0,000

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,590	0,331	0,001	
2013	0,744	0,334	0,001	
2016	0,828	0,326	0,001	
Anstieg von 2013 zu 2016	0,085	0,391	0,002	0,000
Anstieg von 2010 zu 2013	0,154	0,383	0,002	0,000
Änderung der Anstiegs- raten	-0,069	0,658	0,003	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	-0,002	0,748

Anteil der Generikaverordnungen im Bereich Thrombozytenaggregationshemmer (ohne Ticagrelor):

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,648	0,316	0,004	

2013	0,743	0,313	0,004	
2016	0,829	0,318	0,004	
Anstieg von 2013 zu 2016	0,086	0,365	0,004	0,000
Anstieg von 2010 zu 2013	0,094	0,375	0,004	0,000
Änderung der Anstiegsraten	-0,008	0,624	0,007	0,262

Hessen	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,677	0,322	0,006	
2013	0,783	0,321	0,006	
2016	0,820	0,332	0,006	
Anstieg von 2013 zu 2016	0,038	0,381	0,007	0,000
Anstieg von 2010 zu 2013	0,106	0,373	0,007	0,000
Änderung der Anstiegsraten	-0,068	0,631	0,011	0,000

Bund	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,671	0,324	0,001	
2013	0,764	0,320	0,001	
2016	0,832	0,322	0,001	
Anstieg von 2013 zu 2016	0,068	0,368	0,002	0,000
Anstieg von 2010 zu 2013	0,093	0,368	0,002	0,000
Änderung der Anstiegsraten	-0,025	0,617	0,003	0,000

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,048	0,000

Bayernspezifische Indikatoren – Leitsubstanzziele

Anteil der verordneten Leitsubstanz-DDDs im Bereich MS-Therapeutika:

Bayern	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,304	0,404	0,009	
2013	0,270	0,392	0,009	
2016	0,227	0,364	0,008	
Anstieg von 2013 zu 2016	-0,043	0,466	0,011	0,000
Anstieg von 2010 zu 2013	-0,034	0,470	0,011	0,002
Änderung der Anstiegs- raten	-0,009	0,784	0,018	0,618

Hessen	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,422	0,441	0,018	
2013	0,394	0,428	0,018	
2016	0,299	0,394	0,016	
Anstieg von 2013 zu 2016	-0,095	0,464	0,019	0,000
Anstieg von 2010 zu 2013	-0,028	0,498	0,021	0,176
Änderung der Anstiegs- raten	-0,067	0,803	0,033	0,045

Bund	Mittelwert	Standard- abweichung	Standardfehler des Mittelwertes	Signifikanz
2010	0,334	0,409	0,004	

2013	0,296	0,393	0,004	
2016	0,236	0,358	0,004	
Anstieg von 2013 zu 2016	-0,060	0,444	0,005	0,000
Anstieg von 2010 zu 2013	-0,038	0,474	0,005	0,000
Änderung der Anstiegsraten	-0,023	0,773	0,008	0,005

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
Anstieg von 2013 zu 2016	0,052	0,018

Anteil der verordneten Leitsubstanz-DDDs im Bereich TNF- α -Inhibitoren:

Bayern	Mittelwert	Standardabweichung
2016	0,051	0,173

Hessen	Mittelwert	Standardabweichung
2016	0,063	0,196

Bund	Mittelwert	Standardabweichung
2016	0,041	0,152

Vergleich Bayern Hessen	Mittlere Differenz	Signifikanz
2016	-0,012	0,277

2.4 Ergebnistabellen und -bilder des Strukturbruchmodells nach Chow

Durchschnittliche Kosten je Verordnung

Für die durchschnittlichen Kosten je Verordnung wurde das Zeitreihen-Regressionsmodell unter Einbeziehung der fünf beschriebenen Adjustierungs- und Differenzierungsvariablen

$$Kosten_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Basierend auf dieser Schätzung wurde zunächst auf ein Vorliegen von Heteroskedastizität mithilfe des Breusch-Pagan-Tests getestet und anschließend der Chow Test auf Strukturbruch durchgeführt.

Es folgen die einzelnen Schätz- und Testergebnisse.

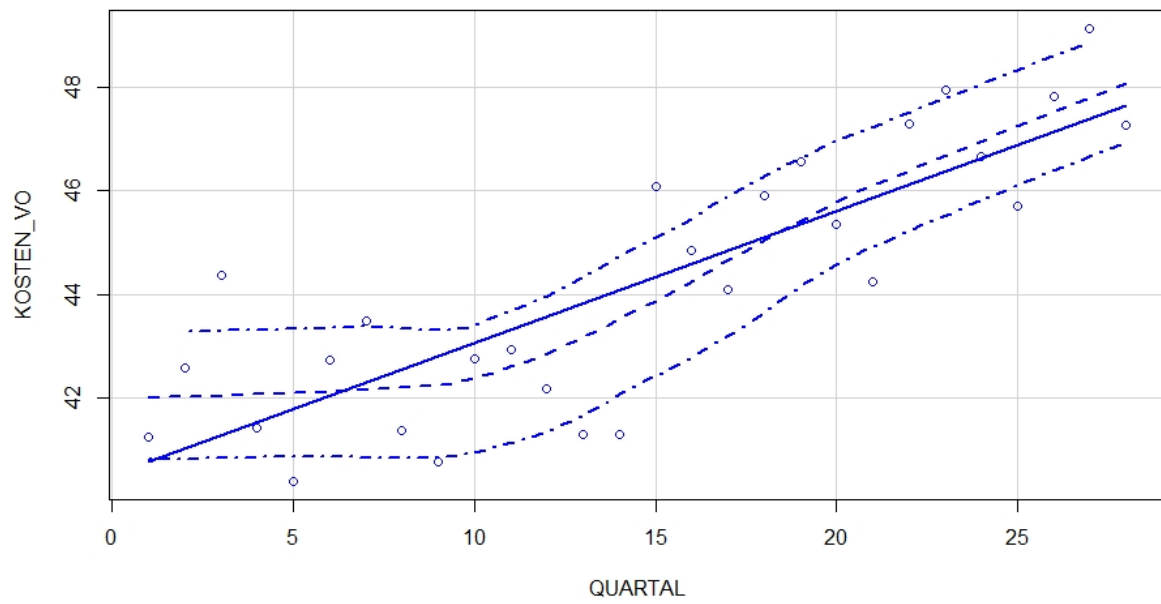
	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	20,204929	0,4392**
Anteil der Patient*innen im Alter ab 70	-16,450000	-0,01286*
Anteil der weiblichen Patient*innen	89,114874***	0,5582**
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	48,604704*	0,0177**
Anteil der Polypharmazie-Patient*innen	17,477730	0,01116
Anzahl der Fälle je Quartal	-0,023142**	-0,02397***

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,838	0,9475
F-Statistik	28,93 bei 5 und 22 Freiheitsgraden	98,47 bei 5 und 22 Freiheitsgraden
p-Wert	5,633 * 10 ⁻⁹	2,661 * 10 ⁻¹⁴

Bayern:

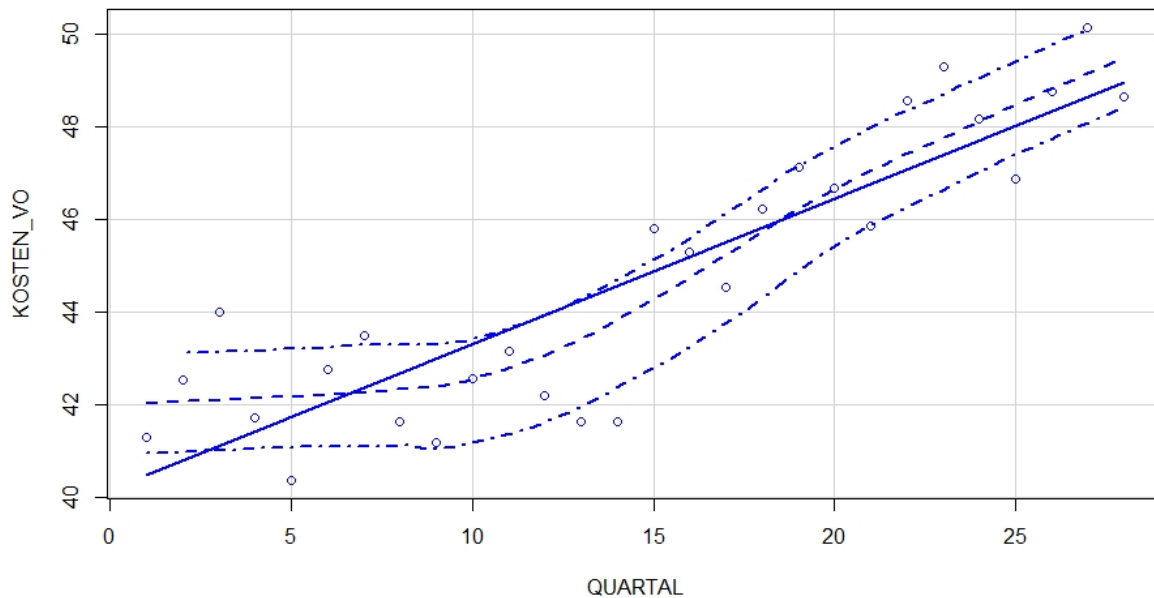
Für die KV Bayerns hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	7,4142 bei 5 Freiheitsgraden	0,1916
Chow-Test	1,1143	0,397

Hessen:

Für die KV Hessen hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	6,271 bei 5 Freiheitsgraden	0,2807
Chow-Test	0,22154	0,964

Durchschnittliche Kosten je Patient*in

Für die durchschnittlichen Kosten je Patient*in wurde analog das Zeitreihen-Regressionsmodell

$$KostenPat_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Basierend auf dieser Schätzung wurde auch hier auf ein Vorliegen von Heteroskedastizität getestet und anschließend der Chow Test auf Strukturbruch durchgeführt.

Es folgen die einzelnen Schätz- und Testergebnisse der beiden KVEn.

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	121,26150*	251,13263***

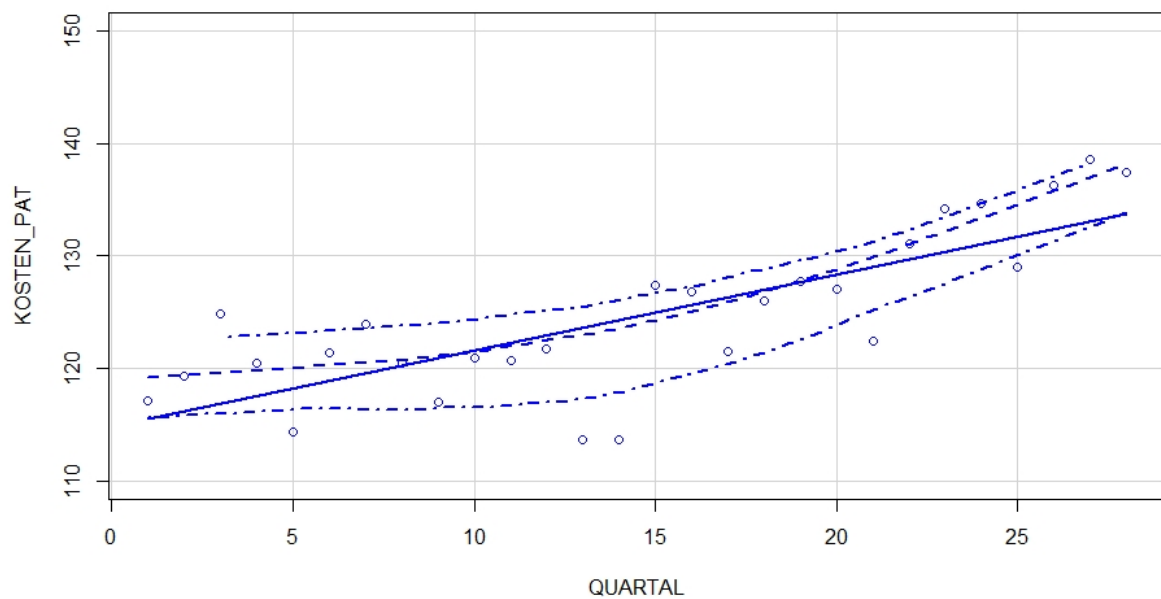
Anteil der Patient*innen im Alter ab 70	-189,99764.	-467,48184.
Anteil der weiblichen Patient*innen	215,61373**	137,52302.
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	209,46505*	548,08774***
Anteil der Polypharmazie-Patient*innen	52,49540	157,58247
Anzahl der Fälle je Quartal	-0,08138**	-0,12892***

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,7772	0,8943
F-Statistik	19,84 bei 5 und 22 Freiheitsgraden	46,71 bei 5 und 22 Freiheitsgraden
p-Wert	$1,728 * 10^{-7}$	$5,485 * 10^{-11}$

Bayern:

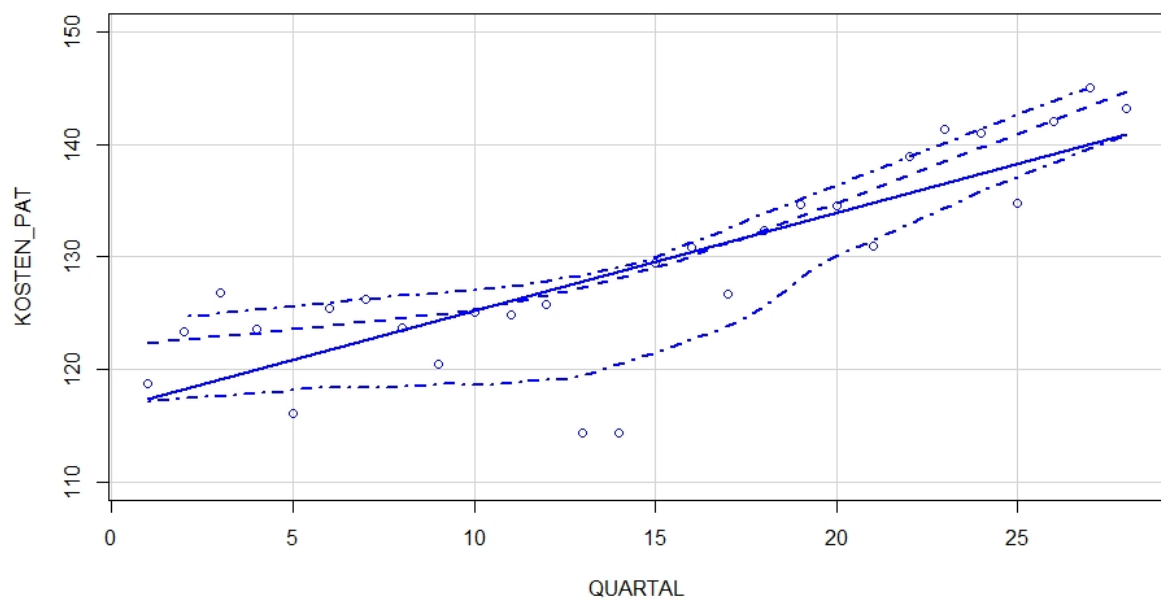
Für die KV Bayerns hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	6,8122 bei 5 Freiheitsgraden	0,235
Chow-Test	0,97118	0,4753

Hessen:

Für die KV Hessen hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	7,1472 bei 5 Freiheitsgraden	0,2099
Chow-Test	0,84761	0,5522

Verordnungsanteil von Generika am Gesamtmarkt

Für den durchschnittlichen Verordnungsanteil von Generika am Gesamtmarkt wurde das Zeitreihen-Regressionsmodell

$$Generika_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Für dieses Modell wird auf ein Vorliegen von Heteroskedastizität getestet und anschließend der Chow Test auf Strukturbruch durchgeführt.

Es folgen die Schätzergebnisse der beiden KVen und anschließend die einzelnen Testergebnisse.

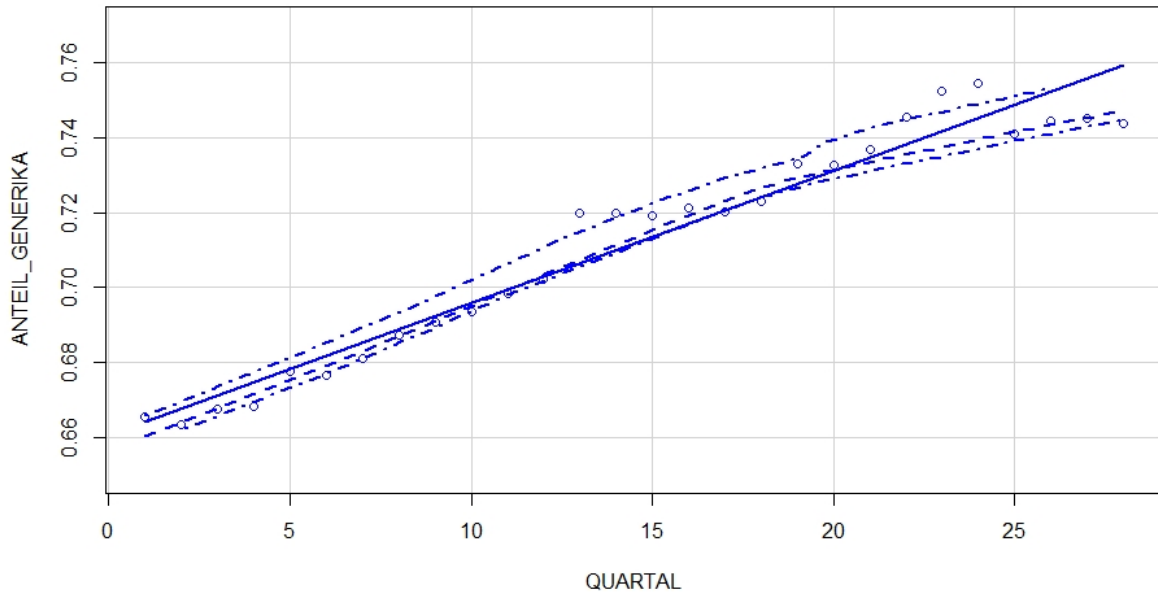
	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	0,1108	0,2839.
Anteil der Patient*innen im Alter ab 70	-0,4962*	1,295**
Anteil der weiblichen Patient*innen	0,8155***	7,209*
Anteil der Patient*inne mit einem Charlson-Index über dem 90%-Perzentil	-0,05580	8,254*
Anteil der Polypharmazie-Patient*innen	0,6940*	-2,334**
Anzahl der Fälle je Quartal	$1,346 * 10^{-4}$ *	$5,758 * 10^{-5}$

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,9441	0,8409
F-Statistik	92,23 bei 5 und 22 Freiheitsgraden	29,55 bei 5 und 22 Freiheitsgraden
p-Wert	$5,274 * 10^{-14}$	$4,619 * 10^{-9}$

Bayern:

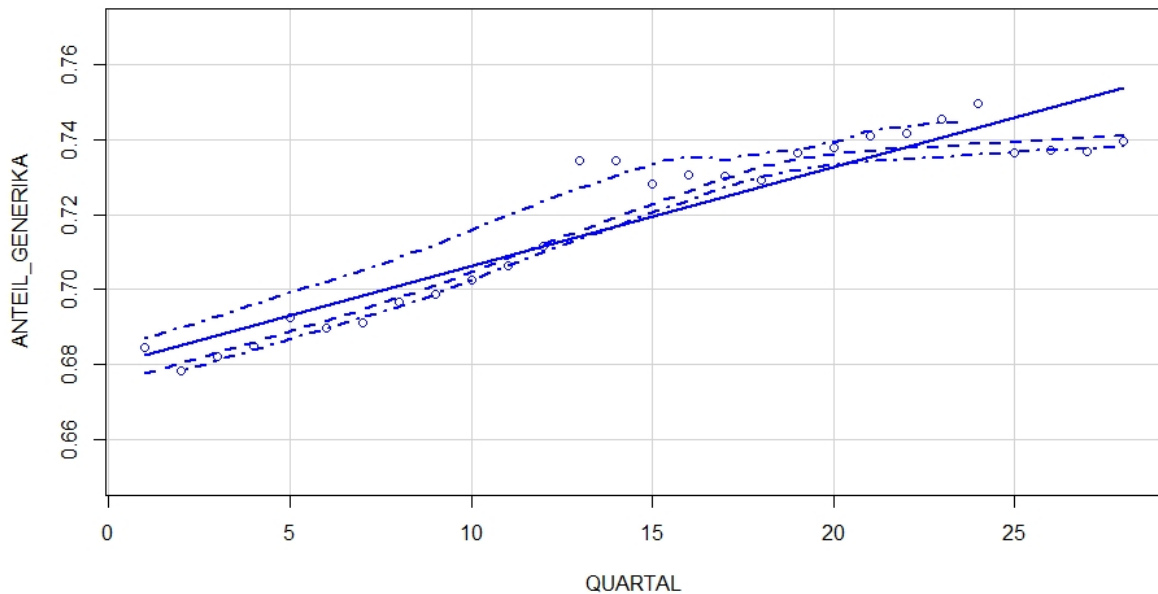
Für die KV Bayerns hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	2,9689 bei 5 Freiheitsgraden	0,7048
Chow-Test	4,1874	0,01015

Hessen:

Für die KV Hessen hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	11,466 bei 5 Freiheitsgraden	0,04289

Zur Überprüfung auf ein Vorliegen von Heteroskedastizität wurde der Breusch-Pagan-Test durchgeführt, der mit einer Teststatistik von BP=11,466 bei 5 Freiheitsgraden und einem p-Wert von 0,04289 die Nullhypothese der Homoskedastizität verwirft. Somit wird die Alternativhypothese der Heteroskedastizität bestätigt. Da hier die Voraussetzung des Chow-Tests nicht erfüllt ist, muss das Modell zunächst transformiert werden, um die für den Test vorausgesetzte Homoskedastizität zu erhalten, indem eine robuste Kovarianzmatrixschätzung durchgeführt wird. Das transformierte Modell hat die folgende Form:

	Koeffizientenschätzung
Intercept	0,43372
Anteil der Patient*innen im Alter ab 70	1,2897**
Anteil der weiblichen Patient*innen	0,61287
Anteil der Patienten mit einem Charlson-Index über dem 90%-Perzentil	0,75272
Anteil der Polypharmazie-Patient*innen	-2,0608*
Anzahl der Fälle je Quartal	$9,2669 * 10^{-6}$

	Teststatistik	p-Wert
Chow-Test	1,7564	0,172

Patientenanteil mit DOAK-Verordnung an allen Patient*innen mit Antikoagulantienverordnung

Für den Patientenanteil mit DOAK-Verordnung an allen Antikoagulationspatienten wurde das Zeitreihen-Regressionsmodell

$$DOAK_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Basierend auf dieser Schätzung wurde auch hier auf ein Vorliegen von Heteroskedastizität getestet und anschließend der Chow Test auf Strukturbruch durchgeführt.

Es folgen nun die Schätzergebnisse der beiden KVen gefolgt von den jeweiligen Testergebnissen der beiden KVen. Die Parameterschätzung ergibt:

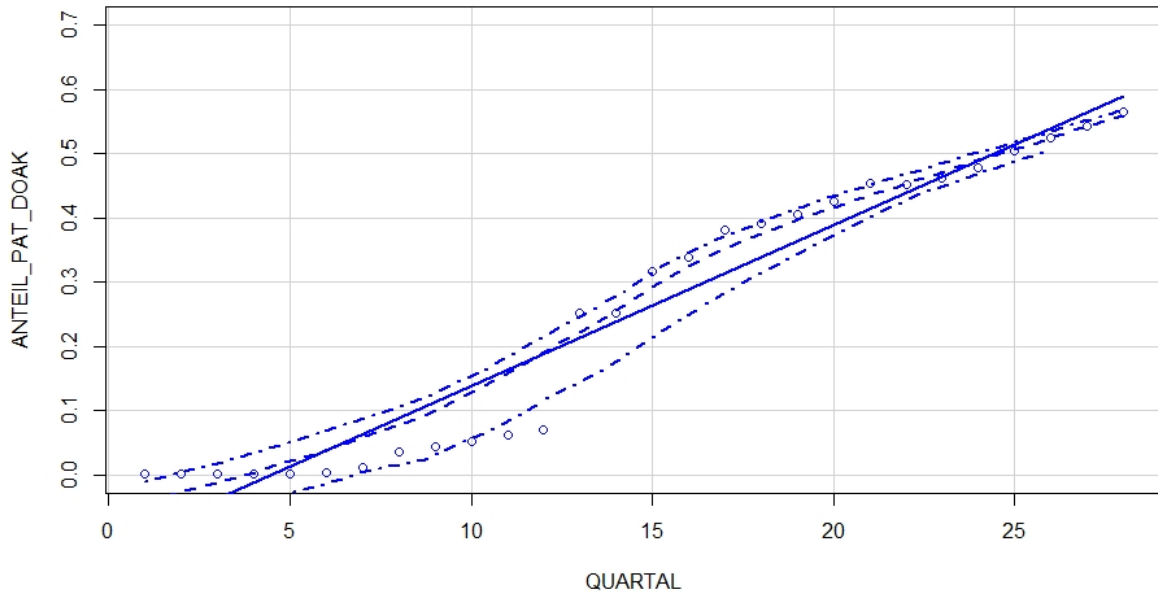
	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	-3,6902166 ***	-3,1382893***
Anteil der Patient*innen im Alter ab 70	-5,7808405**	2,0067579
Anteil der weiblichen Patient*innen	8,2601024	5,1706422***
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	0,1947542***	10,0022220***
Anteil der Polypharmazie-Patient*innen	2,9279215	-9,0827056*
Anzahl der Fälle je Quartal	0,0003739	0,0004759

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0, 9431	0, 9708
F-Statistik	90,57 bei 5 und 22 Freiheitsgraden	180,05 bei 5 und 22 Freiheitsgraden
p-Wert	6,376 * 10 ⁻¹⁴	< 2,2e-16

Bayern:

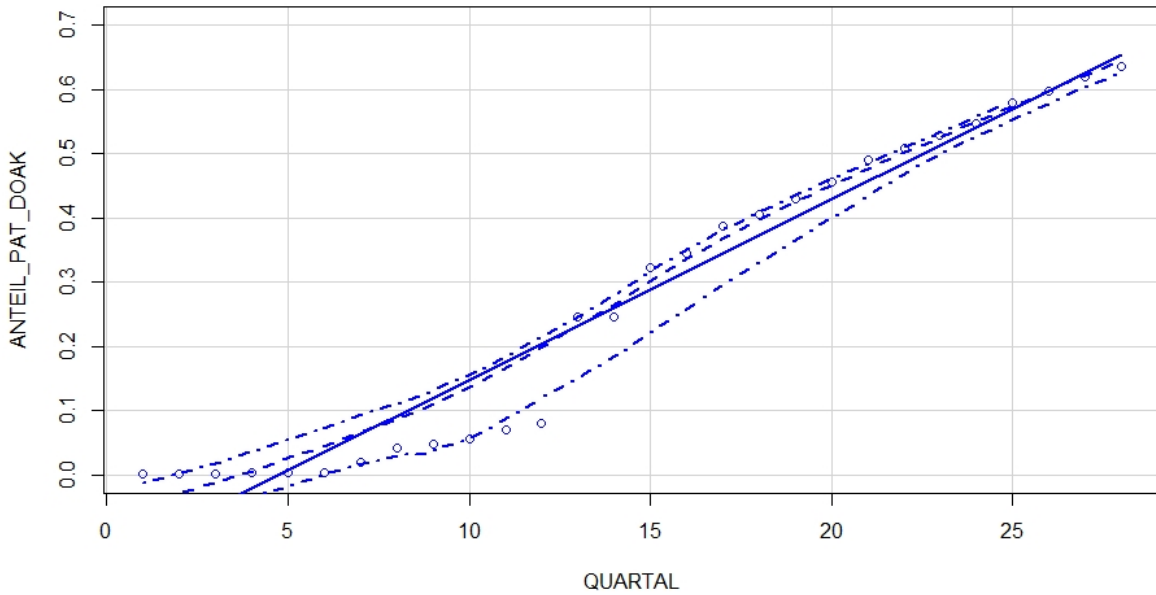
Für die KV Bayerns hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	4,1513 bei 5 Freiheitsgraden	0,5278
Chow-Test	5,0541	0,004387

Hessen:

Für die KV Hessen hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	16,074 bei 5 Freiheitsgraden	0,006637

Zur Überprüfung auf ein Vorliegen von Heteroskedastizität wurde der Breusch-Pagan-Test durchgeführt, der mit einer Teststatistik von BP=16,074 bei 5 Freiheitsgraden und einem p-Wert von 0,006637 die Nullhypothese der Homoskedastizität verwirft. Somit wird die Alternativhypothese der Heteroskedastizität bestätigt. Da hier die Voraussetzung des Chow-Tests nicht erfüllt ist, muss das Modell zunächst transformiert werden, um die für den Test vorausgesetzte Homoskedastizität zu erhalten, indem eine robuste Kovarianzmatrixschätzung durchgeführt wird. Für das transformierte Modell ergibt sich die folgende Koeffizientenschätzung:

	Koeffizientenschätzung
Intercept	-3,13828929***
Anteil der Patient*innen im Alter ab 70	2,00675787
Anteil der weiblichen Patient*innen	5,17064216***
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	10,00222204***
Anteil der Polypharmazie-Patient*innen	-9,08270558*

Anzahl der Fälle je Quartal	0,00047592
-----------------------------	------------

	Teststatistik	p-Wert
Chow-Test	0,19412	0,9739

Erreichung von Leitsubstanzquoten

Für die Erreichung der allgemeinen Leitsubstanzquoten (d.h. Anteil aller Leitsubstanzverordnungen über alle Leitsubstanzgruppen der eigenen KV hinweg) wurde das Zeitreihen-Regressionsmodell

$$\text{AnteilLeitsubstanzen}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Basierend auf dieser Schätzung wurde auch hier auf ein Vorliegen von Heteroskedastizität getestet und anschließend der Chow Test auf Strukturbruch durchgeführt.

Es folgen die Schätzergebnisse der beiden KVen gefolgt von den jeweiligen Testergebnissen. Die Parameterschätzung ergibt:

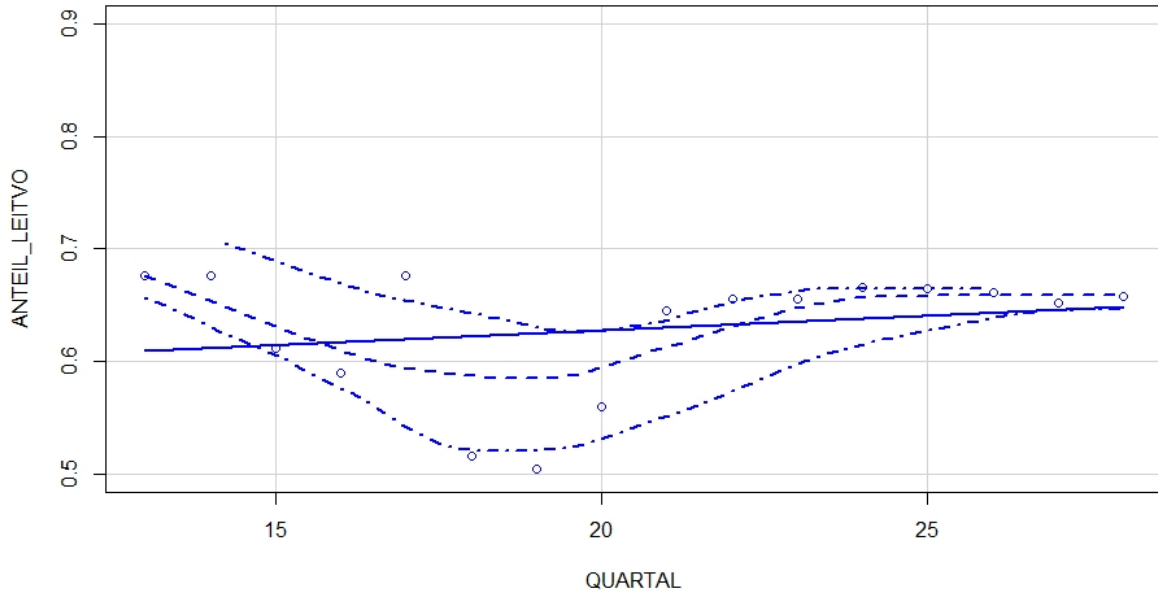
	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	3,1669606.	0,3698**
Anteil der Patient*innen im Alter ab 70	-3,2463783	0,4544e
Anteil der weiblichen Patient*innen	-0,1553461	0,02849
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	12,2743642	-1,867*
Anteil der Polypharmazie-Patient*innen	-6,3863766	1,160
Anzahl der Fälle je Quartal	-0,0014100.	$2,358 * 10^{-4}***$

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,006738	0,6011
F-Statistik	1,02 bei 5 und 10 Freiheitsgraden	6,727 bei 5 und 14 Freiheitsgraden
p-Wert	0,4551	0,002161

Bayern:

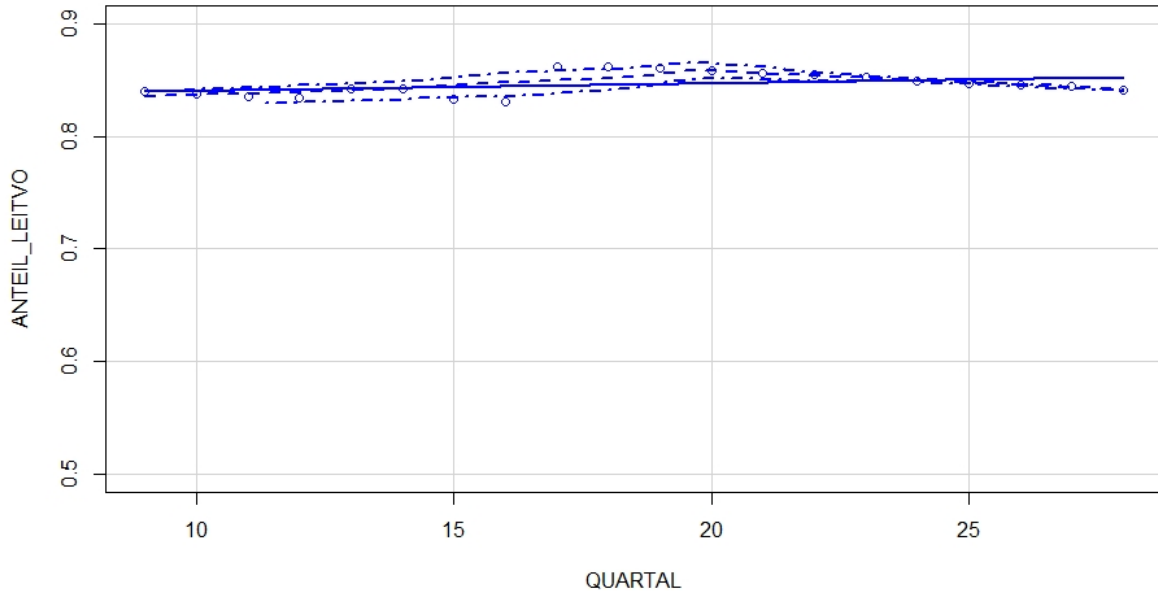
Für die KV Bayerns hat die betrachtete Zeitreihe den folgenden Verlauf:



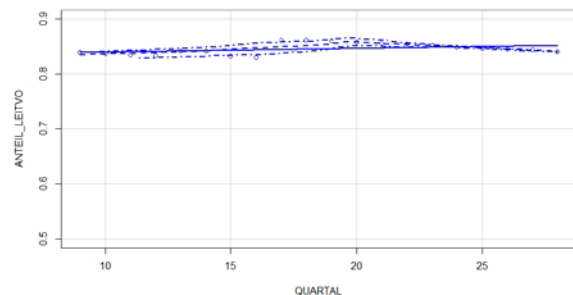
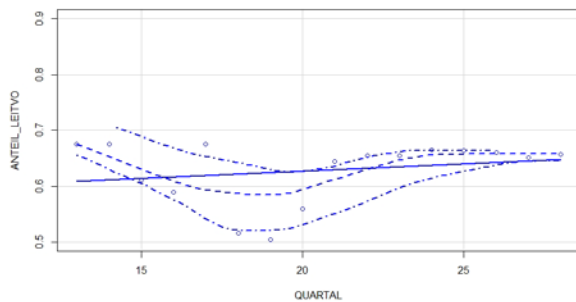
	Teststatistik	p-Wert
Breusch-Pagan-Test	2,9689 bei 5 Freiheitsgraden	0,7048
Chow-Test	50,836	0,0009881

Hessen:

Für die KV Hessen hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	1,7575 bei 5 Freiheitsgraden	0,8816
Chow-Test	5,3741	0,01659



Im direkten Vergleich von Bayern und Hessen liegt eine erheblich größere Schwankung in der KV Bayerns vor. In der KV Hessen hingegen findet die Bewegung in einem sehr schmalen Bereich statt.

Verordnungsanteil von generischen DDDs im Bereich Analgetika (außer BTM-pflichtige Opioide)

Für den Verordnungsanteil von generischen DDDs im Bereich Analgetika ohne BTM-pflichtige Opioide wurde das Zeitreihen-Regressionsmodell

$$\text{Anteil generischer DDDs}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Dieses Modell wird auf Heteroskedastizität getestet und anschließend wird der Chow Test auf Strukturbruch durchgeführt.

Es folgen zunächst die Schätzergebnisse der beiden KVen. Im Anschluss werden die Testergebnisse der einzelnen KVen präsentiert.

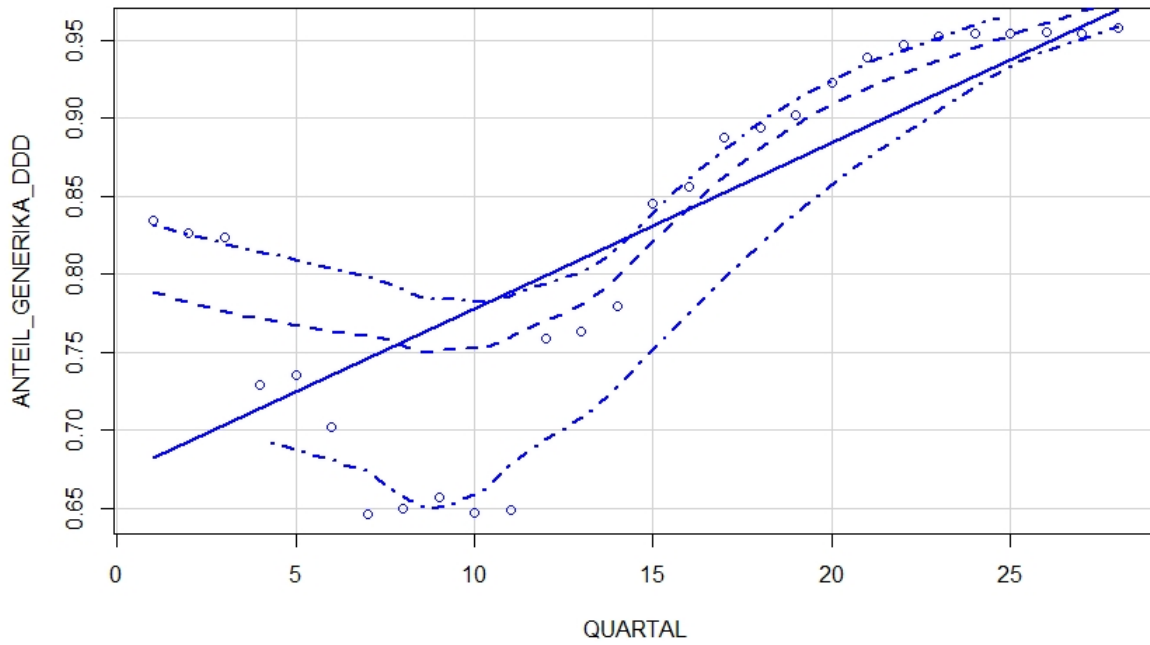
	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	-0,5946	-0,7559
Anteil der Patient*innen im Alter ab 70	-6,9337**	-4,7609.
Anteil der weiblichen Patient*innen	6,5679***	6,0212***
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	0,5357	0,8798
Anteil der Polypharmazie-Patient*innen	1,9305	0,1481
Anzahl der Fälle je Quartal	-0,0005	-0,0003

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,7309	0,9082
F-Statistik	15,67 bei 5 und 22 Freiheitsgraden	54,39 bei 5 und 22 Freiheitsgraden
p-Wert	$1,298 * 10^{-6}$	$1,195 * 10^{-11}$

Bayern:

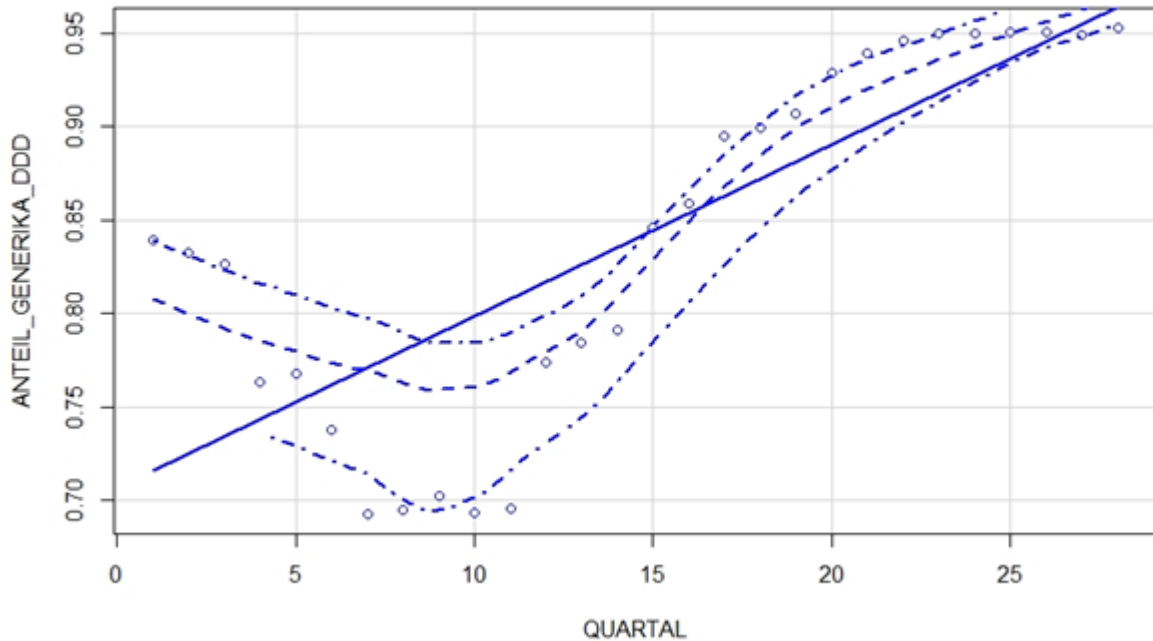
Für die KV Bayerns hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	7,4137 bei 5 Freiheitsgraden	0,1916
Chow-Test	0,57067	0,748

Hessen:

Für die KV Hessen hat die betrachtete Zeitreihe den folgenden Verlauf:



Zudem wurden folgende Tests durchgeführt:

	Teststatistik	p-Wert
Breusch-Pagan-Test	8,926 bei 5 Freiheitsgraden	0,1121
Chow-Test	0,59412	0,7308

Verordnungsanteil von generischen DDDs im Bereich der Mittel bei obstruktiven Atemwegserkrankungen

Für den Verordnungsanteil von generischen DDDs im Bereich der Mittel bei obstruktiven Atemwegserkrankungen wurde das Zeitreihen-Regressionsmodell

$$\text{Anteil generischer DDDs}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

geschätzt. Dieses Modell wird auf Heteroskedastizität getestet und anschließend wird der Chow Test auf Strukturbruch durchgeführt. Es folgen die einzelnen Schätz- und Testergebnisse der beiden KVen.

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	0,6835***	0,3540.
Anteil der Patient*innen im Alter ab 70	-0,5112*	0,3861

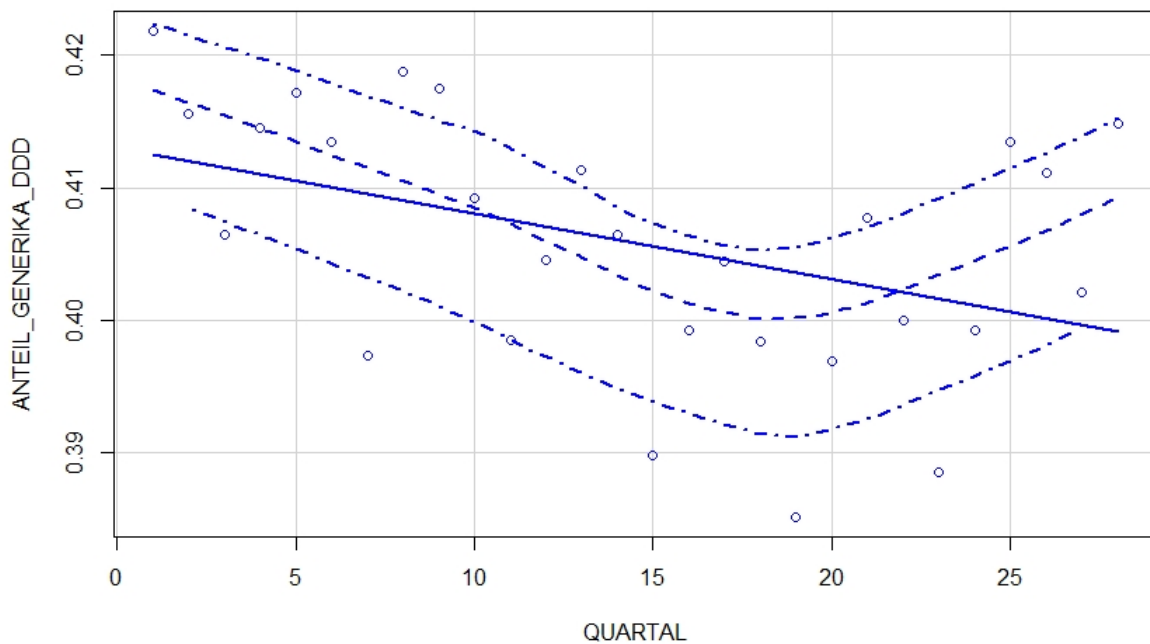
Anteil der weiblichen Patient*innen	-0,2264	-0,2311
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	0,2049	-0,8450.
Anteil der Polypharmazie-Patient*innen	-0,2283	1,301
Anzahl der Fälle je Quartal	$-1,321 * 10^{-5}$	$1,438 * 10^{-4}$.

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,4098	0,6344
F-Statistik	4,749 bei 5 und 22 Freiheitsgraden	10,37 bei 5 und 22 Freiheitsgraden
p-Wert	0,004297	$3,281 * 10^{-5}$

Bayern:

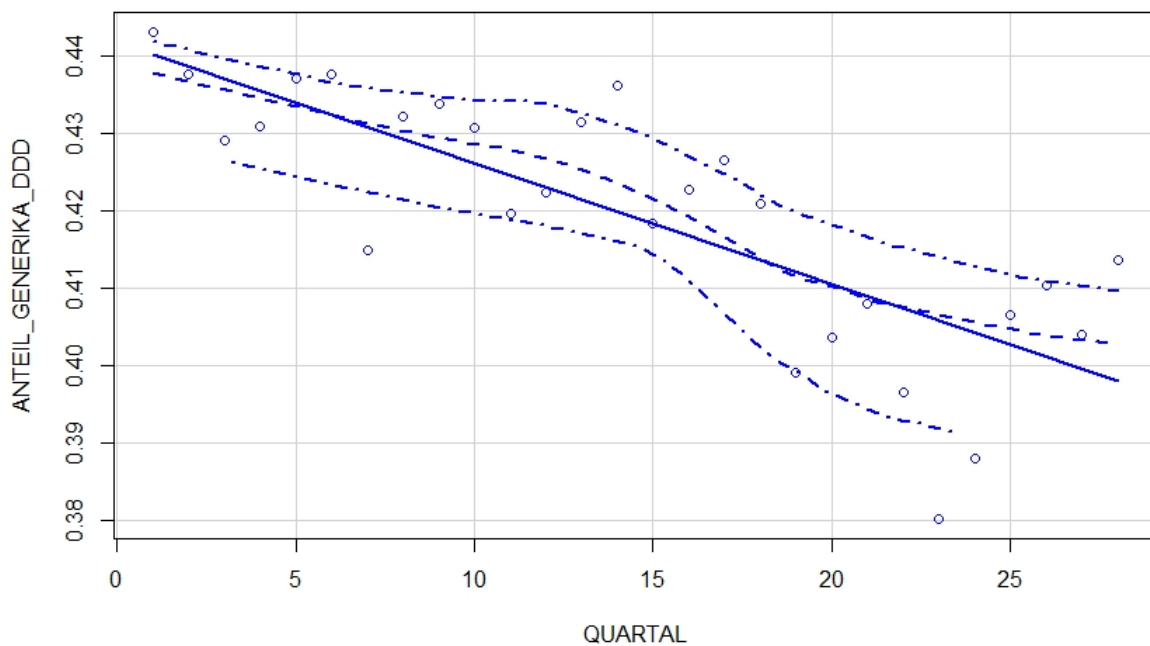
Für die KV Bayerns hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	4,8534 bei 5 Freiheitsgraden	0,434
Chow-Test	5,6671	0,002537

Hessen:

Für die KV Hessen hat die betrachtete Zeitreihe den folgenden Verlauf:



Zudem wurden folgende Tests durchgeführt:

	Teststatistik	p-Wert
Breusch-Pagan-Test	6,994 bei 5 Freiheitsgraden	0,2211
Chow-Test	4,7782	0,00568

Verordnungsanteil von generischen DDDs im Bereich der Mittel mit Wirkung auf das Renin-Angiotensin-System

Für den Verordnungsanteil von generischen DDDs im Bereich der Mittel mit Wirkung auf das Renin-Angiotensin-System wurde ebenfalls das Zeitreihen-Regressionsmodell

$$\text{Anteil generischer DDDs}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

geschätzt. Auch dieses Modell wird auf Heteroskedastizität getestet und es wird anschließend der Chow Test auf Vorliegen eines Strukturbruchs durchgeführt. Es folgen die Schätzergebnisse der beiden KVen gefolgt von den jeweiligen Testergebnissen. Die Parameterschätzung ergibt:

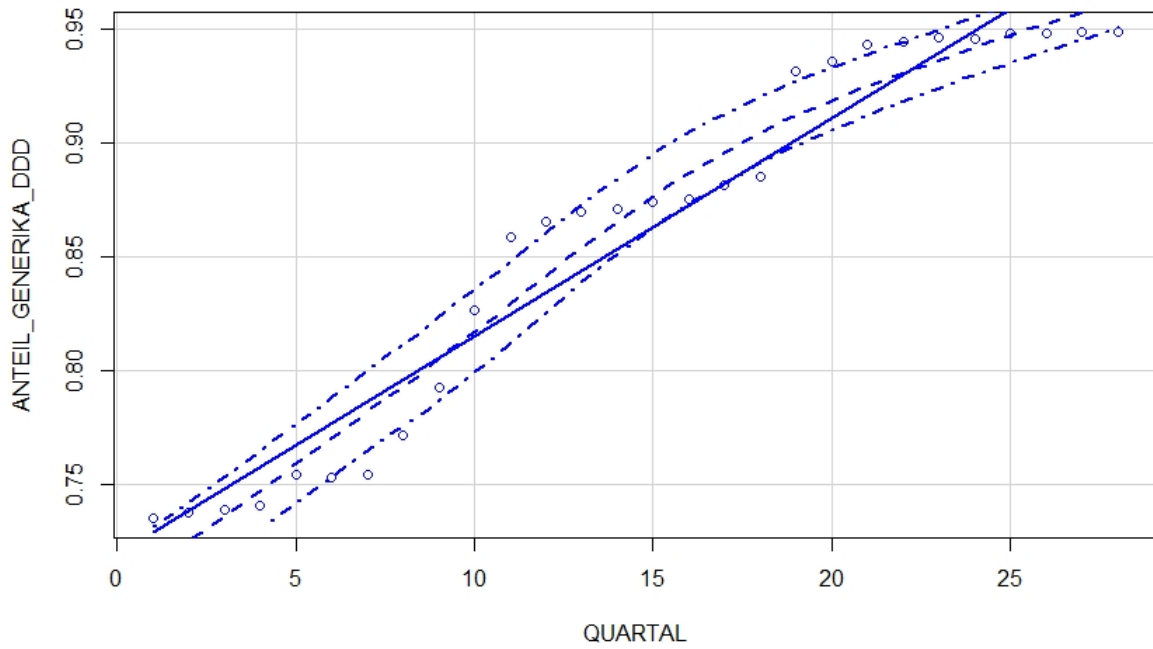
	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	-0,7076524.	0,1584395
Anteil der Patient*innen im Alter ab 70	-2,1491029*	-0,6463397
Anteil der weiblichen Patient*innen	1,6324935*	0,0305873
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	-0,2685773	3,2094736.
Anteil der Polypharmazie-Patient*innen	3,3621441**	0,0344748
Anzahl der Fälle je Quartal	0,0005104*	0,0003144

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,8939	0,8348
F-Statistik	46,48 bei 5 und 22 Freiheitsgraden	28,28 bei 5 und 22 Freiheitsgraden
p-Wert	$5,763 * 10^{-11}$	$6,975 * 10^{-9}$

Bayern:

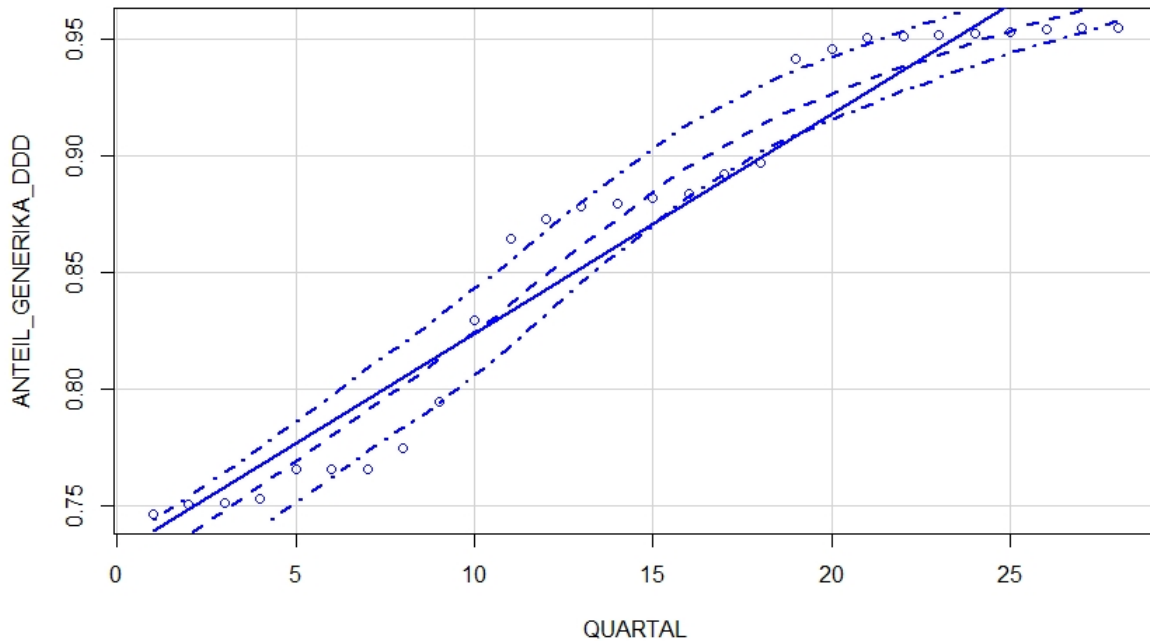
Für die KV Bayerns hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	5,0795 bei 5 Freiheitsgraden	0,4063
Chow-Test	3,7695	0,01566

Hessen:

Für die KV Hessen hat die betrachtete Zeitreihe den folgenden Verlauf:



Zudem wurden folgende Tests durchgeführt:

	Teststatistik	p-Wert
Breusch-Pagan-Test	3,457 bei 5 Freiheitsgraden	0,6299
Chow-Test	10,684	$7,715 \cdot 10^{-5}$

Verordnungsanteil von generischen DDDs im Bereich der Thrombozytenaggregationshemmer (ohne Ticagrelor)

Für den Verordnungsanteil von generischen DDDs im Bereich der Mittel der Thrombozytenaggregationshemmer ohne Ticagrelor wurde das Zeitreihen-Regressionsmodell

$$\text{Anteil generischer DDDs}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

geschätzt. Nach Prüfung auf Heteroskedastizität wird anschließend der Chow Test auf Strukturbruch durchgeführt. Es folgen die Schätzergebnisse der beiden KVen gefolgt von den Testergebnissen:

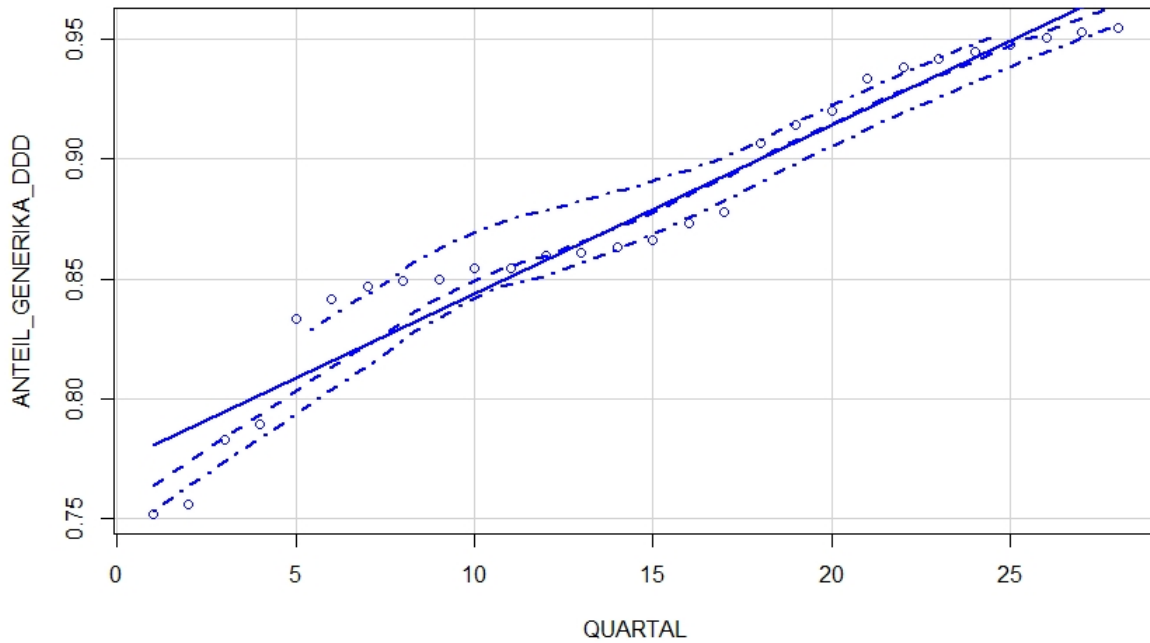
	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	0,4697**	0,8789395*
Anteil der Patient*innen im Alter ab 70	-0,1102	3,2188130.
Anteil der weiblichen Patient*innen	$7,219 * 10^{-2}$	-2,3322277***
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	1,113***	3,8771934**
Anteil der Polypharmazie-Patient*innen	$1,520 * 10^{-2}$	-3,1776266
Anzahl der Fälle je Quartal	$1,541 * 10^{-4}$.	0,0003754*

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,9638	0,8072
F-Statistik	145 bei 5 und 22 Freiheitsgraden	23,62 bei 5 und 22 Freiheitsgraden
p-Wert	$4,488 * 10^{-16}$	$3,662 * 10^{-8}$

Bayern:

Für die KV Bayerns hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	11,578 bei 5 Freiheitsgraden	0,04105

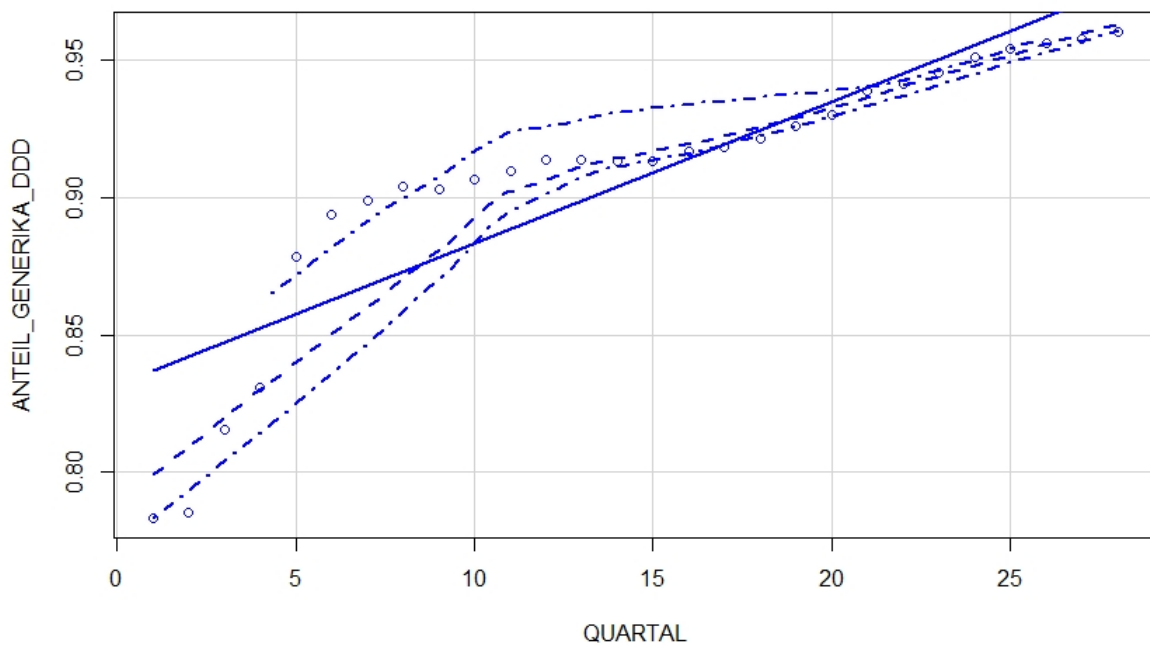
Der zur Überprüfung auf ein Vorliegen von Heteroskedastizität durchgeführte Breusch-Pagan-Test liefert eine Teststatistik von BP=11,578 bei 5 Freiheitsgraden und einem p-Wert von 0,04105. Dies bedeutet, dass bei einem Signifikanzniveau von 5% die Nullhypothese verworfen und die Alternativhypothese der Heteroskedastizität angenommen wird. Da hier die Voraussetzung des Chow-Tests nicht erfüllt ist, muss das Modell zunächst transformiert werden, um die für den Test vorausgesetzte Homoskedastizität zu erhalten, indem eine robuste Kovarianzmatrixschätzung durchgeführt wird. Für das transformierte Modell ergibt sich die folgende Koeffizientenschätzung:

	Koeffizientenschätzung
Intercept	0,46969321***
Anteil der Patient*innen im Alter ab 70	-0,11018968
Anteil der weiblichen Patient*innen	0,07219358
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	1,11289488***
Anteil der Polypharmazie-Patient*innen	0,01520035
Anzahl der Fälle je Quartal	0,00015405

	Teststatistik	p-Wert
Chow-Test	0,74674	0,6207

Hessen:

Für die KV Hessen hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	11,482 bei 5 Freiheitsgraden	0,04263

Der zur Überprüfung auf ein Vorliegen von Heteroskedastizität durchgeführte Breusch-Pagan-Test liefert eine Teststatistik von $BP=11,482$ bei 5 Freiheitsgraden und einem p-Wert von 0,04263. Aufgrund dessen wird die Nullhypothese bei einem Signifikanzniveau von 5% verworfen und die Alternativhypothese der Heteroskedastizität bestätigt. Da hier die Voraussetzung des Chow-Tests nicht erfüllt ist, wird das Modell nun zunächst transformiert, um die für den Test vorausgesetzte Homoskedastizität zu erhalten. Für das transformierte Modell ergibt sich die folgende Koeffizientenschätzung:

	Koeffizientenschätzung
Intercept	0,87893949*
Anteil der Patient*innen im Alter ab 70	3,21881297.
Anteil der weiblichen Patient*innen	-2,33222768**
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	3,87719339**
Anteil der Polypharmazie-Patient*innen	-3,17762659
Anzahl der Fälle je Quartal	0,00037536.

	Teststatistik	p-Wert
Chow-Test	1,6728	0,1918

Verordnungsanteil von Leitsubstanz-DDDs innerhalb der Gruppe der MS-Therapeutika

Für den Verordnungsanteil von Leitsubstanz-DDDs innerhalb der Gruppe der MS-Therapeutika wurde das Zeitreihen-Regressionsmodell

$$\text{Anteil Leitsubstanz} - \text{DDDs}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

geschätzt. Dieses Modell wird zunächst auf Heteroskedastizität getestet, um anschließend den Chow Test auf Strukturbruch durchzuführen. Es folgen nun die Schätz- und Testergebnisse der beiden KVen.

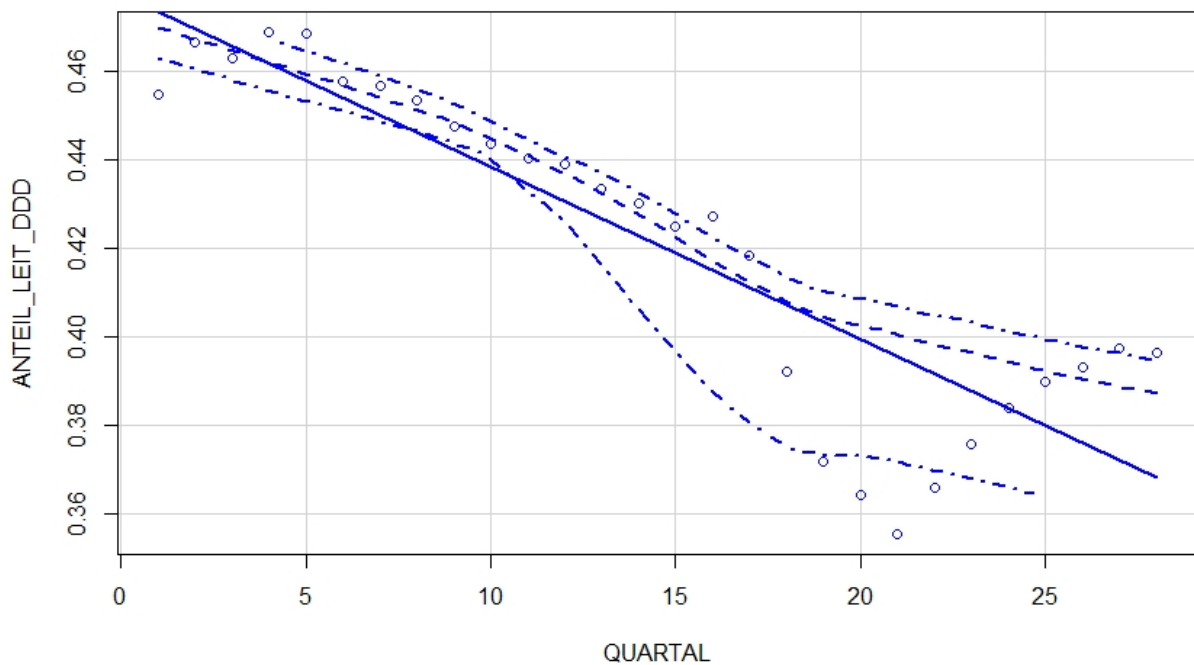
	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	0,6146*	0,9046*
Anteil der Patient*innen im Alter ab 70	0,3752	-0,06849
Anteil der weiblichen Patient*innen	-0,5915	-0,2712
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	-0,6965	-4,237***
Anteil der Polypharmazie-Patient*innen	0,4749	2,101
Anzahl der Fälle je Quartal	$3,868 * 10^{-5}$	$-7,316 * 10^{-5}$

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,7227	0,9194
F-Statistik	15,07 bei 5 und 22 Freiheitsgraden	62,63 bei 5 und 22 Freiheitsgraden
p-Wert	$1,784 * 10^{-6}$	$2,863 * 10^{-12}$

Bayern:

Für die KV Bayerns hat die betrachtete Zeitreihe den folgenden Verlauf:

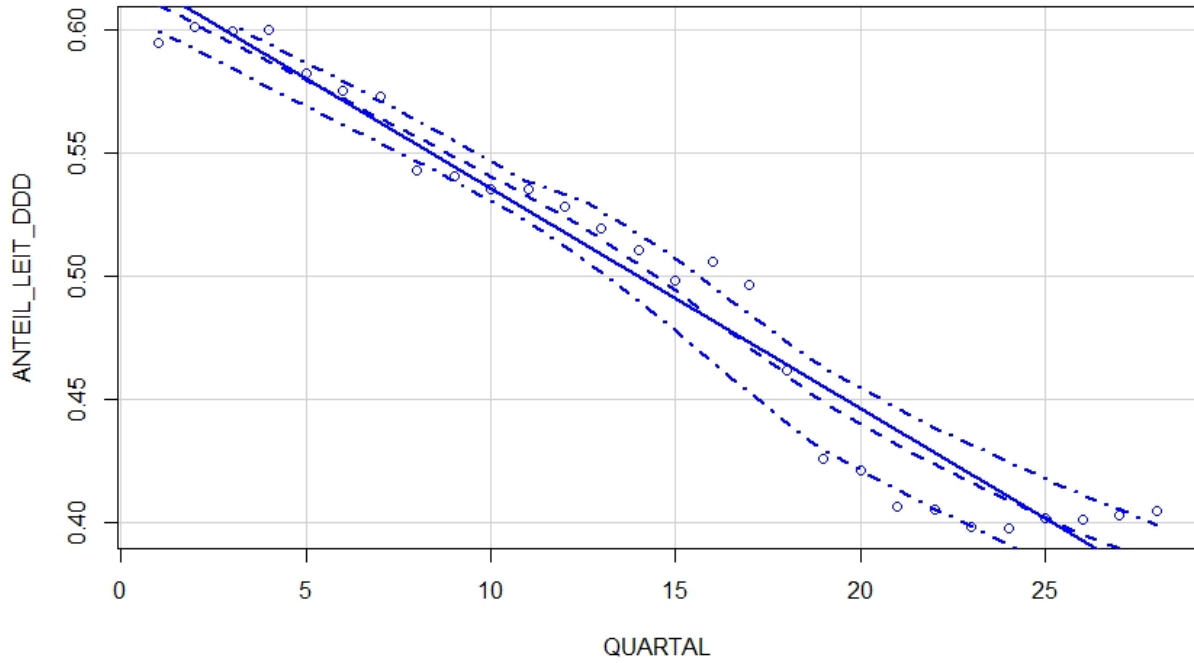


Zudem wurden folgende Tests durchgeführt:

	Teststatistik	p-Wert
Breusch-Pagan-Test	10,017 bei 5 Freiheitsgraden	0,07476
Chow-Test	4,549	0,007082

Hessen:

Für die KV Hessen hat die betrachtete Zeitreihe den folgenden Verlauf:



	Teststatistik	p-Wert
Breusch-Pagan-Test	7,9228 bei 5 Freiheitsgraden	0,1605
Chow-Test	5,9755	0,001951

Anhang 3: Teilprojekt C

3.1 Fokusgruppen-Leitfaden



1. Einleitung

Zunächst möchten wir uns nochmals herzlich bedanken, dass Sie sich zur Teilnahme an unserer Studie bereiterklärt haben. Bevor wir beginnen, möchten wir uns gerne vorstellen und den Ablauf der Fokusgruppe erklären. Wir kommen vom Uniklinikum Erlangen aus dem Fachbereich der Allgemeinmedizin und aus der Abteilung für Allgemeinmedizin der Uni in Marburg. Uns interessiert, wie sie bereits in dem Schreiben vorweg erfahren haben, das Thema Arzneimittelverordnung im Praxisalltag von niedergelassenen Ärzten.

Dabei interessieren uns besonders Ihre persönlichen Erfahrungen, Meinungen und Einschätzungen zu verschiedenen Themenbereichen. Der Schwerpunkt unseres Vorhabens liegt im Bereich der Arzneimitteltherapie bzw. dem ärztlichen Ordnungsverhalten. Uns interessieren Ihre Erfahrung mit der bayerischen Wirkstoffvereinbarung und die Hintergründe, Herausforderungen und ihr Vorgehen bei der Verordnung von Medikamenten.

Wie Sie bereits auch gelesen haben, wird die Fokusgruppe etwa 2 Stunden dauern und digital aufgezeichnet. Anschließend verschriftlichen wir das Interview und anonymisieren alle personenbezogenen Daten. Die Angaben werden vertraulich behandelt und letztlich nur als zusammengefasster ausgewerteter Datensatz publiziert. Es wird also streng darauf geachtet, dass keine direkten oder auch indirekten Rückschlüsse auf Ihre Person möglich sind. Nach Abschluss der Studie werden die Audiodateien gelöscht.

Falls es keine Fragen Ihrerseits vorab gibt, können wir mit der Fokusgruppe beginnen, am Ende der Fokusgruppe haben wir noch Raum für Fragen und Anregungen Ihrerseits eingeplant.

2. Allgemeine Informationen zu Person und Praxiserfahrung

Zunächst möchten wir Sie bitten, sich der Reihe nach vorzustellen. Das hilft uns natürlich, um Sie alle kennenzulernen und zugleich ist es für die Verschriftlichung später wichtig, eine Sprachaufnahme von jedem von Ihnen zu Beginn zu bekommen.

Stellen Sie sich mit ein paar Worten vor und erzählen Sie uns doch, wo Ihre Praxis ist und wie typische Patientenfälle bei Ihnen aussehen.

- Zum Beispiel Alter, Krankheitsbild

3. Allgemeines zum Verordnungsverhalten

Nun geht es um Therapieoptionen und Medikationsverordnungen.

Bsp.: Ganz allgemein, wenn ein Patient in Ihre Praxis kommt, Sie eine Diagnose stellen und er ein Medikament bekommen soll:

Welche Kriterien sind für Sie wichtig, wenn Sie zwischen verschiedenen Arzneimittel-Therapien auswählen?

Beispiele vorlegen, nachdem der Arzt seine ersten Antworten erklärt hat:

- Wirkmechanismus
- Wirkung
- Vorgabe der Leitlinien
- Hersteller
- Preise
- Persönliche Erfahrungswerte
- Pharmavertreter
- Fortbildungen
- Kongresse
- Patientenwünsche
- ehemalige Regressansprüche
- Begleitmedikationen
- Komorbidität
- Studien

Inwiefern haben Patienten und ihre Wünsche Einfluss auf Ihre Arzneimittelverordnung?

- *Was für Wünsche sind das zum Beispiel?*
- *Wie gehen Sie damit im Alltag um? (Stichwort: Aut idem)*
- *Wie reagieren Ihre Patienten, wenn Sie nicht das Präparat verschreiben, welches der Patient haben möchte? (Stichwort: Adhärenz-Probleme)*
- *Wie oft kommt es vor?*

Es werden ja regelmäßig neue Medikamente auf dem Markt zugelassen, bei dem Sie entscheiden müssen, ob Sie es Ihren Patienten geben oder nicht. Wie gehen Sie dabei vor?

Und wie gehen Sie bei vom Krankenhaus / von anderen Fachärzten initiierten Therapien vor?

- *Welche Auswirkungen haben von anderen Fachärzten initiierte Therapien auf Ihr Verordnungsverhalten?*
- *Werden diese weitergeführt?*

4. Gesamtsituation und Bayerische Wirkstoffvereinbarung

Es gab einen Artikel von der KV Hamburg, der diese Art von Wirkstoffvereinbarung zusammenfassend wie folgt beschreibt:

Im ehemaligen System der Richtgrößenprüfung hatten die Ärzte die Verantwortung für die Arzneimittelkosten – ein grundlegender Konstruktionsfehler.

Bei dem neuen Modell geht es nicht mehr um absolute Kosten, sondern um DDD-Quoten und Anteile von Generika und Leitsubstanzen.

Mit Wendung: Der Arzt verantwortet lediglich, was er selbst beeinflussen kann – er wählt den richtigen Wirkstoff aus.

„Die eigentliche Kostenfrage verschiebt das Modell weg von den Ärzten und dahin, wo sie hingehört, nämlich zu den Krankenkassen“

Inwiefern passt das zu den Erfahrungen, die Sie mit der Bayerischen Wirkstoffvereinbarung gemacht haben? Wie beurteilen Sie diese Aussage?

- *Inwiefern hat sich an Ihrem Ordnungsverhalten seitdem etwas verändert?*
- *Welche Auswirkungen nehmen Sie von der Bayerischen Wirkstoffvereinbarung auf Ihre
ärztliche Therapiefreiheit/Therapiehoheit wahr?*
- *Vorteile/Nachteile zu früheren Verfahren?*
- *Gibt es Erkrankungen/Indikationen, die Sie anders therapieren würden, als es die Wirkstoffvereinbarung vorgibt?*
- *Wie schaffen Sie es, Ihre Ziele im Rahmen der Bayerischen Wirkstoffvereinbarung einzuhalten?*
 - *Haben Sie schon einmal ein Beratungsangebot zur Verordnungshilfe wahrgenommen?*
 - *Wenn ja, wie lief dieses ab?*

Wie zufrieden sind Sie aktuell/jetzt im Moment mit Ihrer Verordnungssituation?

- *Wie schätzen Sie die Zufriedenheit Ihrer Kollegen/anderer Fachärzte mit der Verordnungssituation ein?*

Was erwarten Sie von der KVB? Wird die Kassenärztliche Vereinigung diesen Erwartungen gerecht?

Was würde Ihnen das Verordnen erleichtern? Wie könnte man das System der bayerischen Wirkstoffvereinbarung verbessern?

In der medizinischen Versorgung wird viel von Kosten und Wirtschaftlichkeit gesprochen, die auch laut Gesetz überprüft werden soll. Auf der anderen Seite soll natürlich ein bestimmter Qualitätsstandard auch eingehalten werden

Sind in Ihren Augen im Bereich der Arzneimitteltherapie die Begriffe „Wirtschaftlichkeit“ und „Qualität“ vereinbar?

5. Abschluss & Fazit

Zuletzt möchten wir noch Raum für weitere Aspekte geben, die Ihnen möglicherweise noch wichtig sind.

Gibt es denn weitere Aspekte, welche Sie gerne noch ansprechen möchten?

Nachdem wir diese verschiedenen Facetten beleuchtet haben, möchten wir gerne ein Fazit ziehen.

Nehmen Sie sich doch jetzt noch eine Minute Zeit und überlegen, wie Sie die Bayerische Wirkstoffvereinbarung abschließend beschreiben würden. *(Positive wie negative Beschreibungen/Adjektive; einschränkend, hilfreich, unterstützend, absichernd) Fazit der Reihe nach.*

Ausklang: Herzlichen Dank für die Teilnahme!

3.2 Interview-Leitfaden



Anmerkungen.

Pro Interview werden ca. 2 Vignetten vorbereitet und je nach Fachgebiet konstruiert.

*. Gelb markierte Fragen wurden ausschließlich in den Einzelinterviews gestellt.

Allgemeine Informationen zu Person und Praxiserfahrung

Zunächst möchten wir Sie bitten, sich der Reihe nach vorzustellen. Das hilft uns natürlich, um Sie alle kennenzulernen und zugleich ist es für die Verschriftlichung später wichtig, eine Sprachaufnahme von jedem von Ihnen zu Beginn zu bekommen.

Stellen Sie sich mit ein paar Worten vor und erzählen Sie uns doch, wo Ihre Praxis ist und wie typische Patientenfälle bei Ihnen aussehen.

- Zum Beispiel Alter, Geschlecht, Krankheitsbild, Verhalten

Allgemeines zum Verordnungsverhalten

Nun geht es um Therapieoptionen und Medikationsverordnungen.

Bsp.: Ganz allgemein, wenn ein Patient in Ihre Praxis kommt, Sie eine Diagnose stellen und er ein Medikament bekommen soll:

Welche Kriterien sind für Sie wichtig, wenn Sie zwischen verschiedenen Arzneimittel-Therapien auswählen?

Beispiele vorlegen, nachdem der Arzt seine ersten Antworten erklärt hat:

- Wirkmechanismus
- Wirkung
- Vorgabe der Leitlinien
- Hersteller
- Preise
- Persönliche Erfahrungswerte
- Pharmavertreter
- Fortbildungen
- Kongresse
- Patientenwünsche
- ehemalige Regressansprüche
- Begleitmedikationen
- Komorbidität
- Studien

Inwiefern haben Patienten und ihre Wünsche Einfluss auf Ihre Arzneimittelverordnung?

- *Was für Wünsche sind das zum Beispiel?*
- *Wie gehen Sie damit im Alltag um? (Stichwort: Aut idem)*
- *Wie reagieren Ihre Patienten, wenn Sie nicht das Präparat verschreiben, welches der Patient haben möchte? (Stichwort: Adhärenz-Probleme)*
- *Wie oft kommt es vor?*

Wenn sich bei einem Medikament konkrete Fragen ergeben, wie gehen Sie dann vor?*

Hier vor allem: Wirksamkeit, Sicherheit, Wirtschaftlichkeit – wie informieren Sie sich?

Es werden ja regelmäßig neue Medikamente auf dem Markt zugelassen, bei dem Sie entscheiden müssen, ob Sie es Ihren Patienten geben oder nicht. Wie gehen Sie dabei vor?

Und wie gehen Sie bei vom Krankenhaus / von anderen Fachärzten initiierten Therapien vor?

- *Welche Auswirkungen haben von anderen Fachärzten initiierte Therapien auf Ihr Verordnungsverhalten?*
- *Werden diese weitergeführt?*
-

Gesamtsituation und Bayerische Wirkstoffvereinbarung

Es gab einen Artikel von der KV Hamburg, der diese Art von Wirkstoffvereinbarung zusammenfassend wie folgt beschreibt:

Im ehemaligen System der Richtgrößenprüfung hatten die Ärzte die Verantwortung für die Arzneimittelkosten – ein grundlegender Konstruktionsfehler.

Bei dem neuen Modell geht es nicht mehr um absolute Kosten, sondern um DDD-Quoten und Anteile von Generika und Leitsubstanzen.

Mit Wendung: Der Arzt verantwortet lediglich, was er selbst beeinflussen kann – er wählt den richtigen Wirkstoff aus.

„Die eigentliche Kostenfrage verschiebt das Modell weg von den Ärzten und dahin, wo sie hingehört, nämlich zu den Krankenkassen“

Inwiefern passt das zu den Erfahrungen, die Sie mit der Bayerischen Wirkstoffvereinbarung gemacht haben? Wie beurteilen Sie diese Aussage?

- *Inwiefern hat sich an Ihrem Verordnungsverhalten seitdem etwas verändert?*
- *Welche Auswirkungen nehmen Sie von der Bayerischen Wirkstoffvereinbarung auf Ihre*

ärztliche Therapiefreiheit/Therapiehoheit wahr?

- *Vorteile/Nachteile zu früheren Verfahren?*
- *Mit welchen Adjektiven kann man sie beschreiben? (Positive wie negative Beschreibungen; einschränkend, hilfreich, unterstützend, absichernd)*
- *Gibt es Erkrankungen/Indikationen, die Sie anders therapieren würden, als es die Wirkstoffvereinbarung vorgibt?*
- *Wie schaffen Sie es, Ihre Ziele im Rahmen der Bayerischen Wirkstoffvereinbarung einzuhalten?*
 - o *Haben Sie schon einmal ein Beratungsangebot zur Verordnungshilfe wahrgenommen?*
 - o *Wenn ja, wie lief dieses ab?*

Wie zufrieden sind Sie aktuell/jetzt im Moment mit Ihrer Verordnungssituation?

- *Wie schätzen Sie die Zufriedenheit Ihrer Kollegen/anderer Fachärzte mit der Verordnungssituation ein?*

Was erwarten Sie von der KVB? Wird die Kassenärztliche Vereinigung diesen Erwartungen gerecht?

Was würde Ihnen das Verordnen erleichtern? Wie könnte man das System der bayerischen Wirkstoffvereinbarung verbessern?

Die Bayerische Wirkstoffvereinbarung wurde ja aus einem bestimmten Grund eingeführt und basiert auf den Paragraphen zum Wirtschaftlichkeitsgebot (Paragraphen vorlegen). **Wo führt Ihrer Meinung nach diese Entwicklung in den nächsten fünf bis zehn Jahren hin?**

Vorlegen der Paragraphen:

12 SGB V Wirtschaftlichkeitsgebot

Die Leistungen müssen ausreichend, zweckmäßig und wirtschaftlich sein; sie dürfen das Maß des Notwendigen nicht überschreiten. Leistungen, die nicht notwendig oder unwirtschaftlich sind, können Versicherte nicht beanspruchen, dürfen die Leistungserbringer nicht bewirken und die Krankenkassen nicht bewilligen.

§ 106 SGB V Wirtschaftlichkeitsprüfung

Die Krankenkassen und die Kassenärztlichen Vereinigungen überwachen die Wirtschaftlichkeit der vertragsärztlichen Versorgung durch Beratungen und Prüfungen.

In der medizinischen Versorgung wird viel von Kosten und Wirtschaftlichkeit gesprochen, die auch laut Gesetz überprüft werden soll. Wie geht es Ihnen eigentlich persönlich damit?*

- *Arbeitsmotivation*
- *Emotion*

Sind in Ihren Augen im Bereich der Arzneimitteltherapie die Begriffe „Wirtschaftlichkeit“ und „Qualität“ vereinbar?

Abschluss & Fazit

Zuletzt möchten wir noch Raum für weitere Aspekte geben, die Ihnen möglicherweise noch wichtig sind.

Was war Ihre Motivation hier bei uns teilzunehmen?*

Gibt es denn weitere Aspekte, welche Sie gerne noch ansprechen möchten?

Nachdem wir diese verschiedenen Facetten beleuchtet haben, möchten wir gerne ein Fazit ziehen.

Wenn Sie jetzt an Ihre Praxisarbeit denken, welche Aspekte sehen Sie kritisch / Was würden Sie gerne ändern, wenn Sie könnten?*

Und was bereitet Ihnen besonders Freude?*

Ausklang: Herzlichen Dank für die Teilnahme!

Begleitfragebogen zu Interviews und Fokusgruppen

Demographie-Fragebogen

Wir möchten Sie um folgende Angaben zu Ihrer Person und Praxis bitten. Alle Angaben werden streng vertraulich behandelt.

Alter: 30-40 Jahre 41-50 Jahre 51-60 Jahre 61 oder älter

Geschlecht: Weiblich Männlich

Fachrichtung: _____

Facharzt seit: _____ Arbeitserfahrung: _____
(Seit wann sind Sie ärztlich tätig?)

Wie lange haben Sie in der Klinik gearbeitet? _____

Wie lange sind Sie schon ambulant tätig? _____

Seit wann sind Sie in dieser Praxis tätig? _____

Um welche Praxisart handelt es sich?

Einzelpraxis Gemeinschaftspraxis MVZ an Klinik angegliedertes MVZ

Sonstiges: _____

Lage der Praxis:

ländlich (< 5.000 Einwohner) Klein-/Mittelstadt (< 100.000 EW) Großstadt (≥ 100.000 EW)

Wie viele ÄrztInnen arbeiten in dieser Praxis? _____

Arbeiten Sie: Vollzeit Teilzeit Anzahl Stunden: _____; zusätzlich in der Klinik

Wie lange besteht die ärztliche Zusammensetzung dieser Praxis? _____
(Falls mehrere Facharztgruppen in Ihrer Praxis arbeiten, bezogen auf Ihre Facharztgruppe)

Hat die ärztliche Zusammensetzung der Praxis in den Jahren 2017 und 2018 gewechselt?
(Falls mehrere Facharztgruppen in Ihrer Praxis arbeiten, bezogen auf Ihre Facharztgruppe)

ja nein seit 01.01.18

Falls ja: wann und wie oft? _____

Interview/Fokusgruppe geführt am: _____

Von: _____

Anhang 4: Teilprojekt D

4.1 Datenschutzkonformer Rekrutierungsweg

Um den Datenschutz zu wahren und keine Informationen über Auffälligkeiten, die KV-intern vorliegen, an die Universitäten weiterzuleiten, wurde die Rekrutierung der Ärzt*innen über die KV Bayerns durchgeführt. Diese verschickte im Auftrag der Universitäten die Anschreiben und Einladungen zur Studienteilnahme, interessierte Ärzt*innen konnten sich bei Bereitschaft zur Teilnahme an den Interviews bei der Universität Marburg zurückmelden. Auf diese Weise entschieden Ärzt*innen, ob sie teilnehmen wollten und damit ihre Auffälligkeit zumindest gegenüber den Universitäten Marburg und Erlangen bekunden, und die KV Bayerns erhielt gleichzeitig keine Informationen darüber, welche Ärzt*innen an den Interviews teilnehmen.

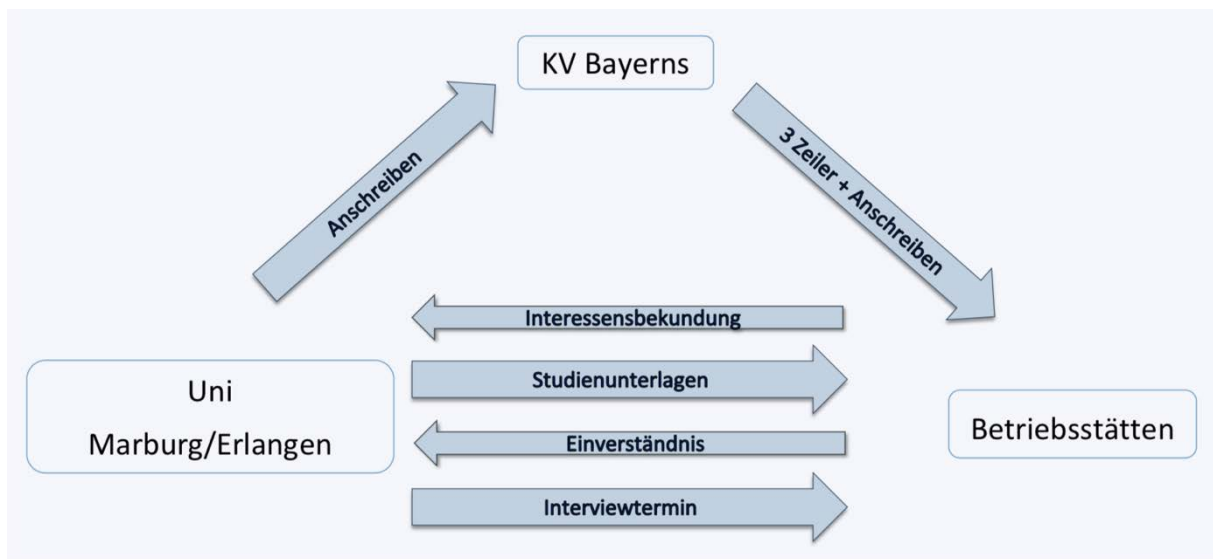


Abbildung 20: entwickelter, datenschutzkonformer Rekrutierungsweg

4.2 Studienunterlagen



Empfängeranschrift

Fachbereich Medizin

Abteilung für Allgemeinmedizin,
Präventive und Rehabilitative Me-
dizin

Prof. Dr. med.

**Norbert Donner-Banzhoff,
M.H.Sc.**

WiMi.: Julia

Gollnick Tel.:
06421 / 28

25010

Fax.: 06421 / 28 65121

E-Mail: gollnicj@staff.uni-

marburg.de Anschrift: Karl-

von-Frisch-Str. 4

35043 Marburg

Web: <http://www.uni-marburg.de/fb20/allgprmed>

Marburg/Erlangen, 31. Au-
gust 2018

Teilnahme an der „WirtMed“-Studie

Lieber Frau...../Lieber Herr.....,

herzlichen Dank für Ihr Interesse an unserer Studie zur Wirtschaftlichkeit und Qualität bei der Verordnung von Medikamenten!

Durch Ihre Teilnahme leisten Sie einen wichtigen Beitrag zur Verbesserung des Versorgungssystems, Förderung der Verordnungsqualität und vor allem auch zur Verbesserung der Arbeitszufriedenheit niedergelassener Ärzte.

Mit diesem Schreiben übersenden wir Ihnen weitere Informationen zur Studie inkl. Aufklärung über Aspekte des Datenschutzes nach der neuen DSGVO vom 25.05.2018 (*Studieninformationsblatt*), eine *Einverständniserklärung* und das Formular für die Überweisung Ihrer *Aufwandsentschädigung*.

Wir bitten Sie, diese auszufüllen und an uns zurückzufaxen (Fax: 06421-28 65121)!

Gerne können Sie auf dem Einverständniserklärung-Formular zwei bis drei Terminwünsche für die Zeit ab Mitte November bis Ende Februar angeben, die wir bei der Findung der Interviewtermine bestmöglich berücksichtigen werden.

Für die endgültige Terminabsprache und die Möglichkeit Ihnen weitere Fragen zu beantworten, werden wir – die wissenschaftlichen Mitarbeiter und Interviewer – uns demnächst telefonisch bei Ihnen melden.

Wir bedanken uns für Ihr Interesse, Ihre Zeit und Ihre wichtige Kooperation! Mit freundlichen Grüßen

Prof. Dr. med. Norbert Donner-Banzhoff

Prof. Dr. med. Thomas Kühlein

Zum Verbleib beim Arzt

Studien-ID:

TPD1819

**„Die Verordnung von Arzneimitteln: Prüfung und Steuerung von
Wirtschaftlichkeit und Qualität
- WirtMed-Studie“**

Liebe Kolleginnen und Kollegen,

der Frage, wie der Spagat zwischen guter Patientenversorgung und wirtschaftlichem Verordnungsverhalten ohne Angst vor Regressen gemeistert werden kann, haben wir – zwei universitäre Einrichtungen innerhalb der Allgemeinmedizin – uns zur Aufgabe gemacht.

Dabei kooperieren wir, die Abteilung/das Institut für Allgemeinmedizin in Marburg (Projektleiter/Konsortialführung Prof. Dr. med. Norbert Donner-Banzhoff) und Erlangen (Leitung Prof. Dr. med. Thomas Kühlein) in einem vom Innovationsausschuss des G-BA geförderten Konsortialprojekt mit weiteren Partnern verschiedenster Expertisen.

Wir freuen uns sehr über Ihr Interesse und Ihre Bereitschaft an der Teilnahme der Interviews im

Rahmen eines der fünf Teilprojekte der „WirtMed –Studie“.

Studienziel:

In diesem Teilprojekt soll die Meinung niedergelassener Kolleginnen und Kollegen zum Thema

„Wirtschaftlichkeit und Qualität bei der Verordnung von Medikamenten“ eingeholt werden. Im Fokus stehen die individuellen Erfahrungen, aber auch Probleme und/oder die möglichen Alternativen zu den derzeitigen Steuerungs- und Prüfmodellen. Diese Befragung wendet sich speziell an Kolleginnen und Kollegen, die Schwierigkeiten (gehabt) haben, die Vorgaben zur Wirtschaftlichkeit zu erreichen. Das Projekt soll helfen, die Situation und die Meinungen dieser Gruppe zu verstehen. Ziel ist die Entwicklung von alternativen Modellen, welche zum einen die Qualität der Verordnungen fördern, zum anderen die Arbeitszufriedenheit niedergelassener Ärzte verbessern sollen.

Die Modelle sollen derart konzipiert und den verantwortlichen Körperschaften an die Hand gegeben werden, dass möglichst eine Steuerung und keine Prüfung im herkömmlichen Sinne stattfindet, die Regressangst verringert und die Vereinbarkeit von Wirtschaftlichkeit und Qualität in der Arzneimitteltherapie erreicht wird.

Ablauf Interview:

Das Interview kann bei Ihnen in der Praxis oder gegebenenfalls in einer unserer Räumlichkeiten stattfinden und wird voraussichtlich zwei Stunden in Anspruch nehmen. Das Gespräch wird von wissenschaftlichen Mitarbeiterinnen der Philipps-Universität Marburg/des Universitätsklinikums Erlangen geführt und mit einem digitalen Audiogerät aufgezeichnet.

Aufwand Arzt/Ärztin:

Einmalige Interviewteilnahme von circa zwei Stunden wahlweise in Ihrer Praxis oder der Abteilung für Allgemeinmedizin der Philipps-Universität Marburg bzw. dem Institut für Allgemeinmedizin des Uniklinikums Erlangen.

Aufwandsentschädigung

Für die einmalige Interviewteilnahme erhalten Sie eine Aufwandsentschädigung i.H.v. 90,00 €.

Datenschutz

Für die Durchführung der Studie und die Erhebung Ihrer Daten benötigen wir Ihr Einverständnis. Wir sichern Ihnen zu, dass Ihre persönlichen und praxisbezogenen Daten nur anonymisiert erfasst und verarbeitet werden. Die abschließende Auswertung erfolgt qualitativ. Nach Beendigung der Studie werden die anonymisierten Daten 10 Jahre in den beiden universitären Einrichtungen gespeichert. Alle übrigen personen- und praxisbezogenen Daten werden nach Abschluss der Studie vernichtet.

Bitte unterschreiben Sie die beiliegende Einverständniserklärung, um an der Studie teilzunehmen.

Die Teilnahme an der Studie ist freiwillig; Sie können Ihr Einverständnis jederzeit und ohne Angabe von Gründen zurückziehen, ohne dass Ihnen daraus Nachteile entstehen.

Bis zur vollständigen Anonymisierung Ihrer Daten haben Sie zusätzlich das Recht, jederzeit Auskunft (*einschließlich unentgeltlicher Überlassung einer Kopie*) über die Sie betreffenden personen- und praxisbezogenen Daten zu bekommen sowie deren Berichtigung oder Löschung zu verlangen.

Für Rückfragen stehen Ihnen die wissenschaftlichen Mitarbeiterinnen Julia Gollnick Abteilung für Allgemeinmedizin, Fachbereich Medizin der Philipps-Universität Marburg, 35043 Marburg, Tel. 06421-28-25010, sowie Nikoletta Lippert, Allgemeinmedizinisches Institut, Universitätsklinikum Erlangen, 91054 Erlangen, Tel. 09131-85-45774, jederzeit zur Verfügung.

Wir danken Ihnen für Ihre Mitarbeit!

Arzt

Ort, Datum

Unterschrift (Studienteam)

Ich, Frau/Herr _____ **(Namen bitte deutlich schreiben!)**

wurde heute vollständig über Wesen und Bedeutung der oben genannten wissenschaftlichen Studie aufgeklärt. Ich habe das Studieninformationsblatt sowie die nachfolgend abgedruckten Rechte über meinen Datenschutz gelesen und verstanden. Ich hatte ausreichend Möglichkeit, Fragen zu stellen und habe die Antworten verstanden.

Ich weiß, dass die Teilnahme freiwillig ist. Meine Zustimmung kann ich jederzeit ohne Angabe von Gründen widerrufen, ohne dass sich dieser Entschluss nachteilig auswirken wird.

Ich wurde über die neuen Richtlinien der Datenschutz-Grundverordnung (EU-DSGVO) in Kenntnis gesetzt. Mir ist zudem bekannt, dass bei dieser Studie personen- und praxisbezogene Daten über mich digital aufgezeichnet, anonymisiert gespeichert und qualitativ ausgewertet werden. Die Sprachaufzeichnung aus dem Interview mit den personen- und praxisbezogenen Daten werden nach Studienende gelöscht. Die anonymisierten Transkripte und die Auswertung dieser Daten werden nach Beendigung der Studie 10 Jahre in der Abteilung für Allgemeinmedizin der Philipps-Universität Marburg/ im Allgemeinmedizinischen Institut des Universitätsklinikums Erlangen gespeichert.

Ich erkläre mich hiermit einverstanden, an der Studie „WirtMed“ teilzunehmen.

Ein Duplikat des Studieninformationsblattes und dieser Einverständniserklärung habe ich erhalten.

Ort, Datum Unterschrift (Ärztin/Arzt)

Ich bestätige, dass ich oben genannte(n) **Ärztin/Arzt**, die/der ihr/sein Einverständnis für die Teilnahme gegeben hat, bezüglich Zweck, Nutzen und Art der Studie aufgeklärt habe.

Marburg/Erlangen, 12. Oktober 2018

**** Datenschutz-Pflichtinformation hier nicht aufgeführt ***

Zurück ins
Studienzentrum

Studien-ID:

TPD1819

Ich, Frau/Herr _____ (**Namen bitte deutlich schreiben!**)

wurde heute vollständig über Wesen und Bedeutung der oben genannten wissenschaftlichen Studie aufgeklärt. Ich habe das Studieninformationsblatt sowie die nachfolgend abgedruckten Rechte über meinen Datenschutz gelesen und verstanden. Ich hatte ausreichend Möglichkeit, Fragen zu stellen und habe die Antworten verstanden.

Ich weiß, dass die Teilnahme freiwillig ist. Meine Zustimmung kann ich jederzeit ohne Angabe von Gründen widerrufen, ohne dass sich dieser Entschluss nachteilig auswirken wird.

Ich wurde über die neuen Richtlinien der Datenschutz-Grundverordnung (EU-DSGVO) in Kenntnis gesetzt. Mir ist zudem bekannt, dass bei dieser Studie personen- und praxisbezogene Daten über mich digital aufgezeichnet, anonymisiert gespeichert und qualitativ ausgewertet werden. Die Sprachaufzeichnung aus dem Interview mit den personen- und praxisbezogenen Daten werden nach Studienende gelöscht. Die anonymisierten Transkripte und die Auswertung dieser Daten werden nach Beendigung der Studie 10

Jahre in der Abteilung für Allgemeinmedizin der Philipps-Universität Marburg/im Allgemeinmedizinischen Institut des Universitätsklinikums Erlangen gespeichert.

Ich erkläre mich hiermit einverstanden, an der Studie „WirtMed“ teilzunehmen.

Ein Duplikat des Studieninformationsblattes und dieser Einverständniserklärung habe ich erhalten.

Ort, Datum Unterschrift (Ärztin/Arzt)

Ich bestätige, dass ich oben genannte(n) **Ärztin/Arzt**, die/der ihr/sein Einverständnis für die Teilnahme gegeben hat, bezüglich Zweck, Nutzen und Art der Studie aufgeklärt habe.

Marburg/Erlangen, 12. Oktober 2018

Zurück ins Studienzentrum

Studien-ID:

WirtMed-Studi€

Aufwandsentschädigung

Projekt: Die Verordnung von Arzneimitteln: Prüfung und Steuerung von Wirtschaftlichkeit und Qualität - WirtMed-Studie

Projektleiter: Prof. Dr. med. Norbert Donner-Banzhoff, MSc
Kostenstelle: _____

Aufwandsentschädigung für Ärztinnen und Ärzte:

Die Ärztin / der Arzt hat im Rahmen der Studie an einem Interview teilgenommen. Sie/Er erhält eine Aufwandsentschädigung in Höhe von:

90 €

Der Betrag ist zu überweisen an:

Name:

Straße:

PLZ, Ort:

Name des Kreditinstituts:

IBAN: | DE | | | | |

(ab hier Bearbeitung durch Abteilung für Allgemeinmedizin der Uni Marburg/Uniklinik Erlangen)

Sachlich und rechnerisch richtig:

Marburg/Erlangen, den

Unterschrift:

Zur Überweisung an Drittmittelabteilung gegeben am:

Unterschrift:

Interviewleitfaden

Der Interviewleitfaden dient lediglich zur inhaltlichen Orientierung und Vorbereitung. Die genauen Formulierungen etc. ergeben sich aus dem Gesprächsverlauf und sind hier als beispielhaft anzusehen.

Pro Interview werden ca. 2 Vignetten vorbereitet und je nach Fachgebiet konstruiert.

Einleitung

Zunächst möchten wir uns nochmals herzlich bedanken, dass Sie sich zur Teilnahme an unserer Studie bereiterklärt haben. Bevor wir beginnen, möchten wir uns gerne vorstellen und den Ablauf des Interviews erklären. Wir sind von der Universität *Marburg bzw. Erlangen* aus dem Fachbereich der Allgemeinmedizin. Uns interessiert, wie sie bereits in dem Schreiben vorweg erfahren haben, das Verordnungsverhalten von Ärzten.

Dabei interessieren uns besonders Ihre persönlichen Erfahrungen, Meinungen und Einschätzungen zu verschiedenen Themenbereichen. Zur Vereinfachung haben wir auch kleine Fallbeispiele vorbereitet. Der Schwerpunkt unseres Vorhabens liegt im Bereich der Arzneimitteltherapie bzw. dem ärztlichen Verordnungsverhalten. Dabei geht es vor allem darum, die Ärzte zu Wort kommen zu lassen, die nicht immer die von der KV vorgegebenen Ziele der Wirkstoffvereinbarung erreicht haben. Uns interessieren die Hintergründe, Herausforderungen und ihr Vorgehen bei der Verordnung von Medikamenten.

Wie sie bereits auch gelesen haben, wird das Interview maximal 2 Stunden dauern und digital aufgezeichnet. Anschließend verschriftlichen/transkribieren wir das Interview und anonymisieren alle Daten zu Personen, Ort oder weitere sensible Informationen. Die Angaben werden vertraulich behandelt und letztlich nur als zusammengefasster ausgewerteter Datensatz publiziert. Es wird also streng darauf geachtet, dass keine Rückschlüsse auf Ihre Person möglich sind. Nach Abschluss der Studie werden die Audiodateien gelöscht.

Falls es keine Fragen Ihrerseits vorab gibt, können wir mit dem Interview beginnen, am Ende des Interviews haben wir noch Raum

für Fragen und Anregungen Ihrerseits eingeplant

1. Allgemeine Informationen zu Person und Praxiserfahrung

Sie haben uns ja schon einen kleinen Fragebogen ausgefüllt, aber wir würden Sie und ihr Arbeitsumfeld gerne noch ein bisschen weiter kennenlernen

Praxisalltag beschreiben lassen

Wenn Sie an Ihre Arbeit als XY denken, was ist der schönste Aspekt in Ihrem Beruf? Welche Aspekte sind weniger schön?

Wenn Sie an Ihre Arbeit in der Praxis denken, wie kann ich mir einen typischen Patienten von Ihnen vorstellen? Beschreiben Sie mir doch in wenigen Worten diesen typischen Patienten.

2. Rolle als Arzt

Selbstbild als Arzt

Uns interessiert, was einen guten Arzt ausmacht. Wenn Sie also einen jungen Mediziner sehen,

was würden Sie ihm mit auf dem Weg geben, wenn er fragt, welche Eigenschaften einen guten Arzt ausmachen?

Wir haben Kärtchen mitgebracht, auf denen Eigenschaften bzw. Verhaltensweisen stehen. Diese können Sie sich erst einmal in Ruhe ansehen (Pause) –

Nun nehmen Sie bitte drei Karten heraus, die Sie für wichtig erachten und uns erzählen warum.

Eigenschaften/Verhaltensweisen eines guten Arztes

Handelt wirtschaftlich	Arbeitet auf dem aktuellen Stand der Wissenschaft
Verantwortungsbewusst	Hat Führungsqualitäten
Starker moralischer Charakter	Zusammenarbeit mit Kollegen
Toleranz von Unsicherheit bei Entscheidungen	Kritisches Hinterfragen von Regularien (z.B. Krankenkassen)
Trifft Entscheidungen mit dem Patienten zusammen	Unabhängigkeit von Industrieinteressen
<input type="checkbox"/> Welche Begriffe haben Sie ausgewählt und warum? <input type="checkbox"/> Gibt es eine Eigenschaft, welche Ihnen hier fehlt?	

3. Allgemeines zum Verordnungsverhalten

Nun geht es um Therapieoptionen und Medikationsverordnungen.

Bsp.: Ganz allgemein, wenn ein Patient in Ihre Praxis kommt, Sie eine Diagnose stellen und er ein Medikament bekommen soll:

Welche Kriterien sind für Sie wichtig, wenn Sie zwischen verschiedenen Arzneimittel-Therapien auswählen?

Beispiele:

- *Wirkmechanismus*
- *Wirkung*
- *Vorgabe der Leitlinien*
- *Hersteller*
- *Preise*
- *Persönliche Erfahrungswerte*
- *Pharmavertreter*
- *Fortbildungen*
- *Kongresse*
- *Patientenwünsche*
- *ehemalige Regressansprüche*
- *Begleitmedikationen*
- *Komorbidität*
- *Studien*

Es werden ja regelmäßig neue Medikamente auf dem Markt zugelassen, bei dem Sie entscheiden müssen, ob Sie es Ihren Patienten geben wollen oder nicht.

Wie gehen Sie dabei vor?

Wie gehen Sie bei der Auswahl des Wirkstoffes bzw. des konkreten Herstellers vor?

*Verschreiben Sie vorwiegend Wirkstoffe oder eher Fertigarzneimittel? Welche Rolle spielt der Preis eines Arzneimittels für Sie?
– Warum?*

Was fällt Ihnen zum Ausdruck „die Therapiehoheit des Arztes“ ein?

Wie setzen Sie das im Alltag um?

Welche Gründe führen dazu, dass Sie Patienten an einen anderen Facharzt überweisen?

- *Wann überweisen Sie für die Rezeptierung von Medikamenten?*
- *Was für Verordnungen sind das überwiegend?*

4. Patientenbezogenes Ordnungsverhalten/Fallbeispiele

Nun soll etwas mehr auf Patienten bzw. bestimmte Patienten eingegangen werden, die in Ihre Praxis kommen.

☐ Vorlegen des Fallbeispiels

Uns interessiert Ihre Vorgehensweise bei diesem Patienten. Es geht nicht um Richtig oder Falsch, vielmehr Ihre persönlichen Überlegungen in einem Bereich, in dem es viele verschiedene Meinungen gibt.

Welchen Wirkstoff würden Sie hier ansetzen/verordnen?

Wir haben hier noch Kärtchen mit einer Auswahl an Wirkstoffen vorbereitet. Bitte bringen Sie diese in eine Reihenfolge, entsprechend Ihrer Präferenz und erläutern uns Ihre Entscheidung.

- Manche Medikamente sind teuer und der Zusatznutzen ist nicht gegeben. In welchem Fall würden Sie es trotzdem ansetzen?

Beispiel: Ein Patient mit Dauertherapie war vor ca. 6 Wochen da und hatte eine 100er Packung seines Medikaments bekommen. Nun kommt er zu Ihnen in die Praxis und möchte ein Rezept für eine neue Packung haben.

Wie handhaben Sie in Ihrer Praxis Folgeverordnungen?

- Wer bereitet das Rezept für Folgeverordnungen vor? (MFA, Arzt selbst?)
- Wird geprüft, ob schon alle Tabletten aufgebraucht sein können?

Wie gehen Sie im Alltag mit expliziten Patientenwünschen zu bestimmten Firmen oder Rezepten um?

- *Wie oft kommt es vor?*
- *Was für Wünsche sind das zum Beispiel?*
- *Wie wird damit umgegangen? (Bspw. XY)*

5. Gesamtsituation

Jetzt möchten wir gerne allgemein mehr zu Ihrer aktuellen Situation / Gesamtsituation erfahren.

Wie zufrieden sind Sie aktuell/jetzt im Moment mit Ihrer Verordnungssituation?

- *Wie schätzen Sie die Zufriedenheit Ihrer Kollegen/anderer Fachärzte mit der Verordnungssituation ein?*

Welche Verordnungen bzw. Bereiche führten bei Ihnen bisher zu Schwierigkeiten?

Was würde Ihnen das Verordnen erleichtern?

6. Bayerische Wirkstoffvereinbarung

In dem nächsten Abschnitt möchten wir uns vor allem auf die bayerische Wirkstoffvereinbarung beziehen.

Die bayerische Wirkstoffvereinbarung sagt Ihnen vermutlich deutlich mehr als uns als Nicht- Mediziner. Diese hat für verschiedenste Bereiche Ziele festgelegt, die man erreichen soll.

- So wird z.B. die Verordnung einer bestimmten Leitsubstanz oder eines Mindestanteils an Generika angestrebt.

Welche Erfahrungen haben Sie denn mit der bayerischen Wirkstoffvereinbarung gemacht?

- *Inwiefern hat sich an Ihrem Ordnungsverhalten seitdem etwas verändert?*
- *Vorteile/Nachteile zu früheren Verfahren?*

Vorlegen der Paragraphen:

12 SGB V Wirtschaftlichkeitsgebot

(1) Die Leistungen müssen ausreichend, zweckmäßig und wirtschaftlich sein; sie dürfen das Maß des Notwendigen nicht überschreiten. Leistungen, die nicht notwendig oder unwirtschaftlich sind, können Versicherte nicht beanspruchen, dürfen die Leistungserbringer nicht bewirken und die Krankenkassen nicht bewilligen.

§ 106 SGB V Wirtschaftlichkeitsprüfung

(1) Die Krankenkassen und die Kassenärztlichen Vereinigungen überwachen die Wirtschaftlichkeit der vertragsärztlichen Versorgung durch Beratungen und Prüfungen.
[...]

Wo führt Ihrer Meinung nach diese Entwicklung in den nächsten fünf bis zehn Jahren hin?

Wie stehen Sie generell zu den Sie vertretenden Körperschaften, wie bspw. der KV,

Berufsverbände oder der Ärztekammer?
Was <u>erwarten</u> Sie von den Sie vertretenden Körperschaften (KV, Berufsverbände)?
7. „Up-to-date“-bleiben
In diesem Themenbereich geht es jetzt darum, wie Sie sich Informationen beschaffen.
Welche Informationsquellen nutzen Sie, um im Praxisalltag bezüglich der Medikamentenversorgung regelmäßig „auf dem Laufenden“ zu bleiben?
<i>„Regelmäßig“ erläutern: täglich/wöchentlich/monatlich</i>
Wenn Sie bei einem Medikament konkrete Fragen haben, wie gehen Sie dann vor? Hier vor allem: Wirksamkeit, Sicherheit, Wirtschaftlichkeit – wie informieren Sie sich?
Welche Rolle spielt es für Sie, Informationen zu Medikamenten von Pharmavertretern zu bekommen?
8. Abschluss
Der letzte Abschnitt bezieht sich auf unsere Studie und bietet Raum für weitere Aspekte, die Ihnen noch wichtig sind.
Da es uns ja vor allem darum geht, die Ärzte mal zu Wort kommen zu lassen, die nicht immer die vorgegebenen Ziele der KV erreicht haben, liegt uns eine Frage ganz besonders am Herzen: Wie geht es Ihnen eigentlich persönlich damit?
- Arbeitsmotivation?
Welche Aspekte würden Sie abschließend gerne noch ansprechen?
Was war Ihre Motivation zur Teilnahme an unserer Studie?
Vielleicht noch eine etwas provokante Frage zum Schluss – Sind in Ihren Augen im Bereich der Arzneimitteltherapie die Begriffe „Wirtschaftlichkeit“ und „Qualität“ vereinbar?
Ausklang: Vielen Dank für die Teilnahme!

Nachfragen:

Wie laufen Gespräche mit Pharmavertretern ab?

- *Werden Studien zu Produkten vorgestellt?*
- *Werden Vor-/Nachteile zu anderen Produkten vorgestellt?*
 - o *Auf welche Medikamentengruppen gehen Sie am meisten ein?*
Wovon hängt ab, ob ein Pharmavertreter Sie überzeugen kann, das Medikament auszuprobieren?
- *Was nimmt man aus solchen Gesprächen so mit?*
- *Bringen Ihnen die Gespräche was?*
- *Was machen solche Gespräche mit Ihnen?*

Inwiefern haben Sie Erfahrungen mit Adherence Problemen bei Patienten bezüglich der Wirkstoffverordnung gemacht?

- *Wie äußern sich Patienten zu den Verordnungen? Kommt viel Widerspruch oder Einwände bezüglich der Auswahl?*
- *Verweigern Patienten die Einnahme, wenn Sie nicht das Präparat verschreiben, was der Patient haben möchte?*

Werden Medikamente auch ausversehen vertauscht, was zu Zwischenfällen oder ggf. sogar Notfällen führen kann?

Dokumentation

Unterbrechungen/ Besondere Ereignisse

Subjektive Einschätzung:

4.4 Untersuchte Fachgruppen für weitere Verordnungsanalysen

Für genauere Analysen der Verordnungsdaten wurden acht Fachgruppen ausgewählt, die nach unserer Einschätzung ein umfangreiches oder auffälliges Verordnungsverhalten aufweisen, oder in der Stichprobe der Interviews auffällig häufig oder mit häufig genannten problematischen Arzneimitteln vertreten waren.

Tabelle: betrachtete Fachgruppen der weiteren Verordnungsanalysen

Betrachtete Fachgruppen
Orthopädie
Pädiatrie
Internisten ohne Schwerpunkt
Hausärzte
Neurologen/Nervenärzte/Psychiater
Internisten mit Schwerpunkt Kardiologie
Internisten mit Schwerpunkt Gastroenterologie
Internisten mit Schwerpunkt Pneumologie

4.5 Stichprobenbeschreibung

Fachrichtung	Anzahl interviewter Ärzte
Augenarzt*in	1
Gastroenterologie	2
Hämatologie-Onkologie	2
Neurochirurgie	1
Neurologie	2
Innere Medizin-Kardiologie	1
Kinder- und Jugendpsychiatrie	1
Chirurgie/Unfallchirurgie	2
Hausarzt*innen	3

Orthopädie (und Unfallchirurgie)	9
Pädiatrie, Kinderrheumatologie	2

Die interviewten 26 Ärzt*innen gehörten folgenden Fachgruppen an:

Tabelle 32: Stichprobe: interviewte Fachgruppen

Besonders häufig befragt wurde die Fachgruppe der Orthopäden.

Interviewte Ärzt*innen arbeiteten in folgenden Praxisarten:

*Tabelle: Stichprobe: Praxisarten der interviewten Ärzt*innen*

Praxisart	Anzahl
Einzelpraxis	11
Gemeinschaftspraxis	8
MVZ	3
an Klinik angegliedertes MVZ	1
Sonstiges:	3
Praxisgemeinschaft	1
Ermächtigungsambulanz	2

Die Praxen der interviewten Ärzt*innen lagen vorwiegend im klein- und mittelstädtischen Raum, wenige Betriebsstätten lagen in städtischen oder ländlichen Gegenden.

Das Alter der interviewten Ärzt*innen lag zwischen 41 und über 61 Jahre. In wenigen Fällen bestehen neben der Arbeit im ambulanten Sektor ein stationärer Stellenanteil und die Arbeitserfahrung im Rahmen der ärztlichen Tätigkeiten beläuft sich auf 13 bis 36 Jahre, während Ärzt*innen seit zwei bis 31 Jahren in der aktuellen Praxis arbeiten. Eine Person hat die Arzneimittelverordnung im Rahmen der vor der Bayerischen Wirkstoffvereinbarung gültigen Richtgrößenprüfung nicht miterlebt.

4.6 Ergebnisse: qualitativ - Details

Rolle als Ärzt*in, wichtige Ärzt*innen-Eigenschaften:

Aus einer Auswahl von zehn Eigenschaften (siehe Tabelle) identifizierten interviewte Ärzt*innen die drei in ihren Augen relevantesten Eigenschaften einer guten Ärzt*in.

*Tabelle: Eigenschaften, die Ärzt*innen aufweisen sollten*

Eigenschaft	Häufigkeit der Nennung
Arbeitet auf dem aktuellen Stand der Wissenschaft	18
Verantwortungsbewusst	16,5
Trifft Entscheidungen mit der/m Patient*in zusammen	15,5
Starker moralischer Charakter	7
Zusammenarbeit mit Kolleg*in	5,5
Kritischen Hinterfragen von Regularien (z.B. Krankenkassen)	4
Hat Führungsqualitäten	4
Toleranz von Unsicherheit bei Entscheidungen	3
Handelt wirtschaftlich	2,5
Unabhängigkeit von Industrieinteressen	2

Anmerkung: Da sich zwei Ärzt*innen bei der Auswahl der dritten Eigenschaft nicht zwischen zwei Eigenschaften entscheiden konnten, wurden die jeweils an dritter Prioritätenstelle aufgeführten beiden Eigenschaften mit 0,5 Wertepunkten aufgenommen (geteilter Punkt).

Neben oben genannten und ausgewählten Eigenschaften, wurde folgende ebenso als wichtig erachtete Eigenschaften ergänzt, die laut den Ärzt*innen in obiger Auswahl fehlten oder nicht ausreichend zum Ausdruck kamen.

Tabelle: ärztliche Eigenschaften, die zusätzlich als wichtig erachtet werden

Weitere wichtige Eigenschaften
Hohe Motivation zur Arbeit
Wie begleite ich eine/n Patient*in
Menschliches Miteinander (sollte auch in der med. Ausbildung besser sein)
Eisernes Nervenkostüm (bezogen auf Arbeitsalltag, stabile Psyche)
Einfühlungsvermögen
Liebt seinen Beruf
Ethos=Leitlinie von allem
Empathie (dreimal genannt)
Psych. Stabilität und auf sich achten (Leben, Aktivitäten, Freunde/Partner)

Zufriedenheit mit der eigene Verordnungssituation

Die interviewten Ärzt*innen sind auf Grund des Zwiespaltes einerseits ihre Patient*innen nach bestem Wissen und Gewissen versorgen zu wollen, andererseits mit ihrem Verordnungsverhalten nicht auffällig zu werden, mit ihrer Verordnungssituation überwiegend mäßig zufrieden.

Tabelle: Zufriedenheit mit der eigenen Verordnungssituation

Zufriedenheit	Häufigkeit
Unzufrieden	5
Mäßig bis relativ zufrieden	12
Zufrieden	9

Fachgruppenverteilung

Die Zielverfehlen-Betriebsstätten verteilen sich nicht gleichmäßig auf die verschiedenen Fachgruppen, sondern es gibt Fachgruppen, die einen besonders hohen oder überdurchschnittlich hohen Anteil an Zielverfehlen-Betriebsstätten aufweisen und Fachgruppen, die anteilig weniger Zielverfehlen-Betriebsstätten aufweisen als der Durchschnitt.

Fachgruppe	Anzahl BS (ZE+ZV)	Anzahl BS (ZE)	Anteil ZE	Anzahl BS (ZV)	Anteil ZV
MKG-Chirurgen	48	19	39,6%	29	60,4%
Orthopäden	635	525	82,7%	110	17,3%
Chirurgen	479	408	85,2%	71	14,8%
fachärztliche Internisten mit SP Gastroenterologie	192	166	86,5%	26	13,5%
FA Physikalische und Rehabilitative Medizin	123	107	87,0%	16	13,0%
Anästhesisten	129	115	89,1%	14	10,9%
fachärztliche Internisten ohne SP	347	315	90,8%	32	9,2%
fachärztliche Internisten mit SP Endokrinologie	28	26	92,9%	2	7,1%
Augenärzte	502	468	93,2%	34	6,8%
fachärztliche Internisten mit SP Kardiologie	272	256	94,1%	17	6,3%
Kinder- und Jugendpsychiater	113	107	94,7%	6	5,3%
fachärztliche Internisten mit SP Pneumologie	139	132	95,0%	7	5,0%
Nervenärzte/Neurologen/Psychiater	611	585	95,7%	26	4,3%
fachärztliche Internisten mit SP Angiologie	27	26	96,3%	1	3,7%
Psychotherap. tätige Ärzte	231	223	96,5%	8	3,5%
Urologen	290	280	96,6%	10	3,4%
Neurochirurgen	117	113	96,6%	4	3,4%
fachärztliche Internisten mit SP Nephrologie	161	155	96,3%	5	3,1%
Kinder- und Jugendärzte	676	657	97,2%	19	2,8%
Frauenärzte	1212	1195	98,6%	17	1,4%
Hautärzte	392	387	98,7%	5	1,3%
Allgemeinärzte/hausärztlich tätige Internisten	6241	6192	99,2%	56	0,9%
fachärztliche Internisten mit SP Hämatologie/Onkologie	117	116	99,1%	1	0,9%
HNO-Ärzte	444	442	99,5%	2	0,5%
fachärztliche Internisten mit SP Rheumatologie	62	62	100,0%	0	0,0%
Gesamt	12486	12104	96,9%	499	4,0%

In Dunkelblau hinterlegt sind die fünf Fachgruppen mit dem höchsten Anteil an Zielverfehlern. Alle Fachgruppen, die hellerblau hinterlegt sind, liegen oberhalb des Zielverfehlenanteils in der Grundgesamtheit der Ärzte bzw. Betriebsstätten, der bei 4,0 Prozent liegt. Diese Fachgruppen weisen somit einen überdurchschnittlich hohen Anteil an Zielverfehlen-Betriebsstätten auf,

während die weiß hinterlegten Betriebsstätten einen unterdurchschnittlich hohen Anteil an Zielverfehlen-Betriebsstätten zeigen.

Da in einer Betriebsstätte mehrere Fachgruppen vertreten sein können, die jedoch alle unter der gleichen Betriebsstättennummer laufen, gehen manche Betriebsstättennummern mehrfach in die Auswertung ein und führen zu kleinen Abweichungen hinsichtlich der errechneten Zahlen an Betriebsstätten auf der Gesamtebene.

Vergleich ZE-ZV hinsichtlich Charakteristika mit potenziellem Einfluss auf Auffälligkeit

Zielerreicher und Zielverfehlen unterscheiden sich hinsichtlich verschiedener Charakteristika. Auf Arzteebene zeigt der hohe Anteil an männlichen Ärzten und der hohe Anteil an Ärzt*innen ohne Zusatzbezeichnungen in der Gruppe der Zielverfehlen einen Hinweis auf den Einfluss auf die Auffälligkeit.

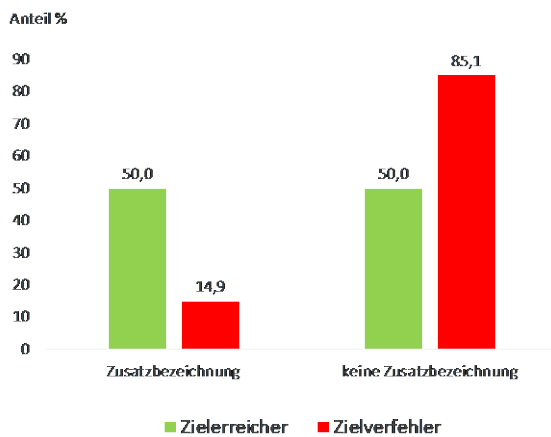


Abbildung: Vergleich ZE-ZV hinsichtlich der geführten Zusatzbezeichnungen

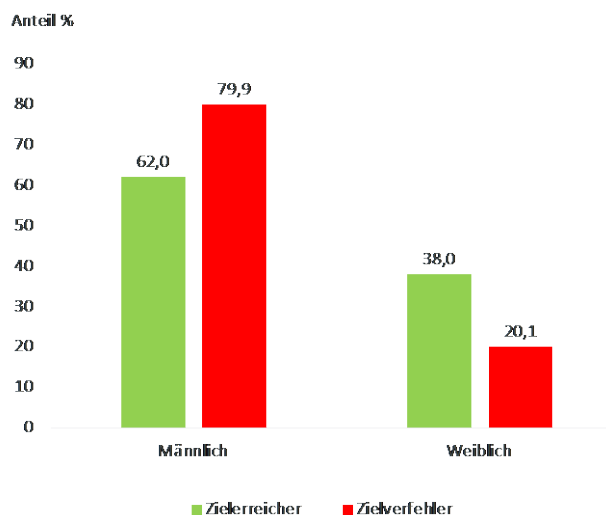


Abbildung: Vergleich ZE-ZV hinsichtlich des Geschlechtes

Auf Praxisebene zeigen die städtische Praxislage sowie kleine Betriebsstätten, gemessen an der Anzahl abgerechneter Fälle pro Quartal einen Einfluss auf die Auffälligkeit.



Abbildung: Vergleich ZE-ZV hinsichtlich der Praxislage

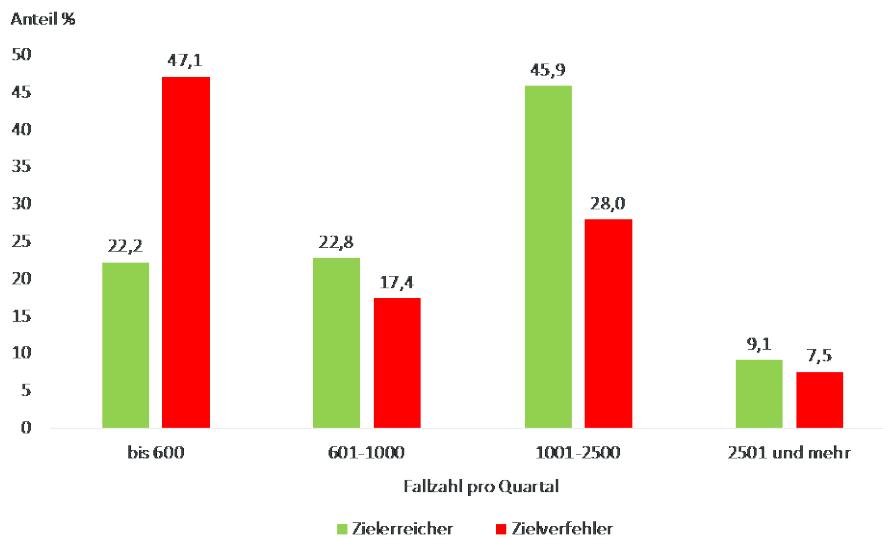


Abbildung: Vergleich ZE-ZV hinsichtlich der Praxisgröße, abgebildet über die abgerechnete Fallzahl pro Quartal

Ebene mit potenziellen Einflussfaktoren auf die Auffälligkeit

Auf vier verschiedenen Ebenen werden Faktoren identifiziert, die einen Einfluss auf die Auffälligkeit haben können.

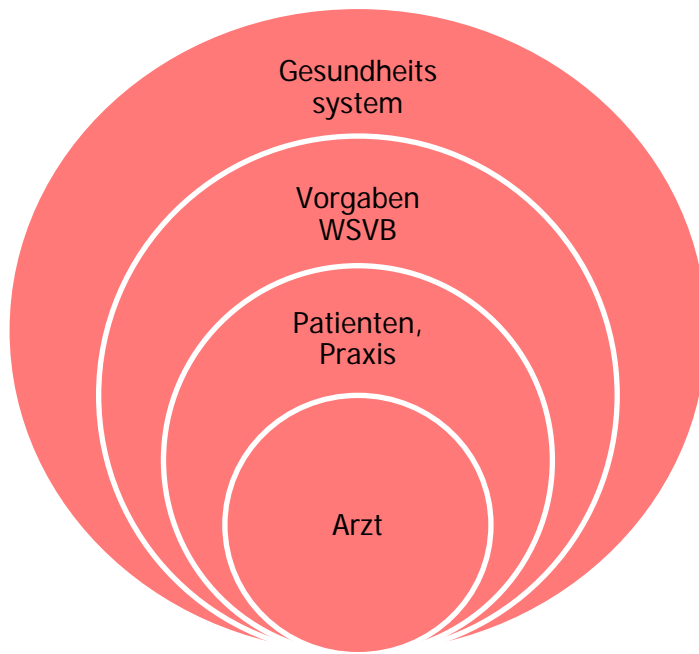


Abbildung: Ebenen mit potenziellen Einflussfaktoren auf die Auffälligkeit

4.7 Ziele und Zielerreichung der WSV

Welche Einzelziele werden von der Fachgruppe nicht erreicht?

Einzelziele, die deutlich verfehlt wurden (Zielerreichung <90%) sind insbesondere Einzelziele, die nur ein geringes Verordnungsvolumen aufweisen und bezogen auf die Indikation eher Arzneimittel darstellen, die nicht das typische Verordnungsspektrum der jeweiligen Fachgruppe darstellen.

Tabelle: Einzelziele, bei den die Zielvorgabe von der gesamten Fachgruppe deutlich verfehlt wurde (<90% der vorgegebenen Zielwerte).

Fachgruppe	Nicht erreichte Einzelziele	Art des Ziels	DDD-Menge	Zielerreichung
Orthopäden	Antiparkinsonmittel	Generikaziel	557,67	64,66%
	DOAKs	Leitsubstanzziel	16.704,32	55,78%
	Ophthalmologika	Generikaziel	2.726,75	85,76%
Gastroenterologie	MS-Therapeutika	Leitsubstanzziel	366,29	34,79%
	Andere Antianämika	Leitsubstanzziel	802,46	36,29%
	Antikoagulantien	Leitsubstanzziel	32.796,84	69,69%
Internisten ohne SP	MS-Therapeutika	Leitsubstanzziel	645,7	42,76%
	Antikoagulantien	Leitsubstanzziel	341.732,96	68,93%
	Rhinologika	Generikaziel	11.893,26	78,67%
Kardiologie	Antikoagulantien	Leitsubstanzziel	728.220,06	58,89%
	Rhinologika	Generikaziel	3.213,34	87,65%
	Orale Antidiabetika	Generikaziel	227.441,22	87,68%
Pneumologie	Antikoagulantien	Leitsubstanzziel	32.005,96	35,71%
	Ophthalmika	Generikaziel	3.443,79	61,61%
	Antiparkinsonmittel	Generikaziel	2.092,67%	81,85%
	Antiepileptika	Generikaziel	2.879,33	89,93%
Nervenärzte/	Antikoagulantien	Leitsubstanzziel	7.990,90	16,58%

Neurologen/ Psychiater	Urologika	Generikaziel	51.366,27	85,14%
	DOAKs	Leitsubstanzziel	7.288,90	89,54%
Kinder- und Jugendärzte	Koloniestimulierende Faktoren	Leitsubstanzziel	151,3	27,97%
	TNF-alpha-Blocker	Leitsubstanzziel	22.932,42	35,81%
	MS-Therapeutika	Leitsubstanzziel	323,96	39,34%
	Psycholeptika	Generikaziel	52.183,11	62,62%
	Andere Antianämika	Leitsubstanzziel	3.126,44	66,89%
	Antiparkinsonmittel	Generikaziel	2.084,84	78,17%
	Lipidsenker	Generikaziel	8.141,90	79,35%
	Kombigruppe kardiovaskuläres System	Generikaziel	64.311,94	84,07%
Hausärzte	Antikoagulantien	Leitsubstanzziel	22.223.847,10	71,54%
	Andere Antianämika	Leitsubstanzziel	56.880,63	74,72%

Welche Ziele sind die verordnungsmäßig umfangreichsten Ziele?

Bezogen auf das Verordnungsvolumen stellen die Einzelziele mit dem höchsten Verordnungsvolumen pro Fachgruppe jeweils fachgruppentypische Indikationsgebiete dar. Die Einzelziele werden überwiegend erreicht und nur wenige Einzelziele werden knapp verfehlt (Zielerreichung >90%).

In der folgenden Tabelle sind die jeweils drei Ziele mit dem höchsten Verordnungsvolumen pro Fachgruppe aufgeführt und nach der Art des Ziels, dem jeweiligen Verordnungsvolumen dem erreichten Zielwert dargestellt.

Tabelle: Drei Ziele mit dem jeweils höchsten Verordnungsvolumen pro Fachgruppe

Fachgruppe	Ziel mit hohem Verordnungsumfang	Art des Ziels	DDD-Menge	Ziel-erreichung
Orthopäden	Antiphlogistika	Generikaziel	3.673.056,94	107,91%
	Mittel zur Behandlung von Knochenerkrankungen	Generikaziel	1.864.718,78	103,38%
	Analgetika (ohne BTM)	Generikaziel	634.758,24	101,72%

Gastro- enterologie	PPI	Mengenziel	1.320.355,50	111,82%
	TNF-alpha-Blocker	Leitsubstanz- ziel	508.012,62	146,99%
	Mittel mit Wirkung auf das RAS-System	Generikaziel	366.334,52	100,85%
Internisten ohne SP	Mittel mit Wirkung auf das RAS-System	Generikaziel	3.738.192,80	100,42%
	Kombigruppe kardiovas- kuläres System	Generikaziel	2.885.445,07	99,91%
	PPI	Mengenziel	2.148.875,00	106,59%
Kardiologie	Mittel mit Wirkung auf das RAS-System	Generikaziel	3.478.049,40	101,91%
	Kombigruppe kardiovas- kuläres System	Generikaziel	2.460.292,26	101,05%
	Lipidsenker	Generikaziel	1.654.913,32	94,80%
Pneumologie	Mittel bei obstruktiven Atemwegserkrankungen	Generikaziel	10.045.618,00	103,45%
	Systemische Corticoide	Generikaziel	614.152,03	103,51%
	PPI	Mengenziel	254.028,00	120,10%
Nervenärzte/ Neurologen/ Psychiater	Psychoanaleptika (ohne ADHS, ohne Antidem- tativa)	Generikaziel	24.755.047,47	101,23%
	Psycholeptika	Generikaziel	7.072.299,75	99,46%
	Antiepileptika	Generikaziel	5.319.697,57	96,82%
Kinder- und Jugendärzte	Antiphlogistika	Generikaziel	2.098.457,85	101,97%
	Mittel bei obstruktiven Atemwegserkrankungen	Generikaziel	2.061.194,71	99,61%
	Ophthalmologika	Generikaziel	1.222.314,17	121,37%
Hausärzte	Mittel mit Wirkung auf das RAS-System	Generikaziel	290.176.258,3 9	100,69%
	Kombigruppe kardiovas- kuläres System	Generikaziel	198.480.232,0 4	100,46%
	PPI	Mengenziel	104.936.679,0 0	106,70%

Welche TOP3 Arzneimittel stehen einer Zielerreichung jeweils entgegen?

Arzneimittel, die einer Zielerreichung entgegenstehen sind besonders Originalpräparate, Mee-too-Präparate und Generika, die preislich oberhalb des Festbetrags liegen. Ebenso stellt Lithium ein auffälliges Arzneimittel dar, das der Zielerreichung entgegensteht.

Tabelle: Arzneimittel, die einer Zielerreichung am meisten entgegenstehen

TOP 1 Arzneimittel	DDD-Menge	TOP 2 Arzneimittel	DDD-Menge	TOP 3 Arzneimittel	DDD-Menge
Internisten ohne Schwerpunkt					
Mittel mit Wirkung auf das RAS-System (3.738.192,80 DDD)					
Exforge® HCT	21.770,00	Preterax®/ Bipreterax®	8.890,00	Dafiro® HCT	8.218,00
Kombigruppe kardiovaskuläres System (2.885.445,07 DDD)					
Nebivolol STADA	8.680,00	Nepresol®	8.456,621	Carmen®	7.600,00
Kardiologie					
Lipidsenker (1.654.913,32 DDD)					
Ezetrol®	115.970,00	Inegy®	42.750,00	Atozet®	27.440,00
Neurologie					
Psycholeptika (7.072.299,75 DDD)					
Quilonum®	354.153,602	Fluanxol	148.831,497	Tavor®	86.836,00
Antiepileptika (5.319.697,57 DDD)					
Orfiril®	141.923,323	Ergenyl®	123.383,735	Briviact	38.674,20
Kinder- und Jugendmedizin					
Mittel bei obstruktiven Atemwegserkrankungen (2.061.194,71 DDD)					
Flutide®	82.400,00	Viani®	79.650,00	Atrovent®	36.218,708

Wie viele relevante Ziele haben die jeweiligen Fachgruppen?

Die Anzahl relevanter Einzelziele variiert zwischen den verschiedenen Fachgruppen stark. So gibt es Fachgruppen mit nur zwei relevanten Einzelzielen wie beispielsweise in der Augenheilkunde oder 23 Einzelziele für Hausärzte. Relevante Einzelziele sind Ziele, die fachgruppentypische Arzneimittel umfassen und somit den Handlungsalltag der Fachgruppe abbilden.

Tabelle: Untersuchte Fachgruppen mit relevanten Einzelzielen

Fachgruppe	Anzahl Einzelziele
Augenärzte	2
Psychotherapeutisch tätige Ärzte	3
Hautärzte	4
Kinder- und Jugendpsychiater	5
Gynäkologen	5
Internisten SP Pneumologen	7
Orthopäden	7
Internisten SP Endokrinologie	7
Internisten SP Kardiologie	7
Internisten SP Rheumatologie	7
Urologie	7
Internisten SP Gastroenterologen	8
Pädiatrie	8
Neurologen	10
Internisten SP Onkologie	11
Hausärzte	23
Anästhesie	
Chirurgie	
MKG-Chirurgen	
HNO-Ärzte	
Internisten ohne SP	keine Daten vorliegend
Internisten SP Nephrologie	
Internisten SP Angiologie	
Neurochirurgie	
Fachärzte für physikalische und rehabilitative Medizin	

Um welche Art von Zielen handelt es sich jeweils?

Der Anteil von Leitsubstanzzielen ist bei den nicht erreichten Einzelzielen deutlich höher. Im Gegensatz dazu stellen die Einzelziele mit dem hohen Verordnungsvolumen überwiegend Generikaziele dar. Details sind den vorangehenden Tabellen xxx [hier bitte die Tabellen-Nr auflisten, in denen Generika/Leitsubstanzen erwähnt werden] zu entnehmen.

Welche Auffälligkeiten ergeben sich jeweils?

Verschiedene Facharztgruppen haben unterschiedliche Anzahlen an relevanten Einzelzielen, die ihre typischen Verordnungsalltag widerspiegeln. Gleichzeitig stellen die Einzelziele, bei den weniger als 90 Prozent der geforderten Zielwerte erreicht wurden, Ziele dar, die nicht den jeweiligen fachgruppentypischen Verordnungsrahmen abbilden und auch nur einen geringen Verordnungsumfang umfassen. Die Einzelziele, die hingegen ein sehr hohes Verordnungsvolumen im typischen Verordnungsalltag umfassen, werden entweder erreicht oder nur knapp verfehlt, so dass davon auszugehen ist, dass die bayerische Wirkstoffvereinbarung den jeweiligen Verordnungsalltag der Fachgruppen relativ gut abbildet. Während einige Arzneimittel nur mit einem geringen Verordnungsvolumen der Zielerreichung entgegenstehen, fallen wenige Arzneimittel wie beispielsweise Ezetrol® deutlich mehr ins Gewicht. Prolia® und Humira® stellen weitere Arzneimittel mit einem hohen Verordnungsvolumen dar, die einer wirtschaftlichen Verordnung im Rahmen der bayerischen Wirkstoffvereinbarung entgegenstehen.

Das Einzelziel der Antikoagulantien wurde von fast allen betrachteten Fachgruppen nicht erreicht.

Anhang 5: Teilprojekt E - Ergebnistabellen und Abbildungen

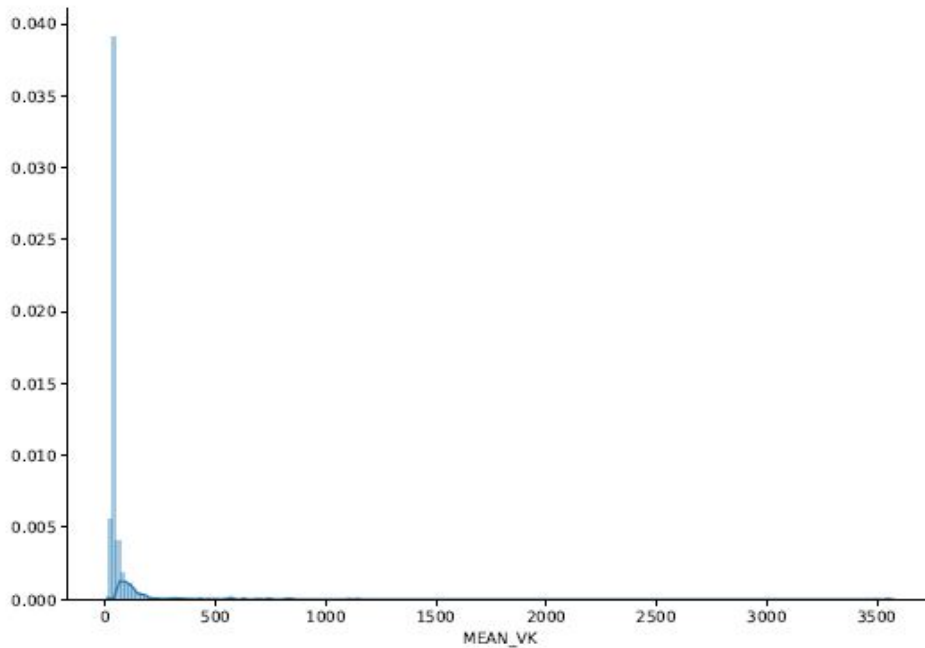


Abbildung: Univariate Verteilung (hellblaues Histogramm) der Kosten je Verordnung (mean_vk) im Quartal 4/2016 und der dazugehörige Kernaldichteschätzer (dunkelblaue stetige Funktion) – dabei ist auf der y-Achse die Wahrscheinlichkeitsdichte aufgetragen

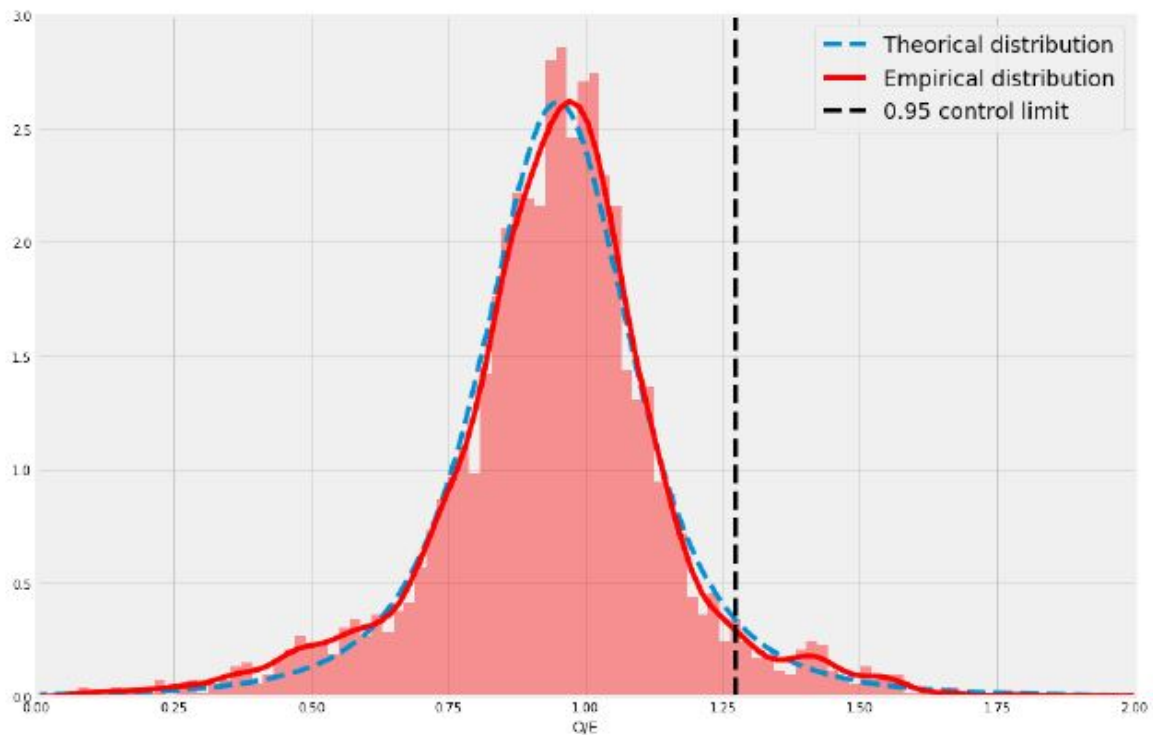


Abbildung: O/E-Indikator der Kosten je Verordnung im Quartal 4/2016. Rot abgetragen ist die empirische Verteilung (empirical distribution), die mittels Kerndichteschätzung (kernel density estimation) geschätzt wurde – blau abgetragen ist die theoretische Verteilung (theoretical distribution). Zudem ist schwarz abgetragen die 95%-Kontrollgrenze (control limit), die mit der Bayesschen Inferenz erzeugt wurde.



Abbildung: SHAP Values der vorhergesagten Verordnungskosten (mittels Machine-Learning-Modell) einer betrachteten Betriebsstätte

Tabelle: Tabelle mit Mittelwerten aller Einflussvariablen des Indikators Kosten je Verordnung im Quartal 4/2016 und die beobachteten Werte und SHAP Values des dazugehörigen Regression Forests der betrachteten Betriebsstätte - inkl. eines Rankings der Einflussnahme der Variablen jeweils in Klammern ()

Einflussvariable	Mittelwert (mean value)	Beobachteter Wert (observed value)	SHAP Value
Verordnungstiefe	0,6526	0,6837	-13,139 (1)
Anteil der Patient*innen mit Diagnoseanzahl über dem 90%-Perzentil	0,197	0,171	-7,062 (3)
Anteil der Polypharmaziepatienten	0,1622	0,1486	-1,390 (7)
Patientenanzahl je Ärzt*in	590	609	-1,564 (6)
Anteil von männlichen Patienten	0,418	0,342	3,286 (5)
Durchschnittsalter der Patient*innen	60,36	59,31	1,169 (9)
Anteil der Patient*innen, die in einem ländlichen Gebiet wohnen	0,155	0,0	1,270 (8)
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	0,491	0,002	3,564 (4)
Neurologisch-psychiatrische Praxis-Kennung	0,036	0	-0,180 (10)
Kardiologische Praxis-Kennung	0,096	1	8,236 (2)

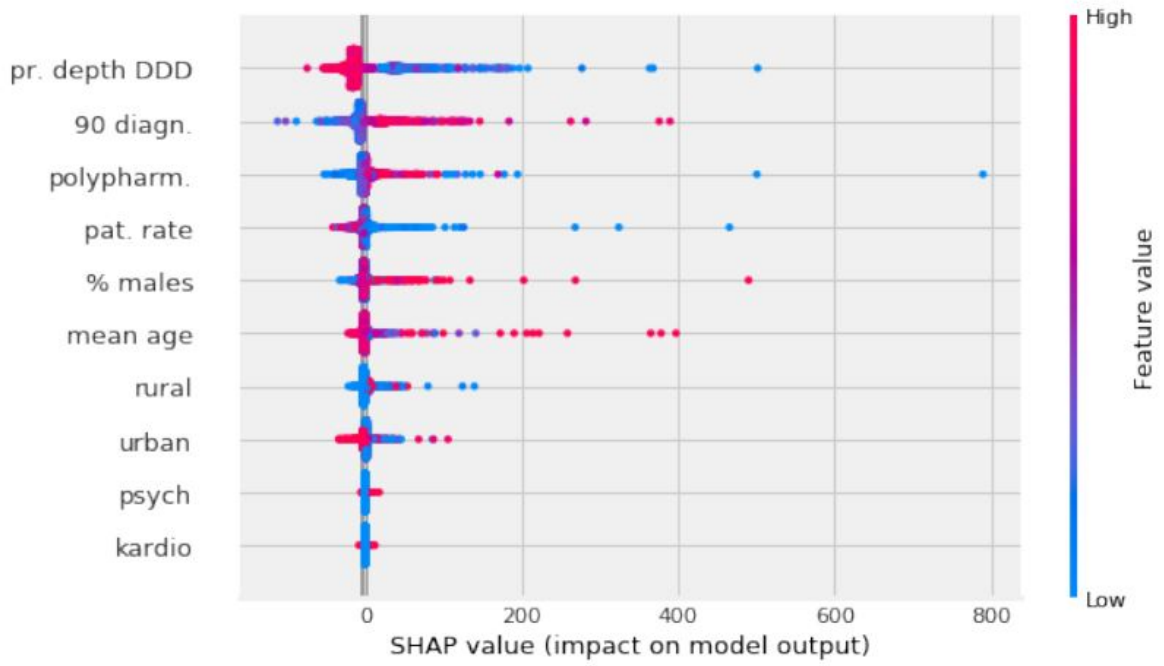


Abbildung: SHAP summary plot der Kosten je Verordnung

Verordnungsanteil von Generika am Gesamtmarkt

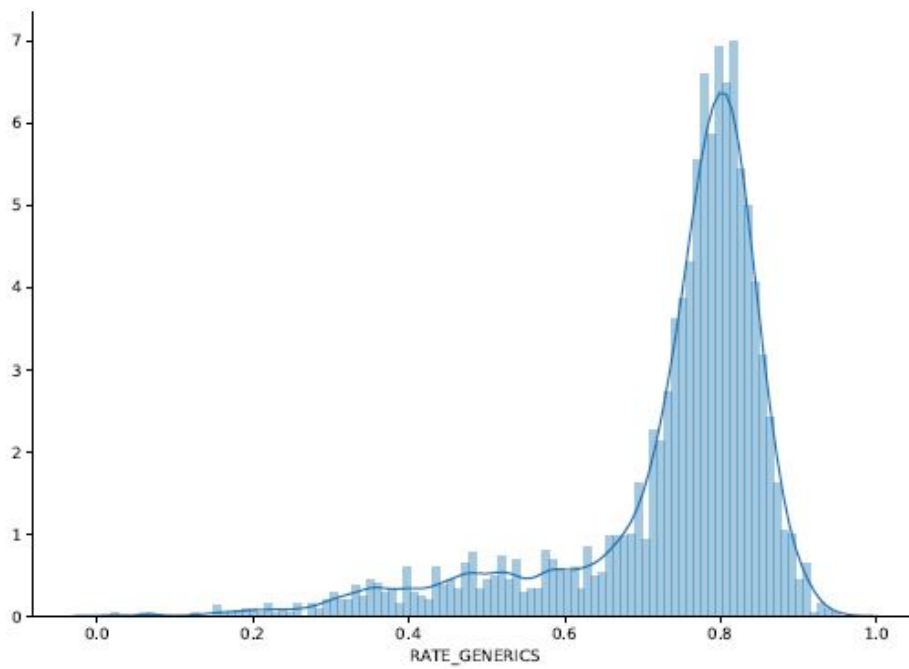


Abbildung: Univariate Verteilung (hellblaues Histogramm) des Verordnungsanteils von Generika am Gesamtmarkt (rate generics) im Quartal 4/2016 und der dazugehörige Kernedichteschätzer (dunkelblaue stetige Funktion) – dabei ist auf der y-Achse die Wahrscheinlichkeitsdichte aufgetragen

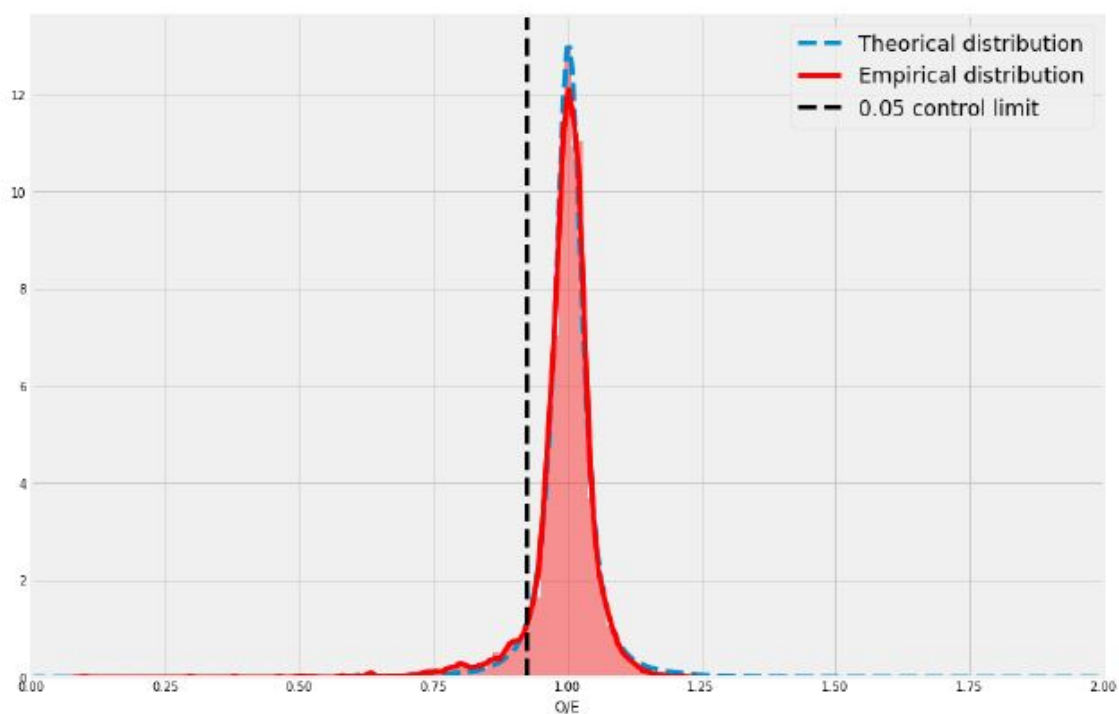


Abbildung: O/E-Indikator des Verordnungsanteils von Generika am Gesamtmarkt im Quartal 4/2016. Rot abgetragen ist die empirische Verteilung (empirical distribution), die mittels Kerndichteschätzung (kernel density estimation) geschätzt wurde – blau abgetragen ist die theoretische Verteilung (theoretical distribution). Zudem ist schwarz abgetragen die 95%-Kontrollgrenze (control limit), die mit der Bayesschen Inferenz erzeugt wurde.

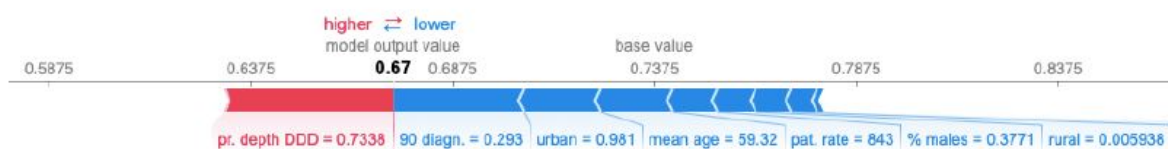


Abbildung: SHAP Values des vorhergesagten Verordnungsanteils von Generika am Gesamtmarkt (mittels Machine-Learning-Modell) einer betrachteten Betriebsstätte

Tabelle: Tabelle mit Mittelwerten aller Einflussvariablen des Indikators Verordnungsanteil von Generika am Gesamtmarkt im Quartal 4/2016 und die beobachteten Werte und SHAP Values des dazugehörigen Regression Forests der betrachteten Betriebsstätte - inkl. eines Rankings der Einflussnahme der Variablen jeweils in Klammern ()

Einflussvariable	Mittelwert (mean value)	Beobachteter Wert (observed value)	SHAP Value
Verordnungstiefe	0,6526	0,7338	0,0418 (1)
Anteil der Patient*innen mit Diagnoseanzahl über dem 90%-Perzentil	0,197	0,293	-0,0324 (2)
Anteil der Polypharmaziepatienten	0,1622	0,1791	-0,0064 (8)
Patientenanzahl je Ärzt*in	590	843	-0,0112 (5)
Anteil von männlichen Patienten	0,418	0,377	-0,0094 (6)
Durchschnittsalter der Patient*innen	60,36	59,31	-0,0179 (4)
Anteil der Patient*innen, die in einem ländlichen Gebiet wohnen	0,155	0,0059	-0,0088 (7)
Anteil der Patient*innen, die in einem städtischen Gebiet wohnen	0,491	0,981	-0,0189 (3)
Neurologisch-psychiatrische Mischpraxis-Kennung	0,036	0	-0,0016 (9)
Kardiologische Mischpraxis-Kennung	0,096	0	0,0001 (10)

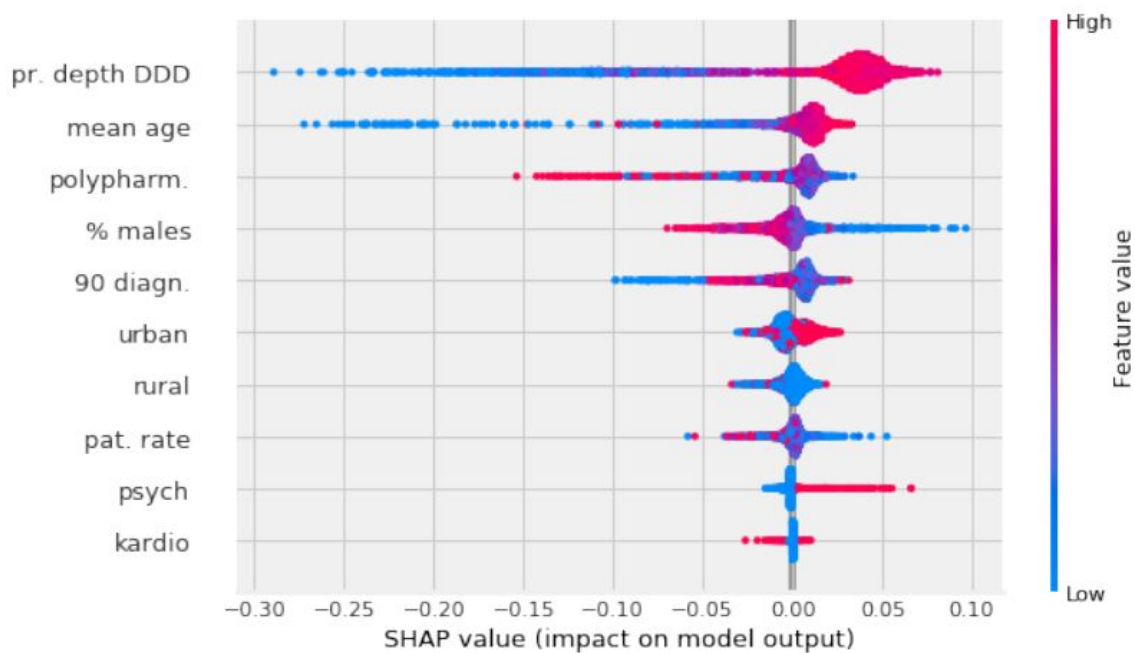


Abbildung: SHAP summary plot des Verordnungsanteils von Generika am Gesamtmarkt

Verordnete PPI-Mengen (DDD je Arzneimittelpatient)

Tabelle: Prozentuale Aufteilung aller Betriebstätten in die vier Auffälligkeitskategorien für den Indikator der durchschnittlich verordneten PPI-Mengen

Zeitraum	Anzahl Pra- xen gesamt	Anteil Kategorie 1	Anteil Kategorie 2	Anteil Kategorie 3	Anteil Kategorie 4
1/2016	6205	93,1%	3,0%	1,1%	2,8%
2/2016	6205	92,7%	3,1%	1,1%	3,0%
3/2016	6205	93,4%	2,9%	1,2%	2,5%
4/2016	6205	93,3%	3,0%	1,1%	2,6%
2016 gesamt	6205	88,3% (5478)	1,2% (73)	0,2% (10)	0,4% (22)

1-2016

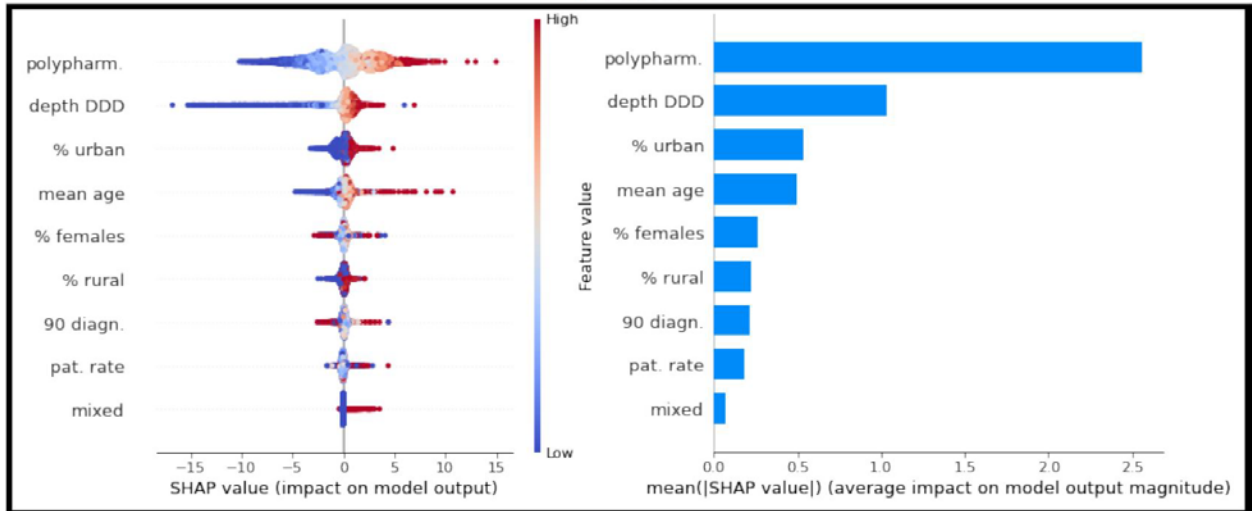


Abbildung: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 1/2016

2-2016

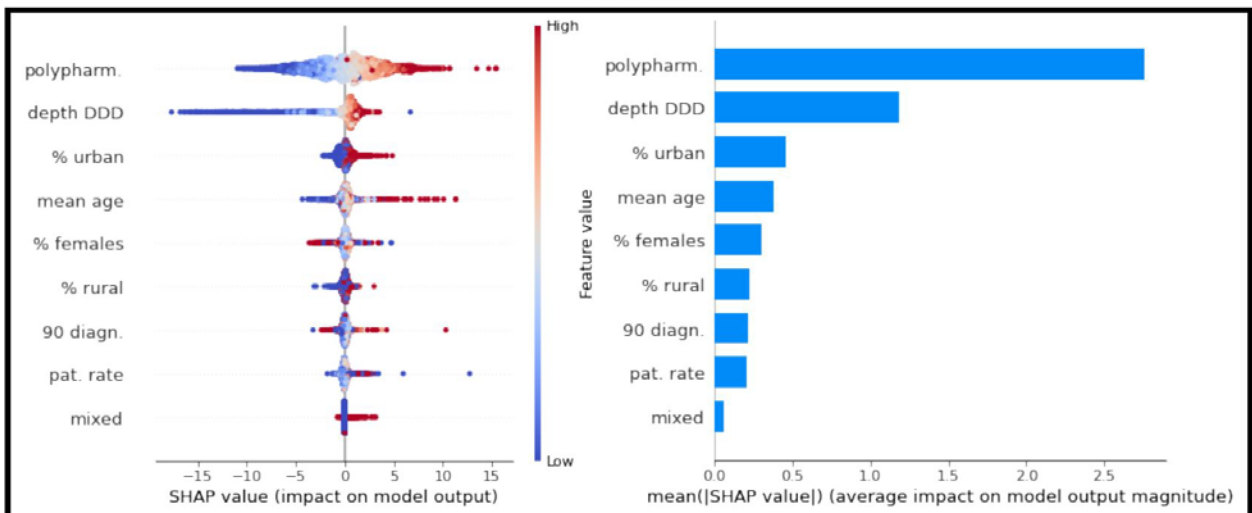


Abbildung: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 2/2016

3-2016

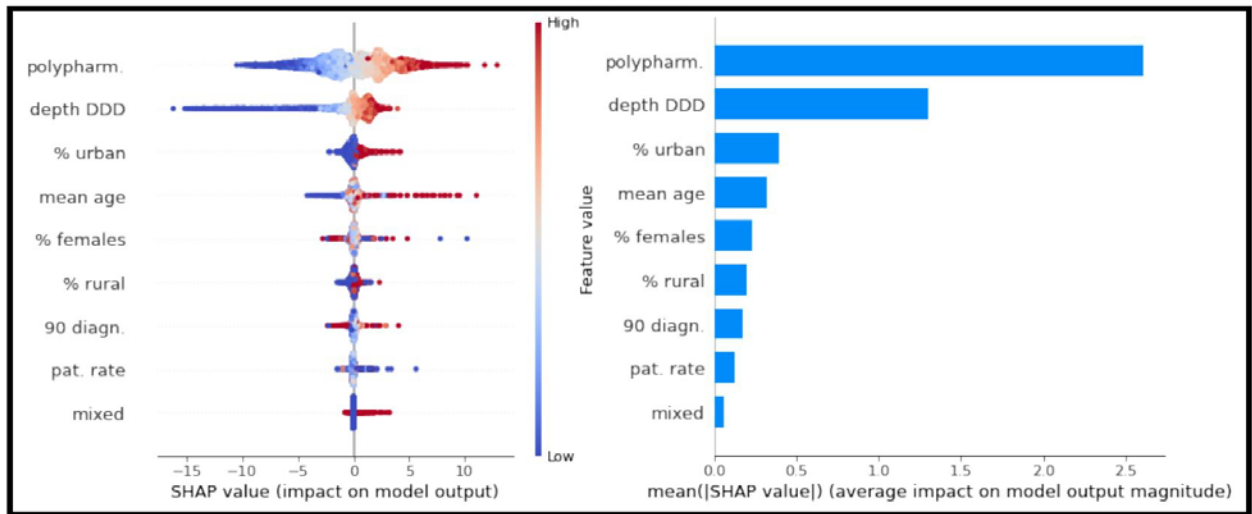


Abbildung: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 3/2016

4-2016

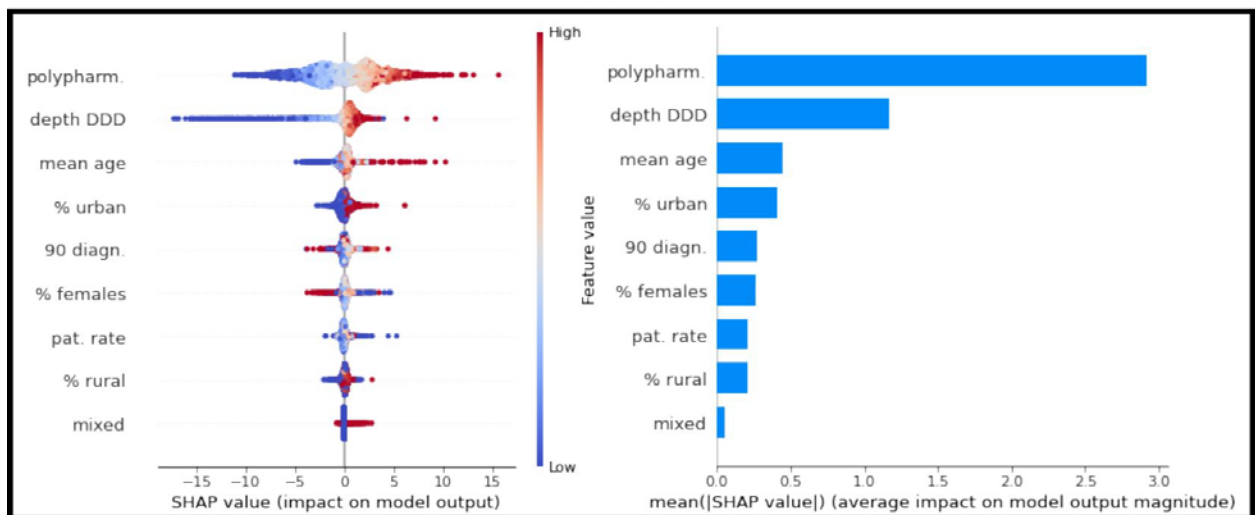


Abbildung: SHAP summary plot der durchschnittlich verordneten PPI-Mengen im Quartal 4/2016

Verordnungsanteil von DOAKs an allen Antikoagulantien (DDDs)

Tabelle: Prozentuale Aufteilung aller Betriebstätten in die vier Auffälligkeitskategorien für den Indikator des Verordnungsanteils von DOAK an allen Antikoagulantien

Zeitraum	Anzahl Praxen gesamt	Anteil Kategorie 1	Anteil Kategorie 2	Anteil Kategorie 3	Anteil Kategorie 4
1/2016	5878	94,7%	2,9%	0,5%	1,9%
2/2016	5878	94,7%	2,6%	0,6%	2,1%
3/2016	5878	94,7%	2,0%	0,8%	2,4%
4/2016	5878	94,6%	2,2%	0,8%	2,4%
2016 gesamt	5878	88,7% (5214)	0,4% (23)	0,1% (4)	0,1% (4)

1-2016

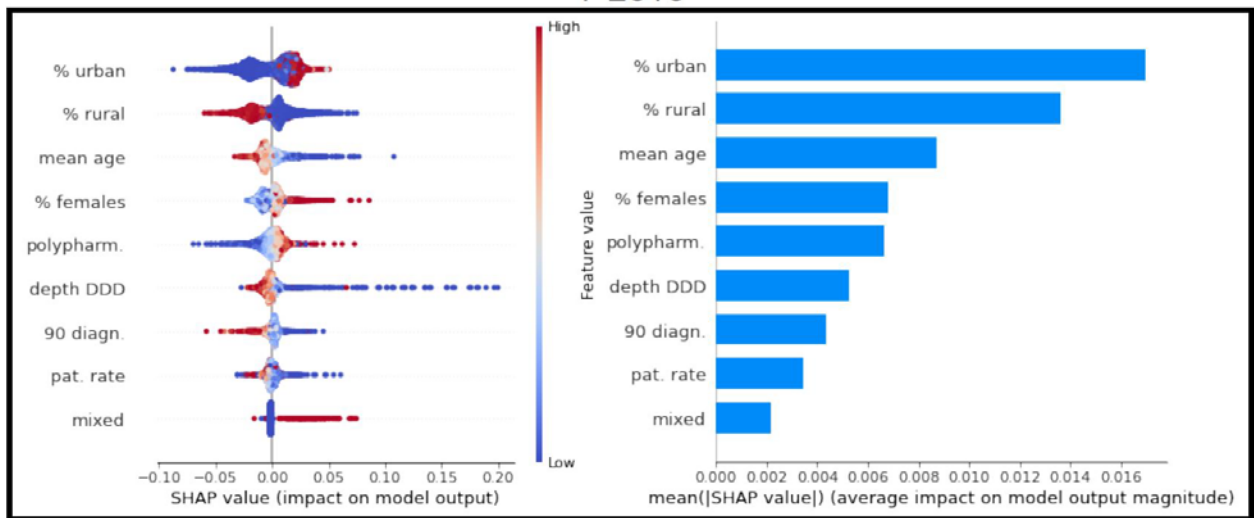


Abbildung: SHAP summary plot des Verordnungsanteils der DOAKs an allen Antikoagulantien im Quartal 1/2016

2-2016

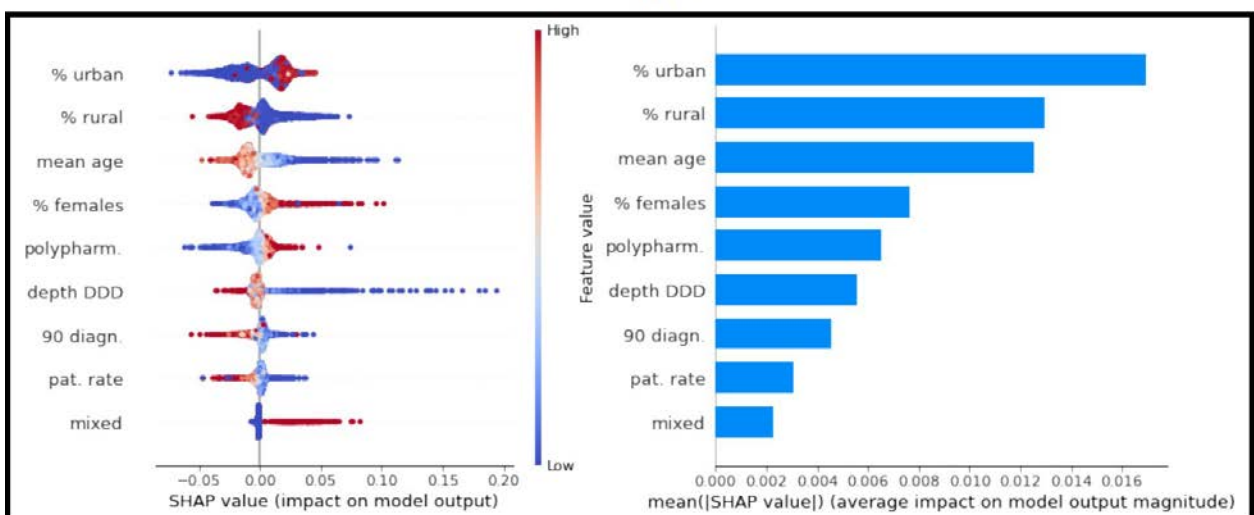


Abbildung: SHAP summary plot des Verordnungsanteils der DOAKs an allen Antikoagulantien im Quartal 2/2016

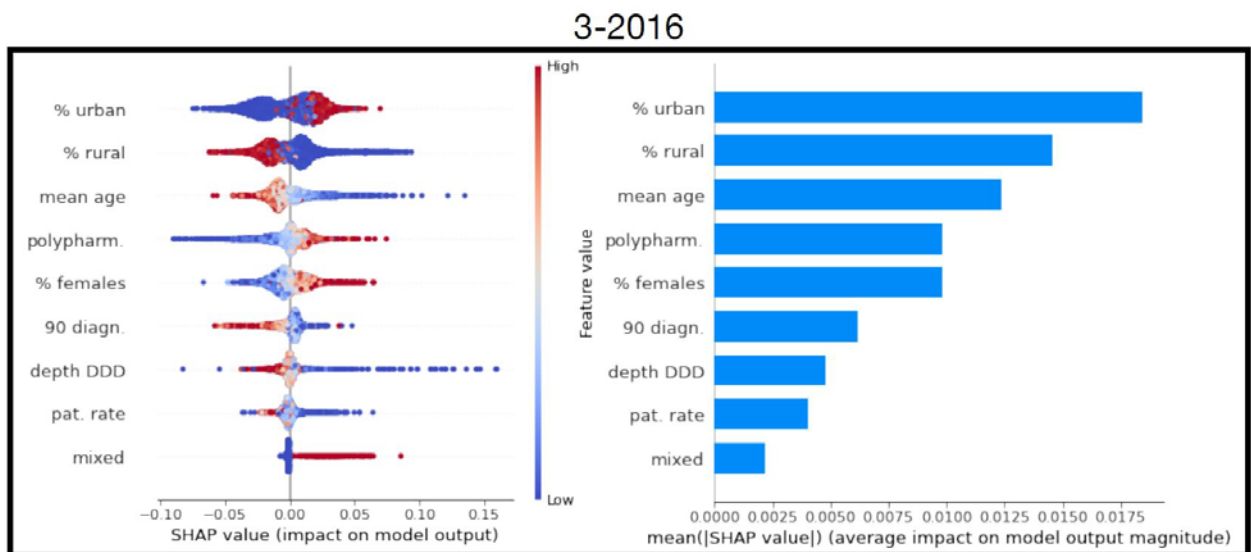


Abbildung: SHAP summary plot des Verordnungsanteils der DOAKs an allen Antikoagulantien im Quartal 3/2016

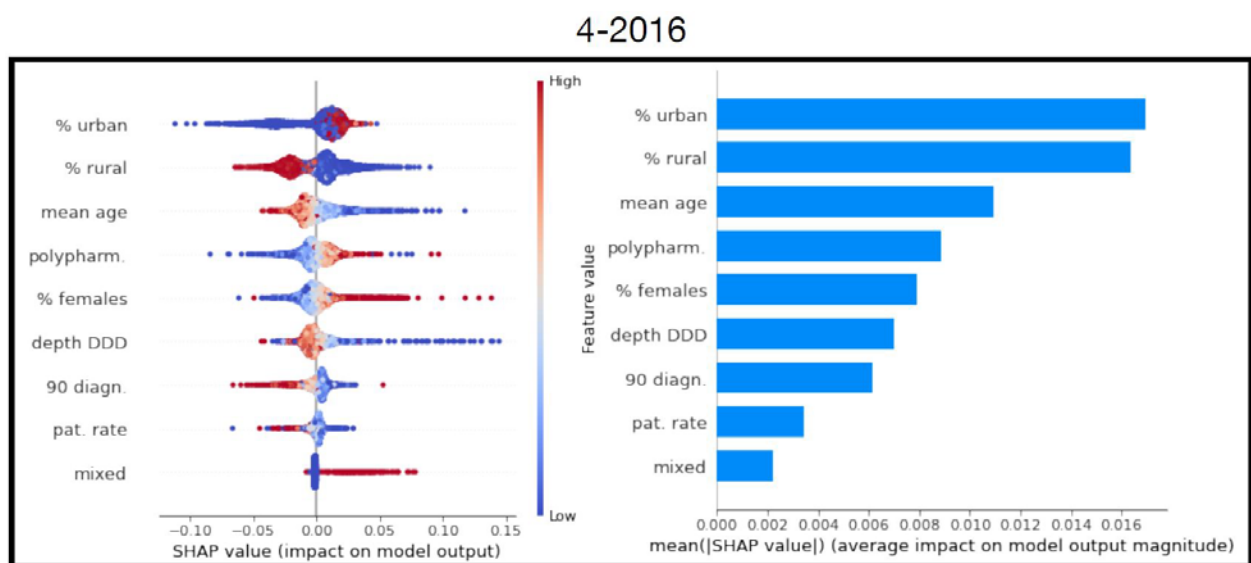


Abbildung: SHAP summary plot des Verordnungsanteils der DOAKs an allen Antikoagulantien im Quartal 4/2016

Anteil von generischen DDDs im Bereich der Lipidsenker

Tabelle: Prozentuale Aufteilung aller Betriebstätten in die vier Auffälligkeitskategorien für den Indikator des Ordnungsanteils von generischen DDDs im Bereich der Lipidsenker

Zeitraum	Anzahl Pra- xen gesamt	Anteil Kategorie 1	Anteil Kategorie 2	Anteil Kategorie 3	Anteil Kategorie 4
1/2016	6139	92,0%	1,7%	0,0%	6,3%
2/2016	6139	92,0%	1,4%	0,0%	6,5%
3/2016	6139	92,3%	1,4%	0,0%	6,3%
4/2016	6139	92,3%	1,5%	0,1%	6,1%
2016 gesamt	6139	85,4% (5240)	0,5% (30)	0,0% (0)	1,7% (79)

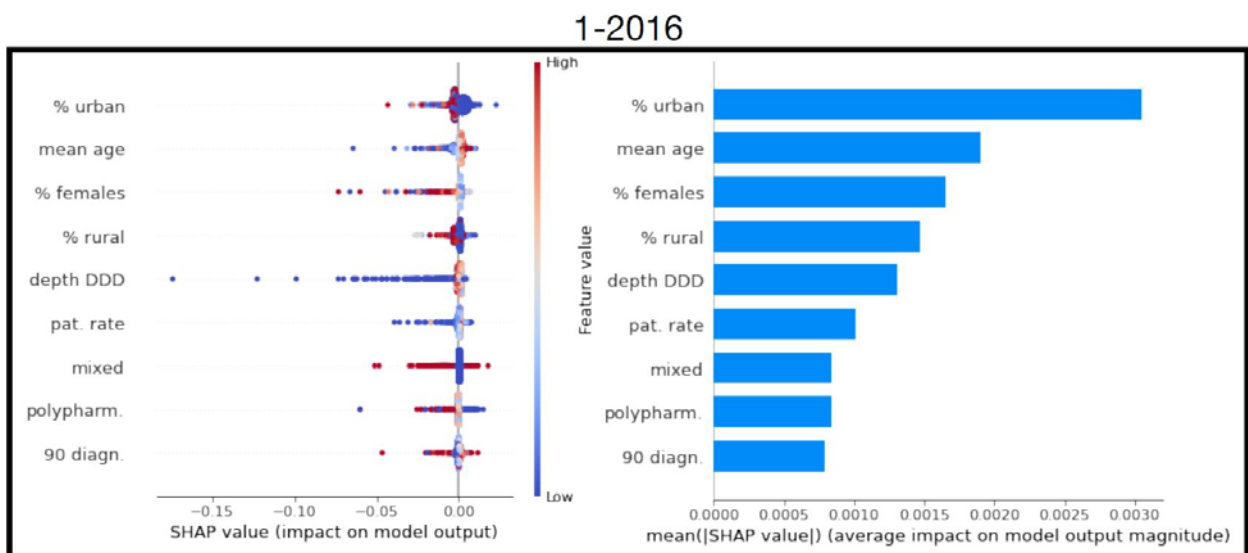


Abbildung: SHAP summary plot des Ordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 1/2016

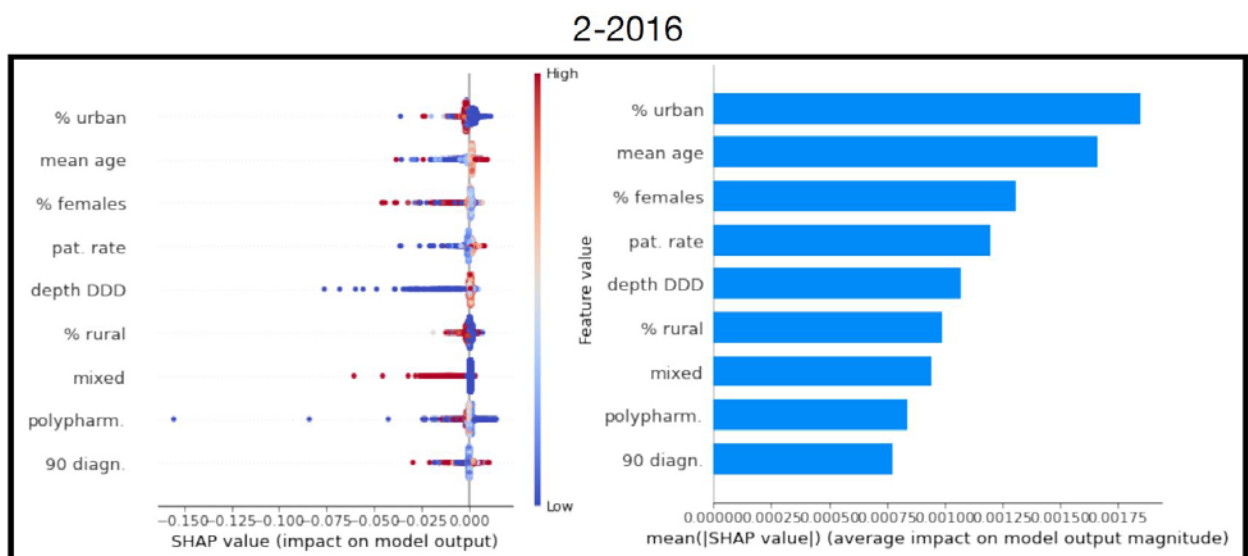


Abbildung: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 2/2016

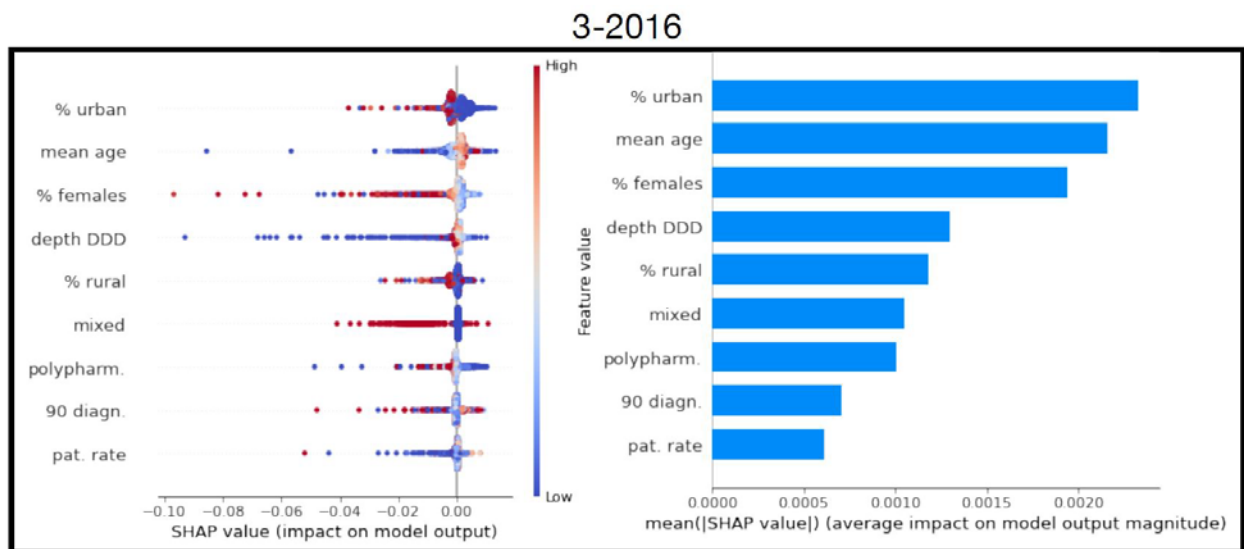


Abbildung: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 3/2016

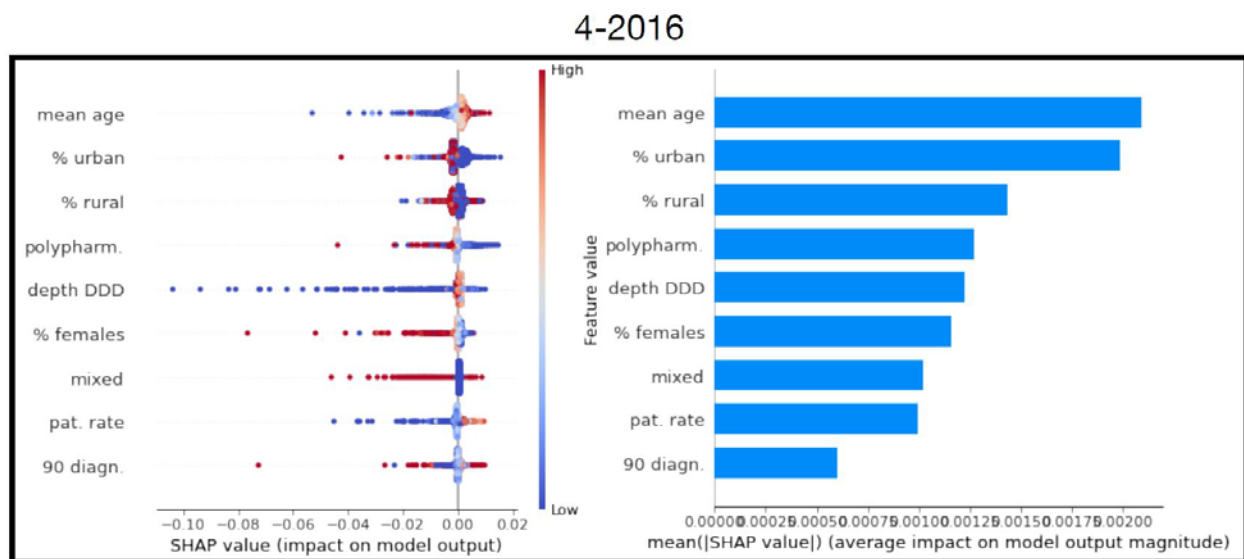


Abbildung: SHAP summary plot des Verordnungsanteils von generischen DDDs im Bereich der Lipidsenker im Quartal 4/2016

Anlage 1: Teilprojekt B – Ergebnistabellen der Hauptergebnisse

Durchschnittliche Kosten je Verordnung

Für die durchschnittlichen Kosten je Verordnung wurde das Zeitreihen-Regressionsmodell unter Einbeziehung der fünf beschriebenen Adjustierungs- und Differenzierungsvariablen

$$\text{Kosten}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Basierend auf dieser Schätzung wurde zunächst auf ein Vorliegen von Heteroskedastizität mithilfe des Breusch-Pagan-Tests getestet und anschließend der Chow Test auf Strukturbruch durchgeführt.

Es folgen die einzelnen Schätz- und Testergebnisse.

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	20,204929	0,4392**
Anteil der Patient*innen im Alter ab 70	-16,450000	-0,01286*
Anteil der weiblichen Patient*innen	89,114874***	0,5582**
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	48,604704*	0,0177**
Anteil der Polypharmazie-Patient*innen	17,477730	0,01116
Anzahl der Fälle je Quartal	-0,023142**	-0,02397***

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,838	0,9475
F-Statistik	28,93 bei 5 und 22 Freiheitsgraden	98,47 bei 5 und 22 Freiheitsgraden
p-Wert	5,633 * 10 ⁻⁹	2,661 * 10 ⁻¹⁴

Bayern:

	Teststatistik	p-Wert
Breusch-Pagan-Test	7,4142 bei 5 Freiheitsgraden	0,1916
Chow-Test	1,1143	0,397

Hessen:

	Teststatistik	p-Wert
Breusch-Pagan-Test	6,271 bei 5 Freiheitsgraden	0,2807
Chow-Test	0,22154	0,964

Durchschnittliche Kosten je Patient*in

Für die durchschnittlichen Kosten je Patient*in wurde analog das Zeitreihen-Regressionsmodell

$$\text{KostenPat}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Basierend auf dieser Schätzung wurde auch hier auf ein Vorliegen von Heteroskedastizität getestet und anschließend der Chow Test auf Strukturbruch durchgeführt.

Es folgen die einzelnen Schätz- und Testergebnisse der beiden KVen.

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	121,26150*	251,13263***
Anteil der Patient*innen im Alter ab 70	-189,99764.	-467,48184.
Anteil der weiblichen Patient*innen	215,61373**	137,52302.
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	209,46505*	548,08774***
Anteil der Polypharmazie-Patient*innen	52,49540	157,58247
Anzahl der Fälle je Quartal	-0,08138**	-0,12892***

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,7772	0,8943
F-Statistik	19,84 bei 5 und 22 Freiheitsgraden	46,71 bei 5 und 22 Freiheitsgraden
p-Wert	1,728 * 10 ⁻⁷	5,485 * 10 ⁻¹¹

Bayern:

	Teststatistik	p-Wert
Breusch-Pagan-Test	6,8122 bei 5 Freiheitsgraden	0,235
Chow-Test	0,97118	0,4753

Hessen:

	Teststatistik	p-Wert
Breusch-Pagan-Test	7,1472 bei 5 Freiheitsgraden	0,2099
Chow-Test	0,84761	0,5522

Verordnungsanteil von Generika am Gesamtmarkt

Für den durchschnittlichen Verordnungsanteil von Generika am Gesamtmarkt wurde das Zeitreihen-Regressionsmodell

$$Generika_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Für dieses Modell wird auf ein Vorliegen von Heteroskedastizität getestet und anschließend der Chow Test auf Strukturbruch durchgeführt.

Es folgen die Schätzergebnisse der beiden KVen und anschließend die einzelnen Testergebnisse.

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	0,1108	0,2839.
Anteil der Patient*innen im Alter ab 70	-0,4962*	1,295**
Anteil der weiblichen Patient*innen	0,8155***	7,209*
Anteil der Patient*inne mit einem Charlson-Index über dem 90%-Perzentil	-0,05580	8,254*
Anteil der Polypharmazie-Patient*innen	0,6940*	-2,334**
Anzahl der Fälle je Quartal	$1,346 * 10^{-4}$ *	$5,758 * 10^{-5}$

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,9441	0,8409

F-Statistik	92,23 bei 5 und 22 Freiheitsgraden	29,55 bei 5 und 22 Freiheitsgraden
p-Wert	$5,274 * 10^{-14}$	$4,619 * 10^{-9}$

Bayern:

	Teststatistik	p-Wert
Breusch-Pagan-Test	2,9689 bei 5 Freiheitsgraden	0,7048
Chow-Test	4,1874	0,01015

Hessen:

	Teststatistik	p-Wert
Breusch-Pagan-Test	11,466 bei 5 Freiheitsgraden	0,04289

Zur Überprüfung auf ein Vorliegen von Heteroskedastizität wurde der Breusch-Pagan-Test durchgeführt, der mit einer Teststatistik von BP=11,466 bei 5 Freiheitsgraden und einem p-Wert von 0,04289 die Nullhypothese der Homoskedastizität verwirft. Somit wird die Alternativhypothese der Heteroskedastizität bestätigt. Da hier die Voraussetzung des Chow-Tests nicht erfüllt ist, muss das Modell zunächst transformiert werden, um die für den Test vorausgesetzte Homoskedastizität zu erhalten, indem eine robuste Kovarianzmatrixschätzung durchgeführt wird. Das transformierte Modell hat die folgende Form:

	Koeffizientenschätzung
Intercept	0,43372
Anteil der Patient*innen im Alter ab 70	1,2897**
Anteil der weiblichen Patient*innen	0,61287
Anteil der Patienten mit einem Charlson-Index über dem 90%-Perzentil	0,75272
Anteil der Polypharmazie-Patient*innen	-2,0608*
Anzahl der Fälle je Quartal	$9,2669 * 10^{-6}$

Auf das transformierte Modell wurde der Chow-Test durchgeführt:

	Teststatistik	p-Wert
Chow-Test	1,7564	0,172

Patientenanteil mit DOAK-Verordnung an allen Patient*innen mit Antikoagulantienverordnung

Für den Patientenanteil mit DOAK-Verordnung an allen Antikoagulationspatienten wurde das Zeitreihen-Regressionsmodell

$$DOAK_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Basierend auf dieser Schätzung wurde auch hier auf ein Vorliegen von Heteroskedastizität getestet und anschließend der Chow Test auf Strukturbruch durchgeführt.

Es folgen nun die Schätzergebnisse der beiden KVen gefolgt von den jeweiligen Testergebnissen der beiden KVen. Die Parameterschätzung ergibt:

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	-3,6902166 ***	-3,1382893***
Anteil der Patient*innen im Alter ab 70	-5,7808405**	2,0067579
Anteil der weiblichen Patient*innen	8,2601024	5,1706422***
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	0,1947542***	10,0022220***
Anteil der Polypharmazie-Patient*innen	2,9279215	-9,0827056*
Anzahl der Fälle je Quartal	0,0003739	0,0004759

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,9431	0,9708
F-Statistik	90,57 bei 5 und 22 Freiheitsgraden	180,05 bei 5 und 22 Freiheitsgraden
p-Wert	6,376 * 10 ⁻¹⁴	< 2,2e-16

Bayern:

	Teststatistik	p-Wert
Breusch-Pagan-Test	4,1513 bei 5 Freiheitsgraden	0,5278
Chow-Test	5,0541	0,004387

Hessen:

	Teststatistik	p-Wert
Breusch-Pagan-Test	16,074 bei 5 Freiheitsgraden	0,006637

Zur Überprüfung auf ein Vorliegen von Heteroskedastizität wurde der Breusch-Pagan-Test durchgeführt, der mit einer Teststatistik von BP=16,074 bei 5 Freiheitsgraden und einem p-Wert von 0,006637 die Nullhypothese der Homoskedastizität verwirft. Somit wird die Alternativhypothese der Heteroskedastizität bestätigt. Da hier die Voraussetzung des Chow-Tests nicht erfüllt ist, muss das Modell zunächst transformiert werden, um die für den Test vorausgesetzte Homoskedastizität zu erhalten, indem eine robuste Kovarianzmatrixschätzung durchgeführt wird. Für das transformierte Modell ergibt sich die folgende Koeffizientenschätzung:

	Koeffizientenschätzung
Intercept	-3,13828929***
Anteil der Patient*innen im Alter ab 70	2,00675787
Anteil der weiblichen Patient*innen	5,17064216***
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	10,00222204***
Anteil der Polypharmazie-Patient*innen	-9,08270558*
Anzahl der Fälle je Quartal	0,00047592

Auf das transformierte Modell wurde der Chow-Test angewandt:

	Teststatistik	p-Wert
Chow-Test	0,19412	0,9739

Erreichung von Leitsubstanzquoten

Für die Erreichung der allgemeinen Leitsubstanzquoten (d.h. Anteil aller Leitsubstanzverordnungen über alle Leitsubstanzgruppen der eigenen KV hinweg) wurde das Zeitreihen-Regressionsmodell

$$\text{AnteilLeitsubstanzen}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Basierend auf dieser Schätzung wurde auch hier auf ein Vorliegen von Heteroskedastizität getestet und anschließend der Chow Test auf Strukturbruch durchgeführt.

Es folgen die Schätzergebnisse der beiden KVen gefolgt von den jeweiligen Testergebnissen. Die Parameterschätzung ergibt:

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	3,1669606.	0,3698**
Anteil der Patient*innen im Alter ab 70	-3,2463783	0,4544e
Anteil der weiblichen Patient*innen	-0,1553461	0,02849
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	12,2743642	-1,867*
Anteil der Polypharmazie-Patient*innen	-6,3863766	1,160
Anzahl der Fälle je Quartal	-0,0014100.	$2,358 * 10^{-4***}$

Mit den Gesamtregressionswerten:

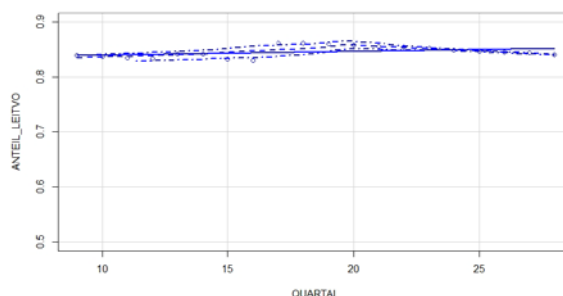
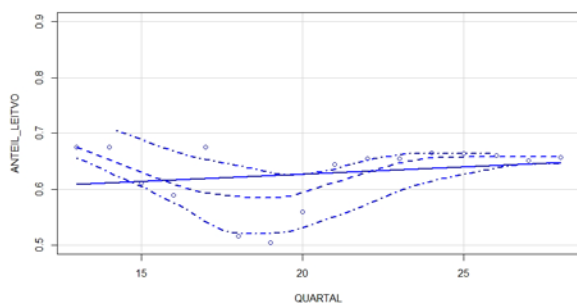
	Bayern	Hessen
Adjustiertes R ²	0,006738	0,6011
F-Statistik	1,02 bei 5 und 10 Freiheitsgraden	6,727 bei 5 und 14 Freiheitsgraden
p-Wert	0,4551	0,002161

Bayern:

	Teststatistik	p-Wert
Breusch-Pagan-Test	2,9689 bei 5 Freiheitsgraden	0,7048
Chow-Test	50,836	0,0009881

Hessen:

	Teststatistik	p-Wert
Breusch-Pagan-Test	1,7575 bei 5 Freiheitsgraden	0,8816
Chow-Test	5,3741	0,01659



Im direkten Vergleich von Bayern und Hessen liegt eine erheblich größere Schwankung in der KV Bayerns vor. In der KV Hessen hingegen findet die Bewegung in einem sehr schmalen Bereich statt.

Verordnungsanteil von generischen DDDs im Bereich Analgetika (außer BTM-pflichtige Opioide)

Für den Verordnungsanteil von generischen DDDs im Bereich Analgetika ohne BTM-pflichtige Opioide wurde das Zeitreihen-Regressionsmodell

$$\text{Anteil generischer DDDs}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

sowohl für die KV Bayerns als auch die KV Hessen geschätzt. Dieses Modell wird auf Heteroskedastizität getestet und anschließend wird der Chow Test auf Strukturbruch durchgeführt.

Es folgen zunächst die Schätzergebnisse der beiden KVen. Im Anschluss werden die Testergebnisse der einzelnen KVen präsentiert.

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	-0,5946	-0,7559
Anteil der Patient*innen im Alter ab 70	-6,9337**	-4,7609.
Anteil der weiblichen Patient*innen	6,5679***	6,0212***
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	0,5357	0,8798
Anteil der Polypharmazie-Patient*innen	1,9305	0,1481
Anzahl der Fälle je Quartal	-0,0005	-0,0003

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,7309	0,9082
F-Statistik	15,67 bei 5 und 22 Freiheitsgraden	54,39 bei 5 und 22 Freiheitsgraden
p-Wert	1,298 * 10 ⁻⁶	1,195 * 10 ⁻¹¹

Bayern:

	Teststatistik	p-Wert
Breusch-Pagan-Test	7,4137 bei 5 Freiheitsgraden	0,1916
Chow-Test	0,57067	0,748

Hessen:

	Teststatistik	p-Wert
Breusch-Pagan-Test	8,926 bei 5 Freiheitsgraden	0,1121
Chow-Test	0,59412	0,7308

Verordnungsanteil von generischen DDDs im Bereich der Mittel bei obstruktiven Atemwegserkrankungen

Für den Verordnungsanteil von generischen DDDs im Bereich der Mittel bei obstruktiven Atemwegserkrankungen wurde das Zeitreihen-Regressionsmodell

$$\text{Anteil generischer DDDs}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

geschätzt. Dieses Modell wird auf Heteroskedastizität getestet und anschließend wird der Chow Test auf Strukturbruch durchgeführt. Es folgen die einzelnen Schätz- und Testergebnisse der beiden KVen.

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	0,6835***	0,3540.
Anteil der Patient*innen im Alter ab 70	-0,5112*	0,3861
Anteil der weiblichen Patient*innen	-0,2264	-0,2311
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	0,2049	-0,8450.
Anteil der Polypharmazie-Patient*innen	-0,2283	1,301

Anzahl der Fälle je Quartal	$-1,321 * 10^{-5}$	$1,438 * 10^{-4}$.
-----------------------------	--------------------	---------------------

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,4098	0,6344
F-Statistik	4,749 bei 5 und 22 Freiheitsgraden	10,37 bei 5 und 22 Freiheitsgraden
p-Wert	0,004297	$3,281 * 10^{-5}$

Bayern:

	Teststatistik	p-Wert
Breusch-Pagan-Test	4,8534 bei 5 Freiheitsgraden	0,434
Chow-Test	5,6671	0,002537

Hessen:

	Teststatistik	p-Wert
Breusch-Pagan-Test	6,994 bei 5 Freiheitsgraden	0,2211
Chow-Test	4,7782	0,00568

Verordnungsanteil von generischen DDDs im Bereich der Mittel mit Wirkung auf das Renin-Angiotensin-System

Für den Verordnungsanteil von generischen DDDs im Bereich der Mittel mit Wirkung auf das Renin-Angiotensin-System wurde ebenfalls das Zeitreihen-Regressionsmodell

$$\text{Anteil generischer DDDs}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

geschätzt. Auch dieses Modell wird auf Heteroskedastizität getestet und es wird anschließend der Chow Test auf Vorliegen eines Strukturbruchs durchgeführt. Es folgen die Schätzergebnisse der beiden KVen gefolgt von den jeweiligen Testergebnissen. Die Parameterschätzung ergibt:

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	-0,7076524.	0,1584395
Anteil der Patient*innen im Alter ab 70	-2,1491029*	-0,6463397
Anteil der weiblichen Patient*innen	1,6324935*	0,0305873

Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	-0,2685773	3,2094736.
Anteil der Polypharmazie-Patient*innen	3,3621441**	0,0344748
Anzahl der Fälle je Quartal	0,0005104*	0,0003144

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,8939	0,8348
F-Statistik	46,48 bei 5 und 22 Freiheitsgraden	28,28 bei 5 und 22 Freiheitsgraden
p-Wert	5,763 * 10 ⁻¹¹	6,975 * 10 ⁻⁹

Bayern:

	Teststatistik	p-Wert
Breusch-Pagan-Test	5,0795 bei 5 Freiheitsgraden	0,4063
Chow-Test	3,7695	0,01566

Hessen:

	Teststatistik	p-Wert
Breusch-Pagan-Test	3,457 bei 5 Freiheitsgraden	0,6299
Chow-Test	10,684	7,715 * 10 ⁻⁵

In beiden KVen ist dabei ein Abflachen der Zuwächse nach dem angezeigten Strukturbruch durch die schon sehr hohen generischen Anteile im Bereich der Mittel mit Wirkung auf das Renin-Angiotensin-System zu verzeichnen (sowohl in der KV Bayerns als auch in der KV Hessen liegen die Anteile nach dem Strukturbruch bei ca. 95 %). Insgesamt ist somit ein weiteres Wachstum im generischen Bereich kaum mehr möglich.

Verordnungsanteil von generischen DDDs im Bereich der Thrombozytenaggregationshemmer (ohne Ticagrelor)

Für den Verordnungsanteil von generischen DDDs im Bereich der Mittel der Thrombozytenaggregationshemmer ohne Ticagrelor wurde das Zeitreihen-Regressionsmodell

$$\text{Anteil generischer DDDs}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

geschätzt. Nach Prüfung auf Heteroskedastizität wird anschließend der Chow Test auf Strukturbruch durchgeführt. Es folgen die Schätzergebnisse der beiden KVen gefolgt von den Testergebnissen:

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	0,4697**	0,8789395*
Anteil der Patient*innen im Alter ab 70	-0,1102	3,2188130.
Anteil der weiblichen Patient*innen	$7,219 * 10^{-2}$	-2,3322277***
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	1,113***	3,8771934**
Anteil der Polypharmazie-Patient*innen	$1,520 * 10^{-2}$	-3,1776266
Anzahl der Fälle je Quartal	$1,541 * 10^{-4}$.	0,0003754*

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,9638	0,8072
F-Statistik	145 bei 5 und 22 Freiheitsgraden	23,62 bei 5 und 22 Freiheitsgraden
p-Wert	$4,488 * 10^{-16}$	$3,662 * 10^{-8}$

Bayern:

	Teststatistik	p-Wert
Breusch-Pagan-Test	11,578 bei 5 Freiheitsgraden	0,04105

Der zur Überprüfung auf ein Vorliegen von Heteroskedastizität durchgeführte Breusch-Pagan-Test liefert eine Teststatistik von BP=11,578 bei 5 Freiheitsgraden und einem p-Wert von 0,04105. Dies bedeutet, dass bei einem Signifikanzniveau von 5% die Nullhypothese verworfen und die Alternativhypothese der Heteroskedastizität angenommen wird. Da hier die Voraussetzung des Chow-Tests nicht erfüllt ist, muss das Modell zunächst transformiert werden, um die für den Test vorausgesetzte Homoskedastizität zu erhalten, indem eine robuste Kovarianzmatrixschätzung durchgeführt wird. Für das transformierte Modell ergibt sich die folgende Koeffizientenschätzung:

	Koeffizientenschätzung
Intercept	0,46969321***
Anteil der Patient*innen im Alter ab 70	-0,11018968
Anteil der weiblichen Patient*innen	0,07219358
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	1,11289488***
Anteil der Polypharmazie-Patient*innen	0,01520035
Anzahl der Fälle je Quartal	0,00015405

Auf das transformierte Modell wurde der Chow-Test angewandt:

	Teststatistik	p-Wert
Chow-Test	0,74674	0,6207

Hessen:

	Teststatistik	p-Wert
Breusch-Pagan-Test	11,482 bei 5 Freiheitsgraden	0,04263

Der zur Überprüfung auf ein Vorliegen von Heteroskedastizität durchgeführte Breusch-Pagan-Test liefert eine Teststatistik von BP=11,482 bei 5 Freiheitsgraden und einem p-Wert von 0,04263. Aufgrund dessen wird die Nullhypothese bei einem Signifikanzniveau von 5% verworfen und die Alternativhypothese der Heteroskedastizität bestätigt. Da hier die Voraussetzung des Chow-Tests nicht erfüllt ist, wird das Modell nun zunächst transformiert, um die für den Test vorausgesetzte Homoskedastizität zu erhalten. Für das transformierte Modell ergibt sich die folgende Koeffizientenschätzung:

	Koeffizientenschätzung
Intercept	0,87893949*
Anteil der Patient*innen im Alter ab 70	3,21881297.
Anteil der weiblichen Patient*innen	-2,33222768**
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	3,87719339**
Anteil der Polypharmazie-Patient*innen	-3,17762659
Anzahl der Fälle je Quartal	0,00037536.

Auf das transformierte Modell wurde der Chow-Test angewandt:

	Teststatistik	p-Wert
Chow-Test	1,6728	0,1918

Verordnungsanteil von Leitsubstanz-DDDs innerhalb der Gruppe der MS-Therapeutika

Für den Verordnungsanteil von Leitsubstanz-DDDs innerhalb der Gruppe der MS-Therapeutika wurde das Zeitreihen-Regressionsmodell

$$\text{Anteil Leitsubstanz} - \text{DDDs}_t = \alpha_t + \beta_{1,t}x_{1,t} + \beta_{2,t}x_{2,t} + \beta_{3,t}x_{3,t} + \beta_{4,t}x_{4,t} + \beta_{5,t}x_{5,t} + \varepsilon_t$$

geschätzt. Dieses Modell wird zunächst auf Heteroskedastizität getestet, um anschließend den Chow Test auf Strukturbruch durchzuführen. Es folgen nun die Schätz- und Testergebnisse der beiden KVen.

	Koeffizientenschätzung Bayern	Koeffizientenschätzung Hessen
Intercept	0,6146*	0,9046*
Anteil der Patient*innen im Alter ab 70	0,3752	-0,06849
Anteil der weiblichen Patient*innen	-0,5915	-0,2712
Anteil der Patient*innen mit einem Charlson-Index über dem 90%-Perzentil	-0,6965	-4,237***
Anteil der Polypharmazie-Patient*innen	0,4749	2,101
Anzahl der Fälle je Quartal	$3,868 * 10^{-5}$	$-7,316 * 10^{-5}$

Mit den Gesamtregressionswerten:

	Bayern	Hessen
Adjustiertes R ²	0,7227	0,9194
F-Statistik	15,07 bei 5 und 22 Freiheitsgraden	62,63 bei 5 und 22 Freiheitsgraden
p-Wert	$1,784 * 10^{-6}$	$2,863 * 10^{-12}$

Bayern:

	Teststatistik	p-Wert
Breusch-Pagan-Test	10,017 bei 5 Freiheitsgraden	0,07476
Chow-Test	4,549	0,007082

Hessen:

	Teststatistik	p-Wert
Breusch-Pagan-Test	7,9228 bei 5 Freiheitsgraden	0,1605
Chow-Test	5,9755	0,001951

Anlage 2: Toolbox

Die folgende Tabelle gibt einen Überblick über Struktur der geplanten Toolbox. Es folgen weitere Dokumente, in welchen die einzelnen Themen bearbeitet werden (siehe rechte Spalte). Diese sind lediglich als Orientierung zu verstehen, da sie beständig weiterentwickelt werden müssen. Es ist geplant, die Toolbox nur registrierten Nutzern zugänglich zu machen.

1. Einstieg

	Wirtschaftlichkeit und Qualität der Medikamenten-Verordnung: Instrumente zur Steuerung („Toolbox“)	künftig Verlinkte Dokumente
Grundlagen	<p>Die Ausgaben der GKV für Arzneimittel betragen mehr als 40 Mrd € pro Jahr. Einerseits helfen Arzneimittel, Beschwerden zu lindern, Krankheiten zu heilen und Prognosen zu verbessern; andererseits können unbedachte Verordnungen das System der solidarischen Krankenversicherung ungerechtfertigt belasten.</p> <p>Die Verordnung von Arzneimitteln unterliegt deshalb dem Wirtschaftlichkeitsgebot des SGB V. Krankenkassen und kassenärztliche Vereinigungen haben die gesetzliche Aufgabe, die Wirtschaftlichkeit von Arzneimittel-Verordnungen zu prüfen und zu fördern (§106 SGB V).</p> <p>Diese Toolbox stellt den zuständigen Körperschaften Instrumente zur Verfügung, Wirtschaftlichkeit und Qualität von Arzneimittel-Verordnungen zu bestimmen und zu verbessern. Sie geht auf das Projekt WirtMed zurück, dass 2018 bis 2021 vom G-BA gefördert wurde.</p>	→ Grundlagen
Indikatoren und Störfaktoren	<p>Die Wirtschaftlichkeit und Qualität von Verordnungen kann nicht direkt „gemessen“ werden. Vielmehr müssen Indikatoren für die Auswertung von Verordnungs-Daten definiert werden, die valide Rückschlüsse erlauben und eine Steuerung des Verordnungs-Verhaltens ermöglichen.</p> <p>Indikatoren dienen als Zielgrößen der Prüfung <i>individuellen</i> Verhaltens, beispielsweise um einzelnen Betriebsstätten Feedback über ihre Verordnungen zu geben (Profiling); sie können aber auch genutzt werden, um innovative Steuerungs- und Prüf-Systeme zu evaluieren oder übergreifende Einflüsse auf des Verordnungsgeschehen zu erfassen. Störfaktoren (confounder) können Vergleiche verzerren. Bei individuellem Profiling und der Evaluation von System-Interventionen kann die Berücksichtigung von Störfaktoren helfen, einen Vergleich aussagekräftiger zu machen.</p>	→ Indikatoren
Evaluation von Strategien: Quantitative Analyse von	Die Umsetzung der gesetzlichen Aufgabe, die Wirtschaftlichkeit von Arzneimittel-Verordnungen zu prüfen und zu fördern (§106 SGB V), wird von den kassenärztlichen Vereinigungen mithilfe des Konzepts der Wirtschaftlichkeitsprüfung geleistet. In den dazugehörigen KV-spezifischen Prüfverordnungen wird festgelegt,	→ Strukturbruch

Verordnungsdaten	<p>welches Ordnungsverhalten hinsichtlich definierter Aspekte als auffällig betrachtet wird und in welchen Fällen dies eine Beratung oder Prüfung nach sich ziehen kann.</p> <p>Während es bei der Prüfung nach §106 SGB V um das Ordnungsverhalten einzelner Betriebsstätten geht, kann das hier dargestellte Instrumentarium auch für die Evaluation von Prüf- und Steuer-Strategien genutzt werden. So kann innerhalb einer KV nach dem Wechsel der Prüfordnung ein Vorher-Nachher-Vergleich der Ordnungen von Interesse sein; oder zwei KVen mit unterschiedlichen Prüfordnungen sollen parallel miteinander verglichen werden. Da es sich bei den gemessenen Größen um Zeitreihen handelt (meist Quartale), stellt das Strukturbruchmodell nach Chow eine angemessene statistische Methodik dar.</p>	
Evaluation von Strategien: qualitative Erhebungen	Für eine Evaluation von Prüf- und Steuer-Strategien (z.B. Vergleich von KVen mit verschiedenen Prüfordnungen; Vorher-Nachher-Vergleiche) sind jedoch nicht nur die Ordnungsdaten von Interesse. Zu einer umfassenden Evaluation gehört auch eine qualitative Befragung der niedergelassenen Ärztinnen in Einzelinterviews und/oder Fokusgruppen. Dabei kann die Einschätzung der Ärztinnen zur Sprache kommen, die Auswirkungen im Alltag, aber auch die Barrieren, Probleme und unerwünschte, vielleicht auch unerwartete Effekte.	→ Evaluation qualitativ
Zielverfehlen	Bei der Prüfung der Wirtschaftlichkeit und der Qualität von Medikamenten-Ordnungen werden einzelne Ärzte oder Betriebsstätten auffällig, weil sie Ziele verfehlen oder außerhalb der festgelegten Kontrollgrenzen liegen. Für prüfende Instanzen ist es von Interesse, mehr über diese Gruppe zu erfahren. Damit wird zum einen das Prüfverfahren auf seine Validität untersucht: wenn unter diesen Betriebsstätten ein hoher Anteil mit legitimen Praxisbesonderheiten zu finden ist, kann dies ein wertvolles Feedback für eine Optimierung des Prüfverfahrens sein. Zum anderen können auf der Basis solcher Analysen maßgeschneiderte Interventionen mit dem Ziel besserer Ordnungen entwickelt werden (z.B. spezifische Information oder Unterstützung).	→ Zielverfehlen
Individuelles Profiling	International sind sehr differenzierte Instrumente zum sog. Profiling entwickelt worden. Darunter versteht man im Kontext medizinischer Versorgung die Identifizierung von Einheiten (Ärzten, Krankenhäusern, Praxen usw.), welche in nicht mehr akzeptablem Maße von einer Norm abweichen. Auch die herkömmliche Wirtschaftlichkeitsprüfung ist somit ein „Profiling“. Diese Aufgabe ist sehr anspruchsvoll, da beim Profiling die Charakteristika der versorgten Patienten (z.B. Morbidität, „case mix“), Eigenschaften der Leistungserbringer und Zufallsschwankungen zu berücksichtigen sind.	→ Profiling basic; Profiling advanced

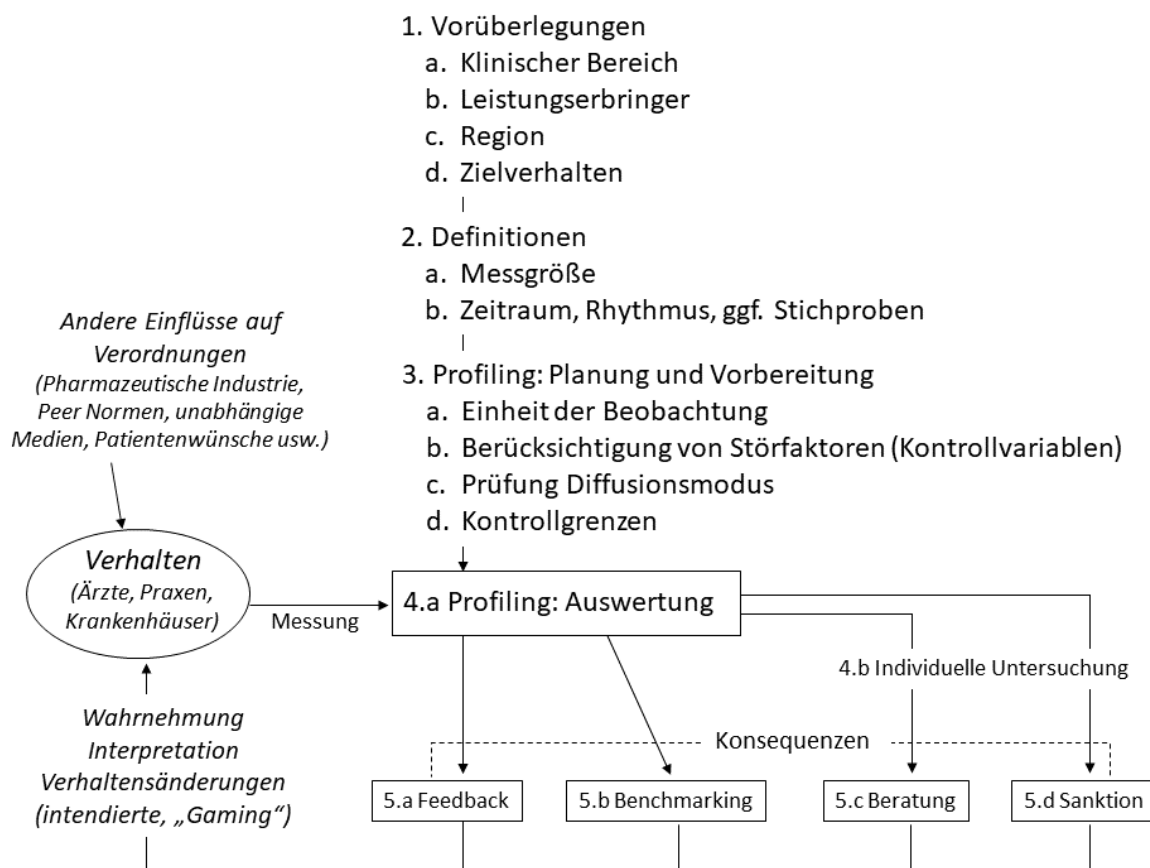
<p>Diffusionsanalysen</p>	<p>Ergebnisse der Diffusions- und Implementierungs-Forschung zeigen, dass die Verbreitung von Innovationen durch verschiedene Faktoren beeinflusst wird. Dabei spielen Vernetzungen der ÄrztInnen eine wesentliche Rolle. Um einen validen Eindruck vom Verordnungsgeschehen zu erhalten und mögliche Interventionen zielgerecht umzusetzen, kann die zeitlich-(räumliche) Ausbreitung der Verordnung von Arzneimitteln analysiert werden. Dabei können sogenannte Verordnungsnetze von ÄrztInnen als identifiziert werden, die das Verordnungsgeschehen beeinflussen. Beziehungen zwischen den ÄrztInnen (Anzahl geteilter Patienten) sind hier von Interesse. Neben den verordnenden ÄrztInnen nehmen aber auch PatientInnen und ihre Pfade (z. B. Folgeverordnungen bei anderen ÄrztInnen oder Krankenhausaufenthalte) Einfluss auf die Verbreitung neuer Arzneimittel genauso wie regionale Faktoren.</p>	<p>→ Diffusion</p>
<p>Das WirtMed-Projekt</p>	<p>Philipps-Universität Marburg Abteilung für Allgemeinmedizin, Präventive und Rehabilitative Medizin 35043 Marburg [Konsortialführung Prof. Dr. med. Norbert Donner-Banzhoff, MHSc]</p> <p>Friedrich-Alexander-Universität Erlangen-Nürnberg Universitätsklinikum Erlangen Allgemeinmedizinisches Institut 91054 Erlangen</p> <p>Technische Universität München Lehrstuhl für Gesundheitsökonomie Fakultät für Sport-und Gesundheitswissenschaft 80992 München</p> <p>Zentralinstitut für die kassenärztliche Versorgung in der Bundesrepublik Deutschland 10587 Berlin</p> <p>Politecnico di Milano MOX - Modeling and Scientific Computing Dipartimento di Matematica "F. Brioschi"</p>	

	<p>Mailand, Italien</p> <p>Kassenärztliche Vereinigung Bayerns 80687 München</p> <p>Kassenärztliche Vereinigung Hessen 60486 Frankfurt</p> <p>AOK Bayern - Die Gesundheitskasse 81739 München</p> <p>AOK – Die Gesundheitskasse in Hessen 65760 Eschborn</p> <p>AOK Plus – Die Gesundheitskasse für Sachsen und Thüringen 01067 Dresden</p>	
--	---	--

2. Grundlagen - Prüfung und Steuerung von Wirtschaftlichkeit und Qualität der Arzneimittel-Verordnungen

Unter einem Profiling versteht man die Identifizierung von Leistungserbringern (Krankenhäuser, Abteilungen, Praxen, Ärzte/innen) mit „auffälliger“ Leistung im Vergleich zu anderen Einheiten oder zu einem Referenzstandard. Mit „auffällig“ können sowohl positive (bessere Ergebnisse) als auch negative Abweichungen (schlechtere Ergebnisse) gemeint sein. Ein Profiling kann in einem „formativen“ Sinne genutzt werden, d.h. Leistungserbringer können dadurch Schwachstellen identifizieren und Verbesserungen implementieren. Es kann aber auch eine „summative“ Funktion haben mit negativen administrativen Konsequenzen oder finanziellen Sanktionen. - Die in Deutschland etablierte Wirtschaftlichkeitsprüfung von Medikamenten-Verordnungen entspricht der o.g. Definition eines Profilings.

Eine Prüfung von Wirtschaftlichkeit oder Versorgungsqualität besteht aus den in der Abbildung dargestellten Komponenten, auch wenn die entsprechenden Annahmen und Entscheidungen häufig implizit gemacht bzw. getroffen werden.



Die Vorüberlegungen (1.a-d) sind keine fachlich-wissenschaftlichen i.e.S., sondern sind von pragmatischen und wirtschaftlichen Gesichtspunkten geleitet. Dabei geht es z.B. um die Beschränkung auf die Behandlung einer bestimmten Erkrankung (z.B. Diabetes mellitus Typ 2), eine Gruppe von Leistungserbringern (z.B. hausärztliche Versorgung), Regionen (z.B. Land / KV / Teile davon), und schließlich das interessierende Zielverhalten (z.B. Orientierung an Leitsubstanzen, Verordnungskosten, Versorgungsqualität, VO von problematischen Verordnungen).

Je nach Zielgröße ist die Verfügbarkeit von Daten eingeschränkt (z.B. Verordnungskosten bei rabattierten Arzneimitteln).

Im nächsten Schritt (2.a-b) sind Messgrößen (Indikatoren) zu definieren, z.B. Anteil von Leit-substanzen an Verordnungen, Verordnungskosten, Übereinstimmung mit Qualitätssystemati-ken (z.B. Medikationskatalog der KBV), Anteil von problematischen Wirkstoffen. Außerdem sind der Zeitraum der Analyse und der zeitliche Rhythmus festzulegen. Weiterhin ist zu ent-scheiden, ob eine Stichprobe (z.B. Zufallsauswahl oder einzelne Region) oder die Grundge-samtheit (gesamtes Land/KV-Bezirk) untersucht werden sollen.

Ein Profiling zur Analyse von Prozess und Outcome-Daten stellt hohe Ansprüche an das me-thodische Verständnis und die Umsetzung (einschließlich Software, Programmierung). Vor der eigentlichen Auswertung müssen methodische Entscheidungen getroffen werden (3.a-d): wel-ches soll die Einheit der Beobachtung sein, z.B. Betriebsstätte (BS) oder individuelle Ärzte, ggf. auch Patienten und ihre Charakteristika in einer Mehr-Ebenen-Analyse? Welche Störvariablen (Confounder) sind relevant und sollen bei der Auswertung berücksichtigt werden (z.B. Alter, Morbidität); dies entspricht einer Vorweg-Berücksichtigung von Praxisbesonderheiten, damit kann die Prüfung insgesamt effizienter gestaltet werden. Die Verordnungstiefe (Anteil der Ver-ordnungen oder DDDs, welche eine BS für ihre Patienten übernimmt) kann ebenfalls von Be-deutung sein. Um Rückmeldungen oder Interventionen effizient zu fokussieren, kann der Dif-fusionsmodus der Verordnungen eines neuen Medikaments untersucht werden (z.B. Verord-nungen initiiert von wenigen Zentren oder eher diffus mit multiplen Erst-Verordnern). Schließ-lich sind Kontrollgrenzen festzulegen, jenseits derer eine Einheit als „auffällig“ gilt.

Nach dem eigentlichen Profiling (4.a) sind abhängig von der Zielgenauigkeit des angewandten Verfahrens in unterschiedlichem Ausmaß „händische“ Einzelprüfungen der identifizierten Ein-heiten erforderlich (4.b - „Praxisbesonderheiten“). Mögliche Konsequenzen (5.a-d) aus den auf diesem Wege erhobenen Befunden bestehen in Feedback an die Verordner, Benchmar-king, Beratung und/oder Sanktionen. Ziel ist eine – möglichst präventive – Rückwirkung auf das Ordnungsverhalten.

Dabei ist realistisch zu bedenken, dass das Ordnungs-Verhalten vielfältigen Einflüssen un-terliegt. Auch die Feedback-Rückmeldungen aus einem hier skizzierten System der Steuerung können in verschiedener Weise verstanden und umgesetzt werden. Die Reaktionen auf das Feedback müssen nicht immer der Zielsetzung des Programms entsprechen; Gegen- und Aus-weichstrategien (sog. gaming) sind nie auszuschließen.

Je präziser und transparenter die Analysen und Rückmeldungen sind, desto eher können die betroffenen Ordner die Ergebnisse akzeptieren und künftig nutzen. Idealerweise sollten Prüfungen dieser Art weniger bedrohlich sein, sondern ein differenziertes und transparentes Steuerungssystem darstellen (entsprechend einem Regelkreis bzw. physiologischer Homöo-stase).

Die hier für ein individuelles Profiling dargestellten Überlegungen und notwendigen methodi-schen Entscheidungen gelten auch für Systemvergleiche. Bei diesen geht es nicht darum, ein-zelne Ordner zu identifizieren. Vielmehr soll ein System der Prüfung bzw. Steuerung evalu-iert werden. Die Einführung eines neuen Systems der Wirtschaftlichkeitsprüfung ist ein typi-scher Anlass für eine solche Evaluation. Als Vergleich kann das Ordnungsgeschehen im be-troffenen Land (bzw. andere geografische oder politische Einheit) im vorangehenden Zeit-raum dienen. Außerdem kommt ein paralleler Vergleich mit einem Land in Frage, in welchem

ein anderes bzw. herkömmliches System implementiert ist. Beide Vergleichs-Möglichkeiten schließen sich nicht gegenseitig aus.

3. Indikatoren und Störfaktoren

Indikatoren (Zielgröße)

Das Verhalten der Verordner von Medikamenten, deren Einstellungen, Motive oder Kompetenz können wir nie direkt beobachten oder messen. Wir gehen jedoch davon aus, dass es in Routine-Verordnungsdaten Hinweise auf diese Sachverhalte gibt. Diese können dazu dienen, die Verordnungen zu beschreiben, zu analysieren, zu bewerten und ggf. Interventionen zu begründen bzw. zu steuern.

Diesem Zweck dienen Indikatoren, d.h. in Routinedaten abgebildete Variablen, die wir als Hinweise auf die o.g. Sachverhalte verstehen. Diese wählen wir nach inhaltlichen und methodischen Gesichtspunkten aus. Wegen der Relevanz des Themas und der Schwere der Konsequenzen ist hier besondere Sorgfalt angebracht.

Indikatoren dienen als Zielgrößen der Prüfung *individueller* Verhaltens, beispielsweise um einzelnen Betriebsstätten Feedback über ihre Verordnungen zu geben (Profiling); sie können aber auch genutzt werden, um Steuerungs- und Prüf-Systeme zu evaluieren („Systemevaluation“, z.B. Wirkstoff-Vereinbarung der KV Bayerns) oder übergreifende Einflüsse auf des Verordnungsgeschehen zu erfassen.

Einige Indikatoren sind klar der Prüfung auf *Wirtschaftlichkeit* eines Verhaltens zuzuordnen, wie z.B. „Kosten pro VO“, die meisten haben jedoch auch bzw. überwiegend *Qualitäts*-Implikationen, wie z.B. VO bestimmter Wirkstoffe oder VO von Analogpräparaten. Die Relevanz dieser Zielsetzungen wird in der Tabelle (3. und 4. Spalte) entsprechend angegeben.

Soweit nicht anders angegeben, beziehen sich die Indikatoren auf das VO-Verhalten von Betriebsstätten. Die Analysen können die VOen auf dem Niveau der Betriebsstätte aggregieren, entsprechend auch Störfaktoren berücksichtigen; andererseits sind Mehrebenen-Analysen bzw. hierarchische Auswertungen möglich, welche auch die einzelnen Patienten und deren Charakteristika in der Auswertung (mit-)berücksichtigen.

Die hier vorgeschlagenen Indikatoren werden für die Auswertung von Routinedaten genutzt (z.B. entsprechend §300 Abs. 2 SGB V). Sie bilden den postulierten Sachverhalt niemals perfekt ab. So ist ein Teil der PPI-Verordnungen gerechtfertigt (z.B. Peptisches Ulkus, schwere Refluxsymptomatik); die VO eines DOAK mag gerechtfertigt sein wegen einer klaren und selbständigen Patientenpräferenz usw.

Die hier vorgeschlagenen Indikatoren stellen sämtlich Brüche (Anteile [engl. proportion], Quotienten [engl. ratio], Raten) dar. Im „Zähler“ wird das untersuchte Verschreibungsverhalten quantifiziert; dieses muss jedoch auf einen „Nenner“ bezogen werden, um interpretierbar zu sein. Dieser Nenner kann sämtliche in einer Betriebsstätte mit einer Medikamenten-VO behandelten Patienten umfassen, z.B. beim Anteil der Generika-VOen. Bei anderen Indikatoren ist ein engerer Bezug sinnvoll, z.B. bei Oxycodon bzw. Targin®: hier sind sämtliche Opioid-VO ein geeigneter Nenner. Indirekt geht damit auch die Betriebsstätte-Größe in die Analyse ein (z.B. große Betriebsstätte haben mehr Opioid-VOen).

Um unterschiedliche Wirksubstanzen innerhalb einer Indikationsgruppe miteinander vergleichen zu können, ist die Normierung mit Hilfe der „defined daily doses“ (DDD) entwickelt worden. Mit dieser Größe lassen sich etwa Verordnungen einer Wirkstoff-Gruppe (z.B. verschiedene PPI) übergreifend auswerten oder miteinander vergleichen.

In die Definition des Nenners ist immer die Zeit miteinzubeziehen, z.B. ein Quartal oder ein Jahr. Damit hängt zusammen, wie die behandelten Patienten berücksichtigt werden. Werden jeweils Quartale betrachtet, ist meist der Behandlungsfall die beobachtete Einheit. Dies bedeutet, dass die erfolgten Behandlungen zusammengefasst werden, auch wenn eine Patientin sie an verschiedenen Terminen im betrachteten Quartal erhält. Werden mehrere Quartale betrachtet, muss die fehlende Unabhängigkeit der Patienten bzw. Fälle berücksichtigt werden: ein Patient sucht die Praxis u.U. mehrere Male und in mehr als einem Quartal auf. Er trägt also mehrere Beobachtungen bei, die voneinander nicht mehr unabhängig sind.

Die von Betriebsstätte und Krankenhäusern kodierte Behandlungsdiagnosen sind zurückhaltend zu interpretieren. Hier ist von einer großen Variabilität in Bezug auf Diagnosestellung und Verschlüsselung auszugehen. Neben dem „non-differential“ Error ist auch ein Bias anzunehmen, da ICD-Verschlüsselungen (auch) der Rechtfertigung von abgerechneten Leistungen dienen können; mit anderen Worten: bei Grenzfällen dürfte eher eine gravierende Erkrankung bzw. Schwere eines Problems kodiert werden.

Trotzdem sind ICD-Diagnosen nicht wertlos. So ist etwa die Kodierung einer Multiplen Sklerose in mehreren aufeinanderfolgenden Quartalen belastbarer als verschieden kodierte Atemwegsinfekte. Wenn möglich sollte eine Abgleichung dadurch erfolgen, dass die Diagnose-Kodierung einer Patientin in einer Betriebsstätte mit der Kodierung durch mindestens eine andere Betriebsstätte abgeglichen wird. Eine weitere Möglichkeit besteht darin, eine Diagnose nur bei der VO typischer Medikamente anzunehmen. Diese Strategien sind allerdings immer noch fehlerbehaftet (z.B. MS weiter behandlungsbedürftig, auch wenn krankheits-modifizierende Medikamente abgesetzt sind bzw. primär nicht verordnet worden sind).

Jeder Indikator und jede Auswertung von Routinedaten haben ihre Grenzen und nicht vermeidbaren Fehler. Es ist jeweils zu entscheiden, ob die zu erwartenden inhaltlichen Ergebnisse einer Auswertung die Fehlermöglichkeiten aufwiegen.

Tabelle 1: Indikatoren für die Wirtschaftlichkeit und Qualität von Medikamenten-Verordnungen

	Beispiel	Relevanz Qualität	Relevanz Kosten	Nenner	Anwendung	Kommentar	
VO eines Original-Präparates	VO von Targin®; Prolia®	++	++	Zeit; VOen der Wirkstoffgruppe (z.B. Anteil an Opioiden; an Osteoporosemitteln)	Profiling		
VO eines Wirkstoffes bzw. -gruppe	VO von PPI; DOAK	+++	+	Zeit; VO-Patienten bzw. Wirkstoffgruppe	Profiling		
Anteil der Generika-VO		+	+++	Zeit; VO-Patienten	Profiling Systemevaluation		
Anteil der Generika-DDDs		+	+++	Zeit; VO-Patienten	Profiling Systemevaluation		
Anteil Analogpräparate		++	++	Zeit; VO-Patienten	Profiling Systemevaluation		
Anteil Biosimilars-VO		∅	+++	Zeit; VO des Wirkstoffes (z.B. Insulin glargin, Ethanercept)	Profiling Systemevaluation		
Anteil Biosimilars-DDDs		∅	+++	Zeit; VO des Wirkstoffes (z.B. Insulin glargin, Ethanercept)	Profiling Systemevaluation		

Erfüllung des Kriteriums	Übereinstimmung mit Medikations-Katalog der KBV	++	++		Profiling Systemevaluation		
Kosten pro VO		∅	+++	Zeit; VO-Patienten	Profiling Systemevaluation	Anfällig für Confounding durch Case-mix	
VO-Kosten pro Patient		∅	+++	Zeit; VO-Patienten	Profiling Systemevaluation	Anfällig für Confounding durch Case-mix	
VO-Kosten pro Betriebsstätte		∅	+++	Zeit; VO-Patienten	Profiling Systemevaluation	Anfällig für Confounding durch Case-mix und Betriebsstätte-Größe	

Erläuterungen: case-mix – Zahl und Schwere behandlungsrelevanter Erkrankungen

Störfaktoren (confounder)

Bei der individuellen Prüfung von Betriebsstätten (Profiling) dient die Berücksichtigung von Störfaktoren (Adjustierung) dazu, einen „fairen Vergleich“ zu ermöglichen. So soll beispielsweise eine Betriebsstätte nicht dadurch benachteiligt werden, dass dort eher alte Menschen behandelt werden, wodurch die Kosten pro Patient höher sind als in Betriebsstätten mit einem niedrigeren Durchschnittsalter der Patienten. Die Berücksichtigung von Störgrößen kann eine Prüfung vereinfachen, da die „händische“ Prüfung von Praxis-Besonderheiten reduziert werden kann. Idealerweise werden die aktuell bei der Wirtschaftlichkeits-Prüfung relevanten Charakteristika von Betriebsstätte bereits in das Profiling-Modell systematisch eingebracht und machen damit eine „händische“ Prüfung - jedenfalls dieser Größen - überflüssig.

Bei der Evaluation von System-Interventionen kann die Berücksichtigung von Störfaktoren helfen, einen Vergleich aussagekräftiger zu machen. Wenn sich eine Interventions- und eine Kontrollgruppe sich in Bezug auf demografische Charakteristika und Morbidität unterscheiden, ist die Berücksichtigung (Ausschaltung) von Störfaktoren geboten, um falsche Schlussfolgerungen zu vermeiden.

Bei den hier diskutierten Analysen sind unsere Interventionsmöglichkeiten begrenzt; die eleganteste Methode, mit Störfaktoren (sogar mit unbekanntem bzw. undefinierten) umzugehen, die randomisierte Zuteilung, ist kaum möglich. Es handelt sich also durchweg um Beobachtungs-Studien bzw. Analogien dazu.

Störfaktoren (Confounder) zu berücksichtigen, ist in der empirischen Forschung eine alltägliche Aufgabe. Bei der Prüfung der Wirtschaftlichkeit gemäß §106b SGB V, wir sprechen hier von regulatorischer im Gegensatz zu einer empirischen Zielsetzung, müssen jedoch Besonderheiten beachtet werden. Hier geht es nicht vorrangig darum, die Wirklichkeit zu verstehen, sondern um die Einhaltung von Normen wie z.B. des Wirtschaftlichkeitsgebots nach §12 SGB V, um die Feststellung von Verstößen mit der Konsequenz von Feedback, Benchmarking, vielleicht sogar Sanktionen.

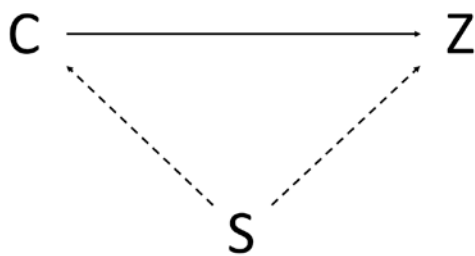


Abbildung 2

Diese unterschiedlichen Zielsetzungen wirken sich auf die Auswahl von Störfaktoren aus. Ein Störfaktor (engl. confounder) ist dadurch definiert, dass er sowohl mit Charakteristika der Betriebsstätte (in Abbildung 3 mit ‚C‘ bezeichnet; dies kann auch die nominale Zuordnung sein) wie auch mit der Zielgröße assoziiert sind. Dies gilt beispielsweise für das mittlere Alter der Patienten einer Betriebsstätte (in Abbildung 3 mit ‚S‘ bezeichnet), wenn unsere Zielgröße die Verordnungskosten pro Patient sind (‚Z‘). Praxen mit einem höheren Durchschnittsalter ihrer Patienten haben höhere VO-Kosten, obwohl sie vielleicht sehr kostenbewusst sind. Eine Berücksichtigung des Störfaktors „mittleres Alter der Patienten einer Betriebsstätte“ macht einen fairen Vergleich mit Praxen möglich, die eher jüngere Patienten betreuen.

Andere Überlegungen ergeben sich bei der Prüfung eines Kriteriums der Verordnungsqualität, wie etwa der Übereinstimmung mit dem Medikationskatalog der KBV oder der VO von PPI. Hier ist möglicher Weise das Kriterium mit dem Alter der Patienten empirisch assoziiert; für ein regulatorisches Vorgehen hätte eine entsprechende Adjustierung jedoch keine Bedeutung, da eine Abweichung vom Medikationskatalog oder höhere PPI-VOen in Abhängigkeit vom Alter nicht sinnvoll begründet werden können.

Empirisch mag es so sein, dass Betriebsstätte in der Nähe von Krankenhäusern höhere Kosten pro Patient oder VO aufweisen, vielleicht auch seltener Leitsubstanzen (vorzuziehende Wirkstoffe in Vereinbarungen von KVen und Krankenkassen) verordnen. Dies sollte jedoch nicht vorschnell dazu führen, für den Faktor „Nähe zum Krankenhaus“ zu adjustieren. Inhaltlich gibt es keine Rechtfertigung dafür, dass in der Nähe von Krankenhäusern mit höheren Kosten und/oder geringerer Qualität behandelt wird. Eine Adjustierung würde ein reales Problem nur verdecken, das mit konkreten, eher strukturellen Maßnahmen angegangen werden sollte (z.B. Erarbeitung eines Konsens/Behandlungspfades von Krankenhaus und Praxen der einer Region).

Die Größe einer Praxis mag sich positiv auf die Verordnungsqualität auswirken (z.B. weil größere Praxen eher interne Behandlungspfade vereinbaren und durchsetzen). Würde man für diesen Faktor bei einer regulatorischen Prüfung adjustieren, würde man den größeren Betriebsstätten ihren Vorsprung an Qualität bzw. Effizienz nehmen, für den sie sich (möglicherweise) die strukturellen Voraussetzungen geschaffen haben. Hier handelt es sich also um einen Wirkmechanismus, für den weder in analytischen noch in regulatorischen Analysen adjustiert werden sollte („w“ in Abbildung 4).

Dass in empirischen Analysen eine Eigenschaft von Patienten oder Leistungserbringern mit dem interessierenden Qualitätskriterium assoziiert ist, sollte also nicht zwangsläufig dazu führen, dass für diese Größen bei einer regulatorischen Auswertung adjustiert wird. Vielmehr ist jeweils zu überlegen, ob für die Berücksichtigung des Confounders eine Rechtfertigung besteht. Rein mathematisch verhalten sich die Größen S in Abbildung 3 und W in Abbildung 4 identisch. Es ist vor der Auswertung zu entscheiden, ob die Berücksichtigung einer Störgröße begründet ist.

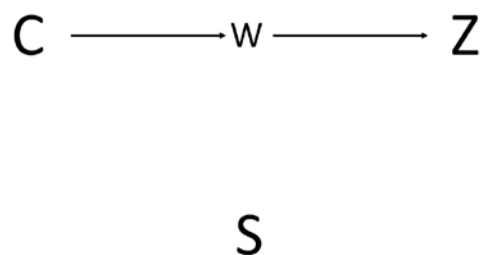


Abbildung 3

Störfaktoren können in ein mathematisches Modell eingebracht werden, wie z.B. beim klassischen Profiling. Hierbei wird aus den Charakteristika der Praxis ein „erwarteter“ Wert der Zielgröße berechnet. Aus der Differenz zum „tatsächlichen“ Wert einer Betriebsstätte kann dann geschlossen werden, ob es sich um eine interventions- oder zumindest untersuchungsbedürftige Auffälligkeit handelt.

Eine weitere Möglichkeit, Störfaktoren zu berücksichtigen, besteht darin, die Auswertung auf eine Gruppe (=Ausprägung eines Merkmals) zu beschränken (sog. Restriktion). Dieses Vorgehen wird meist in Bezug auf eine Fachgruppe gewählt. So macht es wenig Sinn, Hausärzte mit Rheumatologen zu vergleichen. Diese Gruppen werden deshalb getrennt ausgewertet und beurteilt.

Gerade bei verursachten Kosten hat die Berücksichtigung des case-mix der untersuchten Betriebsstätte eine besondere Bedeutung. Allerdings bietet dies große Schwierigkeiten. Adjustierung auf der Basis kodierter Diagnosen, seien dies einzelne oder zusammengesetzte Indizes, leiden unter den o.g. Problemen bei Nenner-Definitionen. Das Alter der behandelten Patienten ist im Grunde genommen ebenfalls eine Proxy-Variable. Denn das Alter per se ist nicht behandlungsbedürftig; die meisten chronischen Erkrankungen sind jedoch mit dem Alter assoziiert (KHK, Schlaganfall, Hypertonie, Malignome usw.).

Tabelle 2: Störfaktoren (Confounder) bei der Prüfung von Wirtschaftlichkeit und Qualität von Medikamenten-Verordnungen

	Operationalisierung(en)	Ebenen	Verfahren	Kommentar
Alter der Patienten		Individuell; aggregiert (z.B. Betriebsstätte)	Geschichtete oder multivariate Auswertung	
Case-mix	Multimorbiditäts-Indices	Individuell; aggregiert (z.B. Betriebsstätte)	Geschichtete oder multivariate Auswertung	
Polypharmazie	z.B. > 5 Dauermedikamente	Individuell; aggregiert (z.B. Betriebsstätte)		Im Einzelfall prüfen, ob sachlich begründet
Einzelne Diagnosen (ICD)		Individuell; aggregiert (z.B. Betriebsstätte)	Prüfung getrennt nach Erkrankungen bzw. -Gruppen (Restriktion)	
Arztfachgruppe(n)		Bezieht sich auf Betriebsstätte bzw. einzelne Ärzte	Prüfung getrennt nach Fachgruppe (Restriktion)	Vergleiche über Fachgruppen bzw. Sektoren-Grenzen hinweg meist nicht sinnvoll
Lage	z.B. regionalstatistische Raumtypen	Bezieht sich auf Betriebsstätte		
Betriebsstätte-Größe	Behandelte Patienten/Fälle Arztsitze	Bezieht sich auf Betriebsstätte		
Verordnungstiefe	Anteil der VO durch die Betriebsstätte an sämtlichen VO, die ein Patient erhält	Bezieht sich auf Betriebsstätte		

Evaluation von Strategien: Quantitative Analyse von Verordnungsdaten Die vertragsärztliche Versorgung von gesetzlich versicherten PatientInnen soll ausreichend, zweckmäßig und wirtschaftlich durchgeführt werden (Wirtschaftlichkeitsgebot - §12 SGB V). Dabei steht die Verordnung von Arzneimitteln - neben den durchgeführten ärztlichen Leistungen – in einem besonderen Fokus.

Um die Entwicklung der Arzneimittelverordnungen besser erfassen, verstehen und unter wirtschaftlichen Aspekten steuern zu können, sind Arzneimittelverordnungsdaten von großer Relevanz. Mit deren Hilfe können Analysen durchgeführt, Verordnungsmuster oder auch -unterschiede (z.B. zwischen Arztgruppen oder verschiedenen KV-Regionen) aufgedeckt und Entscheidungshilfen bzgl. einer Steuerung entwickelt werden. Die Verordnungsdaten liegen dabei zumeist in der Zeit fortlaufend als Zeitreihen oder Paneldaten vor.

Wenn quantitative Vergleiche im Bereich der Verordnungsdaten durchgeführt werden sollen, gibt es dazu verschiedenen Möglichkeiten. In dieser Toolbox wird das Strukturbruchmodell beispielhaft vorgestellt. Zudem sind unter anderem auch Vergleiche von verbundenen oder auch unverbundenen Stichproben sowie allgemeine Zeitreihenanalysen gängige Analysemethoden.

Mit dem Ansatz des *Vergleichs von verbundenen Stichproben* kann das Verordnungsgeschehen für eine vorab definierte Betriebsstätte/einen vorab definierten Arzt oder auch eine Gruppe (z.B. alle Ärzte einer KV oder einer bestimmten Fachgruppe) analysiert werden, indem gemittelte Verordnungsniveaus zu gewählten Zeitpunkten statistisch miteinander verglichen werden. Mit diesem Ansatz kann bestimmt werden, ob sich zwischen den betrachteten Zeitpunkten signifikante Unterschiede ergeben.

Mit dem Ansatz des *Vergleichs von unverbundenen Stichproben* kann der Unterschied beim Verordnungsgeschehen zwischen vorab definierten Betriebsstätte/Ärzten oder auch Gruppen (z.B. Ärzte zweier KVen oder zweier Fachgruppen) analysiert werden. Dazu wird jeweils die gemittelte Differenz des Verordnungsniveaus gebildet und bestimmt, ob diese signifikant ist. Trifft dies zu, liegt ein signifikanter Unterschied im Verordnungsniveau der beiden betrachteten Gruppen vor.

Um Muster oder auch Unterschiede im Ordnungsverhalten - im Zeitverlauf oder auch zwischen verschiedenen Vorordnern - zu analysieren, ist die wohl gängigste und bewährteste statistische Methode, die Anwendung von statistischer Zeitreihen- bzw. Paneldatenanalyse auf die jeweils vorliegenden Verordnungsdaten. Um diese statistischen Analyseverfahren anwenden zu können, sind vorab verschiedene Dateneigenschaften zu überprüfen und ggf. in der Analyse zu berücksichtigen. Grundlegend sollten vor einer Analyse die folgenden Fragen beantwortet werden:

- Sind die Elemente der betrachteten Zeitreihe unabhängig voneinander oder seriell korreliert?
- Ist die Varianz über die Zeit hinweg konstant oder variiert diese?
- Liegt bei vorliegenden Paneldaten ein Fixed Effects Model oder Random Effects Model vor?

- Liegt bei vorliegenden Paneldaten Homo- oder Heteroskedastizität vor?

Für eine umfassende Vorbereitung empfehlen sich die folgenden Quellen:

- (1) Kirchgässer, Wolters, Hassler – Introduction to Modern Time Series Analysis, Springer
- (2) Shumway, Stoffer – Time Series Analysis and Its Application, Springer
- (3) Baltagi – Econometrics, Springer
- (4) Baltagi – Econometric Analysis of Panel Data, John Wiley & Sons
- (5) Wooldridge – Econometric Analysis of Cross Section and Panel Data, MIT Press

4. Das Strukturbruchmodell

Sollen bestimmte Aspekte des Verordnungsgeschehens analysiert werden (z.B. die Verordnungskosten oder der Anteil von generischen Verordnungen in bestimmten Indikationsbereichen), so sind äußere Einflüsse von großer Relevanz und müssen einbezogen werden. Es gilt dabei zwischen langfristigen Einflussvariablen (Confoundern), wie z.B. die Patientenstruktur einer Betriebsstätte, und einzelnen punktuellen Ereignissen, die Einfluss nehmen, zu unterscheiden. Liegt im Betrachtungszeitraum ein einzelnes Ereignis zu einem bekannten Zeitpunkt vor, von welchem man vermutet, dass es einen Einfluss auf das Verordnungsgeschehen ausübt, wie z.B. eine Gesetzesänderung oder eine Einführung einer neuen Prüfverordnung in einer KV, so kann dies mit dem Strukturbruchmodell (nach Chow) überprüft werden.

Das hier präsentierte Modell (Strukturbruchmodell) beschäftigt sich also mit der Fragestellung, ob ein potentiell einflussnehmendes Ereignis den Verlauf von Zeitreihen oder Paneldaten signifikant verändert. Dies bedeutet:

Liegt bei der Analyse von Zeitreihen oder Paneldaten im betrachteten Zeitraum ein (bekanntes) Ereignis vor, das potentiell Einfluss auf die betrachtete Zeitreihe oder das Panelmodell nimmt, so kann mit dem Strukturbruchmodell bestimmt werden, ob tatsächlich eine statistisch signifikante Änderung vorliegt oder nicht. Die messbare Änderung des Verlaufs der Zeitreihe oder der Paneldaten ab dem Ereignis wird auch Strukturbruch genannt. Der Punkt der Änderung ist der Strukturbruchpunkt.

Mögliche Anwendungen im Bereich der Wirtschaftlichkeitsprüfung sind z.B. die Analyse von Verordnungskosten, Analysen zum Anteil der generischen Verordnungen an allen Verordnungen oder der Veränderung der Verordnungen von bestimmten Wirkstoffen oder Wirkstoffgruppen über einen bestimmten Zeitraum hinweg. Dabei kann die gewünschte Untersuchungsvariable (Indikator) als Zeitreihe modelliert, evtl. Confounder (Verweis) einbezogen und die mögliche Beeinflussung durch ein externes Ereignis auf die Untersuchungsvariable untersucht werden. Dieses externe Ereignis kann jedes Ereignis sein, für welches vermutet wird, dass es eine Veränderung der Indikatorvariable auslöst. Für die oben genannten Beispiele könnte es z.B. die Einführung eines neuen Prüfsystems, die Markteinführung von neuen Originalpräparaten oder neuer generischer Präparate oder auch neu eingeführte gesetzlichen

Regelungen sein. Wichtig ist, dass der Zeitpunkt des Ereignisses vor Analysebeginn bekannt ist, um einen Vorher-Nachher-Zeitraum bilden zu können.

Mögliche Fragestellungen für eine Analyse mittels Strukturbruchmodell sind u.a.:

- Verändert die Einführung einer neuen Prüfvereinbarung der Wirtschaftlichkeitsprüfung die zeitliche Entwicklung der Verordnungskosten (z.B. stärkerer oder abgeschwächer Anstieg nach der Einführung)?
- Haben neue Vorgaben bzgl. Generikaquoten einen messbaren Effekt auf den generischen Anteil an allen Verordnungen/ an allen verordneten DDDs?
- Hat eine Empfehlung für das vorrangige Verordnen von bestimmten Wirkstoffen innerhalb dazu passender Wirkstoffgruppen eine Erhöhung des Ordnungsanteils dieser Wirkstoffe bewirkt?

Es folgt nun die Herleitung und der Aufbau des allgemeinen Strukturbruchmodells nach Chow - auch als Chow-Test bekannt - inkl. der nötigen Voraussetzungen und evtl. Beschränkungen. Als zweites wird eine Durchführungsanleitung für den Chow-Test zur Verfügung gestellt (inkl. Hilfestellungen für eine Analyse mittels des Analyseprogramms R).

Diese ist als Hilfestellung für das eigene Durchführen dieser Analyseart gedacht und benennt und beschreibt die Zwischenschritte, die dazu nötig sind.

Das allgemeine Strukturbruchmodell- Test auf Poolfähigkeit (Chow-Test)

Das allgemeine Strukturbruchmodell und der damit verbundene Test auf Poolfähigkeit bzw. Strukturbruch ist anwendbar auf ein Setting, in dem Paneldaten vorliegen (oder auch Zeitreihendaten).

Der Test auf Poolfähigkeit bzw. auf Strukturbruch nach Chow prüft, ob es zwischen zwei Untergruppen von Individuen (mit vorab bekannter Aufteilung) oder alternativ im Zeitverlauf zu einem vorab bekannten Zeitpunkt einen Bruch in den betrachteten Paneldaten gibt. Somit kann mit demselben Ansatz entweder überprüft werden, ob sich eine Gesamtgruppe von Individuen in zwei Untergruppen mit verschiedenen Ausprägungen oder verschiedenem Verhalten unterteilen lässt oder aber auch, ob sich vorliegende Ausprägungen oder Verhaltensweisen zu einem vorab bekannten Zeitpunkt (aufgrund eines bekannten Ereignisses) ändern.

Methodisch bedeutet dies, dass im Fall eines Strukturbruchs nicht ein Gesamtmodell über alle Paneldatenpunkte gelegt werden sollte, sondern dass zwei Modelle die Zusammenhänge der Paneldaten besser abbilden. Dabei sollte jeweils ein Modell für jede Untergruppe bzw. für jeden Zeitabschnitt (d.h. jeweils die Abschnitte vor und nach dem zeitlichen Bruchpunkt) aufgestellt werden. Bei der Visualisierung aller Datenpunkte in einer Grafik, kann sich der Bruch, auch Strukturbruch genannt, z.B. als Sprung oder „Knick“ manifestieren. Der Chow-Test prüft nun, als Nullhypothese, ob über alle Datenpunkte hinweg ein einheitliches Modell angenommen werden kann. Wenn dieser Fall nicht vorliegt, wird die Annahme verworfen und der Strukturbruch bestätigt (Verwurf der Nullhypothese und somit Bestätigung der Alternativhypothese).

Es folgen nun die Modellierung des Chow-Tests.

Das Chow-Test-Modell

Der Chow-Test prüft auf die Poolfähigkeit (Poolability) bzw. auf einen Strukturbruch in den vorliegenden Daten und kann sowohl auf Paneldaten als auch auf Zeitreihendaten angewandt werden. Im Folgenden wird das Modell für Paneldaten präsentiert.

Insgesamt gibt es zwei mögliche Settings, für welche auf Poolfähigkeit (bzw. Strukturbruch) getestet werden kann:

- über alle Individuen des Paneldatensatzes hinweg
- über alle Zeitpunkte hinweg.

Im Falle, dass die Poolfähigkeit abgelehnt wird, liegt ein Strukturbruch vor.

Im Folgenden wird das Setting für den Test auf Poolfähigkeit bzw. Strukturbruch über alle Individuen des Paneldatensatzes hinweg präsentiert. Das Vorgehen beim Setting der Poolfähigkeit über alle Zeitpunkte hinweg ist analog.

Setting

Seien N Individuen und T Zeitpunkte gegeben. Zudem gelte, dass die Regressionsparameter nicht über die Zeit aber möglicherweise über die Individuen hinweg variieren. Ziel sei die Testung auf Poolfähigkeit über alle Individuen hinweg.¹² Dann betrachten wir das folgende Modell:

→ (Unbeschränktes) Modell:

$$\begin{aligned} y_{it} &= \alpha_i + x'_{it}\beta_i + \epsilon_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T \\ \Leftrightarrow y_i &= \underbrace{\alpha_i \mathbf{1}_T + X_i \beta_i}_{=Z_i \theta_i} + \epsilon_i, \quad i = 1, \dots, N, \end{aligned}$$

wobei für Z_i und θ_i gilt:

- $Z_i = [\mathbf{1}_T : X_i]$ ist eine $T \times (K + 1)$ – Matrix und
- $\theta_i = \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix}$ ist ein $(k + 1)$ – Vektor.

¹² Das Vorliegen der Poolfähigkeit über die Zeit kann auf analoge Weise getestet werden.

Die Hypothese der Poolfähigkeit

Die Hypothese der Poolfähigkeit (Nullhypothese) der gegebenen Paneldaten nimmt an, dass kein Strukturbruch vorliegt und somit alle Koeffizienten θ_i konstant sind über alle Individuen. Die Nullhypothese ist gegeben durch:

$$H: \theta_i \equiv \theta \quad \text{für} \quad i = 1, \dots, N.$$

Diese Hypothese korrespondiert zum beschränkten Modell

$$y = \alpha 1_{NT} + X\beta + \epsilon = Z\theta + \epsilon.$$

Das Modell lässt sich in einer kompakten Notation folgendermaßen darstellen:

$$y = \underbrace{\begin{pmatrix} Z_i & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ 0 & \cdots & 0 & Z_N \end{pmatrix}}_{=:Z^*} + \underbrace{\begin{pmatrix} \theta_1 \\ \vdots \\ \theta_N \end{pmatrix}}_{=: \theta^*} + \epsilon. \quad (1.1)$$

Test auf Poolfähigkeit

Für den Test auf Poolfähigkeit werden zwei verschiedene Fälle unterschieden:

Im 1. Fall ist der Fehlerterm normalverteilt mit Erwartungswert null und einer konstanten Varianz (es liegt somit Homoskedastizität vor). Im 2. Fall hingegen ist der Fehlerterm normalverteilt mit Erwartungswert null und einer nicht-konstanten Varianz (es liegt somit Heteroskedastizität vor). Vor Durchführung des Tests auf Strukturbruch ist es daher immer wichtig zu prüfen, welcher der beiden Fälle vorliegt.

Fall 1: $\epsilon \sim N(0, \sigma^2 I_{NT})$

Der Fehlerterm ist normalverteilt mit Erwartungswert null und einer konstanten Varianz, d.h. es liegt somit Homoskedastizität vor. Dann gilt ebenfalls:

- Es liegen keine Individual-Effekte (individual effects) oder Zeit-Effekte (time effects) vor.
- H ist eine lineare Hypothese in einem linearen Modell (1.1).
- Der F-test ist hierbei der gleichmäßig trennschärfste bzw. gleichmäßig mächtigste Test (uniformly most powerful invariant α_0 -test).

F-Test Statistik:

$$F = \frac{(RRSS - RSS)/DF_1}{RSS/DF_2} \underset{H}{\sim} F_{DF_1, DF_2} \quad (1.2)$$

wobei die Residuenquadratsumme - auch (*unrestricted*) *Residual Sum of Squares* (RSS) genannt - gegeben ist durch

$$\begin{aligned} RSS &= \min_{\theta^*} \|y - Z^* \theta^*\|^2 \\ &= \|y - \underbrace{Z^* \hat{\theta}_{OLS}^*}_{=P_{R(Z^*)} \cdot y}\|^2 = \|(I - \underbrace{P_{R(Z^*)}}_{=:M^*}) y\|^2 = y' M^* y \end{aligned}$$

mit $P_{R(Z^*)} = Z^* (Z^{*'} Z^*)^{-1} Z^*$.

Der Residuenvektor in Modell (1.1) sei mit

$$\hat{e}^* = y - Z^* \hat{\theta}_{OLS}^* = M^* y \quad \text{und} \quad \hat{e}_i^* = M_i y_i$$

gegeben durch:

$$\hat{e}^* = \begin{pmatrix} \hat{e}_1^* \\ \vdots \\ \hat{e}_N^* \end{pmatrix}$$

und somit ist

$$RSS = \hat{e}^{*'} \hat{e}^* = \sum_{i=1}^N \hat{e}_i^* \hat{e}_i^*$$

Die beschränkte Residuenquadratsumme - auch *restricted residual sum of squares* (RRSS) genannt - ist (unter der Annahmen H) gegeben durch:

$$RRSS = \min_{\theta} \|y - Z\theta\|^2 = \|y - \underbrace{Z \hat{\theta}_{OLS}}_{=P_{R(Z)} \cdot y}\|^2 = y' M y = \hat{e}' \hat{e}$$

Wobei $\hat{e} = (I_{NT} - P_{R(Z)})y = My$ der Vektor der OLS-Residuen im beschränkten Modell ist (mit $P_{R(Z)} = Z(Z'Z)^{-1}Z$).

Die *Freiheitsgrade* (*degrees of freedom*) sind gegeben durch

$$DF_2 = \text{rank}(M^*) = \text{tr}(M^*) = N(T - K - 1)$$

$$DF_1 = \text{rank}(M) - DF_2 = \text{tr}(M) - N(T - K - 1)$$

$$= (NT - K - 1) - N(T - K - 1) = (K + 1)(N - 1)$$

Damit folgt

$$F = \frac{(\hat{e}'\hat{e} - \sum_{i=1}^N \hat{e}_i^* \hat{e}_i^*) / (K + 1)(N - 1)}{\sum_{i=1}^N \hat{e}_i^* \hat{e}_i^* / N(T - K - 1)} \underset{H}{\sim} F_{(K+1)(N-1), N(T-K-1)} \quad (1.3)$$

→ ein α_0 -Test (F-test) wird abgelehnt, wenn

$$F > F_{(K+1)(N-1), N(T-K-1)}^{1-\alpha_0}$$

Bemerkungen:

- Dieser Test wurde von Chow entwickelt und wird daher Chow-Test genannt.
- Dieser Test ist robust gegen einen Verstoß der Normalverteilungsannahme, d.h. es wird asymptotisch das Signifikanzniveau α_0 erreicht (ohne Vorliegen der Normalverteilung).
- Der F-Test ist nicht robust gegen Verletzung der Homoskedastizitätsannahme, so dass im Falle eines Modells mit $(\epsilon \sim N(0, \Omega), \Omega \neq \sigma^2 I)$ die Anwendung des Chow-Tests modifiziert werden muss.

Fall 2: $\epsilon \sim N(0, \Omega)$

Der Fehlerterm ist normalverteilt mit Erwartungswert null und einer nicht-konstanten Varianz, d.h. es liegt somit Heteroskedastizität vor. Da der F-Test nicht robust gegen Verletzung der Homoskedastizitätsannahme muss eine Modifikation des Modells stattfinden, um den Chow-Test im Anschluss anwenden zu können.

- Idee: Ist Ω bis auf einen Skalarfaktor bekannt (d.h. $\Omega = \sigma^2 \Sigma$ mit bekanntem Σ), kann das Modell (unter der Annahme H und Gegenannahme H^-) transformiert werden. Dies wird durch eine Multiplikation mit $\Sigma^{-1/2}$ erreicht. Auf diese Art erhält man ein transformiertes Modell mit Fehlerkovarianzmatrix $\sigma^2 I_{NT}$.

→ Der F-Test (Chow-Test) kann nun auf das transformierte Modell angewandt werden.

- Ist Ω unbekannt, wird Σ durch einen konsistenten Schätzer $\hat{\Sigma}$ ersetzt (z.B. durch eine konsistente Schätzung der Varianzkomponenten)

→ Für den Test wird die neu erhaltene F-Statistik \hat{F} anstelle von F genutzt, wobei gilt:

$$\hat{F} \underset{H}{\sim} \hat{F}_{(N-1)(K+1), N(T-K-1)} \approx F_{(N-1)(K+1), N(T-K-1)}$$

Somit erhält man einen approximativen F-Test, welcher das Signifikanzniveau asymptotisch erreicht.

Durchführungsanleitung für den Chow-Test (Ablauf der Analyse in R)

Im Folgenden wird der Ablauf für eine Strukturbruchanalyse beschrieben, die auf Paneldaten angewandt wird (in insgesamt 11 Schritten). Bei ausgewählten Schritten und auch abschließend sind zudem Hilfestellungen für die Umsetzung in R angegeben.

Eine Illustration der hier aufgeführten Schritte 1-10 auf dem Foliensatz „Getting Started in Fixed/Random Effects Models using R“ der Stanford University.¹³

Das Verfahren kann in die folgenden 11 Schritte aufgeteilt werden:

1. Gepooltes Panelmodell

Zunächst wird ein gepooltes Panelmodell mittels OLS-Regression aufgestellt – dieses Modell bildet keine Heterogenität zwischen den Untergruppen oder über die Zeit ab.

2. Panelmodell mit Fixed Effects

In einem zweiten Schritt wird ein Panelmodell mit Fixed Effects bestimmt (mithilfe der R-library *plm*) – dies ermöglicht die Abbildung von konstanten Unterschieden zwischen den betrachteten Einheiten bzw. Zeitpunkten und somit einer konstanten Heterogenität über alle Einheiten bzw. Zeitpunkte hinweg.

3. F test for individual effects (Test auf Fixed Effects)

Das in Schritt 2 aufgestellte Panelmodell mit Fixed Effects wird im Anschluss mittels F-Test getestet, um zu überprüfen, ob sich diese Modellwahl für die vorliegenden Daten besser eignet als das gepoolte Panelmodell.

4. Panelmodell mit Random Effects

In einem nächsten Schritt wird ein Panelmodell mit Random Effects geschätzt (ebenfalls mithilfe der R-library *plm*) – dies ermöglicht die Abbildung von nicht-konstanten Unterschieden zwischen den betrachteten Einheiten bzw. Zeitpunkten und somit einer nicht-konstanten Heterogenität über alle Einheiten bzw. Zeitpunkte hinweg.

5. Breusch-Pagan Lagrange Multiplier for random effects (Test auf Random Effects)

Das in Schritt 4 aufgestellte Panelmodell mit Random Effects wird mittels des Breusch-Pagan Lagrange Multiplier ebenfalls auf seine Eignung getestet und bestimmt, ob es sich besser eignet als das gepoolte Panelmodell.

6. Hausman-Test (Entscheidung bzgl. der Modellart: Fixed Effects oder Random Effects)

Wurden sowohl das Fixed-Effects-Modell als auch das Random-Effects-Modell aufgestellt und mittels der eigenen Tests als mögliche Modellierung bestätigt, kann mit dem Hausman-Test überprüft werden, welche der beiden Modellvarianten die vorliegenden Paneldaten besser wiedergibt.

7. Weitere Tests: Korrelationstest

¹³ Siehe Foliensatz in den Quellen.

Nach der Ermittlung der richtigen Modellart (gepoolt, mit Fixed Effects oder Random Effects) kann es sinnvoll sein, die Paneldaten auf Korrelationen zu untersuchen, z.B. nach der Reihenkorrelation (serial correlation). Insbesondere bei großen Paneldatensätzen ist eine Abklärung von möglichen Korrelationen sinnvoll, z.B. mit dem „Breusch-Godfrey/Wooldridge test for serial correlation in panel models“ (dieser ist zu finden in der R-library *lmtest*).

8. Test auf Homoskedastizität bzw. Heteroskedastizität

Essentiell für die Anwendung des Chow-Tests ist die Ermittlung, ob Heteroskedastizität vorliegt oder von doch von Homoskedastizität ausgegangen werden kann. Dies kann mithilfe des Breusch-Pagan-Tests ermittelt werden - wird dieser abgelehnt liegt Heteroskedastizität vor. In diesem Fall muss das Modell zwingend (mittels einer robusten Kovarianzmatrixschätzung) angepasst werden.

9. Robuste Kovarianzmatrixschätzung

Eine robuste Kovarianzmatrixschätzung ist in R implementiert als „Robust covariance matrix estimation (Sandwich estimator)“. Für verschiedene Modellarten gibt es jeweils Anpassungen für die Schätzung (z.B. ob ein Fixed-Effects-Modell vorliegt oder ein Random-Effects-Modell).

10. Transformation des Modells mithilfe der geschätzten Kovarianzmatrix -> man erhält robuste Standardfehler (SE)

Hat man die Art und Weise der robusten Kovarianzmatrixschätzung für die vorliegenden Paneldaten ermittelt, wird die R-Funktion für diese Kovarianzschätzung direkt in die Modellschätzungsfunktion für das gepoolte/ Fixed-Effects-/ Random-Effects- Modell eingesetzt. Das so resultierende Modell ist nun das transformierte Gesamtmodell, das sowohl die Modellart als auch die Kovarianzmatrixschätzung enthält.

11. Chow-Test

Final kann nun der Chow-Test auf Strukturbruch auf das transformierte Gesamtmodell angewandt werden (falls die Transformation durchgeführt werden musste). Dazu wird die R-Funktion für die Gesamtmodellschätzung direkt in den R-Befehl für den Chow-test eingefügt.

Wird der Chow-test abgelehnt, so liegt ein Strukturbruch vor.

Potentielle Probleme in R

- Nicht alle Packages in R sind kompatibel – viele sind nur auf Zeitreihen ausgelegt. Wenn man keine Zeitreihen, sondern Paneldaten verwendet, kann es häufiger zu Problemen kommen.
- Es gibt verschiedene Packages und Ansätze für die Modelltransformation, nicht alle sind mit den späteren Analysen kompatibel.
- Die Transformation muss in den Chow-Test selber (zusammen mit der Schätzgleichung und den Daten) integriert werden.

5. Evaluation von Strategien: qualitative Erhebungen

Aufgabe/Problem

Wie können Steuerungs-/Prüfssysteme von deren Anwendern evaluiert werden?

Untersuchtes Beispiel (Beobachtungsebene)

- Bayerische Wirkstoffvereinbarung (WSV)

Aufgaben/Fragestellungen

- Wie bewerten niedergelassene Ärzt*innen die WSV?
- Welche Verbesserungsvorschläge lassen sich daraus ziehen?
- Wie verändert sich das Ordnungsverhalten durch die WSV?
- Welche Schnittstellenprobleme entstehen bei Krankenhausverordnungen?

Implementierbare Tools

Einzelinterviews

Ein Interview dient zur Informationsermittlung einer bestimmten Forschungsfrage. Es existieren verschiedene Formen des Interviews.

Die folgenden Punkte stellen wesentliche Unterscheidungskriterien dar:

- Grad der Standardisierung:
 - Standardisiert (Wortlaut und Reihenfolge der Interviewfragen sind vorab festgelegt und für alle Teilnehmenden gleich)
 - Halb-Standardisiert (Reihenfolge und Formulierung sind flexibel; das halb-standardisierte Leitfadeninterview stellt die am häufigsten verwendete Interviewform dar)
 - Nicht-Standardisiert (Nur Themenkomplexe sind vorgegeben)
- Anzahl Forschende: Einzel-(1 Interviewer), Tandem-(2 Interviewer) & Boardinterview (>2 Interviewer)
- Modalität: Face-to-Face, telefonisch, online

Für die Evaluation eines Steuerungs-/Prüfsystems eignet sich insbesondere ein halb-standardisiertes Tandeminterview.

Interviewführung und -verlauf

In den meisten Interviews, außer nicht-standardisierten, wird ein Leitfaden verwendet. Dieser kann von den Forschenden genutzt werden, um im Vorfeld Wissen zu organisieren, explizieren und zu diskutieren. Während des Interviews kann er zusätzlich als Anhaltspunkt und am Ende als Checkliste verwendet werden. Ein Leitfaden sollte in der Regel 8-15 Fragen auf maximal 2 Seiten beinhalten.

Das Einzelinterview gliedert sich in Eröffnung, Hauptteil und der Endphase. Bei der Eröffnung sollte den Teilnehmenden erklärt werden, um was es sich in dem Interview handelt und das dieses aufgezeichnet wird. Die Aufzeichnung kann per Audio oder bei der zusätzlichen Auswertung von Gestik und Mimik per Video stattfinden. Sollte die Aufzeichnung nicht möglich

oder gewünscht sein, sollten Inhalt und Ablauf des Gesprächs protokolliert werden. Im Hauptteil werden die Fragen, basierend am Leitfaden, an den Teilnehmenden gestellt. In der Endphase wird das Ende des Gesprächs angekündigt, sich für die Teilnahme bedankt und dem Teilnehmenden die Möglichkeit gegeben ihrerseits wichtige Punkte zu ergänzen.

Experteninterviews als Sonderform

Ärzt*innen sind Expert*innen für ihre Alltagserfahrungen mit z.B. neuen Prüfordnungen. Experteninterviews sind stärker auf die Erhebung von praxis- und erfahrungsbezogenem Wissen ausgerichtet (1). Es werden in der Regel Leitfadeninterviews mit eng gefassten, aber dennoch offenen Fragen durchgeführt. Hier bietet sich auch die Vorlage von Fallvignetten an.

Anwendung: Für die Einzelinterviews wurden entsprechende Fallvignetten vorbereitet. Teilnehmende sollten hier erläutern, wie sie bei bestimmten Krankheitsbildern bzw. Patientenfällen vorgehen würden.

Tipps:

Formulierung von Interviewfragen

- Thematisch eng gefasste Fragen
- Keine Fragen, deren Antworten leicht auch anderweitig zu bekommen sind
- Keine Ja/Nein-Fragen
- Nur eine Frage auf einmal
- Keine Suggestivfragen
- Keine doppelten Verneinungen
- Nicht zu viele Warum-Fragen
- Bei der Formulierung an die Ausdrucksweise der Teilnehmenden anlehnen

Kompetenz beim Führen von Interviews

- Rolle der Interviewenden definieren zum Expertenstatus (Machtrelation!)
- Zuhören und dem Befragten dies auch signalisieren (mittels Körpersprache o. Zuhörersignale)
- Gespür dafür, wenn etwas nicht so läuft wie es sollte und dementsprechend intervenieren können
- Gespräch strukturieren, Antworten zusammenfassen, das Interview nach Ausschweifungen zum eigentlichen Thema zurückbringen
- Keine Angst vor Fehlern

Fokusgruppen

Gruppeninterviews oder auch Fokusgruppen genannt, sind moderierte Diskussionsgruppen zu einem bestimmten Thema. Die Diskussion wird meist von Forschenden unter Rückgriff auf einen thematischen Leitfaden geleitet. Die Forschenden stoßen die Diskussion an und moderieren sie, äußern sich aber nicht inhaltlich zum Thema. Ziel der Fokusgruppe ist es mittels gruppenförmiger kommunikativer Interaktion die Assoziationen, Denkmuster, Meinungen und Wahrnehmungen, die die Teilnehmenden mit dem vorgegebenen Thema verbinden, zu erheben. Daher eignet sich diese Methode besonders dafür, neue Themen zu erkunden.

Interviewverlauf

1. Theoretische Vorüberlegung: bisherige Befunde, psychische Sperren & Überwindungen zum Leitfaden und der Zusammensetzung der Gruppe
2. Erläuterung der Gesprächsregeln: Begrüßung der Teilnehmenden. Einander ausreden lassen oder das Wort durch Moderator*in erteilen lassen
3. Präsentation eines Grundreizes: Mittels Grundreiz das Thema umreißen und auf provokante Weise das Gespräch in Gang bringen
4. Freie Gruppendiskussion: Moderator*in beschränkt sich auf Gesprächssteuerung
5. Reizargumente: Ähnlich wie Grundreiz, um die Diskussion in Gang zu bringen; in Form von Zeitungsartikeln oder Zitaten möglich
6. Metadiskussion: Ende der Gruppendiskussion, in der die Teilnehmenden sich zur Diskussion äußern können

Tipps:

- **Neutrale Gesprächsführung**
Die Moderation sollte nicht ihre eigene Meinung oder ihr Expertenwissen zum Ausdruck bringen, sondern eine thematisch neutrale und unabhängige Position vertreten.
- **Gruppengröße**
Die ideale Gruppengröße umfasst sechs bis zehn Personen. Bei dieser Größe stellen Sie sicher, dass alle Teilnehmenden ausreichend Redezeit bekommen und eine Gruppendynamik entstehen kann.
- **Rekrutierung**
Die Rekrutierung der Teilnehmenden ist mit einem nicht zu unterschätzenden Aufwand verbunden, für den genügend Zeit und Ressourcen bereitgestellt werden sollten. Die nötigen Voraussetzungen wie Alter, Geschlecht, Bildungsstand und thematische Kriterien der Teilnehmenden sollten in einem Rekrutierungsbogen festgelegt sein.
- **Aufwandsentschädigung**
Den Teilnehmenden sollten Sie eine Aufwandsentschädigung für die Teilnahme an der Gruppendiskussion erhalten, die je nach Charakteristika der Teilnehmenden, Vorbereitungs- und Gruppendauer zwischen 50 und 100 € liegen kann. Die Ärzt*innen in unserer Studie erhielten 90 € für ein Einzelinterview bzw. 150 € für die Teilnahme an einer Fokusgruppe.
- **Vorbereitung der Teilnehmenden**
Je nach Kenntnisstand der Teilnehmenden und Komplexität des Diskussionsthemas kann es sinnvoll sein, den Teilnehmenden vorab Vorbereitungsmaterial zum Beispiel in Form eines kurzen Hintergrundpapiers zur Verfügung zu stellen.
- **Zusammensetzung der Gruppe**
Bei der Zusammensetzung der Gruppe müssen vorab Überlegungen zur soziodemografischen Ausgestaltung sowie dem Bekanntheitsgrad der Mitglieder gemacht werden. Soziodemografisch wird zwischen homogenen und heterogenen Gruppen differenziert. Beim Bekanntheitsgrad kennen sich sogenannte natürliche Gruppen bereits, wohingegen Ad-hoc-Gruppen für die Untersuchung zusammengestellt werden.

Online-Fokusgruppen

Online-Fokusgruppen können synchron (zeitgleich) oder asynchron (zeitversetzt) stattfinden. In unserer Studie fanden die Fokusgruppen synchron statt.

Asynchrone Fokusgruppen können beispielsweise über bestehende oder selbst erstellte digitale Rahmen wie Online-Foren, Blogs, Webseiten oder soziale Netzwerke geführt werden. Für synchrone Fokusgruppen kann auf Anwendungen für Videokonferenzen oder Chats zurückgegriffen werden.

Die Gruppengröße kann bei der asynchronen Durchführung größer sein, sollte jedoch nicht 20 Teilnehmende überschreiten. Die Gruppengröße zwischen der realen und der synchronen Online-Durchführung ist ähnlich.

Tipps:

- **Link erst kurz vor dem Termin verschicken**
Den Link zur Videokonferenz erst ein paar Tage vor dem Termin mit der Erinnerung verschicken. So müssen die Teilnehmenden nicht lange suchen.
- **Starten Sie 15min vor dem Termin**
Öffnen Sie den virtuellen Raum bereits 15min vor dem eigentlichen Termin und Sie dies den Teilnehmenden auch. Technische Probleme der Teilnehmenden lassen sich in dieser Zeit gut regeln. Fehlende Studienunterlagen können in dieser Zeit ausgefüllt werden (z.B. Soziodemographische Angaben).
- **Auf die Umgebung achten**
Verwenden Sie einen möglichst ruhigen Bereich und vermeiden Sie Räume, die ein Echo haben. Bevorzugen Sie weiße Wände in der Umgebung, achten Sie auf eine gute Ausleuchtung Ihres Gesichts und richten Sie Ihre Kamera nicht gegen Fenster (Licht nicht im Rücken). Tragen Sie bei Möglichkeit einfarbige und helle Kleidung mit wenig Ausschnitt.
- **Allgemeine technischen Voraussetzungen**
Achten Sie auf ein stabiles Internet (mind. 4 Mbit-Leitung), eine gute Webcam und ein gutes Headset/Mikrofon. Prüfen Sie, ob Ihr Browser auf dem neuesten Stand ist. Die Teilnehmenden sollten zuvor testen, ob alles ordnungsgemäß funktioniert. Erstellen Sie eine Anleitung für die Teilnehmenden.
- **Verwendung eines Headsets**
Verwenden Sie ein gutes Headset mit dem Mikrofon nahe am Mund. Vermeiden Sie die standardmäßig eingebauten Mikrofone des Computers.
- **Videokonferenzsystem**
Verwenden Sie ein Videokonferenzsystem wie z.B. WebEx by Cisco oder GoToMeeting, welches ohne Download und vorherige Anmeldung funktioniert. Dies mindert die Hürde zur Teilnahme.
- **Aufnahme der Fokusgruppe**
Eine Aufnahme im Videokonferenzsystem ist kostenpflichtig. Bei Verwendung eines externen Aufnahmegerätes, nehmen Sie die Fokusgruppe bestenfalls an einem anderen PC/in einem anderen Raum auf. So können Sie weiterhin Ihr Headset verwenden und störende Tipngeräusche während der Protokollführung vermeiden.
- **Stummschaltung und Online-Gesprächsführung**
Richten Sie die Stummschaltung bei Eintritt aller Teilnehmenden ein. Bei kleiner Gruppengröße bis zu fünf Personen kann ggfls. auf die Stummschaltung der Teilnehmenden

während der Durchführung verzichtet werden. Das ermöglicht einen natürlicheren Gesprächsfluss. Nachfragen der Moderator*innen können mit der Aufhebung der eigenen Stummschaltung und non-verbale Zeichen signalisiert werden.

Ethik und Datenschutz

Ethische und datenschutzrechtliche Aspekte sind bei jeder Erhebung zu berücksichtigen. Die Teilnehmenden sollten der Erhebung zustimmen. In diesem Zusammenhang sollten Überlegungen dazu angestellt werden, inwieweit den Teilnehmenden eine völlige Anonymität zugesichert werden kann. Insbesondere bei den Fokusgruppen ist darauf hinzuweisen, dass eine Anonymisierung oder Pseudonymisierung während der Erhebung nicht und diese erst im Rahmen der Datenaufbereitung möglich ist. Die Teilnehmenden selbst sollen sich zur Verschwiegenheit über den Inhalt der Diskussionen und der einzelnen Beiträge verpflichten (DNVF). Das Videokonferenzsystem sollte so ausgewählt und umgesetzt werden, dass es den Anforderungen der DSGVO genügt. Für die Fokusgruppen verwendeten wir unterschiedliche Videokonferenzsysteme, sowohl WebEx by Cisco als auch Zoom nach entsprechenden Neuerungen. In der universitären FAU-Instanz von Zoom sind alle datenschutz- und sicherheitsrelevanten Optionen aktiviert. Bei Aufnahme in beispielsweise WebEx ist eine lokale Dateiabspicherung empfehlenswert. Ein externes Aufnahmegerät wurde ebenfalls verwendet.

Gegenüberstellung von Einzelinterviews und (Online-)Fokusgruppen

Die Ergebnisqualität ist in beiden Erhebungsformen sehr gut und ähnlich. Die Kombination beider Erhebungsformen kann empfehlenswert sein, um z.B. Interviewergebnisse durch Fokusgruppen zu validieren. Einzelne Vor- und Nachteile sind in der folgenden Tabelle gegenübergestellt und können für das jeweilige Untersuchungsziel abgewägt werden.

In unserer Studie wurden synchrone Onlinefokusgruppen und Face-to-Face-Fokusgruppen durchgeführt. Der Vergleich beschränkt sich daher auf diese beiden Durchführungsformen. Beide Varianten nehmen ähnlich viel Zeit in Anspruch. In der digitalen Fokusgruppe muss der Moderierende stärker sichtbar sein und die Diskussion vermehrt leiten, um die Teilnehmenden zu motivieren.

Einzelinterview:

- Interviewdauer ca. 60 - 90 Minuten
- Stärker an einzelnen Fällen ausgerichtet, Ärzt*innen bringen eigene Beispiele länger ein
- Möglichkeit für mehr Fragen, entsprechend mehr Details
- Bringen eigene Materialien mit
- Mehr Vertrauen und Nähe zu Interviewer*innen
- Terminfindung deutlich erleichtert
- Weniger Zeitaufwand für die Ärzt*in (keine Anfahrt, kürzere Interviewdauer)
- mehr logistischer Zeitaufwand für die Moderator*in (Anfahrt)
- Auswertung einfacher

Fokusgruppe:

- Interviewdauer ca. 120 Minuten
- Geeignet, um schnell eine Vielzahl an Meinungen einzuholen
- Austausch und Diskussion von Standpunkten und Meinungs-verschiedenheiten
- Moderator*in stärker gefordert, durch die Session zu führen
- Auf Untersuchungsgegenstand konzentriert (hier: WSV). Mit eigenen Beispielen dargestellt, aber weniger lange ausgeführt
- Unterschiedliche Sprechdauer der Teilnehmenden
- Auswertung komplexer als bei Interviews
- Einbringen bisheriger Interviewerkenntnisse in die Fokusgruppen

Online-Befragung:

- Leichtere Terminfindung als in Präsenz
- Kleinere Gruppen als in Präsenz (≤ 5), da schwierigere Einzelansprache der Teilnehmenden
- Kostengünstiger, da keine Versorgung/Anfahrt
- Vertrauen schwieriger aufzubauen als in Präsenz
- Vollständige Studienunterlagen schwerer zu erhalten; Aufwandsentschädigung als Anreiz

Verwendete Literatur und weitere Literaturempfehlungen

Helferich C. Leitfaden- und Experteninterviews. Handbuch Methoden der empirischen Sozialforschung: Springer; 2019. p. 669-86.

Schreier M. Qualitative Erhebungsmethoden. Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor: Springer; 2013. p. 222-44.

Institut für ökologische Wirtschaftsforschung (IÖW). partizipativ innovativ. Toolbox für Unternehmen: Fokusgruppe 2020 [Available from: <https://www.partizipativ-innovativ.de/toolbox-fuer-unternehmen/?openModal=myModal41>].

Poscheschnik G. Qualitative Erhebungsmethoden. Empirisch forschen Die Planung und Umsetzung von Projekten im Studium (S 100–110) Konstanz ua: UVK/Huter & Roth Bodybuilder, Dandys, Kinderflüsterer. 2010.

Neumayr A, Eigenstuhler J, Baubin M. Qualitative Forschungsmethoden in der präklinischen Notfallmedizin. Notfall+ Rettungsmedizin. 2011;14(3):220-5. Pohontsch, N. J., et al. (2018). "Gruppendiskussionen in der Versorgungsforschung—Teil 1: Einführung und Überlegungen zur Methodenwahl und Planung." Das Gesundheitswesen 80(10): 864-870.

Xyländer, M., et al. (2019). "Gruppendiskussionen in der Versorgungsforschung—Teil 2: Überlegungen zum Begriff der Gruppe, zur Moderation und Auswertung von Gruppendiskussionen sowie zur Methode der Online-Gruppendiskussion." Das Gesundheitswesen 82: 998-1007.

Stamer, M., et al. (2015). "Qualitative Studien in der Versorgungsforschung—Diskussionspapier, Teil 3: Qualität qualitativer Studien." Das Gesundheitswesen 77(12): 966-975.

Mey, G. and K. Mruck (2020). Qualitative interviews. Handbuch Qualitative Forschung in der Psychologie, Springer: 315-335.

6. Zielverfeher

Bei der Prüfung der Wirtschaftlichkeit und der Qualität von Medikamenten-Verordnungen werden einzelne Ärzte oder Betriebsstätten auffällig, weil sie Ziele verfehlen oder mit Ihren Verordnungen außerhalb festgelegter Kontrollgrenzen liegen. Für prüfende Instanzen ist es von Interesse, mehr über diese Gruppe zu erfahren. Damit wird zum einen das Prüfverfahren auf seine Validität untersucht: wenn unter diesen Betriebsstätten ein hoher Anteil mit legitimen Praxisbesonderheiten zu finden ist, kann dies ein wertvolles Feedback für eine Optimierung des Prüfverfahrens sein. Zum anderen können auf der Basis solcher Analysen maßgeschneiderte Interventionen mit dem Ziel besserer Verordnungen entwickelt werden (z.B. spezifische Information oder Unterstützung).

Einführung

Bei der Untersuchung von auffälligem Ordnungsverhalten sind von Interesse:

1. Charakteristika der Verordner und Verordnungen
2. Hintergründe und Motive für auffälliges Ordnungsverhalten

In Deutschland gibt es verschiedene Steuerungs- und Prüfsysteme, die das ambulante Arzneimittelverordnungsverhalten betreffen. Während die vorgegebenen Ziele überwiegend erreicht und eingehalten werden, gibt es immer wieder Betriebsstätten, denen es auch längerfristig nicht gelingt, die Ziele zu erreichen, bzw. deren Verordnungen jenseits definierter Kontrollgrenzen liegen. Aus diesem Grund ist es interessant herauszufinden, ob diese Betriebsstätten und die darin arbeitenden Ärzte besondere Charakteristika aufweisen, die sie von den Zielerreichern unterscheiden. Ebenso interessant ist es, die entsprechenden Ärzte zu ihrem Ordnungsalltag zu befragen und herauszufinden, wodurch die Auffälligkeit begünstigt und beeinflusst wird (Perspektive der Betroffenen).

Charakteristika der Verordner und Verordnungen

Anhand deskriptiver Analysen von Ordnungsdaten lassen sich verschiedene Merkmale und ihre Häufigkeiten untersuchen. Nach der Festlegung eines Beobachtungszeitraums und einer Definition von Zielerreichung bzw. -verfehlung, können anhand der Routinedaten von den Kassenärztlichen Vereinigungen Datensätze erstellt werden, die Charakteristika zu Ärzten bzw. Betriebsstätten enthalten. Damit lassen sich die Häufigkeiten verschiedener Merkmalsausprägungen sowohl für die Zielerreicher als auch die Zielverfeher analysieren und vergleichen.

Für die Auswahl der Variablen und der zu untersuchenden Charakteristika wird ein Analyseleitfaden erstellt. Interessante Parameter stellen hier beispielsweise Alter, Geschlecht, Praxislage, Zusatzbezeichnungen dar. Diese Untersuchungen sollten auch nach Fachgruppen aufgeschlüsselt werden.

Anhand der erhaltenen Häufigkeiten lassen sich beispielsweise Tendenzen ableiten, ob in der Gruppe der Zielverfeher ein höherer Anteil von Betriebsstätten im städtischen oder ländlichen Sektor liegen oder auch, ob es sich eher um Betriebsstätten mit wenigen oder vielen Patienten handelt (Scheinzahl pro Quartal). Ebenso können derart geringe Unterschiede in den Häufigkeitsverteilungen vorliegen, dass die beiden Gruppen eher als vergleichbar anzunehmen sind.

In diesem Kontext dienen statistische Tests der Evaluation zufälliger Schwankungen bzw. Gruppenunterschiede. Zum Einsatz kommen die gängigen Verfahren bei 2-Gruppen-Vergleichen: Chi-Quadrat-Test bei kategorialen Variablen, t-Test bei normalverteilten stetigen Variablen

bzw. Mann-Whitney-U-Test als Alternative, wenn das Kriterium einer Normalverteilung nicht erfüllt ist. Hier ist zu beachten, dass bei kleinen Fachgruppen (=hoher Spezialisierungsgrad) Unterschiede meist nicht statistisch zu sichern sind (mit anderen Worten: auch große Unterschiede „statistisch nicht sigifikant“).

Neben Charakteristika von Ärzten und Betriebsstätten lassen sich auch die Verordnungen selbst untersuchen. Als Ausgangspunkt sollte vergleichend untersucht werden, wie breit das Verordnungsvolumen einzelner Fachgruppen ist, welche Vorgaben/Ziele den Fachgruppen vorgegeben sind und in welchem Ausmaß sie erreicht bzw. verfehlt werden.

Ein Detailvergleich von Zielerreichern und Zielverfehlern ist meist nur innerhalb einer Fachgruppe sinnvoll. Dieser sollte sich an relevanten Zielvorgaben orientieren, z.B. Leitsubstanz- oder Generika-Zielen. Meist lassen sich bestimmte Arzneimittel bzw. -gruppen herausarbeiten, bei denen das Ziel besonders häufig verfehlt wird.

Möglicherweise gibt es bestimmte Ziele und Arzneimittelgruppen, die in mehr als einer Fachgruppe für Probleme sorgen, da das Ziel nicht erreicht wird oder weil entsprechende konkrete Arzneimittel über mehrere Fachgruppen hinweg häufig verordnet werden. Gemessen an den DDD-Mengen lässt sich einschätzen, ob die jeweiligen Arzneimittel einen relevanten Anteil der Verordnungen ausmachen oder nicht. Ebenso kann untersucht werden, ob bestimmte Darreichungsformen, wie z.B. parenterale Injektionen, vermehrt problematisch sind.

Hintergründe und Motive für auffälliges Ordnungsverhalten

Um einen Eindruck zu erhalten, welche Faktoren eine längerfristige Auffälligkeit begründen und beeinflussen könnten, bietet es sich an, mit betroffenen Ärzten zu sprechen. Wichtig ist hierfür ein Vorgehen mit Diskretion, damit sich betroffene Ärzte, die mit ihrem eigenen Verhalten möglicherweise auch sehr unzufrieden sind, nicht noch „schlechter“ oder gar als „schwarzes Schaf“ abgestempelt fühlen. Deshalb empfiehlt es sich die Interviewführung an eine externe Institution wie Universitäten oder Marktforschungsinstitut abzugeben, um eine möglichst neutrale und geschützte Gesprächsatmosphäre zu schaffen. Einige Ärzte schämen sich wegen der Thematik und öffnen sich nur in vertrauter Atmosphäre; andererseits sind sie erfahrungsgemäß häufig dankbar dafür, dass das Thema angesprochen wird und sie sich damit nicht mehr allein gelassen fühlen.

Die zunächst intern (meist auf der Basis von Daten der KV) definierten Zielverfehrer sollten möglichst von einer unabhängigen Institution (z.B. Universität) angeschrieben werden. Die Teilnahme an der Befragung muss freiwillig sein, Vertraulichkeit im Umgang mit den Ergebnissen (Anonymisierung!) ist absolute Voraussetzung. Der Umgang mit den Erhobenen Daten ist ausführlich zu beschreiben.

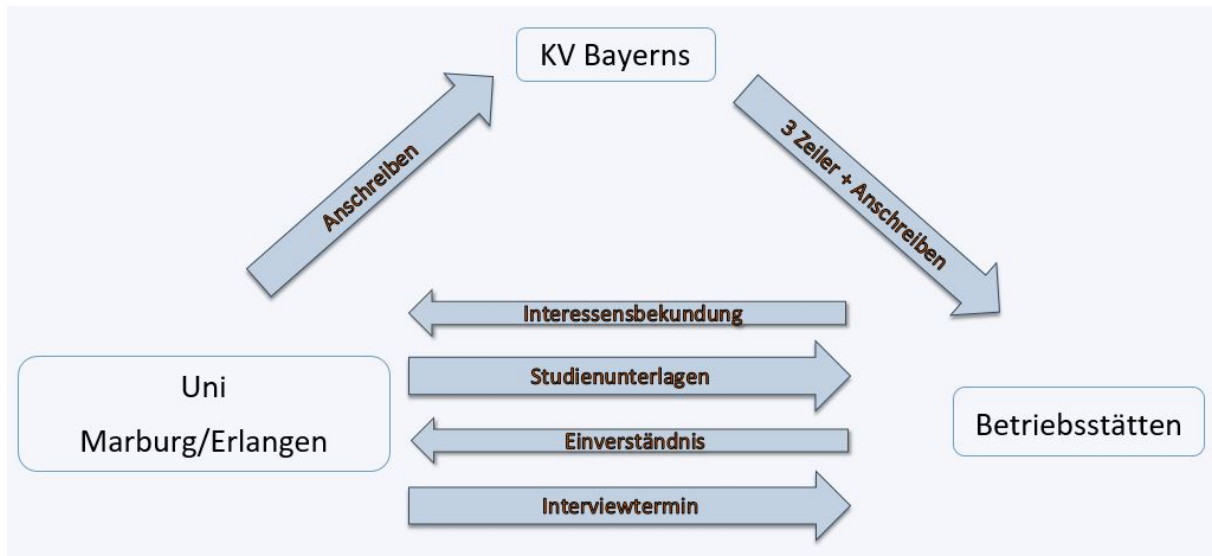


Abbildung: Rekrutierungsstrategie einer sensiblen Gruppe unter Wahrung des Datenschutzes am Beispiel des WirtMed-Projektes

Da es sich um sensible Themen mit starker individueller Betroffenheit handelt, empfiehlt sich das Format von Einzel-Interviews, im Gegensatz zu Fokusgruppen, wo sich die Betroffenen vermutlich zu sehr exponiert fühlen. Für die zu führenden Interviews ist ein Leitfaden zu entwickeln. Gerade bei sensiblen Themen empfiehlt es sich, mit allgemeinen Themen zu beginnen, mit denen man das Gespräch anregen und Vertrauen aufbauen kann. Über allgemeine Fragen zum Praxisalltag, Verordnungsvorgängen, Einflussfaktoren auf die Arzneimittelauswahl und patientenbezogene Aspekte sowie ein Fallbeispiel (jeweils auf die Fachrichtung bezogen) kann sich den gesundheitspolitischen Themen genähert werden. Hier geht es dann vor allem um Zufriedenheit und Probleme mit der aktuellen Verordnungssituation, sowie verschiedene Aspekte rund um das jeweilige Steuerungssystem (in diesem Fall die bayer. Wirkstoffvereinbarung). Gerade die Auswahl von konkreten Arzneimitteln oder Einflussfaktoren auf die Verordnungsauswahl lassen sich anhand von Fallbeispielen oder Mindmaps visualisieren und erarbeiten.

Themenbereiche

Allgemeine Informationen zu Person und Praxiserfahrung

Rolle als Arzt

Allgemeines zum Ordnungsverhalten

Patientenbezogenes Ordnungsverhalten/Fallbeispiele

Gesamtsituation

Bayer. Wirkstoffvereinbarung

„Up-to-date“-bleiben

Abschluss

Tab: Aufbau und Oberthemen des eingesetzten Interviewleitfadens

Für die Interviews im WirtMed-Projekt hat sich ein Tandem-Prinzip bewährt: eine Interviewerin hatte einen Hintergrund in Pädagogik/Soziologie/Psychologie, eine andere war Apothekerin. Die Interviews wurden mit Audiogeräten aufgezeichnet und extern transkribiert. Im Anschluss erfolgte eine Codierung mit dem Programm MAXQDA. Dabei wurde gemischt induktiv-deduktiv vorgegangen. Anschließend wurde das codierte Material code-bezogen abstrahiert und jeweilige Schlussfolgerungen gezogen (qualitative Inhaltsanalyse nach Mayring – hier kommen zahlreiche andere Methoden in Frage; siehe entsprechende Lehrbücher). Auf diese Weise können Ergebnisse erhalten werden zu: Motivation, problematischen Verordnungen, Wahrnehmungen, Einschätzungen und Meinungen sowie alternativen Vorschlägen der Ärzte.

Da es sich um qualitative Interviews mit einer ausgewählten Anzahl von Ärzten handelt, erhält man nicht unbedingt für die Allgemeinheit repräsentatives Bild. Allerdings erhält man wertvolle Einblicke und Hinweise, wo Probleme und Barrieren gegen eine wirtschaftliche Verordnung liegen können.

Interpretation und Kombination beider Studienarme

Die beiden analytischen Zugänge lassen sich miteinander kombinieren. Die quantitativen Daten über die Charakteristika der Zielverfehrer im Vergleich zu den Zielerreichern können Hinweise darauf geben, ob die im qualitativen Interview geäußerten Erfahrungen und Zusammenhänge nur eine individuelle bzw. idiosynkratische Problematik darstellen oder sie doch eine breitere Bedeutung haben. Die qualitative Analyse kann Hinweise geben auf Fragestellungen, die mit einem quantitativen Vergleich der beiden Gruppen vertiefend untersucht werden können.

Kombiniert man Ergebnisse beider Zugänge im Sinne einer Mixed methods Analyse, lassen sich verschiedene Faktoren für Auffälligkeiten auf verschiedene Ebenen herausarbeiten. Dafür bieten sich deskriptive Charakteristika von Arzt und Betriebsstätte an, ebenso wie berichtete Einflussfaktoren aus Sicht der Ärzte. Insgesamt stellen sich so vier verschiedene Ebenen mit Einfluss auf das Ordnungsverhalten heraus, von den drei beeinflussbar bzw. veränderbar sind. Faktoren wie Case-Mix, Praxislage etc. sind eher als fix anzusehen und lassen sich in diesem Kontext kaum verändern bzw. bieten keinen Ansatzpunkt für Verbesserungen und Veränderungen.

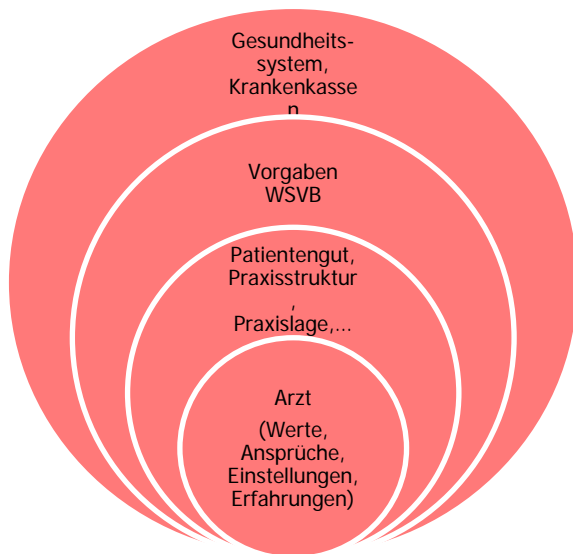


Abbildung: Ebene und Einflussstärke von Faktoren auf die Auffälligkeit

Da jedes Steuerungsinstrument und Regularium einen Teilaspekt des Gesundheitssystems betrifft und auch nur einen bestimmten Bereich des Praxisalltags beeinflusst, werden neben dem Instrument an sich auch andere Ebenen angesprochen und betroffen, die sich auf verschiedenste Weise gegenseitig bedingen und beeinflussen. Für die Evaluation von Teilaspekten im Gesundheitssystem sollte dies immer Bedacht und in die Analyse sowie Auswertung mit einbezogen werden.

7. Individuelles Profiling

Hintergrund und Definition

Die herkömmliche Wirtschaftlichkeitsprüfung nach Richtgrößen, aber auch die meisten anderen Verfahren, sind ein Beispiel für das international so bezeichnete „Profiling“. Darunter versteht man die Identifizierung von Leistungserbringern (Krankenhäuser, Abteilungen, Praxen, Ärzte/innen) mit „auffälliger“ Leistung im Vergleich zu anderen Einheiten oder zu einem Referenzstandard. Mit „auffällig“ sind sowohl positive (bessere Ergebnisse) als auch negative Abweichungen (schlechtere Ergebnisse) gemeint. Dabei ist durchaus nicht unumstritten, wie eine solche Auffälligkeit zu definieren und zu erfassen ist (Christiansen & Morris, 1997). Wenn möglich, sollte eine angemessene Leistung in einem Intervall definiert werden, welches sich aus inhaltlichen und nicht nur aus statistischen Gesichtspunkten herleiten lässt (Burgess et al., 2000). Wegen der damit verbundenen Schwierigkeiten wird allerdings meist nach reinen Verteilungskriterien vorgegangen.

Profiling kann in einem formativen und einem summativen Sinne genutzt werden. Leistungserbringer können mit internen Verfahren Schwachstellen identifizieren und entsprechende Konsequenzen ziehen. Informationen aus Profilinganalysen könnten in klinische Behandlungspfade und Behandlungsprotokolle eingehen und damit die Versorgungssituation verbessern. Dieser formativen Intention steht die summative gegenüber (auch: regulatorisch). Hier prüfen externe Institutionen die Performance von Individuen oder Einheiten in Bezug auf die Einhaltung von Standards. Die Ergebnisse können zu formalen (Entzug einer Zulassung o.ä.), finanziellen (Strafzahlungen, „Regress“) oder edukativen (obligatorische Fortbildung o.ä.) führen.

Profiling kann in Analogie zu einem Screening-Test verstanden werden, bei dem sowohl „falsch positive“ als auch „falsch negative“ Ergebnisse auftreten können, und bei dem im Falle eines positiven Befundes eine „weiterführende Diagnostik“ notwendig ist. Wenn beispielsweise eine Praxis bei der statistischen Wirtschaftlichkeitsprüfung auffällig wird, aber gut begründete Praxisbesonderheiten geltend machen kann, würde es sich hier um einen solchen „falsch positiven“ Befund des Screenings handeln.

Mögliche, im Rahmen eines Profiling zu untersuchende Größen ergeben sich auf der Patienten-, der individuellen Leistungserbringer- und der Organisationsebene (Tucker, 2000). Es können Prozesse, Ergebnisse (klinische Outcomes) und Kosten erfasst werden.

Die meisten international publizierten Beispiele befassen sich mit der Krankenhausmortalität. Die Ergebnisqualität in verschiedenen Institutionen kann aufgrund zufälliger Variationen, systematischer Unterschiede in der Behandlung oder durch Unterschiede in den Patientenpopulationen vor der Behandlung differieren (Marshall & Spiegelhalter, 2001). Daher ist es wichtig, für den case-mix zu adjustieren, der u.a. solche Faktoren wie Alter, Geschlecht, Diagnose, Schwere der Erkrankung umfasst. Marshall und Spiegelhalter konnten zeigen, dass der Vergleich von Chirurgen hinsichtlich kardialer arterieller Bypass-Operationen mittels eines Zwei-Level-Modells nach dem Ansatz von Bayes mit Adjustierung für den case-mix deutlich angemessenere Beurteilungen bezüglich der individuellen Mortalitätsraten ergab als eine herkömmliche logistische Regression (Marshall & Spiegelhalter, 2001). Burgess et al. gelang es, ein entsprechendes Modell für unterschiedliche Krankenhäuser hinsichtlich der Wiedereinweisungen auf die Intensivstation zu entwickeln (Burgess et al., 2000).

Quellen von Variabilität

Für die Prüfung (Prozesse, Outcomes) einer großen Zahl von Leistungserbringern (z.B. Bezirk oder Land) lassen sich in Anlehnung an Marshall et al. (Marshall et al., 2004) folgende Quellen von Variabilität ausmachen:

- A – Zufallsschwankungen: inter- und intraindividuelle Schwankungen (intraindividuell: Schwankungen einer Einheit über mehrere Zeitpunkte)
- B – Bekannte Quellen von Variabilität, die zu Unterschieden zwischen Einheiten, aber auch zu Schwankungen im Zeitverlauf führen können:
 - Demografische Charakteristika von Patienten (Alter, Geschlecht, soziale Schichtzugehörigkeit)
 - Morbidität (case mix)
- C1 – Nicht gemessene (vielfach nicht erfassbare), meist gering wirksame Faktoren, in denen sich Einheiten unterscheiden; in Kombination können diese im Einzelfall zu großen Abweichungen im Outcome führen. Wenn von Dauer, können sie zu zeitlichen Trends (Veränderungen) im Outcome der gesamten Gruppe führen.
 - Schwankungen in der Datenqualität
 - Unterschiedliches Inanspruchnahme-Verhalten
 - Nicht erfasste Morbidität
- C2 – Systematische, (zunächst) nicht erfasste (modellierete) Quellen von Variabilität, die nur einzelne oder eine Minderheit von Einheiten betreffen:

- Unberücksichtigte Case-mix-Faktoren, wie z. B. Praxisschwerpunkte (HIV, Onkologie, Diabetologie) oder ein oder mehrere Altenheime in der versorgten Population
- Ausschläge in der Performance (positiv oder negativ): die Identifizierung dieser Faktoren ist das eigentliche Ziel des Profiling (einschließlich der Wirtschaftlichkeits-Prüfung)

E – Extreme externe Faktoren: diese erfassen die gesamte untersuchte Gruppe, dazu gehören:

- Saisonale Einflüsse (bei den hier untersuchten, überwiegend für chronische Erkrankungen verschriebenen Medikamenten vermutlich nur geringer Effekt – was aber zu untersuchen ist [Asthma & COPD; KHK; Diabetes])
- Epidemien
- System-Veränderungen: dazu gehören Änderungen von Regelungen des regionalen oder nationalen Gesundheitssystems, aber auch die Zulassung neuer Medikamente [unabhängig von ihrer Beurteilung als wirtschaftlich oder nicht]

Unter A fällt ein großer Teil der Faktoren, welche die Komplexität des Versorgungshandelns im Gesundheitswesen ausmachen, wie z.B. die Dynamik zwischen Patient und Arzt. Nicht erfasst, und damit hier subsumiert, sind die Beziehungen handelnder Einheiten (Leistungserbringer) untereinander.

Vor dem Hintergrund der Zielsetzung einer Wirtschaftlichkeits-Prüfung sind alle Quellen von Variabilität „legitim“, bis auf einen Teil der unter C2 (zweiter Spiegelstrich) genannten. Diese Übersicht verdeutlicht die Schwierigkeiten, welche ein Profiling bzw. jegliche Wirtschaftlichkeitsprüfung anhand von Versorgungsdaten einer Region bzw. eines Landes überwinden muss; die untersuchte Zielvariable (hier: Medikamenten-Verschreibung) wird nicht nur von der eigentlich interessierenden Größe, sondern von zahlreichen anderen Quellen von Variabilität beeinflusst, die bei der Auswertung zu isolieren sind bzw. für die adjustiert werden muss.

Exkurs Outlier/Extremwerte

An dieser Stelle soll auf die Problematik von Outliern bzw. Extremwerten aufmerksam gemacht werden, die für Sekundärdaten der hier diskutierten Art charakteristisch sind und einen methodisch angemessenen Umgang bei der weiteren Datenanalyse erfordern (Swart et al., 2014). Zunächst ist zu erwähnen, dass ein herkömmliches Prüfen auf Normalverteilung bei Routinedaten wie hier betrachtet als nicht zielführend erachtet wird, da bereits geringe Abweichungen von der Normalverteilung bei derartigen Stichprobengrößen wie den bei Routinedaten zu erwartenden als Verletzung derselben angezeigt werden (Micceri, 1989).

Von „Outliern“ sprechen wir, wenn in einer empirisch vorgefunden Verteilung eines Merkmals einzelne Beobachtungen stark vom übrigen Feld abweichen. Oft drängt sich der Eindruck auf, dass diese aus inhaltlichen oder methodischen Gründen von der Gruppe abweichen, auf welche sich das eigentliche Forschungsinteresse richtet. Sie werden deshalb bei deskriptiven oder analytischen Fragestellungen häufig aus der Analyse (z.B. Hypothesentestung) ausgeschlossen.

Bei der Prüfung des Verhaltens von Leistungserbringern in Sekundärdaten sind solche Ausreißer jedoch vieldeutig; folgende Erklärungen kommen hier in Frage:

1. Datenverarbeitung: bei der Erfassung oder Verarbeitung der Daten haben sich Fehler ergeben.

2. Eine Betriebsstätte gehört nicht zur interessierenden Grundgesamtheit: beispielsweise wird eine Betriebsstätte administrativ der hausärztlichen Versorgung zugeordnet, betreut aber Patienten aus einem spezialärztlichen Segment (Allergologie o.ä.).
3. Confounder: das Versorgungsspektrum (case mix) weist Besonderheiten auf, für die eine Adjustierung anzustreben ist („Praxisbesonderheiten“).
4. Unwirtschaftliches Ordnungsverhalten bzw. Abweichung vom Medikationskatalog: die Betriebsstätte weist das Verhalten auf, welches tatsächlich das Ziel der regulatorischen Analyse ist.
5. Es handelt sich um Extremwerte, die „noch“ zur Verteilung, d.h. dem Korridor des akzeptablen Verhaltens, gehören (Ieva & Paganoni, 2015).

Regulatorische Analysen im Sinne eines Profiling legen per definitionem ihren Fokus auf die Identifizierung von Ausreißern. Der Interpretation des Phänomens wie auch den Verfahrensweisen im Umgang damit ist wegen der möglicherweise gravierenden Konsequenzen besondere Aufmerksamkeit zu widmen; im Gegensatz dazu ist der Umgang mit Ausreißern etwa beim Testen von Assoziationen weniger kritisch.

Zur Identifikation von Outliern kann die Median absolute deviation (MAD) verwendet werden (Leys et al., 2013):

$$MAD = b(|x_j - (M_j; x_j)|)$$

Dabei stellt b eine Konstante für die Normalverteilung dar (1.4826), x_j ist der individuelle Wert, $M_j; x_j$ der Median der jeweiligen Variablen. Konkret wird der Median der absoluten Abweichungen der individuellen Messwerte vom Median mit der Konstanten multipliziert. Dieses Ergebnis fungiert wiederum im Weiteren als Nenner, während im Zähler für jeden individuellen Wert die absolute Abweichung vom Median steht. Resultierende Werte >2.5 werden formal als Outlier definiert.

Funnel Plots

Funnel Plots werden als graphische Veranschaulichung zum Vergleich der Performance von Institutionen oder auch einzelnen Leistungserbringern im Gesundheitswesen empfohlen (Spiegelhalter, 2005b). Diese Darstellung besteht aus vier Komponenten (van der Willik et al., 2021):

- Indikator: dies ist die interessierende quantitative Größe (Prozesse, Outcomes, Kosten usw.), abzutragen auf der y -Achse. Der Indikator sollte relevant, messbar, veränderbar und mit der Qualität der Gesundheitsversorgung assoziiert sein, und es muss Unterschiede zwischen den Betriebsstätten geben.
- Target: damit ist die normative Erwartung für die untersuchte Population gemeint; diese wird häufig mit dem Mittelwert (oder einem anderen Lagemaß) operationalisiert; eine Alternative stellen extern vorgegebene Ziele dar (z.B. Zahl der Infektionen mit multiresistenten Keimen pro Zeiteinheit; Anteil „rot“ des Medikationskataloges). Die Auswahl eines geeigneten Referenzstandards kann eine Herausforderung sein, da der Referenzstandard ein fairer und praktikabler Vergleichsmaßstab für alle Einheiten (z.B. Betriebsstätten) sein muss.

- Präzisions-Parameter: ein sinnvolles Maß für die Genauigkeit einer Schätzung ist die Zahl der behandelten Patienten (alternativ: Prozeduren [Verordnungen]). Hierbei ist auf eine ausreichende Stichprobengröße und auf mögliche Selektionseffekte zu achten; diese werden auf der x-Achse dargestellt.
- Kontrollgrenzen: Dabei handelt es sich um Grenzen, außerhalb derer eine Einrichtung oder ein Leistungserbringer als „auffällig“ betrachtet wird. Diese können statistisch (etwa entsprechend einer Irrtumswahrscheinlichkeit P) oder inhaltlich definiert sein.

Sind Variablen bekannt, welche mit dem zu untersuchenden Indikator assoziiert sind, kann mit einer indirekten Standardisierung eine Adjustierung vorgenommen werden (Spiegelhalter, 2005b). Bei einem binären Indikator (z.B. Anteil der bei einer Operation verstorbenen Patienten) wird mit Hilfe einer logistischen Regression für jede Einheit die erwartete Zahl von Ereignissen berechnet. Diese kann mit den tatsächlich beobachteten Ereignissen verknüpft werden (Quotient oder Differenz). Durch Multiplikation der standardisierten Rate (Quotient) mit der erwarteten Rate ergibt sich für jede Einheit ein adjustierter Wert des Indikators.

Verburg et al. beschreiben aktuelle Leitlinien zur Erstellung von Funnel plots für Qualitätsindikatoren (Verburg et al., 2018), die sie in 6 Schritte unterteilen. Im ersten Schritt geht es um die Festlegung des Inputs auf inhaltlicher Ebene. Die Entscheidungen auf inhaltlicher Ebene sind Entscheidungen über: (a) den Qualitätsindikator und die zugehörige (b) Datenquelle oder das Register und die Patientenpopulation, einschließlich Ein- und Ausschlusskriterien; (c) der Berichtszeitraum; (d) Kontrollgrenzen und die Frage, ob die Datenanalysten diese erweitern dürfen, um eine overdispersion zu korrigieren. Eine overdispersion liegt vor, wenn eine echte Heterogenität zwischen den Einrichtungen besteht, die über das erwartete Niveau der Variation aufgrund von Zufälligkeiten hinausgeht. Wenn der Wert des Referenzstandards aus den Daten gewonnen wird, empfehlen die Autoren, den Durchschnittswert des Qualitätsindikators über alle eingeschlossenen Patienten zu verwenden und nicht einen Durchschnitt über alle Betriebsstätten. Dies ignoriert allerdings die Korrelation innerhalb der Betriebsstätten und große Betriebsstätten haben einen hohen Einfluss auf das Ergebnis. Im zweiten Schritt erfolgt die Überprüfung der Qualität der für die Case-Mix-Korrektur verwendeten Vorhersagemodelle. Für binäre Outcomes werden der Brier Score und die C-Statistik angeführt, die der area under the curve (AUC) bei der ROC-Analyse entspricht. Im dritten Schritt erfolgt die Prüfung der Frage, ob die Anzahl der Beobachtungen pro Betriebsstätte ausreichend ist. Wichtig ist, durch eine ausreichende Stichprobengröße eine entsprechende statistische Power zu erreichen. Im vierten Schritt wird auf overdispersion getestet. Overdispersion tritt auf, wenn eine echte Heterogenität zwischen Betriebsstätten besteht, die über die zufällige zu erwartende hinausgeht. Die Ursache der Overdispersion ist unklar, aber Heterogenität kann entstehen, wenn Betriebsstätten Patienten mit unterschiedlichen Merkmalen behandeln, für die das Modell keine ausreichende Korrektur vornimmt; z.B. aufgrund von gesundheitspolitischen Entscheidungen oder Variabilität in der tatsächlichen Qualität der angebotenen Versorgung. Im fünften Schritt wird geprüft, ob die Werte der Qualitätsindikatoren mit institutionellen Merkmalen in Zusammenhang stehen. Wenn ein bedeutsamer Zusammenhang zwischen den Werten des Qualitätsindikators und Betriebsstättenmerkmalen besteht, empfehlen die Autoren, die Konstruktion des Funnel plots zu überdenken und eine Verbesserung der Case-Mix-Korrektur in Betracht zu ziehen oder separate Funnel plots für verschiedene Untergruppen von Betriebsstätten. Im sechsten Schritt werden spezifische technische Angaben zur visuellen Gestaltung des Funnel plots festgelegt.

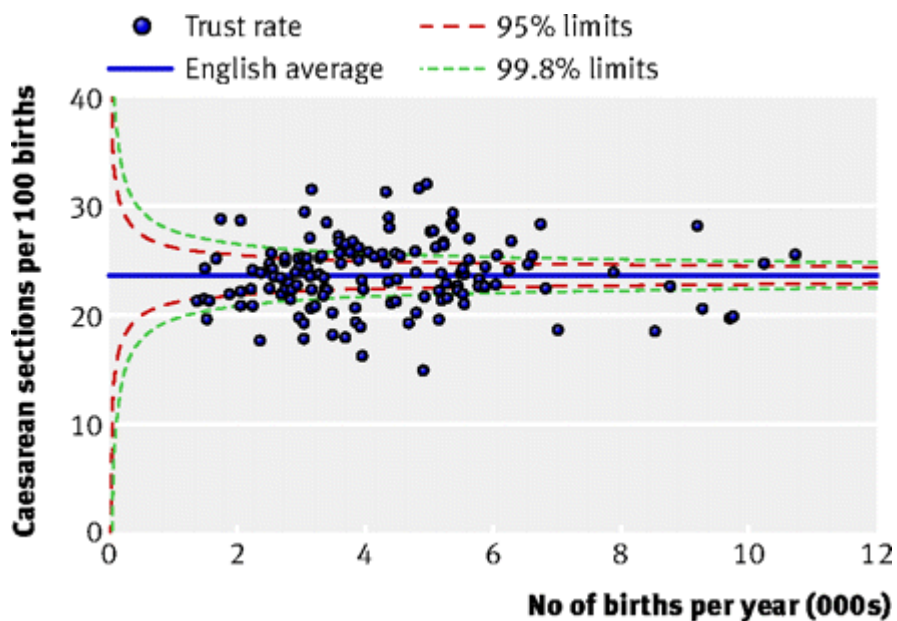


Abbildung: Funnel plot der Rate der Sectio Geburten pro 100 Geburten bei Frauen mit Ein-Kind-Geburten bezogen auf die Anzahl der Geburten pro Jahr in English NHS trusts in 2008, adjustiert für mütterliche Besonderheiten und klinische Risikofaktoren (Bragg et al., 2010).

Das Beispiel in Abbildung zeigt einen Funnel Plot, in dem die Anzahl der Geburten pro Jahr in Tausend der Anzahl der Sectio Geburten pro 100 Geburten gegenübergestellt ist. Die einzelnen Punkte repräsentieren unterschiedliche britische NHS Trusts. Kontrollintervalle auf der Grundlage unterschiedlicher Irrtumswahrscheinlichkeiten sind farbig aufgetragen. Die Trichterform der Verteilung ist hier eingeschränkt, da Einheiten mit einem Volumen von weniger als etwa 1700 Geburten pro Jahr nicht erfasst sind. Vielfach lassen sich in Grafiken dieser Art auch Zusammenhänge von Volumen und Indikator erkennen (Volumeneffekte).

Es zeigt sich, dass eine beträchtliche Anzahl von NHS Trusts sich außerhalb des oberen 99.8%igen Kontrollintervalls befinden und somit als auffällig klassifiziert würden. Dies lässt jedoch Zweifel an der Angemessenheit der Kontrollintervalle aufkommen (Spiegelhalter, 2005b). In diesem Fall ist anzunehmen, dass es sich um das Phänomen der Overdispersion handelt, d.h. relevante Kovariaten haben sich ausgewirkt, die bei der Adjustierung nicht berücksichtigt wurden bzw. werden konnten. Für den Fall, dass eine weitere Aufklärung der Variabilität des Indikators mit weiteren Kovariaten nicht möglich ist, sind verschiedene Verfahren zur Korrektur der Over-Dispersion vorgeschlagen worden (Spiegelhalter, 2005a).

Im vorliegenden Fall stellte sich heraus, dass diese Größe die Anzahl der Notfall-Sectio-Geburten war, so dass sich das oben dargestellte Bild entsprechend relativierte. Problematisch kann sein, dass es mehrere „kleinere“ institutionelle Faktoren geben kann, die zu einer solchen exzessiven Variabilität beitragen können, die an sich jedoch von fraglicher Relevanz für die untersuchte Fragestellung sein können (Spiegelhalter, 2005a). Es ist hier sinnvoll, sowohl die Ebene des Patienten als auch die Ebene der versorgenden Einheit zu betrachten.

Beim Profiling von Leistungserbringern in Bezug auf Gesundheitsoutcomes besteht häufig die Notwendigkeit, für Störvariablen (Confounder) zu adjustieren. Diese Variablen können Charakteristika der behandelten Patienten bezeichnen, wie z.B. das Alter oder die Krankheitslast;

von Bedeutung sind aber auch Eigenschaften der behandelnden Einheiten, wie z.B. geografische Lage, Alter und Geschlecht von Ärzten/innen, Ausstattung oder Behandlungsstile. Die Schwierigkeit liegt dabei in der inhaltlich korrekten Auswahl von Störvariablen. Wenn wir beispielsweise annehmen, dass die Symptombelastung ein Krankheitsmerkmal ist, das einen bestimmten Gesundheitszustand oder die Schwere einer Krankheit widerspiegelt, sollten wir diese als Confounder berücksichtigen. Es ist jedoch auch vorstellbar, dass die Symptombelastung durch Leistungserbringer beeinflusst werden kann (z.B. Ermutigung zur Inanspruchnahme auch bei geringfügigen Beschwerden) und daher auch als eine Folge der Qualität der Gesundheitsversorgung angesehen werden kann, so dass wir keine Korrektur vornehmen sollten (van der Willik et al., 2021). Entsprechend kann sich die Organisation einer Einheit (z.B. Gemeinschafts- statt Einzelpraxis) positiv auf wirtschaftliche oder Qualitäts-Indikatoren auswirken. Auch hier sollte nicht adjustiert werden, da entsprechende Bemühungen von Leistungserbringern um Qualität bzw. Wirtschaftlichkeit konterkariert würden.

Vor dem Hintergrund der großen Variabilität und Komplexität von Verhalten und Outcomes, ist die Aufklärung von Varianz kein realistisches Ziel. Trotzdem muss versucht werden, in statistischen Modellen die verschiedenen Quellen von Variabilität so weit wie möglich zu berücksichtigen. So ist weiterhin davon auszugehen, dass etwa die Patienten einer Einrichtung („Cluster“) einander stärker in Bezug auf interessierende Größen ähneln als solche über die Grenzen einer Einrichtung hinweg. Damit wird eine grundsätzliche Annahme der meisten konventionellen statistischen Verfahren verletzt, nämlich die Unabhängigkeit der Beobachtungen einer untersuchten Stichprobe. – Mit der „Ecological Fallacy“ wird in der Epidemiologie die Diskrepanz von Befunden (Zusammenhängen) zwischen einer Analyse von größeren Einheiten („ökologische Analyse“) einerseits und einer Analyse auf der Ebene von Individuen andererseits bezeichnet. Diese Gefahr besteht auch beim Profiling, wenn aggregierte Maße (in Bezug auf Praxis, Krankenhaus usw.), in eine Analyse eingebracht werden.

In einem einfachen Modell, bei dem für jede Einheit (z.B. Betriebsstätte) eine Abweichung von einem gemeinsamen Mittelwert untersucht wird, wird für jede der gleiche normative Maßstab als gültig angenommen. Das wäre dann angemessen, wenn eine perfekte Adjustierung für alle relevanten Kovariaten (Confounder) möglich wäre – eine unrealistische Annahme. Eine Alternative besteht in der Verwendung sog. Random Effect Modelle, bei denen eine Schwankung um den Mittelwert der Grundgesamtheit durch unbekannte Faktoren als „normal“ angenommen und im statistischen Modell berücksichtigt wird (Jones & Spiegelhalter, 2011).

Vor diesem Hintergrund sind statistische Modelle entwickelt worden, welche verschiedene Niveaus der Aggregation von Beobachtungen erfassen (sog. „hierarchische“, oder „Multilevel“- Analysen von „genesteten“ Einheiten oder „Clustern“). Für Profiling-Analysen in Zusammenhang mit der Wirtschaftlichkeits-Prüfung sind folgende Niveaus von Interesse:

- Verordnung
- Patient
- Betriebsstätte

Für hierarchische Analysen im Kontext des Profiling und die Adjustierung für Confounder auf mehreren Ebenen schlagen Normand et al. 1997 (Normand et al., 1997) eine Erweiterung des logistischen Modells vor.

Statistische Modelle dienen der Informationsreduktion von Daten, die aus der Beobachtung einer Gruppe von Einheiten gewonnen worden sind. Sie bewegen sich grundsätzlich zwischen dem Ziel, diese Daten möglichst genau abzubilden („goodness of fit“). Dieses Ziel wäre etwa

durch ein hochkomplexes Modell erreicht, das jede einzelne Beobachtung durch einen entsprechenden Parameter berücksichtigt. Solche Modelle erweisen sich jedoch meist als ungeeignet, Aussagen über Fälle zu machen, die nicht in der Ableitungs-Stichprobe vertreten waren, d.h. sie sind wenig robust (das sog. „Bias-Variance-Dilemma“). Im Kontext des Profiling ergibt sich außerdem ein weiteres Paradox: hier besteht das Ziel darin, in Bezug auf ein Qualitäts- oder Kostenkriterium divergente Einheiten zu identifizieren. Je eher jedoch deren Ergebnisse in die Formulierung des statistischen Modells eingehen, desto weniger wird es gelingen, problematische Leistungserbringer zu identifizieren (Ohlssen et al., 2007). Die Konsequenz daraus besteht in der Auswahl einfacher, robuster Modelle, die von den vorhandenen Daten nicht zu sehr beeinflusst werden.

Vielfach wird bei der vergleichenden Darstellung von Outcomes mit rangbasierten Modellen gearbeitet („league tables“). Diese sind in unserem Zusammenhang jedoch nicht angemessen, als dass sie keine Aussage treffen, ob eine Performance tatsächlich unangemessen ist, da sie lediglich einen relativen Vergleich abbilden (Ohlssen et al., 2007). Würde dieses Prinzip ernstgenommen, würde regelmäßig ein gewisser Anteil der untersuchten Leistungserbringer einer näheren Analyse oder gar Sanktion zugeführt werden.

Genauere Software-Algorithmen zur Umsetzung der Funnel Plot-Analysen in Form von R codes aus der Publikation von Ieva und Paganoni können bei den Autorinnen angefragt und für eigene Analysen verwendet werden (Ieva & Paganoni, 2015). In weiteren Publikationen werden ebenso R codes zur Erstellung von Funnel plots im Rahmen von Profiling-Analysen zur Verfügung gestellt (Kuhrij et al., 2021; Verburg et al., 2018). Des Weiteren existiert ein spezielles R package „FunnelPlotR“ (<https://rdrr.io/cran/FunnelPlotR/>), welches die Methoden von Spiegelhalter implementiert (Spiegelhalter, 2005b).

Beispiele

Auf einer höheren Datenaggregationsebene (Summierung und Mittelwertbildung innerhalb von Variablen auf Ebene der Betriebsstätten) können herkömmliche Regressionsverfahren bei der Erkennung auffälliger Performanz sinnvoll verwendet werden. Beim Profiling von britischen Krankenhäusern anhand von Sekundärdaten hinsichtlich standardisierter Mortalitätsraten (SMR) für das akute Koronarsyndrom zeigten sich zwischen den Kliniken kaum Unterschiede zwischen einer hierarchischen (multilevel) Datenanalyse und einer nicht-hierarchischen Datenanalyse, die sämtliche Patienten in eine Stichprobe überführt (Bottle et al., 2013). Die Autoren empfehlen, für Verlegungen, Komorbidität (Charlson-Index) und die Interaktion zwischen Komorbidität und Alter zu adjustieren. Des Weiteren sollten die Todesfälle nach Entlassung in das Modell aufgenommen werden. Ferner ist zu berücksichtigen, dass auch für den case mix zu adjustieren sein dürfte. Es wurde der Quotient aus der beobachteten und der erwarteten Mortalitätsrate gebildet, diejenigen Krankenhäuser als auffällig identifiziert, die sich außerhalb von 99.8% der Verteilung dieses Quotienten befinden und ein funnel plot zur graphischen Veranschaulichung angefertigt. Auch in einer weiteren Studie, die explizit die logistische Regression mit der Multilevelanalyse bei der Vorhersage von Sterbefällen während Krankenhausbehandlungen verglichen hat, konnte die Angemessenheit von Regressionen auf Patientenebene bei der Entdeckung von auffälligen Krankenhäusern anhand der krankenhausspezifischen standardisierten Mortalitätsrate festgestellt werden (Alexandrescu et al., 2011). Die meisten der Variablen, die durch die logistischen Regressionsanalysen als signifikante Prädiktoren ausgewählt wurden, fanden sich auch als solche in den Multilevelanalysen wieder. Die Anzahl der positiv wie negativ als auffällig erkannten Krankenhäuser unterschied sich nicht wesentlich zwischen beiden Methoden. Andere Autoren präsentieren Analysen, nach denen

die Kombination aus funnel plots und hierarchischen Modellen die geeignete Methode darstellt, um auffällige Krankenhäuser bezüglich der Sterblichkeit nach akutem Myokardinfarkt während der stationären Behandlung zu erkennen (Ieva & Paganoni, 2015).

Bernard et al. betrachteten französische Krankenhäuser hinsichtlich der Qualität der Pankreatektomie im Zeitraum 2009-2018 mittels Funnel plots (Bernard et al., 2021). Zwischen den Zeitabschnitten 2009-2014 und 2015-2018 zeigten sich keine bedeutsamen Unterschiede in den standardisierten Mortalitätsraten (SMR) und die Anzahl der Kliniken, die sich außerhalb der Kontrollgrenzen befand, war in beiden Zeiträumen ähnlich. Es wurde hier geschlussfolgert, dass die Pankreaschirurgie in Frankreich einer Neuorganisation bedarf.

Ein weiterer Anwendungsbereich stellt die Risikokommunikation bei Krebserkrankungen dar. Die erwarteten Erkrankungen wurden der standardisierten Inzidenzrate (SIR) in Gemeinden um Palermo gegenüber gestellt, so dass unmittelbar auffällige Gebiete identifiziert und einer weiteren Untersuchung zugeführt werden konnten (Mazzucco et al., 2017).

Beim Kostenprofiling bei ambulant tätigen Ärzten in den USA wurde der Frage nachgegangen, inwieweit der Fokus auf den individuellen Arzt oder Arztgruppen gelegt werden sollte, die sich zusammengeschlossen haben, um Ressourcen zu teilen und gemeinsam abzurechnen (Mehrotra et al., 2010a). Es konnte gezeigt werden, dass die Reliabilität individueller Kostenprofile unzureichend ist und zu Fehlklassifikationen von mehr als 20% führt (Adams et al., 2010). Empirisch konnte gezeigt werden, dass die Zusammenschlüsse mit einer größeren Anzahl von Ärzten (ab 4-9) deutlich bessere Reliabilitäten in ihren Kostenprofilen aufwiesen als Einzelpraxen (Mehrotra et al., 2010a). Innerhalb der Gruppen waren die Kosten jedoch heterogen verteilt. Dieselbe Forschergruppe konnte demonstrieren, dass die Anwendung unterschiedlicher Zuschreibungsregeln bei der Kostenzuweisung zu sehr unterschiedlichen Einstufungen einzelner Ärzte führt (Mehrotra et al., 2010b).

Profiling wurde ebenfalls eingesetzt, um die Verordnung von Opioiden, Benzodiazepinen, Stimulanzien und Appetitzüglern durch mehrere Ärzte an einen Patienten in einem 30-Tages-Intervall auf der Basis des Controlled Substance Utilization Review and Evaluation System (CURES) in Kalifornien aufzudecken (Wilsey et al., 2010). Dabei konnte gezeigt werden, dass die Wahrscheinlichkeit, multiple Verordner aufzusuchen („doctor shopping“) mit der Anzahl der gleichzeitig konsumierten Medikamente deutlich anstieg. Jüngeres Alter war assoziiert mit multiplen Verordnern für Opioide und Benzodiazepine, ältere Personen wiesen multiple Verordner eher bei Stimulanzien und Appetitzüglern auf.

Bei der Verwendung des Indikators (γ -Achse) hat offensichtlich dessen Skalierung Einfluss auf die erhaltenen Ergebnisse. Weniger Krankenhäuser wurden als outlier identifiziert, wenn eine kontinuierliche Variable dichotomisiert wurde, im Vergleich zu einem als kontinuierlich belassenen Indikator, z.B. dem Median (Kuhrij et al., 2021). Dies macht den Informationsverlust deutlich, wenn eine kontinuierliche Variable in ein geringeres Skalenniveau transferiert wird. Beide Funnel plots können wertvolle Informationen liefern, je nachdem, wo die Grenze zur Dichotomisierung gewählt wurde. Dadurch kann entschieden werden, ob die Leistung für alle Patienten oder beispielsweise für eine Extremgruppe verbessert werden sollte.

Indikatoren zur Kosteneffizienz sollten auf standardisierten Kostendifferenzen basieren, welche die Anzahl der Patienten, demographische Merkmale und den case mix berücksichtigen (im Gegensatz zum Quotienten der beobachteten/erwarteten Kosten) (Thomas et al., 2004).

Neben dieser eher graphischen Methode des Profiling existieren komplexere statistische Methoden in diesem Bereich. Dabei werden mittels Multilevelanalysen und Markov Chain Monte

Carlo-Verfahren Performanceindikatoren im Rahmen des Ansatzes nach Bayes geschätzt (Normand et al., 1997). Hierbei ist grundsätzlich zwischen einem estimation approach und einem hypothesis testing approach zu unterscheiden (Ohlssen et al., 2007). Ein Nachteil eines großen Teils der Literatur zur Bayes-Multilevelanalyse besteht darin, dass in dieser kaum auf die konkrete Anwendung eingegangen wird. Eine Ausnahme bildet bisher lediglich die Publikation von Finch, Bolin und Kelley (Finch et al., 2019), die konkrete Anleitungen für das Programm R darstellt.

Im folgenden Beispiel, welches auf dem Beispielcode des R package „FunnelPlotR“ und des Syntax von Ieva und Paganoni basiert (Ieva & Paganoni, 2015), soll deutlich gemacht werden, dass eine Verbindung der Analysen zur Korrektur von Overdispersion und Multilevelanalysen, die die Clusterstruktur der Daten (Patienten in Praxen bzw. Krankenhäusern) berücksichtigen, sofern diese als solche nicht-aggregierten Daten vorliegen, zu empfehlen ist.

Der Datensatz „medpar“ enthält die Daten von 1495 Patienten einer speziellen diagnostischen Gruppe aus 54 Krankenhäusern. Die erwartete stationäre Aufenthaltsdauer (Variable „los“ für „length of stay“) wurde als Präzisions-Parameter und der Quotient aus Dauer der stationären Behandlung und erwarteter Dauer der stationären Behandlung wurde als Indikator verwendet.

Zunächst wurde ein verallgemeinertes lineares Modell (glm) mit der Länge der stationären Behandlung als abhängiger Poisson-Variable gerechnet, wobei hier die Clusterung der Daten nicht berücksichtigt wird. Die Prädiktoren waren Mitglied einer Health Maintenance Organization (HMO; dichotom), Patient verstorben ja/nein (dichotom), Patient ist 80 Jahre und älter oder jünger (dichotom), Notaufnahme ja/nein (factor 2; dichotom), geplante Aufnahme ja/nein (factor 3; dichotom).

Die resultierenden Koeffizienten sind im Folgenden dargestellt:

```

glm
Coefficients:
      Estimate Std. Error z value      Pr(>|z|)
(Intercept)  2.26875   0.01246 182.011 < 0.0000000000000002 ***
hmo          -0.07637   0.02393  -3.192    0.00142 **
died         -0.24574   0.01826 -13.458 < 0.0000000000000002 ***
age80        -0.02141   0.02050  -1.045    0.29617
factor(type)2 0.24921   0.02099  11.871 < 0.0000000000000002 ***
factor(type)3 0.74869  0.02627  28.496 < 0.0000000000000002 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Das R^2 des Modells ist mit .104 als eher gering zu beurteilen. Alter 80 Jahre und älter oder jünger ist kein signifikanter Prädiktor für die stationäre Aufenthaltsdauer, während alle anderen Variablen eine statistische Signifikanz erreichen. Die absolute Höhe des Koeffizienten der Variable „Mitglied einer Health Maintenance Organization (HMO)“ ist allerdings als gering einzustufen. Den höchsten Beitrag leistet die Variable geplante Aufnahme (factor 3), wobei eine geplante stationäre Behandlung den stationären Aufenthalt um durchschnittlich 0.75 Tage verlängert. Dieser betrug im Durchschnitt 9.85 Tage (SD 8.83).

Der resultierende Funnel plot (Abbildung) ergibt 25 Krankenhäuser, die als outlier zu bezeichnen wären.

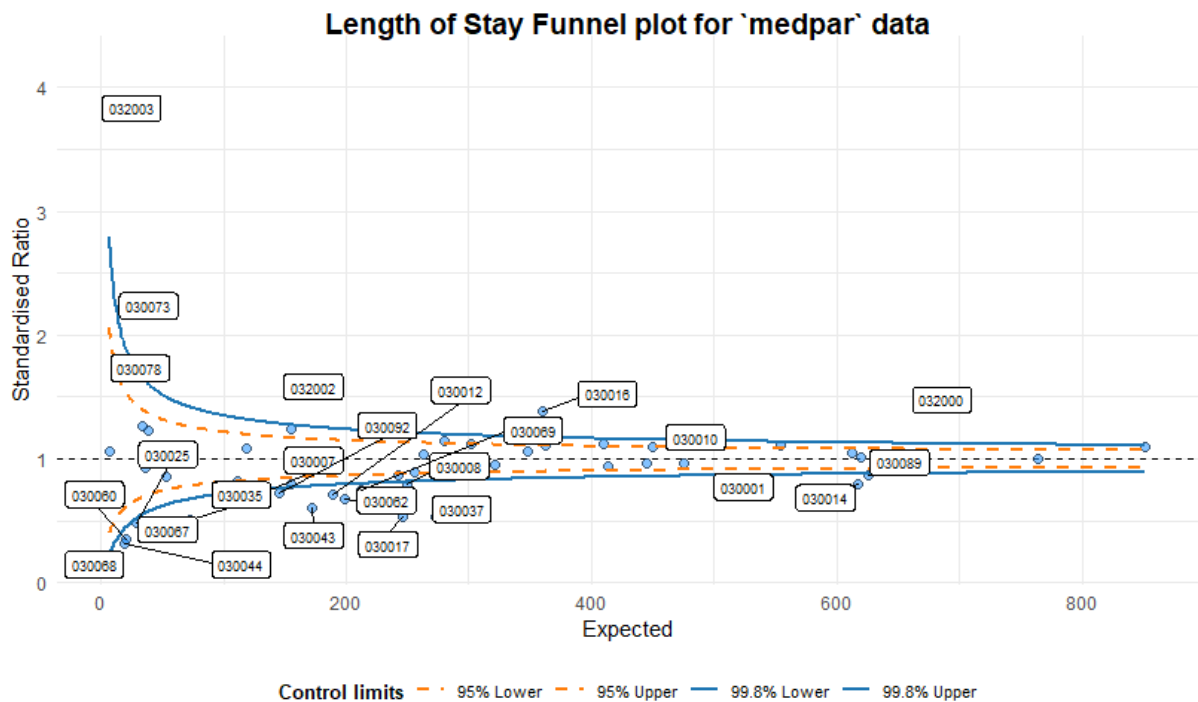


Abbildung: Funnel plot der erwarteten stationären Behandlungstage bezogen auf den Quotienten zwischen den stationären Behandlungstagen und den zu erwartenden stationären Behandlungstagen im Datensatz „medpar“.

Die Prüfung auf overdispersion ergab einen Wert von 6.24, der den Grenzwert von 1 somit deutlich überschreitet (Ieva & Paganoni, 2015). Daher sollte eine Korrektur für overdispersion vorgenommen werden. Das daraus resultierende Funnel plot (Abbildung 3) zeigt noch 9 Krankenhäuser, die als outlier zu bezeichnen wären.

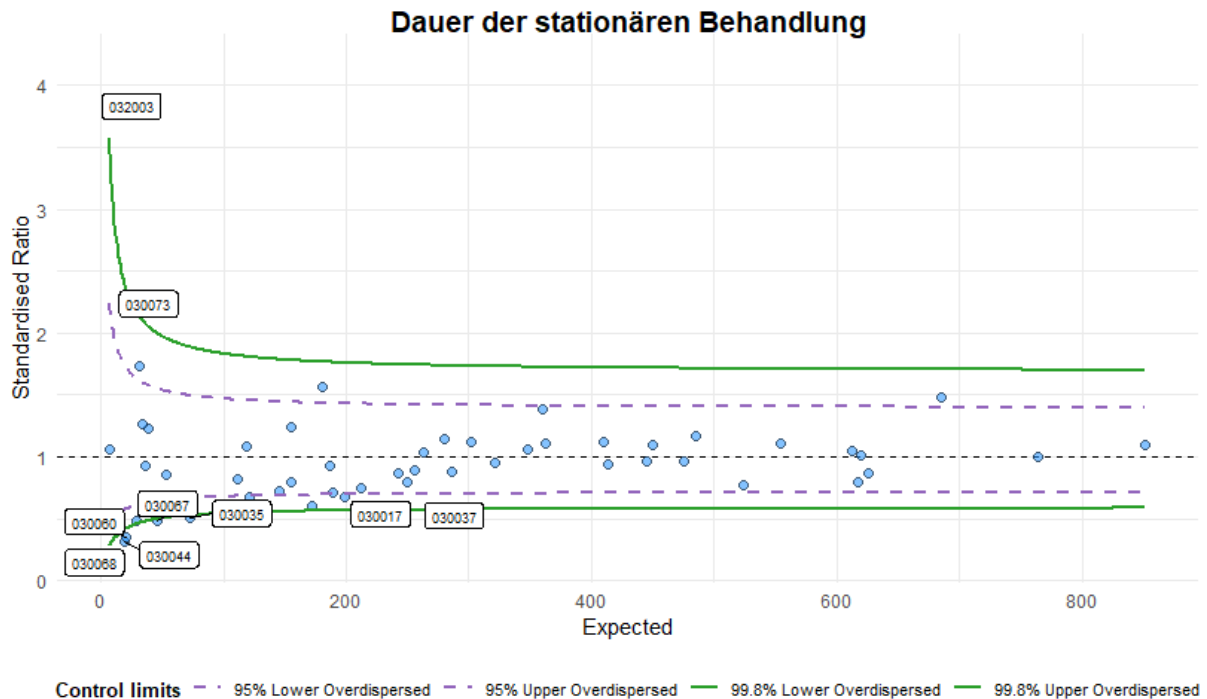


Abbildung: Funnel plot der erwarteten stationären Behandlungstage bezogen auf den Quotienten zwischen den stationären Behandlungstagen und den zu erwartenden stationären Behandlungstagen im Datensatz „medpar“ nach Korrektur für overdispersion.

Anschließend wurde ein Multilevelmodell gerechnet, welches die Clusterstruktur der Daten berücksichtigt (Patienten in Krankenhäusern).

Fixed effects:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	2.24722	0.05880	38.218	< 0.0000000000000002 ***
hmo	-0.09473	0.02594	-3.652	0.00026 ***
died	-0.27186	0.01865	-14.576	< 0.0000000000000002 ***
age80	-0.01927	0.02098	-0.918	0.35840
factor(type)2	0.26223	0.02472	10.606	< 0.0000000000000002 ***
factor(type)3	0.15513	0.04886	3.175	0.00150 **

 Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Es ergibt sich ein ähnliches Ergebnis. Das R^2 des Modells ist mit .094 als eher gering zu beurteilen. Der Koeffizient für die geplanten Eingriffe ist allerdings bei Berücksichtigung der Clusterstruktur deutlich geringer, so dass ein entsprechend geringerer Einfluss auf die stationäre Behandlungsdauer vorhanden ist. Im folgenden Boxplot, welches die Testung auf 5% (blaue Linien) und 1% Signifikanz (rote Linien) hinsichtlich Outlier darstellt, ist ersichtlich, dass nun lediglich 2 Krankenhäuser auf dem 5%-Niveau als Outlier bezeichnet werden können und ein weiteres sich knapp an diesem Niveau befindet.

Boxplot of point estimates of random intercept

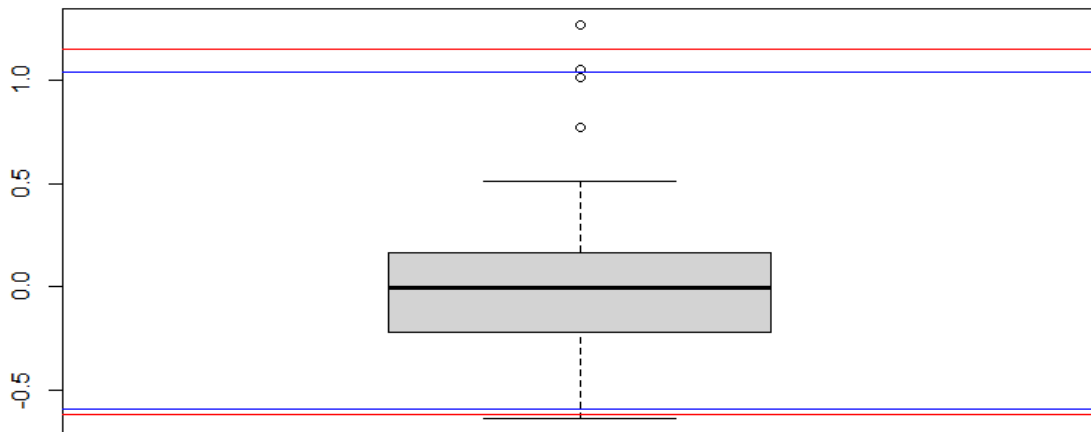


Abbildung: Boxplot der Punktschätzung für den random intercept des Mehrebenenmodells. Die Horizontalen blauen Linien entsprechen den Quantilen von 2,5% und 97,5%. Die horizontalen roten Linien entsprechen den Quantilen von 1% und 99%.

Dies wird auch im folgenden Q-Q-Plot deutlich (Abbildung).

Normal Q-Q Plot

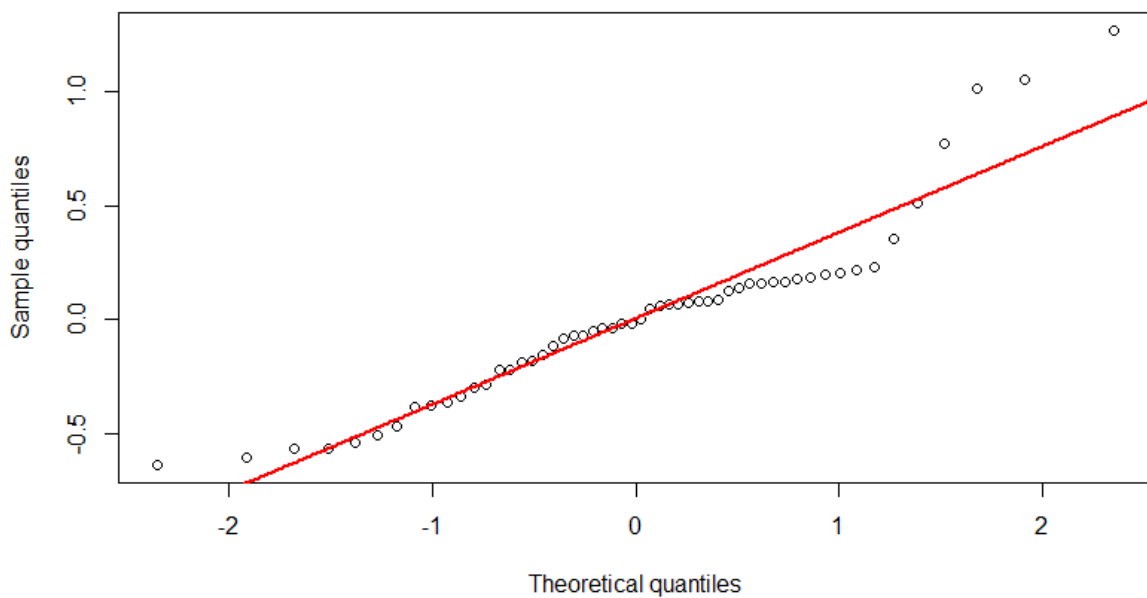


Abbildung: Q-Q-Plot der Punktschätzungen für den random intercept des Mehrebenenmodells zur Visualisierung von Outliern.

Anhand dieses Beispiels ist deutlich geworden, dass durch die Aggregation der Daten auf Betriebsstätten-Ebene Informationen verloren gehen könnten, die durch die Berücksichtigung

der Clusterstruktur der Daten erkannt werden können. Ein ähnliches methodisches Vorgehen wird in der Publikation von Hirsch et al. erläutert (Hirsch et al., 2018).

Literatur

Adams, J. L., Mehrotra, A., Thomas, J. W., & McGlynn, E. A. (2010). Physician Cost Profiling — Reliability and Risk of Misclassification. *The New England Journal of Medicine*, 362, 1014–1021.

Alexandrescu, R., Jen, M.-H., Bottle, A., Jarman, B., & Aylin, P. (2011). Logistic Versus Hierarchical Modeling: An Analysis of a Statewide Inpatient Sample. *Journal of the American College of Surgeons*, 213(3), 392–401. <https://doi.org/10.1016/j.jamcollsurg.2011.06.423>

Bernard, A., Cottenet, J., Aho, S., Doussot, A., Mariet, A.-S., Facy, O., & Quantin, C. (2021). Detecting Hospital Outliers in Post-Pancreatectomy Care Using Funnel Plots from 2009-2018 Based on Nationwide Medico-Administrative Data. *World Journal of Surgery*, 45(7), 2210–2217. <https://doi.org/10.1007/s00268-021-06078-4>

Bottle, A., Sanders, R. D., Mozid, A., & Aylin, P. (2013). Provider profiling models for acute coronary syndrome mortality using administrative data. *International Journal of Cardiology*, 168(1), 338–343. <https://doi.org/10.1016/j.ijcard.2012.09.048>

Bragg, F., Cromwell, D. A., Edozien, L. C., Gurol-Urganci, I., Mahmood, T. A., Templeton, A., & van der Meulen, Jan H (2010). Variation in rates of caesarean section among English NHS trusts after accounting for maternal and clinical risk: Cross sectional study. *BMJ (Clinical Research Ed.)*, 341, c5065. <https://doi.org/10.1136/bmj.c5065>

Burgess, J. F., Christiansen, C. L., Michalak, S. E., & Morris, C. N. (2000). Medical profiling: improving standards and risk adjustments using hierarchical models. *Journal of Health Economics*, 19(3), 291–309. [https://doi.org/10.1016/S0167-6296\(99\)00034-X](https://doi.org/10.1016/S0167-6296(99)00034-X)

Christiansen, C. L., & Morris, C. N. (1997). Improving the Statistical Approach to Health Care Provider Profiling. *Annals of Internal Medicine*, 127, 764–768.

Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel modeling using R* (Second edition). *Statistics in the social and behavioral sciences series*. CRC Press Taylor & Francis Group. <https://www.taylorfrancis.com/books/9781351062268>
<https://doi.org/10.1201/9781351062268>

Hirsch, O., Donner-Banzhoff, N., Schulz, M., & Erhart, M. (2018). Detecting and Visualizing Outliers in Provider Profiling Using Funnel Plots and Mixed Effects Models-An Example from Prescription Claims Data. *International Journal of Environmental Research and Public Health*, 15(9). <https://doi.org/10.3390/ijerph15092015>

Ieva, F., & Paganoni, A. M. (2015). Detecting and visualizing outliers in provider profiling via funnel plots and mixed effect models. *Health Care Management Science*, 18(2), 166–172. <https://doi.org/10.1007/s10729-013-9264-9>

Jones, H. E., & Spiegelhalter, D. J. (2011). The Identification of “Unusual” Health-Care Providers From a Hierarchical Model. *The American Statistician*, 65(3), 154–163. <https://doi.org/10.1198/tast.2011.10190>

Kuhrij, L., van Zwet, E., van den Berg-Vos, R., Nederkoorn, P., & Marang-van de Mheen, P. J. (2021). Enhancing feedback on performance measures: The difference in outlier detection using a binary versus continuous outcome funnel plot and implications for quality improvement. *BMJ Quality & Safety*, 30(1), 38–45. <https://doi.org/10.1136/bmjqs-2019-009929>

- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766. <https://doi.org/10.1016/j.jesp.2013.03.013>
- Marshall, C., Best, N., Bottle, A., & Aylin, P. (2004). Statistical issues in the prospective monitoring of health outcomes across multiple units. *Journal of the Royal Statistical Society. Series a*, 167, 541–559. <http://www.jstor.org/stable/pdf/3559780.pdf?acceptTC=true>
- Marshall, E. C., & Spiegelhalter, D. J. (2001). Institutional Performance. In A. H. Leyland & H. Goldstein (Eds.), *Wiley series in probability and statistics. Multilevel modelling of health statistics* (pp. 127–142). Wiley.
- Mazzucco, W., Cusimano, R., Zarcone, M., Mazzola, S., & Vitale, F. (2017). Funnel plots and choropleth maps in cancer risk communication: A comparison of tools for disseminating population-based incidence data to stakeholders. *BMJ Open*, 7(3), e011502. <https://doi.org/10.1136/bmjopen-2016-011502>
- Mehrotra, A., Adams, J. L., Thomas, J. W., & McGlynn, E. A. (2010a). Cost Profiles: Should The Focus Be On Individual Physicians Or Physician Groups? *Health Affairs*, 29(8), 1532–1538. <https://doi.org/10.1377/hlthaff.2009.1091>
- Mehrotra, A., Adams, J. L., Thomas, J. W., & McGlynn, E. A. (2010b). The Effect of Different Attribution Rules on Individual Physician Cost Profiles. *Annals of Internal Medicine*, 152, 649–654.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105(1), 156–166. <https://doi.org/10.1037//0033-2909.105.1.156>
- Normand, S.L.T., Glickman, M. E., & Gatsonis, C. A. (1997). Statistical Methods for Profiling Providers of Medical Care: Issues and Applications. *Journal of the American Statistical Association*, 92(439), 803–814.
- Ohlssen, D. I., Sharples, L. D., & Spiegelhalter, D. J. (2007). A hierarchical modelling framework for identifying unusual performance in health care providers. *Journal of the Royal Statistical Society. Series a*, 170(4), 865–890.
- Spiegelhalter, D. J. (2005a). Handling over-dispersion of performance indicators. *Quality & Safety in Health Care*, 14(5), 347–351. <https://doi.org/10.1136/qshc.2005.013755>
- Spiegelhalter, D. J. (2005b). Funnel plots for comparing institutional performance. *Statistics in Medicine*, 24(8), 1185–1202. <https://doi.org/10.1002/sim.1970>
- Swart, E., Ihle, P., Gothe, H., & Matusiewicz, D. (Eds.). (2014). *Programmbereich Gesundheit. Routinedaten im Gesundheitswesen: Handbuch Sekundärdatenanalyse: Grundlagen, Methoden und Perspektiven* (2., vollst. überarb. Aufl.). Verlag Hans Huber.
- Thomas, J. W., Grazier, K. L., & Ward, K. (2004). Economic Profiling of Primary Care Physicians: Consistency among Risk-Adjusted Measures. *Health Services Research*, 39(4), 985–1003.
- Tucker, J. L. (2000). The theory and methodology of provider profiling. *International Journal of Health Care Quality Assurance*, 13(7), 316–323. <https://doi.org/10.1108/09526860010378771>
- van der Willik, E. M., van Zwet, E. W., Hoekstra, T., van Ittersum, F. J., Hemmelder, M. H., Zoccali, C., Jager, K. J., Dekker, F. W., & Meuleman, Y. (2021). Funnel plots of patient-reported

outcomes to evaluate health-care quality: Basic principles, pitfalls and considerations. *Nephrology*, 26(2), 95–104. <https://doi.org/10.1111/nep.13761>

Verburg, I. W., Holman, R., Peek, N., Abu-Hanna, A., & Keizer, N. F. de (2018). Guidelines on constructing funnel plots for quality indicators: A case study on mortality in intensive care unit patients. *Statistical Methods in Medical Research*, 27(11), 3350–3366. <https://doi.org/10.1177/0962280217700169>

Wilsey, B. L., Fishman, S. M., Gilson, A. M., Casamalhuapa, C., Baxi, H., Zhang, H., & Li, C. S. (2010). Profiling multiple provider prescribing of opioids, benzodiazepines, stimulants, and anorectics. *Drug and Alcohol Dependence*, 112(1-2), 99–106. <https://doi.org/10.1016/j.drugalcdep.2010.05.007>

8. Profiling - Advanced

Zielsetzung:

Das primäre Ziel einer Analyse mittels Profiling ist die Identifikation von auffälligen Einheiten (z.B. Krankenhäuser oder ärztliche Betriebsstätten), wobei sich die Auffälligkeit auf eine vorab definierte Größe bezieht, die als Indikator modelliert wird und die es zu untersuchen gilt (z.B. Krankenhaussterblichkeit oder Aspekte des ärztlichen Ordnungsverhaltens).

Etabliert hat sich diese Analyseart bereits für die Identifikation von auffälligen Einheiten (zumeist Krankenhäuser), wo Auffälligkeiten, wie z.B. eine erhöhte Sterberate der Patienten, eine erhöhte Rückkehrquote von bereits entlassenen Patienten zurück ins Krankenhaus oder auch stark erhöhte Gesamtkosten im Vergleich zu anderen Krankenhäusern, gesucht und untersucht wurden. Dabei besticht die Profiling-Analyse durch die ihr innewohnende Flexibilität, Unterschiede zwischen den zu betrachtenden Einheiten zu berücksichtigen und diese in das Modell als Adjustierungsvariablen einfließen zu lassen. Somit können verschieden relevante Faktoren wie z.B. (Krankenhaus- oder Praxis-) Größe oder auch die Patientenstruktur (Altersverteilung, Morbidität etc.) berücksichtigt werden (Case Mix). Auf diese Weise wird verhindert, dass zum Beispiel Praxen, die nur auf Grund der Morbiditätsstruktur der Patienten – nicht aber durch das Ordnungsverhalten an sich – höhere Ordnungskosten aufweisen als auffällig betrachtet werden. Die Adjustierung erlaubt es vielmehr, die tatsächlich auffälligen Einheiten zuverlässig zu identifizieren.

Aufbau der Analyseprozedur:

Vorgestellt wird eine Analyseprozedur, die gut interpretierbar ist und künftigen Anwendern in der Entscheidungsfindung bzgl. „echt auffälligen“ Einheiten hilft. Diese Analyseprozedur besteht aus einer Kombination eines multivariaten Ansatzes der Ausreißerbestimmung, Regressionsverfahren und der Bayesschen Inferenz. Für Grundbegriffe und klassische Methoden des Profiling siehe [Verweis Profiling basic]

Die Analyseprozedur setzt sich aus den folgenden Schritten:

1. Bestimmung der Bayesschen Inferenz und der dazugehörigen Verteilung
2. Bestimmung von auffälligen Betriebsstätten mithilfe der Verteilung der Bayesschen Inferenz
3. Regression Forest

4. Bestimmung der auffälligen Betriebsstätten auf Basis des Regression Forests mithilfe eines Observed/Expected-Indikators
5. Interpretation der Ergebnisse mithilfe von Shapley Values

In einem ersten Schritt wird für jeden betrachteten Wirtschaftlichkeitsindikator ein Wahrscheinlichkeitsmodell mithilfe der Bayesschen Inferenz aufgestellt. Mithilfe der so entstehenden Verteilung wird ein Benchmark definiert. Im direkten Vergleich dazu können dann auffällige Einheiten ermittelt werden.

In einem weiteren Schritt wird die jeweilige Indikatorvariable mithilfe des Ansatzes des Regression Forests als endogene (d.h. abhängige) Variable modelliert, wobei sie von mehreren ausgewählten, exogenen Variablen abhängt. Die so mittels Regression-Forest-Ansatz berechneten Variablenwerte werden genutzt, um einen Observed/Expected-Indikator - hier als Differenz der beobachteten und der vom Modell vorhergesagten Werte - zu bestimmen. Mit diesem Vorgehen können ebenfalls (analog zum ersten Schritt) auffällige Betriebsstätten ermittelt werden, indem extreme Residuenwerte des Observed/Expected-Indikators identifiziert werden (indem zunächst die Differenz des Observed- und Expected-Wertes gebildet und dann identifiziert wird, ob die gebildete Differenz den gesetzten Grenzwert überschreitet.).

Im Anschluss werden diese beiden Ansätze der Identifikation von (möglichen) auffälligen Einheiten kombiniert, indem der mittels Random Forest Regression geschätzte Indikator in das Modell der Bayesschen Wahrscheinlichkeit eingebettet wird. Somit erfolgt eine kombinierte Analyse und Entscheidungsfindung, die die Vorteile beider Ansätze verbindet: zum einen ein Wahrscheinlichkeitsmodell, das die zugrundeliegenden Unsicherheiten möglichst vollständig abbildet und vorab bekannte Informationen und Annahmen beinhaltet und zum anderen die Möglichkeit exogene Faktoren zu berücksichtigen, welche die zu analysierende Größe maßgeblich beeinflussen.

Um die Analyseergebnisse final interpretieren zu können, was bei der Anwendung von Machine Learning Modellen nicht trivial ist, wird die Methode der Shapley Values aus der Spieltheorie genutzt. Diese hilft die Einflussnahme der exogenen Variablen angemessen einzuordnen.

Die so durchgeführte Identifikation von (möglichen) auffälligen Einheiten soll zum einen Einheiten ermitteln, die offensichtlich auffällig sind, und zum anderen Einheiten, die auffällig erscheinen, wo aber im Nachgang noch eine individuelle Überprüfung anzuraten ist. So können in der Gruppe der potentiell auffälligen Einheiten die echt-auffälligen von den nicht-auffälligen bestimmt werden, ohne die gesamte Gruppe aller Einheiten untersuchen zu müssen. Dies kann helfen, echt auffällige von scheinbar auffälligen Einheiten gezielt zu trennen.

Schritte der Analyseprozedur:

Im Weiteren folgt eine genauere Beschreibung der einzelnen Schritte der Analyseprozedur.

Bayessche Inferenz:

Der Bayessche Wahrscheinlichkeitsbegriff unterscheidet sich wesentlich von der klassischen (frequentistischen) Wahrscheinlichkeitsdefinition: beim Bayesschen Ansatz kann der Wahrscheinlichkeitsbegriff als Grad persönlicher Überzeugung angesehen werden. In die Bayessche

Inferenz (oder auch Bayes-Statistik) fließen neben vorab bekannten Informationen auch zusätzlich neu erhaltene Informationen oder auch individuelle Annahmen ein. Dies wird ausgedrückt durch den Satz von Bayes:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Dabei steht $P(A)$ für die A-priori-Wahrscheinlichkeit, also die Wahrscheinlichkeitsverteilung für das Ereignis A, welche auf vorhergehenden Annahmen und vorhergehendem Wissen basiert. $P(B)$ steht analog für die A-priori-Wahrscheinlichkeit der Evidenz. $P(B|A)$ ist die Likelihood-Wahrscheinlichkeit der Evidenz, wenn das Ereignis A gegeben ist, und $P(A|B)$ ist die A-posteriori-Wahrscheinlichkeit der angepassten Überzeugung, wenn die Evidenz als gegeben gilt.

Mit diesem Ansatz kann im Folgenden ein Wahrscheinlichkeitsmodell für die Auffindung von auffälligen Einheiten aufgestellt werden, das verschiedene Informationen, Wissensstände und getroffene Annahmen und sogar Expertenmeinungen in die Ermittlung der einzelnen Wahrscheinlichkeiten mitberücksichtigt. Dabei kann insbesondere die Berücksichtigung von Expertenmeinungen das Modellfitting weiter verbessern. Weitere Vorteile dieses Ansatzes sind gut interpretierbare Ergebnisse, die Möglichkeit des Einfließens von subjektiven Inputs, wie z.B. medizinisches Wissen, und die Quantifizierung der jeweiligen Unsicherheiten.

Bestimmung von auffälligen Betriebsstätten mithilfe der Verteilung der Bayesschen Inferenz

Um einen ersten Benchmark zu erhalten, werden Anomalien analysiert, die zunächst vorab um den Einfluss von Confoundern bereinigt wurden. Für die Zielvariable (*target variable*) soll mit dem Ansatz der Bayesschen Inferenz eine Verteilung modelliert werden, für welche potentielle Ausreißer (*Outlier*), also auffällige Einheiten, ermittelt werden können.

Die Wahl des Ausgangsmodells und der Ausgangsverteilung hängt dabei stark von den Eigenschaften der zu betrachtenden Variable ab. Mithilfe von abgeleiteten Parametern kann für diese Variable eine Verteilung approximiert werden, die jeder Beobachtung bzw. jedem Ereignis eine natürliche kumulative Wahrscheinlichkeit zuordnet. Mithilfe der auch so entstehenden Tails der approximierten Verteilung können Extremwerte abgebildet werden.

Regression Forest

Der Ansatz des Regression Forests kommt aus dem Bereich des Machine Learnings und bietet die Möglichkeit, multiple Entscheidungsbäume zu bilden, d.h. eine Kombination aus mehreren verschiedenen Bäumen (Regression Trees).

Regression Forests nutzen dabei verschiedene B-Bootstrap-Stichproben, um auf jeder Stichprobe einen Baum zu erzeugen (Bagging). Grundsätzlich hängt die Anzahl der möglichen Bäume von verschiedenen Parametern ab, wie z.B. von der Anzahl der möglichen Klassifikationen (Einteilung bzw. Gruppierung der Baumendpunkte in „Klassen“) und der Anzahl der betrachteten Merkmale. Dabei führt eine große Anzahl an Bäumen zu keinem Overfitting und ist deshalb immer zu bevorzugen.

Der Vorteil eines Waldes gegenüber einem einzelnen Baum ist, dass jede Variable, die einen Beitrag liefert, auch verwendet werden kann und sollte. Der Nachteil eines Waldes im Gegensatz zu einem einzigen Baum ist jedoch, dass die Verständlichkeit verloren geht, da die für einzelne Bäume verwendeten Klassifikationsregeln (Entscheidungsregeln entlang eines Baumes, die zu bestimmten Ergebnissen und damit zu einer Klassifikation führen) nicht mehr ablesbar sind. Somit gewinnt man mit der Anwendung eines Regression Forests eine gesteigerte Vorhersagbarkeit des Modells, reduziert allerdings im Gegenzug die Interpretierbarkeit der Ergebnisse.

Konstruktion eines Waldes

Für einen Wald werden die einzelnen Bäume nicht aus allen zur Verfügung stehenden Daten bestimmt, sondern es wird für jeden Baum eine Stichprobe (mit Zurücklegen) aus den Beobachtungen gezogen (Bootstrap). Wenn dabei mit vielen unterschiedlichen Stichproben gearbeitet wird, spricht man auch vom *Bagging*. Beim Bagging werden bei der Konstruktion eines Waldes viele Klassifikationsregeln der einzelnen Bäume zusammenfassend gewichtet.

Für die Bestimmung der Verzweigungsregeln der Bäume werden nicht alle zur Verfügung stehenden Variablen (predictors) genutzt, sondern es wird zufällig eine (vorher festgelegte) Anzahl von ihnen aus der Menge aller Variablen gezogen. Nur diese werden dann für das Vorhaben genutzt.

Der Ansatz der Regression Forests ist attraktiv, da diese Methode ein echtes Regressionsmodell approximiert, wobei aber eine höhere Flexibilität und Vorhersagegenauigkeit erreicht wird. Zudem kann ein angenommener nichtlinearer Zusammenhang zwischen exogenen und der jeweils betrachteten endogenen Variable widergespiegelt werden.

Insgesamt soll die Anwendung des Regression Forests zum einen auftretende Extremwerte bzw. Ausreißer des Indikators erhalten, d.h. Extremwerte und Ausreißer sollen auch nach Anwendung des Regression Forests als solches identifizierbar sein, und zum anderen Adjustierungen durch gewählte exogene Variablen ermöglichen. Der Indikator wird innerhalb eines Bayesschen Modells analysiert, um nicht nur zu bestimmen, ob eine betrachtete Einheit ein echter Ausreißer (Outlier) und damit auffällig ist, sondern auch zu welchem Grad man davon überzeugt ist (degree of belief). Somit können auch subjektive Einschätzungen und vor allem auch Expertenmeinungen in das Modell einfließen und einen realen Anteil an der Vorhersage leisten.

Bestimmung der auffälligen Betriebsstätten auf Basis des Regression Forests mithilfe eines Observed/Expected-Indikators

Um ein risikoadjustiertes Wahrscheinlichkeitsmodell für die Ermittlung von Outliern zu entwickeln, werden durch den Regression Forest berechnete Indikatorwerte (predicted values) genutzt. Nimmt man die beobachteten Indikatorwerte (observed values) hinzu, kann das Verhältnis der vorab berechneten zu den beobachteten Werte genutzt werden, um einen Observed/Expected-Indikator (O/E-Indikator) zu bilden. Mit steigender Anpassungsgüte (goodness-of-fit) des aufgestellten Modells konzentriert sich die empirische Wahrscheinlichkeitsverteilung des O/E-Indikators symmetrisch um den Wert 1. Aus diesem Grunde ist eine um den Wert 1 zentrierte Student-t-Verteilung sehr gut geeignet, um in diesem Setting die Outlier zu bestimmen. Da in dieser Analyse alle Variablen ausschließlich positive Werte annehmen, sind

auch die O/E-Indikatorwerte positiv und es wird nur der positive Teil der um den Wert 1 zentrierten Student-t-Verteilung genutzt.

Es werden dann für die Zielvariable - die in diesem Fall der O/E-Indikator ist - mit der entstehenden A-posteriori-Wahrscheinlichkeitsverteilung für jede Beobachtung eigene kumulative Wahrscheinlichkeiten zuordnet. Mithilfe dieser lassen sich dann auffällige bzw. abweichende Werte ermitteln.

Interpretation der Ergebnisse mithilfe von Shapley Values

Wenn es zum Vergleich der durch beide Methoden oder eine der beiden Methoden gefundenen, auffälligen Einheiten kommt, so können die folgenden Fälle unterschieden werden:

- a. Eine Einheit ist nach keiner der beiden Methoden auffällig, somit handelt es sich nicht um einen Outlier.
- b. Eine Einheit ist nach beiden Methoden auffällig, sie kann daher als Outlier gewertet werden.
- c. Eine Einheit ist ein Outlier nach der Methode der Verteilung mittels Bayesscher Inferenz, aber nicht nach der Methode der Verteilung des O/E-Indikators. In diesem Fall hat das Machine-Learning-Modell entschieden, dass der vorliegende Extremwert als solcher gerechtfertigt ist, z.B. durch Confounder („Praxisbesonderheiten“).
- d. Eine Einheit ist kein Outlier bzgl. der Verteilung mittels Bayesscher Inferenz, aber dafür im Hinblick auf die Verteilung des O/E-Indikators. In diesem Fall klassifiziert das Machine-Learning-Modell einen möglichen Extremwert nicht als solchen einen, sondern als echten Ausreißer.

Während die ersten beiden Fälle klar in der Interpretation sind, sind es die beiden unteren Fälle (c und d) nicht - sowohl hinsichtlich der Entscheidungsfindung als auch der späteren Interpretation. Vor allem der letzte Fall sollte im Nachgang zur Analyse genauer betrachtet werden und die dort auffälligen Einheit genauer untersucht werden. Man könnte hier auch formulieren, dass diese Einheiten „Ausreißer-Potential“ besitzen.

Mithilfe dieses Analyseverfahrens, kann im Rahmen einer finalen Feinjustierung der Kontrollgrenzen die Menge der unauffälligen, auffälligen und potentiell auffälligen Einheiten (hier: Betriebsstätten) bestimmt werden. Diese Feinjustierung hängt vom jeweils vorliegenden Setting innerhalb einer Region (z.B. KV) ab und kann auch wirtschaftliche respektive politische Implikationen, wie die Festlegung eines Schwellenwertes ab welchem Betriebsstätten als auffällig gelten sollen, beinhalten. Daher müsste sie individuell für jeden Indikator und jede Region durchgeführt werden. Die jeweiligen Adjustierungen werden ermittelt, indem für eine Referenzmenge an zu betrachtenden Einheiten (z.B. alle Betriebsstätten einer Fachgruppe in einem Quartal in einer KV) die mit dem Analyseverfahren ermittelten auffälligen Betriebsstätten mit den in einer KV tatsächlich auffälligen Betriebsstätten abgeglichen werden. Diejenige Kontrollintervallgröße, welche beide Gruppen möglichst deckungsgleich aufeinander abbildet, stellt die eigentliche Kontrollintervallgröße (für den jeweiligen Indikator) dar.

Da zudem die Entscheidungsfindung und spätere Interpretation innerhalb der Machine-Learning-Modellierung kaum möglich ist, soll zum Schluss noch eine Methode vorgestellt werden, die die spätere Interpretation der erzeugten Ergebnisse erleichtert.

Shapley Values

Für die Interpretierbarkeit der Ergebnisse kann ein spieltheoretischer Ansatz mittels einer Metrik genutzt werden, die als Shapley Values bekannt ist.

Dieser spieltheoretische Ansatz geht davon aus, dass jede Vorhersage des Modells durch die Annahme erklärt werden kann, dass jede einflussgebende (exogene) Variablen als „Spieler“ interpretiert wird und die Modellvorhersagen den „Gewinn“ des Spiels darstellen, den sich alle „Spieler“ teilen. Die Shapley Values bilden dabei eine Metrik aus der kooperativen Spieltheorie und geben an, wie der Gewinn gerecht auf alle Spieler aufgeteilt wird, abhängig davon welchen Beitrag die einzelnen Spieler zum Gesamtgewinn geleistet haben. Dabei kooperieren Spieler und erhalten dadurch Vorteile bzw. einen höheren anteiligen Gewinn.

Angewandt auf das Setting dieser Analyse bildet die Machine-Learning-Modellierung das „Spiel“, die eigentlichen Ergebnisse bzw. Vorhersagen der Modellierung den „Gewinn“ und die einflussgebenden Variablen, die miteinander kooperieren, um einen Gewinn zu erhalten, die „Spieler“. Insgesamt stellen die Shapley Values somit den *durchschnittlichen marginalen Beitrag einer Variable* über alle möglichen Kooperationen dar.

Der Unterschied zu anderen spieltheoretischen Metriken ist dadurch gegeben, dass zusätzlich zu allen einzelnen Variablen auch alle Variablenkombinationen betrachtet werden, da so die Kooperationen untereinander mitberücksichtigt werden. So können die Anteile der einzelnen Variablen an der Vorhersage bzw. dem jeweiligen Ergebnis gut bestimmt werden. Zudem haben Shapley Values die Eigenschaft, dass die so bestimmten Werte bzw. Anteile eindeutig und einfach zu interpretieren sind.

Um die Shapley Values zu implementieren wird die SHAP¹⁴-Methode genutzt. Es handelt sich dabei um ein additives Modell, das aufgrund seiner additiven Form sehr einfach zu interpretieren ist:

$$g(z') = \varphi_0 + \sum_{j=1}^M \varphi_j z'_j.$$

Dabei ist g das erklärende Modell, $z' \in \{0,1\}^M$ ein Vektor, der die Kooperationen beschreibt (mit dem Wert 1, wenn eine Kooperation vorliegt, und dem Wert 0, wenn keine vorliegt), M die maximale Kooperationsgröße (d.h. die Anzahl aller Variablen, die vom Machine-Learning-Modell genutzt werden) und $\varphi_j \in \mathbb{R}$ die Variablenattribute für die Variable j .

Insbesondere kann hier TreeSHAP genutzt werden, welcher speziell für Regression Forests entwickelt wurde.

Der gesamte Prozess wird in folgender Abbildung visualisiert:

¹⁴ SHAP - SHaple Additive exPlanation

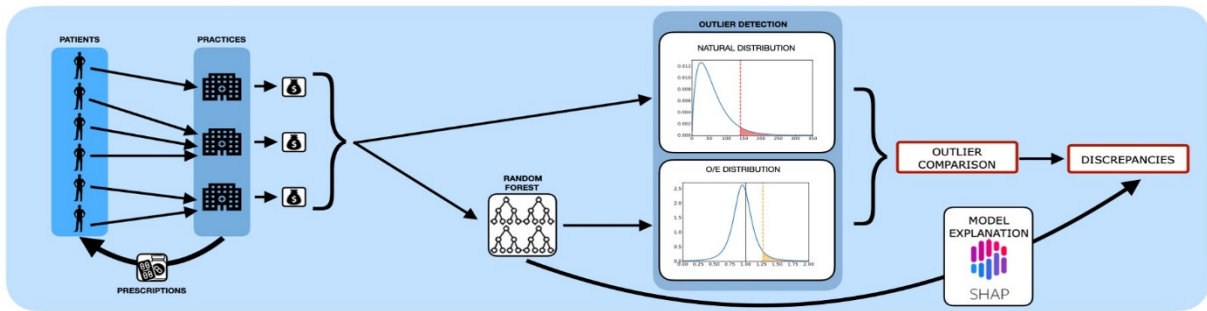


Abbildung: Schema der Analyseprozedur

Fazit

Ziel des hier dargestellten allgemeinen Analyseverfahren ist das Auffinden von auffälligen Einheiten (Betriebsstätten), wobei nur sehr allgemeine Rahmenbedingungen für das statistische Modell vorausgesetzt werden.

Hervorzuheben ist, dass dieses hier dargestellte, neue Verfahren mit seiner kombinierten Nutzung der Bayesschen Inferenz und dem Machine-Learning-Ansatz besticht. Diese innovative Herangehensweise bietet zum einen die Möglichkeit, auf der angenommenen Verteilung der zu bewertenden Indikatorvariable zu arbeiten, und zum anderen trotz Vorliegen einer Nicht-Normalverteilung ein Regressionsmodell aufzustellen, da das Verfahren im Gegensatz zu gängigen Regressionsmodellen nicht von der Normalverteilungsannahme abhängt. Somit wird weder die vorliegende Verteilung durch zu treffende Annahmen verzerrt, noch die Möglichkeit der Ermittlung der einflussgebenden Variablen eingeschränkt. Zusammen mit dem spieltheoretischen Ansatz der Shapley Values kann nicht nur bestimmt werden, welche exogenen Variablen einen Einfluss auf die jeweilige Zielvariable ausüben, sondern auch einordnen, wie groß diese Einflüsse jeweils sind.

Mithilfe dieses Analyseverfahrens kann mittels einer finalen Feinjustierung der Kontrollgrenzen die Menge der auffälligen oder potentiell auffälligen Einheiten (hier: Betriebsstätten) bestimmt werden. Diese Feinjustierung hängt vom jeweils vorliegenden Setting innerhalb einer KV ab (z.B. Kriterien, ab welchem Schwellenwert Betriebsstätten als auffällig gelten sollen) und müsste individuell durchgeführt werden.

Unabhängig von dieser Feinjustierung erhält man durch dieses Analyseverfahren diejenigen Variablen, die die zu betrachtenden Größe der Wirtschaftlichkeitsmessung maßgeblich beeinflussen. Dadurch kann besser unterschieden werden, welche Einheiten tatsächlich wegen ihres Ordnungsverhaltens auffällig sein könnten und welche es vermutlich lediglich auf Grund externer Faktoren (Confounder) zu sein scheinen. Zudem kann aus der Menge der betrachteten Betriebsstätten die Untergruppe der Grenzfälle ermittelt werden, für welche sich eine genauere Betrachtung und Analyse des jeweiligen Ordnungsverhaltens lohnt, um festzustellen, ob bei dieser Praxis eine Optimierung des Ordnungsverhaltens angestrebt werden sollte oder der Ordnungsaufwand - beeinflusst durch (externe) Faktoren – wirtschaftlich ist.

Dieses hier präsentierte Analyseverfahren führt zu gut interpretierbaren Ergebnissen und ist dabei allgemein genug, um auf beliebige Indikatoren und die jeweils gewünschten Adjustierungen anwendbar zu sein. Dabei ist es nicht eingeschränkt auf die klassischen Regressionsannahmen, wie z.B. die Normalverteilungsannahme. Damit kann die Wirklichkeit gut modelliert und analysiert werden, ohne durch nötige Annahmen verzerrt zu werden. Insgesamt ist

dieses Verfahren in der Umsetzung zwar komplex, aber durch die erzeugten Ergebnisse lohnenswert.

Literatur

- Bauder, R., & Khoshgoftaar, T. (2016). A Probabilistic Programming Approach for Outlier Detection in Healthcare Claims. *15th IEEE International Conference on Machine Learning and Applications*, (S. 347-354).
- Chaloner, K., & Brant, R. (1988). A Bayesian approach to outlier detection and residual analysis. *Biometrika*, 75, 4, S. 651-659.
- Holler, M., & Illing, G. (2009). *Einführung in die Spieltheorie*. Springer.
- Ieva, F., & Paganoni, A. M. (2015). Detecting and visualizing outliers in provider profiling via funnel plots and mixed effect models. *Health Care Manag Sci*, 18, S. 166-172.
- Jones, H., & Spiegelhalter, D. (August 2011). The Identification of "Unusual" Health-Care Providers From a Hierarchical Model. *The American Statistician*, Vol. 65, No. 3, S. 154-163.
- Lundberg, S., & al, e. (7. Mar 2019). *Consistent Individualized Feature Attribution for Tree Ensembles*. Von arXiv:1802.03888v3: <https://arxiv.org/abs/1802.03888> abgerufen
- Lundberg, S., & Lee, S. (2017). A Unified Approach to Interpreting Model Predictions., (S. 31st Conference on Neural Information Processing Systems (NIPS 2017)). Long Beach, CA, USA.
- Neuburger, J., & al, e. (2011). Funnel plots for comparing provider performance based on patient-reported outcome measures. *BMJ Quality & Safety*(20), S. 1020-1026.
- Spiegelhalter, D. (2005). Funnel plots for comparing institutional performance. *Statistics in Medicine*, 24, S. 1185-1202.
- Spiegelhalter, D., & al, e. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, 64, part 4, S. 583-639.
- Spiegelhalter, D., & al, e. (2012). Statistical methods for healthcare regulation: rating, screening and surveillance. *Journal of the Royal Statistical Society*, 175 part 1, S. 1-47.

2.3.6 Diffusionsanalysen

2.3.6.1 Einleitung, Hintergrund, Idee

- Ziel des Teilprojekts war die Identifikation von *Diffusionsmustern* von beobachtungswürdigen Medikamenten anhand von Routinedaten
- Dieser Abschnitt der Toolbox soll Varianten aufzeigen, um die Verbreitung bzw. Verwendung von beobachtungswürdigen Medikamenten zu analysieren
- Die Datenbasis stellen Routinedaten dar, die eine retrospektive Analyse ermöglichen und je nach Datenquelle und Zeiträume (Krankenkasse/ KV) unterschiedlich umfassend sein können

- Generell sind verschiedene Faktoren bekannt, die die Verbreitung von Medikamenten beeinflussen

<u>Macro-level</u>	
• Government policies	
<u>Meso-level</u>	
<u>Micro-level</u>	<ul style="list-style-type: none"> • Prescribing characteristics • Marketing efforts targeted at doctors • Drug characteristics • Practice characteristics • Patient characteristics • Contagion through social networks
<ul style="list-style-type: none"> • Physician characteristics: age, gender, medical specialty, professional age, nationality, ... • Scientific orientation 	

(Angelehnt an Lublóy 2014)

- In den präsentierten Ansätzen werden einerseits individuelle Perspektiven (Arzt/ Patient) eingenommen und andererseits globalere (regionale/zeitliche) Perspektiven, um die Verbreitung der beobachtungswürdigen Medikamente zu analysieren
- Der deutsche ambulante Sektor als Besonderheit mit freier Arztwahl → Patientenpfade sind sehr variabel und individuell, insbesondere Medikationen können dadurch von vielen verschiedene ÄrztInnen mit getragen werden

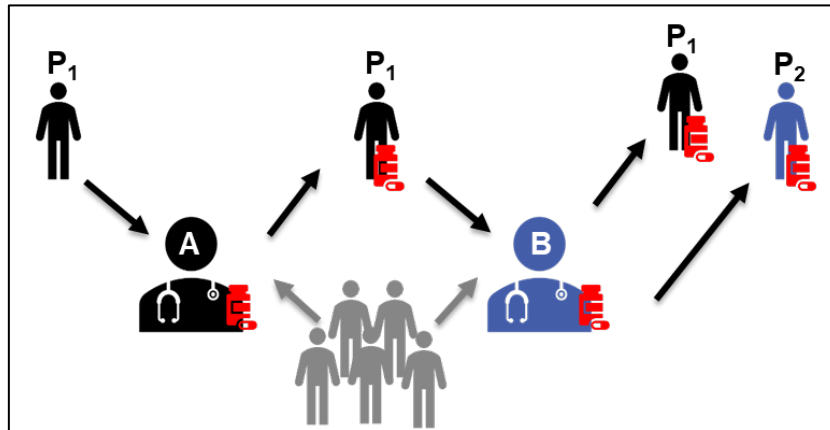
Es werden drei verschiedene Analyseansätze aufgezeigt

- Verbreitung neuer Arzneimittel in empirischen Ärztenetzwerken
- Regionale Ausbreitung von Arzneimitteln
- Untersuchung von Patientengruppen mit unterschiedlichen Verordnungsmengen und die Rolle von Kooperation
- **Kurzer Exkurs zu Netzwerkanalysen** basierend auf Routinedaten (Idee und Ansatz)
 - Empirische Ärztenetze als Resultat faktischer Patientenpfade
 - Ärzte behandeln die gleichen Patienten und werden in Netzen verbunden
 - Patienten tragen Verordnungen in den Netzen weiter, weitere Faktoren können auf Netzwerkebene untersucht werden
 - Innerhalb der Netze sind verschiedene Einflussfaktoren auf Micro- und Mesoebene denkbar

2.3.6.2 Verbreitung neuer Arzneimittel in empirischen Ärztenetzwerken

Mögliche Fragestellungen:

- Identifikation von förderlichen und hemmenden Faktoren bei der Verbreitung von Arzneimitteln innerhalb von empirischen Ärztenetzen in einer Time to Event Analyse.
- Wie beeinflusst der Patientenpfad das Ordnungsverhalten von ÄrztInnen?
- Löst die Behandlung von Patienten mit einer Vormedikation eines gewissen Wirkstoffs eine Weiterverordnung der jeweiligen ÄrztInnen aus? (Anschlussfrage: Und tendieren Ärzte dazu, das neue Medikament dann auch anderen Patienten zu verordnen?)
- Tendieren ÄrztInnen dazu einen neuen Wirkstoff zu verordnen, wenn sie mit vielen anderen den Wirkstoff (bereits) verordnenden Ärzten in empirischen Netzen verbunden sind?



2.3.6.3 Voraussetzungen und Methodik

Was für Voraussetzungen sollten Arzneimittel erfüllen? Welche Daten werden dafür auf welcher Ebene benötigt?

Art der Arzneimittel:

- Das Verfahren eignet sich insbesondere für die Untersuchung von der Verbreitung neuer Arzneimittel und unter denen insbesondere für solche, die keinen/ oder nur einen geringen Zusatznutzen aufweisen.
- Eine mögliche Fragestellung ergibt sich auch in Bezug auf spezifische Subpopulationen, für die eine Verordnung **nicht** empfohlen wird.
-

Benötigte Daten:

- Routinedaten basierend auf individueller Arzt- und Patientenebene:
 - Verordnungsdaten, um relevanten Verordnungen zu identifizieren
 - Ambulante Abrechnungsdaten, um empirische Ärztenetze bilden zu können.
 - Ggf. Diagnosedaten aus ambulantem und stationärem Sektor zur Identifikation der Zielpopulation.
- **Datenumfang:** Alle betroffenen PatientInnen bzw. ÄrztInnen mit relevanter Verordnung + Zufallsstichprobe ihrer PatientInnen, um Ärztenetze zu identifizieren und ÄrztInnen, die das Medikament möglicherweise verordnen könnten (nach Fachgruppe oder vorheriger Verordnungen)

Beobachtungszeitraum:

- Beobachtung ab Markteintritt erleichtert die Identifikation von Neuverordnungen (für ÄrztInnen und PatientInnen)
- Ggf. Berücksichtigung einer AMNOG-Bewertung nach Markteintritt oder weiteren

Analysemethode zusammengefasst: Netzwerkanalyse basierend auf Ärztenetzwerken und anschließendes Time-To-Event Model

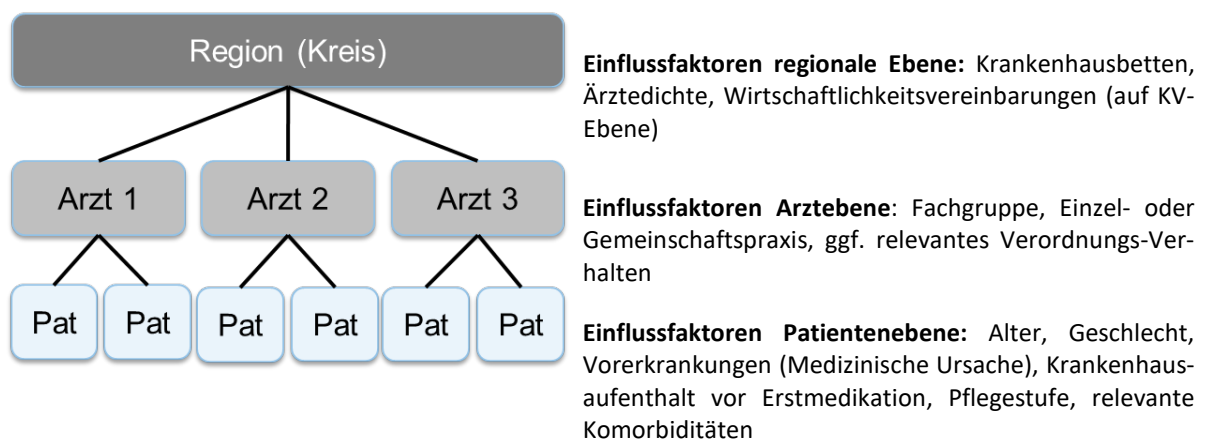
Skizze des Analysevorgehens

- Identifikation von Erstverordnungen (aus Arzt- und/oder Patientensicht) und Herstellen von Monats- oder Quartalsbezug um Zeitreihen zu erstellen (Time-to-Event Analyse)
- Ggf. Identifikation der (Sub-)Population der PatientInnen und Ermittlung relevanter ÄrztInnen (ÄrztInnen, die das Medikament verordnen könnten)
- Ermittlung empirischer Ärztenetze (wer teilt mit wem, wie viele Patienten und wird durch diese miteinander verbunden)
- Erhebung zu untersuchender Einflussfaktoren (zeitkonstant bzw. zeitabhängig)
- Schätzung des Time-To-Event Modells

2.3.6.4 Regionale Verbreitung von Arzneimitteln

Mögliche Fragestellungen:

- Identifikation von Einflussfaktoren auf eine medikamentöse Therapieentscheidung im ambulanten Sektor in Form von Mehrebenen Modellen.
- Welche Einflussfaktoren (auf Mikro-, Meso- und Makroebene) wirken bei einer Therapie-Entscheidung für oder gegen die Verordnung eines bestimmten Wirkstoffs und damit für bzw. gegen die Diffusion? (Mögliche Einflussfaktoren sind Patienten- und Arztcharakteristika, Wirkstoffcharakteristika sowie regionale Faktoren (wie z.B. Wirkstoffvereinbarungen, Ärztedichte, Krankenhausbettendichte etc.))
 - Gibt es auffällige Regionen, in denen zum Beispiel vermehrt das teurere Medikament verordnet wird?
 - Welche Rolle spielen Vorerkrankungen von PatientInnen bei einer Therapieentscheidung?
 - Wie beeinflussen Krankenhausaufenthalte die medikamentöse Therapieentscheidung im ambulanten Sektor?
 - Neigen bestimmte Facharztgruppen bei einer Therapieentscheidung eher dazu ein neueres/teureres Medikament zu verordnen?
 - Führt z.B. eine besonders hohe Krankenhausbettendichte regional im ambulanten Sektor zu einer vermehrten Verordnung eines neuen Medikaments?



2.3.6.5 Voraussetzungen und Methodik

Was für Voraussetzungen sollten Arzneimittel erfüllen? Welche Daten werden dafür auf welcher Ebene benötigt?

Art der Arzneimittel:

Das Verfahren eignet sich insbesondere für die Untersuchung von Therapieentscheidungen. Daher sind solche Wirkstoffe geeignet, die für die grundsätzlich gleiche/ identische Indikation anzuwenden sind.

Eine andere mögliche Fragestellung könnte auch die Entscheidung für oder gegen eine medikamentöse Therapie darstellen.

Benötigte Daten:

- Routinedaten auf individueller Patientenebene (aus den zu analysierenden Regionen):
 - Verordnungsdaten, um relevante Verordnungen zu identifizieren
 - Weitere Versicherten- bzw. Abrechnungsdaten, um relevante Patientencharakteristika erheben zu können (Alter, Geschlecht, Vorerkrankungen, ...)
- Routinedaten für Informationen auf Arztelebene (aus den zu analysierenden Regionen):
 - Praxiszugehörigkeit, Facharztgruppe
 - Ggf. weitere Abrechnungsdaten, um „typisches Verschreibungsverhalten“ zu charakterisieren.
 - Regionaler Bezug (z.B. Praxisort), um Zuordnung zur Region zu ermöglichen.
- Regionale Daten z.B. aus der INKAR-Datenbank des Bundesinstituts für Bau-, Stadt- und Raumforschung und ggf. weitere Informationen zu Wirtschaftlichkeitsverordnungen

Datenumfang: Alle betroffenen PatientInnen mit einer der relevanten Verordnungen

- Bei Therapieentscheidungen zwischen zwei oder mehr Wirkstoffen alle PatientInnen mit mindestens einer Verordnung für einen der Wirkstoffe.
- Bei Therapieentscheidungen für eine Verordnung (in Abgrenzung zu keiner) alle PatientInnen mit relevanter Verordnung sowie passende Vergleichspopulation ohne Verordnung (z.B. anhand von Erkrankungen).

Beobachtungszeitraum:

Je nach Fragestellungen kann der Beobachtungszeitraum frei gewählt werden. Es sollte ausreichend Vor- und Nachbeobachtungszeit eingeplant werden, um ggf. relevante Indikationen und/oder Krankenhausaufenthalte vor und nach der Verordnung zu identifizieren.

Analysemethode zusammengefasst:

Mehrebenenmodell mit binärem oder multinomialem Outcome.

Skizze des Analysevorgehens

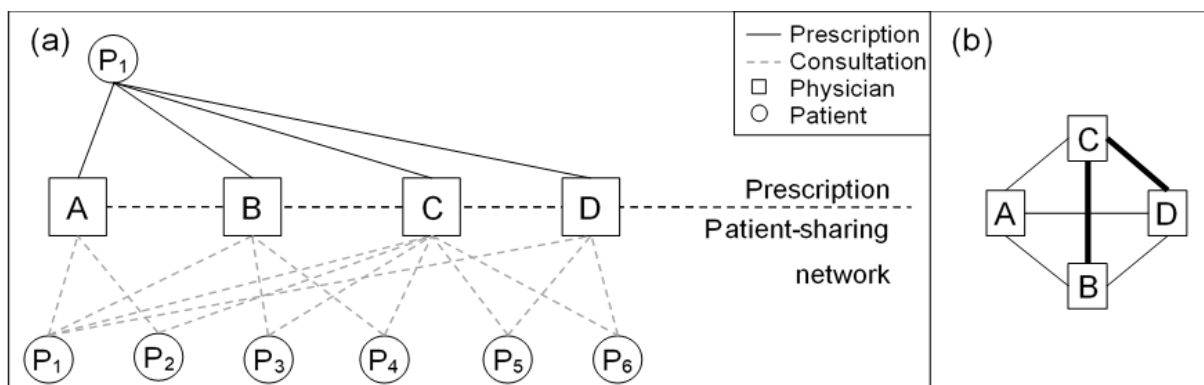
- a) Identifikation PatientInnen mit relevanter Verordnung (oder „Nicht-Verordnung“) und relevanter Charakteristika
- b) Identifikation der VerordnerInnen und ihrer Charakteristika

- c) Ermittlung regionaler Zugehörigkeit und regionaler Einflussfaktoren
- d) Schätzung des Mehrebenenmodells

2.3.6.6 Untersuchung von Patientengruppen mit unterschiedlichen Verordnungsmengen und die Rolle von Kooperation

Mögliche Fragestellungen:

- Identifikation von Unterschieden hinsichtlich ausgewählter Charakteristika auf Patienten-, Arzt-, Netzwerk- oder Umgebungsebene bei verschiedenen Verordnungsmengen (insbesondere bei BTM oder Wirkstoffen mit Abhängigkeitspotenzial).
- Welche Unterschiede bestehen hinsichtlich relevanter Patientencharakteristika (z.B. Alter und Geschlecht der PatientInnen) zwischen den Gruppen?
- Gibt es zum Beispiel Indikationen bei denen besonders häufig große Mengen des Wirkstoffs verordnet werden?
- Gibt es Gruppen an verordnenden ÄrztInnen, die besonders häufig das Medikament verordnen und sind diese auffällig miteinander verbunden?
- Welche Rolle spielen verschiedene Facharztgruppen bei der Erstverordnung oder auch in Bezug auf ihre Beteiligung an der insgesamt verordneten Menge?
- Sind große Mengen des verordneten Wirkstoffs assoziiert mit vielen verordnenden ÄrztInnen und kann es somit ein Indiz für doctor-shopping sein?
- Wie ist die Kooperation gemessen im Sinne der geteilten PatientInnen innerhalb der empirischen Verordnernetze (alle verordnenden ÄrztInnen je PatientIn) zu bewerten?
-



Voraussetzungen und Methodik

Was für Voraussetzungen sollten Arzneimittel erfüllen? Welche Daten werden dafür auf welcher Ebene benötigt?

Art der Arzneimittel:

Das Verfahren eignet sich insbesondere für die Untersuchung von Medikamenten, die entweder als Betäubungsmittel gelten, ein Abhängigkeitspotential aufweisen oder bei sonstigem nicht ordnungsgemäßigem Konsum gefährlich für PatientInnen sein können.

Benötigte Daten:

- Routinedaten auf individueller Patientenebene:
 - Verordnungsdaten, um relevante Verordnungen zu identifizieren inklusive Informationen zur verordneten Menge.
 - Weitere Versicherten- bzw. Abrechnungsdaten, um relevante Patientencharakteristika erheben zu können (Alter, Geschlecht, Vorerkrankungen, sonstige ggf. kontraindizierte medikamentöse Therapie...).
 - Ambulante Abrechnungsdaten, um empirische Verordnungsnetze bilden zu können.
- **Datenumfang:** Alle betroffenen PatientInnen bzw. ÄrztInnen mit relevanter Verordnung + Zufallsstichprobe ihrer PatientInnen, um Verordnungsnetze zu identifizieren.
-

Beobachtungszeitraum:

- Je nach Fragestellungen kann der Beobachtungszeitraum frei gewählt werden. Es sollte ausreichend Vor- und Nachbeobachtungszeit eingeplant werden, um ggf. relevante Indikationen und/oder Krankenhausaufenthalte vor und nach der Verordnung zu identifizieren sowie die empirischen Verordnungsnetze charakterisieren zu können.
-

Analysemethode zusammengefasst:

Ermittlung von PatientInnen- oder Verordnergruppen und vergleichende univariate statistische Tests, um Unterschiede zwischen den Gruppen zu identifizieren.

Skizze des Analyseverfahrens:

- a) Identifikation von Patienten mit relevanter Medikation sowie der insgesamt verordneten Menge des Wirkstoffs über einen definierten Zeitraum.
- b) Einteilung der PatientInnen in inhaltlich relevante Gruppen (z.B. durchschnittliche Verordnungsmenge übersteigt empfohlene DDD oder die Anzahl der ausgestellten Verordnungen ist auffällig hoch, ...)
- c) Ermittlung Verordnernetze auf PatientInnenebene (welche/r verordnende/r ÄrztInnen teilt mit wem, wie viele Patienten) → Charakterisierung der Zusammenarbeit (eng/lose)
- d) Erhebung weiterer relevanter PatientInnen-Charakteristika
- e) Univariate statistische Tests zum Vergleich der Gruppen

Literatur

Zur Netzwerkidee allgemein:

Barnett, M. L. et al. (2011) 'Mapping physician networks with self-reported and administrative data', *Health services research*, 46(5), pp. 1592–1609. doi: 10.1111/j.1475-6773.2011.01262.x.

Coleman, J. S., Katz, E. and Menzel, H. (1966) *Medical Innovation: A Diffusion Study*. New York: Bobbs-Merrill Company.

Donohue, J. M. et al. (2018) 'Influence of peer networks on physician adoption of new drugs', *PLOS ONE*, 13(10), p. e0204826. doi: 10.1371/journal.pone.0204826.

Dugoff, E. H. et al. (2018) 'A scoping review of patient-sharing network studies using administrative data', *Translational Behavioral Medicine*, 8(4), pp. 598–625. doi: 10.1093/tbm/ibx015.

Landon, B. E. et al. (2013) 'Using Administrative Data to Identify Naturally Occurring Networks of Physicians', *Medical Care*, 51(8), pp. 715–721. doi: 10.1038/jid.2014.371.

Liu, Q. and Gupta, S. (2012) 'A Micro-level Diffusion Model for New Drug Adoption', *Journal of product innovation management*, 29(3), pp. 372–384. doi: 10.1111/j.1540-5885.2012.00912.x.

Lublóy, Á. (2014) 'Factors affecting the uptake of new medicines: a systematic literature review', *BMC Health Services Research*, 14(469). doi: 10.1186/1472-6963-14-469.

Stegbauer, C. (2016) *Grundlagen der Netzwerkforschung*. Edited by R. Häußling and Christian Stegbauer. Wiesbaden: Springer.

Zum Analyseansatz der Diffusionsanalyse (im WirtMedProjekt: Sacubitril/Valsartan):

Austin, P. C. (2017) 'A Tutorial on Multilevel Survival Analysis: Methods, Models and Applications', *International Statistical Review*, 85(2), pp. 185–203. doi: 10.1111/insr.12214.A.

van den Bulte, C. and Lilien, G. L. (2001) 'Medical Innovation Revisited: Social Contagion vs. Marketing Effort', *American Journal of Sociology*, 106(5), pp. 1409–1435. doi: 10.1086/320819.

Hastie, T. and Tibshirani, R. (1986) 'Generalized Additive Models', *Statistical Science*, 1(3), pp. 297–318. doi: 10.1214/ss/1177013604.

Iyengar, R., Bulte, C. Van Den and Valente, T. W. (2011) 'Opinion Leadership and Social Contagion in New Product Diffusion', *Marketing Science*, 30(2), pp. 195–212. doi: 10.1287/mksc.1100.0566.

Liu, Y., Yang, Y. K. and Hsieh, C. (2011) 'The determinants of the adoption of pharmaceutical innovation: Evidence from Taiwan', *Social Science & Medicine*, 72(6), pp. 919–927. doi: 10.1016/j.socscimed.2010.12.027.

Lublóy, Á. (2014) 'Factors affecting the uptake of new medicines: a systematic literature review', *BMC Health Services Research*, 14(469). doi: 10.1186/1472-6963-14-469.

Pollack, C. E., Soulos, P. R. and Gross, C. P. (2015) 'Physician's peer exposure and the adoption of a new cancer treatment modality', *Cancer*, 121(16), pp. 2799–2807. doi: 10.1002/cncr.29409

Zur Pregabalin-Analyse:

Ong, M.-S. et al. (2016) 'Provider Patient-Sharing Networks and Multiple-Provider Prescribing of Benzodiazepines.', *Journal of general internal medicine*, 31(2), pp. 164–171. doi: 10.1007/s11606-015-3470-8.

Ong, M.-S. et al. (2017) 'The Impact of Provider Networks on the Co-Prescriptions of Interacting Drugs: A Claims-Based Analysis', *Drug Safety*, 40(3), pp. 263–272. doi: 10.1007/s40264-016-0490-1.

Publizierte Analysen zu Pregabalin und Sacubitril/Valsartan aus dem WirtMed-Projekt:

Flemming, R., Frölich, F., Donner-Banzhoff, N., Sundmacher, L. (2023): Diffusion of a new drug among ambulatory physicians—The impact of patient pathways. *Health Economics*, 1– 13. doi: 10.1002/hec.4650.

Flemming, R. (2022): Patterns of pregabalin prescribing in four German federal states: analysis of routine data to investigate potential misuse of pregabalin. *BMJ Open* 2022;12:e060104. doi:10.1136/bmjopen-2021-060104.