



Pflege und Gesundheit

**Ulrike Famira-Mühlberger,
Christine Mayrhuber, Klaus Nowotny**

Wissenschaftliche Assistenz:
Marion Kogler, Stefan Weingärtner

Oktober 2021
Österreichisches Institut für Wirtschaftsforschung

Pflege und Gesundheit

Ulrike Famira-Mühlberger, Christine Mayrhuber, Klaus Nowotny

Oktober 2021

Österreichisches Institut für Wirtschaftsforschung
Im Auftrag des Dachverbandes der Sozialversicherungsträger

Begutachtung: Matthias Firgo

Wissenschaftliche Assistenz: Marion Kogler, Stefan Weingärtner

Auf Basis innovativer Modelle und detaillierter Gesundheitsdaten untersucht diese Studie – erstmals für Österreich – Zusammenhänge zwischen den bezogenen Gesundheitsleistungen von älteren Personen und dem Pflegegeldsystem. Der erstmalige Pflegegeldbezug lässt sich relativ gut prognostizieren. Dabei korrelieren das Alter, stationäre Aufenthalte in Krankenanstalten, Kontakte zu Allgemeinmedizinerinnen und -medizinern sowie Heilmittel, die das Nervensystem betreffen, am stärksten mit dem erstmaligen Pflegegeldbezug. Beim Übertritt in eine höhere Pflegegeldstufe zählen die bereits bestehende Pflegegeldstufe, das Alter sowie Heilmittel, die mit dem Nervensystem in Verbindung stehen, zu den Faktoren mit dem stärksten statistischen Zusammenhang. Die bestehende Pflegegeldstufe korreliert ebenfalls stark mit dem Eintritt in stationäre Pflege, ebenso die Häufigkeit von Kontakten zu Allgemeinmedizinerinnen und -medizinern sowie das Alter. Mit der Dauer der Krankenhausaufenthalte von Pflegegeldbeziehenden stehen die zuvor erfolgte Gesundheitsdiagnose der essenziellen (primären) Hypertonie sowie rezente Computertomographien von Kopf und Hals bzw. von Abdomen und Becken in einem statistischen Zusammenhang. Die Studie zieht gesundheitspolitische Schlussfolgerungen aus diesen Erkenntnissen.

2021/3/S/WIFO-Projektnummer: 11615

© 2021 Österreichisches Institut für Wirtschaftsforschung

Medieninhaber (Verleger), Herausgeber und Hersteller: Österreichisches Institut für Wirtschaftsforschung,
1030 Wien, Arsenal, Objekt 20 • Tel. (+43 1) 798 26 01-0 • <https://www.wifo.ac.at/> • Verlags- und Herstellungsort: Wien

Verkaufspreis: 50 € • Kostenloser Download: <https://www.wifo.ac.at/wwa/pubid/67194>

Inhalt

Abbildungsverzeichnis	III
Übersichtenverzeichnis	IV
Executive Summary	5
1. Einleitung: Ziele der Studie	8
2. Modul 1: Gesundheitsentwicklung & Pflegegeldbezug	9
2.1 Methodische Vorgangsweise	9
2.2 Daten und deskriptive Statistiken	12
2.3 Schätzungen auf Basis aggregierter Gesundheitsvariablen	16
2.4 Schätzungen auf Basis detaillierter Gesundheitsvariablen	28
2.5 Fazit und Einschätzung des Potentials zur Entwicklung eines kurzfristigen Vorlaufindikators	42
3. Modul 2: Gesundheitliche Gründe für einen Übergang in eine höhere Pflegegeldstufe	44
3.1 Ziel des Moduls 2	44
3.2 Methodik des Moduls 2 und Datenbeschreibung	44
3.3 Analyse der gesundheitlichen Gründe für einen Übergang in eine höhere Pflegegeldstufe	49
3.4 Main Findings und Fazit	55
4. Modul 3: Gesundheitliche Gründe für den Eintritt in stationäre Pflege	57
4.1 Ziel des Moduls 3	57
4.2 Empirisches Modell und Beschreibung der Daten	57
4.3 Analyse des Effekts von Gesundheitsleistungen auf den Eintritt in stationäre Pflege	61
4.4 Main Findings und Fazit	66
5. Modul 4: Häufigkeit von Aufenthaltstagen im Krankenhaus von Pflegegeldbeziehenden	68
5.1 Ziel des Moduls 4	68
5.2 Methodische Vorgangsweise: Empirisches Modell	68
5.3 Krankenhaushäufigkeit und Aufenthaltsdauer von Pflegegeldbeziehenden	70
5.4 Effekte der Diagnosen medizinischen Leistungen auf die Krankenhausaufenthaltsdauer	73
5.5 Medizinische Leistungen und Krankenhausaufenthalte	74
5.6 Main Findings – Kurzzusammenfassung	76
6. Zusammenfassung und Schlussfolgerungen	79
Literatur	85
7. Appendix	86
7.1 ROC-Kurven bei unterschiedlichen Beobachtungs- und Eventzeiträumen	86

7.2	ROC-Kurven bei unterschiedlichen Beobachtungs- und Eventzeiträumen, ohne unmittelbarem Vormonat	90
7.3	Ausführliche Methodenbeschreibung	93
7.3.1	Grundmodell	93
7.3.2	Logistische Regression	93
7.3.3	Trainings- und Testdaten zur Evaluierung der Prognosegüte	94
7.3.4	Wahrheitsmatrix, Sensitivität, Spezifität und ROC-Kurve	95
7.3.5	"Overfitting", Regularisierung und Kreuzvalidierung	97
7.4	Deskriptive Statistiken	100

Abbildungsverzeichnis

Abbildung 1: Schematische Darstellung der Zeiträume in der empirischen Untersuchung	9
Abbildung 2: Schematische Darstellung einer ROC-Kurve	11
Abbildung 3: Altersverteilung bei erstmaligen Pflegegeldbezug, nach Geschlecht	12
Abbildung 4: Pflegegeldstufe bei erstmaligen Pflegegeldbezug, nach Geschlecht	13
Abbildung 5: Erstmalige Pflegegeldbezüge pro Monat für die Jahre 2016—2018, Altersgruppe 60 – 85 Jahre	15
Abbildung 6: ROC-Kurven für Logit-Regression in Übersicht 1	19
Abbildung 7: ROC-Kurven für Logit-Regression in Übersicht 3	23
Abbildung 8: ROC-Kurven auf Basis der Testdaten für Modelle mit unterschiedlichem Beobachtungszeitraum, Eventzeitraum [01/2018, 02/2018]	26
Abbildung 9: ROC-Kurven auf Basis der Testdaten für Modelle mit unterschiedlichem Beobachtungszeitraum, Eventzeitraum [01/2018, 02/2018], Beobachtungszeitraum ohne Dezember 2017	27
Abbildung 10: ROC-Kurven für Modelle in Übersicht 6	29
Abbildung 11: Beta-Koeffizienten der 15 am stärksten korrelierenden Variablen auf Basis der Elastic-Net-Regression in Übersicht 6	33
Abbildung 12: Beta-Koeffizienten der 15 am stärksten korrelierenden Fachgebiets-Codes auf Basis der Elastic-Net-Regression in Übersicht 6	33
Abbildung 13: Beta-Koeffizienten der 15 am stärksten korrelierenden ICD-Diagnosen auf Basis der Elastic-Net-Regression in Übersicht 6	34
Abbildung 14: Beta-Koeffizienten der 15 am stärksten korrelierenden medizinischen Einzelleistungen auf Basis der Elastic-Net-Regression in Übersicht 6	35
Abbildung 15: Beta-Koeffizienten der 15 am stärksten korrelierenden ATC-Codes auf Basis der Elastic-Net-Regression in Übersicht 6	35
Abbildung 16: ROC-Kurven für Modelle in Übersicht 7, Beobachtungszeitraum [10/2017, 12/2017]	36
Abbildung 17: ROC-Kurven für Modelle in Übersicht 8, Eventzeitraum: [11/2018, 12/2018], Beobachtungszeitraum [08/2018, 11/2018]	39
Abbildung 18: ROC-Kurven von Logit-, LASSO- und Ridge Regression-Modellen für den Übertritt in eine höhere Pflegegeldstufe auf Basis der Testdaten	45
Abbildung 19: Veränderung der Pflegegeldstufen nach Geschlecht	46
Abbildung 20: Darstellung der Zielpflegegeldstufe bei Veränderung der Pflegegeldstufe, nach Alter und Geschlecht	47
Abbildung 21: Übertritte der Pflegegeldstufen	48
Abbildung 22: Flow-Diagramm aller Übertritte und Verbleibe von Pflegegeldbeziehenden am Ende des Beobachtungszeitraumes und am Ende des Eventzeitraumes	49
Abbildung 23: Die zehn wichtigsten Diagnosen, die mit einer Höhereinstufung der Pflegegeldstufe in Verbindung stehen	50
Abbildung 24: Medizinische Einzelleistungen und Übergang in eine höhere Pflegegeldstufe	51
Abbildung 25: Fachgebiete ärztlicher Leistungen und Übergang in eine höhere Pflegegeldstufe	52
Abbildung 26: Medikamente/Heilmittel und der Übergang in eine höhere Pflegegeldstufe	53
Abbildung 27: Sonstige persönliche Merkmale und der Übergang in eine höhere Pflegegeldstufe	54
Abbildung 28: Analyse des Übergangs in eine höhere Pflegegeldstufe: Die 15 wichtigsten Merkmale	55
Abbildung 29: Eintritt in stationäre Pflege nach Geschlecht	58
Abbildung 30: Eintritt in stationäre Pflege nach Pflegegeldstufe und Alter	59
Abbildung 31: Eintritt in stationäre Pflege und Höhereinstufung der Pflegegeldstufe ein halbes Jahr und ein Jahr vor Eintritt	60
Abbildung 32: ROC-Kurven von Logit-, LASSO- und Ridge Regression-Modellen für den Eintritt in eine höhere Pflegegeldstufe auf Basis der Testdaten	60
Abbildung 33: Effekte von unterschiedlichen Diagnosen auf Eintritte in stationäre Pflege	61
Abbildung 34: Effekte von medizinischen Einzelleistungen auf Eintritte in stationäre Pflege	62

Abbildung 35: Effekte der Inanspruchnahme von Gesundheitsanbietern nach Fachgebieten auf Eintritte in stationäre Pflege	63
Abbildung 36: Zusammenhänge von bezogenen Medikamenten/Heilmitteln auf Eintritte in stationäre Pflege	64
Abbildung 37: Effekte von sonstigen persönlichen Charakteristika auf Eintritte in stationäre Pflege	65
Abbildung 38: Die 15 wichtigsten Merkmale, die einen Effekt auf Übertritte in stationäre Pflege haben	66
Abbildung 39: Anzahl (linke Achse) und Dauer (rechte Achse) von Krankenhausaufenthalten der Pflegegeldbeziehenden, 2018	72
Abbildung 40: Anzahl von Krankenhausaufenthalten nach Pflegegeldstufen und Alter, 2018	72
Abbildung 41: Dauer von Krankenhausaufenthalten nach Pflegegeldstufen und Alter, 2018	73
Abbildung 42: Die zehn wichtigsten medizinischen Diagnosen von Pflegegeldbeziehenden mit Effekten auf ihre Krankenhausaufenthaltsdauer	74
Abbildung 43: Die zehn wichtigsten medizinischen Einzelleistungen von Pflegegeldbeziehenden mit Effekten auf ihre Krankenhausaufenthalte	75
Abbildung 44: Analyse der Krankenhausaufenthalte: Die 15 wichtigsten Merkmale	76
Abbildung A 1: ROC-Kurven auf Basis der Testdaten für Modelle mit unterschiedlichem Beobachtungszeitraum	87
Abbildung A 2: ROC-Kurven auf Basis der Testdaten für Modelle mit unterschiedlichem Beobachtungszeitraum	90
Abbildung A 3: Schematische Darstellung der Zeiträume in der empirischen Untersuchung	93
Abbildung A 4: Schematische Darstellung einer ROC-Kurve	97

Übersichtenverzeichnis

Übersicht 1: Logit-Schätzung für erstmaligen Pflegegeldbezug im Eventzeitraum [01/2018, 01/2019) auf Basis der Trainingsdaten (75% des Gesamtdatenbestands)	17
Übersicht 2: Modellmetriken für Logit-Regression in Übersicht 1	18
Übersicht 3: Logit-Schätzung für erstmaligen Pflegegeldbezug im Zeitraum [01/2018, 01/2019) auf Basis der Trainingsdaten (75% des Gesamtdatenbestands)	20
Übersicht 4: Modellmetriken für Logit-Regression in Übersicht 3	21
Übersicht 5: Fläche unter der ROC-Kurve (AUC) auf Basis der Testdaten für Modelle mit unterschiedlichen Betrachtungs- und Eventzeiträumen ($t = 01/2018$)	24
Übersicht 6: Modellmetriken für Logit-, LASSO-, Ridge und Elastic Net Regressionen	31
Übersicht 7: Modellmetriken für Logit-, LASSO-, Ridge und Elastic Net Regressionen	38
Übersicht 8: Modellmetriken für Logit-, LASSO-, Ridge und Elastic Net Regressionen	41
Übersicht 9: Durchschnittliche Inanspruchnahme der Gesundheitsdienstleistung Krankenhausaufenthalt nach Alter und Geschlecht, 2019	70
Übersicht A 1: Schematische Darstellung einer Wahrheitsmatrix	95
Übersicht A 2: Neueintritte in Pflegegeldstufen, nach Geschlecht	100
Übersicht A 3: Alter bei Neueintritt in Pflegegeld, nach Geschlecht	100
Übersicht A 4: Veränderung der Pflegegeldstufe, nach Geschlecht	101

Executive Summary

Die demographische Entwicklung der kommenden Jahrzehnte lässt eine stark wachsende Zahl an alten und hochaltrigen Menschen erwarten. Damit wird auch die Zahl an pflegebedürftigen Menschen steigen (Famira-Mühlberger, 2020; Klien et al., 2020). Die Perspektive einer zunehmenden Bedeutung des Pflegesystems für den öffentlichen Haushalt hat den Auftraggeber dazu veranlasst, überprüfen zu lassen, welche empirischen Zusammenhänge zwischen den abgerufenen Gesundheitsleistungen und dem Eintritt in das Pflegegeldsystem (Modul 1), dem Übertritt in eine höhere Pflegegeldstufe (Modul 2), dem Eintritt in die stationäre Pflege (Modul 3) sowie den gesundheitlichen Faktoren für Krankenhausaufenthalte von pflegegeldbeziehenden Personen (Modul 4) zu finden sind. Die in der Studie vorgenommenen Quantifizierungen – die mit neuen und erstmals für diesen Bereich angewendeten Methoden durchgeführt wurden – der vielschichtigen Effekte auf den Pflegegeldbezug zeigen gesundheitspolitische Handlungsfelder. Darüber hinaus überprüfte die Studie die Möglichkeit einer systematischen Vorhersehbarkeit von Pflegegeldbezug in Form eines Vorlaufindikators.

Der Dachverband der Sozialversicherungsträger hat für diese Studie anonymisierte Daten zu den bezogenen Gesundheitsleistungen¹⁾, persönlichen Charakteristika²⁾ und Informationen über stationäre Aufenthalte in Krankenanstalten³⁾ mit ausgewählten Daten der Pflegeinformation des Sozialministeriums (PFIF)⁴⁾ verknüpft und für die Analyse zur Verfügung gestellt.

Die Analyse von Modul 1 zeigt für die Altersgruppe 60 bis 85 Jahre, dass der erstmalige Pflegegeldbezug relativ gut aus den Daten prognostiziert werden kann. Vor allem rezente Gesundheitsleistungen sind für den erstmaligen Pflegegeldbezug von großer Bedeutung. Zu den am stärksten mit dem erstmaligen Pflegegeldbezug korrelierten Faktoren zählen in der multivariaten Analyse das Alter, stationäre Aufenthalte sowie Kontakte zu Allgemeinmedizinerinnen und -medizinern, aber auch Abgaben von Heilmitteln der ATC-Gruppe "N" (Nervensystem). Allerdings zeigt die Analyse auch, dass die Etablierung eines Vorlaufindikators auf individueller Ebene trotz der guten Prognoseleistung weitergehende Forschung (wie z.B. die Evaluierung alternativer fortgeschrittener Klassifizierungsalgorithmen) notwendig machen würde.

Die multivariate Analyse in Modul 2 zeigt die Zusammenhänge zwischen den in Anspruch genommenen Gesundheitsleistungen und den Übertritten in höhere Pflegegeldstufen von pflegegeldbeziehenden Personen ab 60 Jahren. Die stärkste Korrelation mit dem Übergang in eine höhere Pflegegeldstufe kann für die Pflegegeldstufe am Ende des Beobachtungszeitraums beobachtet werden. Eine bereits höhere Pflegegeldeinstufung hat – bei sonst identen

1) Häufigkeit von Kontakten zu Allgemeinmedizinerinnen und -medizinern, Fachärztinnen und -ärzten und anderen Vertragspartnerinnen und -partnern.

2) Alter, Geschlecht, Wohnort auf NUTS-3-Ebene.

3) Dauer, gestellte Haupt- und Nebendiagnosen auf Ebene der ICD-Kategorien sowie erbrachte medizinische Einzelleistungen.

4) Insbesondere Daten zu den Pflegegeldstufen und zum Übergang von Pflegegeldbezugs auf ein Pflegeheim.

Tatbeständen (*ceteris paribus*) – einen negativen Effekt auf die Wahrscheinlichkeit einer Höhereinstufung der Pflegegeldstufe⁵⁾. Ein höheres Alter zeigt – ebenso wie die Verschreibung von mit Demenz verbundenen Medikamenten – einen positiven Zusammenhang mit dem Übertritt in höhere Pflegegeldstufen.

Modul 3 analysiert die Zusammenhänge zwischen den abgerufenen Gesundheitsleistungen⁶⁾ und den Übertritten in stationäre Pflege von pflegegeldbeziehenden Personen ab 60 Jahren⁷⁾. Die multivariate Analyse zeigt, dass die Pflegegeldstufe am Ende des Beobachtungszeitraums die stärkste Erklärungskraft für die Eintritte in die stationäre Pflege hat: Je höher die Pflegegeldstufe, desto höher ist – *ceteris paribus* – die Wahrscheinlichkeit des Eintritts in die stationäre Pflege. Ein ebenfalls starker Zusammenhang, wenn auch schwächer in der Ausprägung, zeigt sich mit den ärztlichen Leistungen von Allgemeinmedizinerinnen und -medizinern⁸⁾ sowie mit einem höheren Alter.

Modul 4 analysiert die Zusammenhänge zwischen der Dauer stationärer Krankenhausaufenthalte und den zuvor in Anspruch genommenen Gesundheitsleistungen von Pflegegeldbeziehenden. Die mit Abstand stärkste Korrelation kann für die Diagnose "essentielle (primäre) Hypertonie" gefolgt von Krankheiten des Harnsystems sowie sonstigen Unfällen im privaten Bereich gefunden werden.

Aus der vorliegenden Analyse können wichtige gesundheitspolitische Schlussfolgerungen gezogen werden. Die Studie hebt die Relevanz des Alters hervor: Jene Personen, die gleiche Charakteristika und dieselben bezogenen Gesundheitsleistungen aufweisen (= *ceteris paribus*), sich aber nur durch ihr Alter von Vergleichspersonen unterscheiden, haben eine höhere Wahrscheinlichkeit, ins Pflegegeldsystem einzutreten, in höhere Pflegegeldstufen überzutreten oder in stationäre Pflege zu wechseln. Dies könnte bedeuten, dass das Alter für die Evaluierung der Pflegebedürftigkeit eine überproportionale Rolle spielt und Personen aufgrund ihres Alters, aber nicht primär aufgrund ihres Gesundheitszustandes eingestuft werden, wobei ein höheres Alter oftmals mit einem Funktionsverlust für die Aktivitäten des täglichen Lebens einhergeht.

Weiters verdeutlicht die Analyse die Bedeutung von Demenz: Sowohl beim Eintritt in das Pflegegeldsystem als auch bei einer Höhereinstufung spielen dementielle Erkrankungen eine wesentliche Rolle. Hier sollten im Rahmen der bestehenden Demenzstrategie die Präventionsfelder deutlich ausgebaut werden.

⁵⁾ Mögliche Erklärungen hierfür sind die Anspruchsvoraussetzungen der Pflegegeldstufen 5 bis 7, die sich nicht im Ausmaß des Pflegeaufwandes unterscheiden, sondern durch die gesundheitlichen Voraussetzungen. Ebenso können die oberen Pflegegeldstufen mit relativ "stabileren" Gesundheitszuständen verbunden sein, die sich vor dem Tod nicht mehr ändern.

⁶⁾ Im Beobachtungszeitraum 2016.

⁷⁾ Im Eventzeitraum 2017.

⁸⁾ In diesem Zusammenhang wohl ein Indiz für einen schlechten Gesundheitszustand.

Allgemeinmedizinerinnen und -mediziner spielen sowohl beim Eintritt in das Pflegegeldsystem als auch beim Eintritt in die stationäre Pflege eine wesentliche Rolle. Dies bedeutet, dass Allgemeinmedizinerinnen und -mediziner für die Einschätzung der Entwicklung des Pflege(geld)systems zentral sind und diese Informationen berücksichtigt werden sollten (Schnittstelle Gesundheit und Pflege) (siehe auch *Famira-Mühlberger et al.*, 2019).

Ergotherapien im Rahmen eines stationären Aufenthalts weisen einen negativen statistischen Zusammenhang mit der Anzahl der Krankenhaustage in der Folgeperiode auf. Es ist zu vermuten, dass Ergotherapie zentral für die Mobilität der Pflegegeldbeziehenden sein könnte und möglicherweise Unfälle im privaten Bereich und damit die Krankenhausbildung verringert. Die Rolle der Ergotherapie für die Resilienz der Pflegegeldbeziehenden sollte näher untersucht werden.

Zwischen manchen Regionen Österreichs unterscheiden sich – *ceteris paribus* – die Wahrscheinlichkeiten, erstmalig Pflegegeld zu beziehen, in eine höhere Pflegegeldstufe zu wechseln oder in stationäre Pflege überzutreten. Zur Erklärung dieser Unterschiede wären weitergehende wissenschaftliche Analysen wünschenswert.

1. Einleitung: Ziele der Studie

Die demographische Entwicklung der kommenden Jahrzehnte lässt eine stark wachsende Zahl an alten und hochaltrigen Menschen erwarten. Damit wird auch die Zahl an pflegebedürftigen Menschen steigen. Mit der steigenden Nachfrage sind – ceteris paribus – neben den privaten Kosten auch höhere Kosten für die öffentliche Hand verbunden (*Famira-Mühlberger, 2020; Klien et al., 2020*). Die Perspektive einer zunehmenden Bedeutung des Pflegesystems für den öffentlichen Haushalt stellt die Frage nach den Zusammenhängen zwischen den abgerufenen Gesundheitsleistungen und dem Eintritt in das Pflegegeldsystem, dem Übertritt in eine höhere Pflegegeldstufe, dem Eintritt in die stationäre Pflege sowie den gesundheitlichen Faktoren für die Krankenhausaufenthalte von pflegegeldbeziehenden Personen stärker ins Zentrum um daraus Ableitungen für eine verbesserte Versorgung und Präventionsansätze zu generieren.

Die Ziele der Studie sind:

- die Analyse der wichtigsten Determinanten für den erstmaligen Bezug von Pflegegeld sowie die Überprüfung einer systematischen Vorhersehbarkeit von Pflegegeldbezug ("Vorlaufindikator") (Modul 1),
- die Analyse der wichtigsten Determinanten für den Übertritt in eine höhere Pflegegeldstufe (Modul 2),
- die Analyse der wichtigsten Determinanten für den Eintritt in die stationäre Pflege (Modul 3), sowie
- die Analyse der wichtigsten Determinanten für die Krankenhausaufenthalte von pflegegeldbeziehenden Personen (Modul 4).

Der Dachverband der Sozialversicherungsträger hat für diese Studie anonymisierte Daten über die bezogenen Gesundheitsleistungen (Diagnosen auf Ebene der ICD-Kategorien, medizinische Einzelleistungen, Häufigkeit von Kontakten zu Allgemeinmedizinerinnen und -medizinern, Fachärztinnen und -ärzten und anderen Vertragspartnerinnen und -partnern), persönlichen Charakteristika (Alter, Geschlecht, Wohnort auf NUTS-3-Ebene) und stationären Aufenthalten in Krankenanstalten mit ausgewählten Daten der Pflegeinformation (PFIF, insbesondere Daten zu den Pflegegeldstufen und zum Übergang von Pflegegeldbezug auf ein Pflegeheim) verknüpft und für die Analyse zur Verfügung gestellt. In Summe stehen 3.614 potenzielle Variablen zur Verfügung.

Diese Studie stellt die erstmalige Verknüpfung von Gesundheitsleistungsdaten und Pflegeinformationsdaten für Österreich dar. Sie präsentiert dementsprechend Ergebnisse, die bislang für die Entwicklung von gesundheitspolitischen Maßnahmen noch nie zur Verfügung standen. Dies ist ein erster Versuch einer gemeinsamen Analyse dieser unterschiedlichen Datensätze und hoffentlich ein Ausgangspunkt für weitere Analysen, die zu einer Verbesserung sowohl des Gesundheitssystems als auch des Pflegesystems beitragen.

2. Modul 1: Gesundheitsentwicklung & Pflegegeldbezug

Ziel dieses Moduls ist es, die von Individuen in Anspruch genommenen Gesundheitsleistungen mit dem erstmaligen Eintritt in den Pflegegeldbezug in Verbindung zu setzen. Dabei sollten die in Anspruch genommenen Gesundheitsleistungen vor dem ersten Pflegegeldbezug systematisch analysiert werden, um die Entwicklung der Pflegegeldbeziehenden in kürzerer Frist zu untersuchen. Auf Basis dieser Ergebnisse soll gezeigt werden, welche Art von Leistungsbezügen mit dem erstmaligen Eintritt in den Pflegegeldbezug signifikant in Verbindung stehen, um in weiterer Folge zu eruieren, ob daraus ein "Vorlaufindikator" für den zukünftigen Pflegegeldbezug auf individueller Ebene entwickelt werden kann.

Dabei handelt es sich um eine explorative Datenanalyse, d.h. es werden keine aus einer Theorie abgeleiteten, im Vorfeld definierten Hypothesen überprüft. Vielmehr kommen geeignete Methoden des statistischen Lernens zur Anwendung, welche geeignet sind, aus einer hohen Anzahl potenzieller Erklärungsfaktoren jene Variablen herauszufiltern, die – gegeben alle anderen Erklärungsfaktoren – den stärksten Zusammenhang mit dem erstmaligen Pflegegeldbezug aufweisen. Dadurch können noch unbekannte, aber in den Daten vorhandene Zusammenhänge zwischen erstmaligem Pflegegeldbezug und persönlichen Eigenschaften bzw. Gesundheitsleistungen aufgedeckt werden. Hierbei ist jedoch darauf hinzuweisen, dass es sich um hypothesengenerierende Verfahren handelt. Die statistisch identifizierten Zusammenhänge sind daher nicht als Kausalzusammenhänge, sondern als Korrelationen (bzw. als Präzedenzen, d.h. dem erstmaligen Pflegegeldbezug vorausgehende Faktoren) zu verstehen. Im Folgenden wird die methodische Vorgangsweise kurz erläutert. Eine ausführlichere Methodenbeschreibung findet sich in Appendix 7.3.

2.1 Methodische Vorgangsweise

Das Ziel der empirischen Analyse ist, auf Basis statistischer Modelle zu ermitteln, welche Wahrscheinlichkeit eines erstmaligen Bezugs von Pflegegeld innerhalb eines Eventzeitraums $[t, t + b)$ sich aus den persönlichen Eigenschaften einer Person i , welche mit X_i bezeichnet werden, sowie aus den von Person i im vorgelagerten Beobachtungszeitraum $[t - a, t)$ konsumierten Gesundheitsleistungen G_i ableiten lässt.^{?)}

Abbildung 1: **Schematische Darstellung der Zeiträume in der empirischen Untersuchung**



Q: Eigene Darstellung.

^{?)} Diese Intervalle sind inklusive dem Startzeitpunkt aber exklusiv dem Endzeitpunkt zu verstehen. Der Zeitpunkt t ist daher z.B. nur im Eventzeitraum enthalten, nicht jedoch im Beobachtungszeitraum.

Eine der gängigsten Methoden zur Modellierung solcher Zusammenhänge ist die logistische Regression (Logit-Regression). Die logistische Regression modelliert den Logarithmus der "Odds" (darunter versteht man das Verhältnis der Wahrscheinlichkeit, dass ein Ereignis eintritt, zur Wahrscheinlichkeit, dass das Ereignis nicht eintritt) als lineare Funktion erklärender Variablen. Die aus einer Stichprobe geschätzten Parameter dieses statistischen Modells können dazu genutzt werden, den zukünftigen Status der abhängigen Variable "erstmaliger Pflegegeldbezug" auch für Beobachtungen zu prognostizieren, die sich nicht innerhalb der Stichprobe befinden.

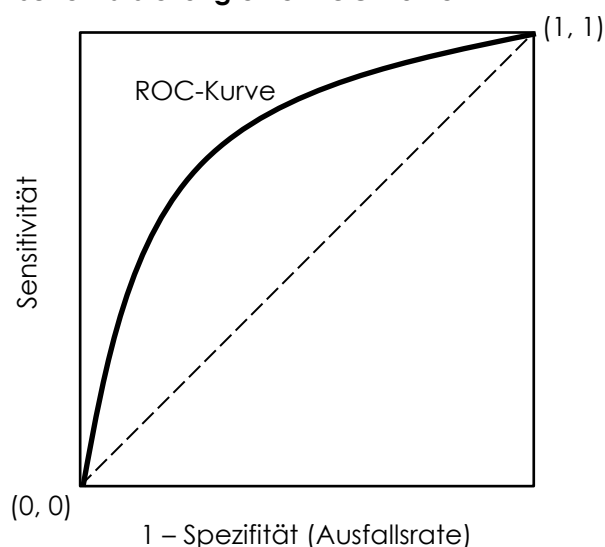
Um die Prognosegüte des Modells zu evaluieren, werden die vorhandenen Beobachtungen im Verhältnis 75% zu 25% in Trainings- und Testdaten geteilt (vgl. *Hastie et al.*, 2009, S. 222f). Die Trainingsdaten werden dazu genutzt, das statistische Modell zu optimieren, d.h. die Modellparameter werden so gewählt, dass sie optimal an die in den Trainingsdaten enthaltenen Beobachtungen angepasst sind und die abhängige Variable in den Trainingsdaten bestmöglich erklären. Dies erlaubt jedoch nur begrenzte Rückschlüsse darauf, wie gut das Modell die abhängige Variable in Daten prognostizieren würde, auf die es nicht "trainiert" ist. Die Prognosegüte wird daher nicht anhand der Trainingsdaten, sondern anhand der Testdaten evaluiert. Dazu werden das aus den Trainingsdaten geschätzte Modell und die in den Testdaten vorhandenen Werte der erklärenden Variablen verwendet, um die abhängige Variable zu prognostizieren. Der Vergleich zwischen der in den Testdaten prognostizierten Wahrscheinlichkeit und den dort tatsächlich beobachteten Werten kann dann als Grundlage für die Evaluierung der Prognosegüte verwendet werden. Eine Möglichkeit, die Prognosegüte zu evaluieren, ist, die Beobachtungen anhand der prognostizierten Wahrscheinlichkeiten zu klassifizieren, indem Beobachtungen, deren prognostizierte Wahrscheinlichkeit einen Schwellenwert τ (z.B. 0,5 = 50%) übersteigen, ein "positiver" prognostizierter Wert zugeteilt wird und Beobachtungen, deren prognostizierte Wahrscheinlichkeit den Schwellenwert nicht überschreitet, ein "negativer" prognostizierter Wert zugeteilt wird.

Die Beurteilung der Prognosegüte kann durch die Konzepte der "Sensitivität" und "Spezifität" durchgeführt werden. Als Sensitivität oder Richtig-Positiv-Rate bezeichnet man den Anteil der "richtig positiv" klassifizierten Beobachtungen an allen tatsächlich "positiven" Beobachtungen. Dies entspricht dem Anteil der Personen, denen vom Modell korrekterweise ein erstmaliger Pflegegeldbezug prognostiziert wird, an allen Personen, die tatsächlich erstmalig Pflegegeld beziehen. Die Spezifität oder Richtig-Negativ-Rate hingegen gibt den Anteil der "richtig negativ" prognostizierten Werte an allen tatsächlich "negativen" Beobachtungen an, d.h. den Anteil der Personen, denen vom Modell korrekterweise kein erstmaliger Pflegegeldbezug vorhergesagt wird, an allen Personen, die tatsächlich nicht erstmalig Pflegegeld beziehen.

Ein gutes Modell sollte bei gegebener Sensitivität eine möglichst hohe Spezifität aufweisen und *vice versa*. Häufig wird anstelle der Spezifität auch die "Ausfallsrate" oder Falsch-Positiv-Rate verwendet. Diese gibt an, wie hoch der Anteil der fälschlicherweise als positiv klassifizierten Beobachtungen an allen tatsächlich negativen Beobachtungen ist.

Um die Wechselwirkung zwischen Sensitivität und Spezifität grafisch darzustellen, wird häufig auf die Darstellung der "Receiver Operating Characteristic" (ROC)-Kurve zurückgegriffen (siehe Winkelmann – Boes, 2009, S. 127f oder James et al., 2013, S. 147f). Diese stellt in einer zweidimensionalen Darstellung mit der Sensitivität auf der vertikalen und der Ausfallsrate ($1 - \text{Spezifität}$) auf der horizontalen Achse die Kombinationen von Sensitivität und $1 - \text{Spezifität}$ für alle Werte im Intervall $[0, 1]$ dar (vgl. Abbildung 2). Die sich daraus ergebende Kurve ist durch die Punkte $(0, 0)$ und $(1, 1)$ begrenzt und liegt üblicherweise über der diagonalen Linie zwischen diesen beiden Punkten. Diese Diagonale stellt die ROC-Kurve eines nichtinformativen "Modells" dar, das Beobachtungen rein zufällig in eine der beiden Klassen einteilt.

Abbildung 2: **Schematische Darstellung einer ROC-Kurve**



Q: Eigene Darstellung

Die Fläche unter dieser ROC-Kurve wird auch AUC ("Area Under the Curve") genannt und stellt eine Maßzahl dafür dar, wie gut Sensitivität und Spezifität eines empirischen Modells sind. Diese variiert zwischen 0,5 (der Fläche unter der diagonalen Linie, welche die ROC-Kurve eines nichtinformativen Modells darstellt) und 1 (der Fläche unter der ROC-Kurve eines optimalen Modells mit einer Sensitivität von 1 und einer Ausfallsrate von 0). Je näher die AUC an 1 liegt, desto besser das empirische Modell. Als Faustregeln für die Einordnung der Prognosegüte anhand der AUC bescheinigen Hosmer – Lemeshow (2000, S. 162) Modellen mit $0,7 \leq AUC < 0,8$ ein "akzeptables" Unterscheidungsvermögen ("acceptable discrimination"), Modellen mit $0,8 \leq AUC < 0,9$ ein "exzellentes" Unterscheidungsvermögen ("excellent discrimination") und Modellen mit $AUC \geq 0,9$ ein "überragendes" ("outstanding discrimination") Unterscheidungsvermögen.

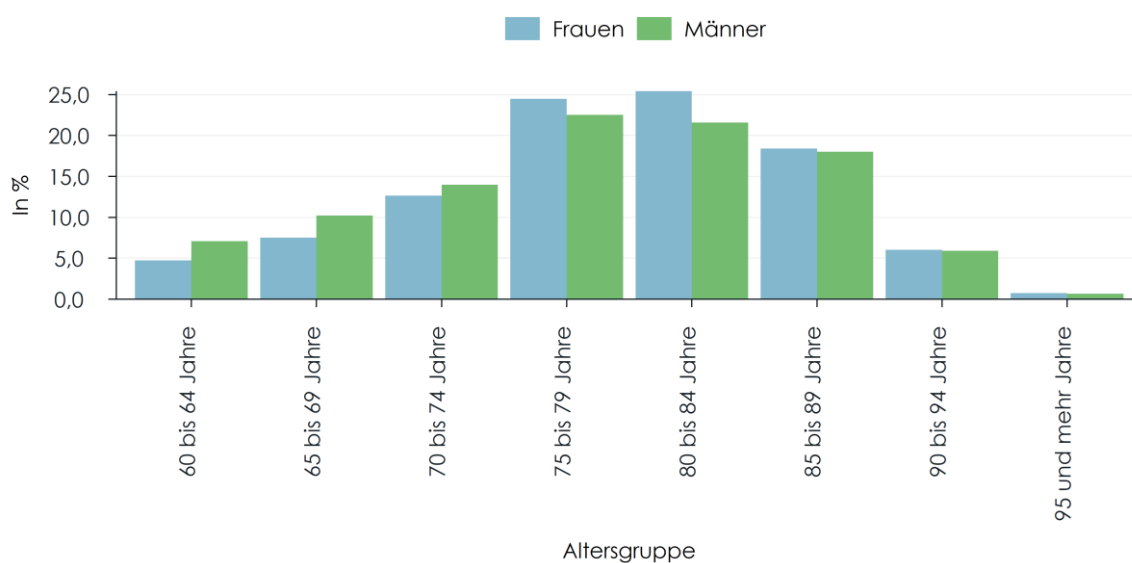
Die Anpassung des statistischen Modells an die Trainingsdaten birgt, insbesondere bei einer hohen Anzahl an erklärenden Variablen, die Gefahr des "Overfitting" (Überanpassung des statistischen Modells an die Trainingsdaten). Eine Möglichkeit, das Risiko des "Overfitting" zu

reduzieren, besteht in der Anwendung von Regularisierungsmethoden. Dabei werden die Parameter des Modells "geschrumpft", sodass eine Überanpassung durch zu starke Einflüsse einzelner Variablen reduziert wird. Zwei der am weitest verbreiteten Regularisierungsmethoden sind die "Ridge Regression" (Hoerl – Kennard, 1970) und der "Least Absolute Shrinkage and Selection Operator" (LASSO, Tibshirani, 1996). Der Unterschied zwischen Ridge Regression und LASSO besteht darin, dass die Regularisierung im LASSO dazu führen kann, dass einige Parameter auf exakt Null geschrumpft werden. In der Ridge Regression werden die Parameter zwar gegen, aber nie exakt auf Null geschrumpft. Die Stärke dieser Schrumpfung wird dabei durch einen "Strafterm" λ reguliert, dessen optimaler Wert durch 5-fache Kreuzvalidierung ermittelt wird (siehe Appendix 7.3). Der LASSO dient demnach auch der Variablenselektion, was insbesondere in Anwendungen mit vielen potenziellen Einflussvariablen von Vorteil ist. Aus diesem Grund wird das LASSO-Regressionsmodell für die Beurteilung der wichtigsten (d.h. am stärksten mit der abhängigen Variable korrelierten) Variablen in Modul 2 bis 4 verwendet.

2.2 Daten und deskriptive Statistiken

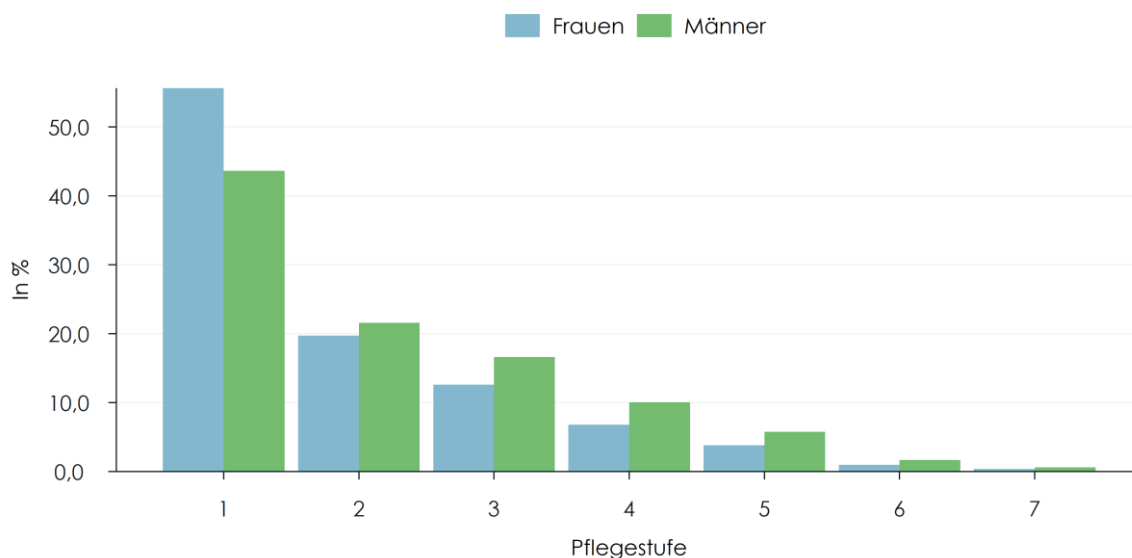
Zur Untersuchung des erstmaligen Pflegegeldbezugs wurden anonymisierte Individualdaten des Dachverbands der Sozialversicherungsträger für die Gesamtheit jener 550.960 Personen zur Verfügung gestellt, welche im Zeitraum 2016 – 2018 60 Jahre oder älter waren und Pflegegeld bezogen. Davon bezogen 388.480 bereits im Jänner 2016 (erstmalig oder wiederholt) Pflegegeld, während 162.480 Personen zwischen Februar 2016 und Dezember 2018 erstmalig Pflegegeld bezogen. Für die Untersuchung des erstmaligen Pflegegeldbezugs innerhalb eines Eventzeitraums $[t, t + b)$ werden nur Beobachtungen verwendet, für welche in den Daten kein Pflegegeldbezug vor dem Beginn des Eventzeitraums (vor dem Zeitpunkt t) registriert wird.

Abbildung 3: Altersverteilung bei erstmaligen Pflegegeldbezug, nach Geschlecht



Q: DVS, WIFO.

Abbildung 4: **Pflegegeldstufe bei erstmaligen Pflegegeldbezug, nach Geschlecht**



Q: DSVS, WIFO.

Im Durchschnitt waren diese bei erstmaligem Pflegegeldbezug 78,2 Jahre alt (Median: 79 Jahre), drei Viertel waren 84 Jahre alt oder jünger.¹⁰⁾ Während Männer im Durchschnitt mit 77,5 Jahren erstmalig Pflegegeld bezogen, sind Frauen beim erstmaligen Pflegegeldbezug über ein Jahr älter (78,7 Jahre, siehe Abbildung 3).¹¹⁾ Die Hälfte der Erstbeziehenden weist dabei Pflegegeldstufe 1 auf, weitere 20% Pflegegeldstufe 2. Allerdings zeigen sich auch hier Unterschiede zwischen Männern und Frauen (Abbildung 4): Während Männer nur zu 43,7% bei erstmaligem Pflegegeldbezug in Stufe 1 sind, beträgt dieser Anteil unter Frauen 55,7%.

Um eine Klassifizierung vornehmen zu können, sind neben den Daten für erstmalige Pflegegeldbezieherinnen und -bezieher auch Beobachtungen für Personen ohne erstmaligem Pflegegeldbezug zwischen 2016 und 2018 notwendig. Weisen etwa Personen mit späterem Pflegegeldbezug Muster in den abgerufenen Gesundheitsleistungen auf, welche sie von Personen ohne späteren Pflegegeldbezug unterscheiden, so könnten diese ohne eine solche "Kontrollgruppe" ohne Pflegegeldbezug nicht identifiziert werden. Zudem würde die ausschließliche Verwendung von Personen mit erstmaligem Pflegegeldbezug bei der Untersuchung der Bestimmungsfaktoren zu einer endogenen Selektion auf Basis der abhängigen (zu erklärenden) Variable und damit zu verzerrten und inkonsistenten Schätzergebnissen führen.¹²⁾ Dies

¹⁰⁾ Das Alter wird in den zur Verfügung gestellten Daten durch das Geburtsjahr approximiert; dabei wird ein Geburtstag zur Jahresmitte angenommen.

¹¹⁾ Die in den übermittelten Daten vorhandenen und in den Abbildungen dargestellten Beobachtungen für Personen unter 60 Jahren werden in der folgenden Untersuchung nicht berücksichtigt.

¹²⁾ Dieses Problem entsteht dadurch, dass die Variablen, welche den Indikator PG_i beeinflussen, damit gleichzeitig beeinflussen, ob die Person sich überhaupt in der Stichprobe befindet. Dadurch kommt es zu einer Korrelation der erklärenden Variablen X_i und G_i mit dem Fehlerterm ε_i , der zu verzerrten und inkonsistenten Schätzergebnissen führt.

kann auch dann zutreffen, wenn der erstmalige Pflegegeldbezug nicht in den Eventzeitraum t bis $t + b$ selbst fällt, sondern erst nach $t + b$ eintritt, wenn sich die erklärenden Variablen zwischen Personen, die im Beobachtungszeitraum nie Pflegegeld beziehen, und Personen, die irgendwann zwischen 2016 und 2018 Beobachtungszeitraum erstmalig Pflegegeld beziehen, systematisch unterscheiden. In diesem Fall ist die externe Validität des Ansatzes nicht mehr gegeben, was zu gravierenden Klassifikationsfehlern führen kann. So ist es z.B. möglich, dass selbst ein Prädiktor, der bei Personen, die irgendwann im Beobachtungszeitraum erstmalig Pflegegeld beziehen, konsistent auftritt, nicht als Indikator erkannt wird, wenn er innerhalb dieser Gruppe eine zu geringe Variation aufweist.

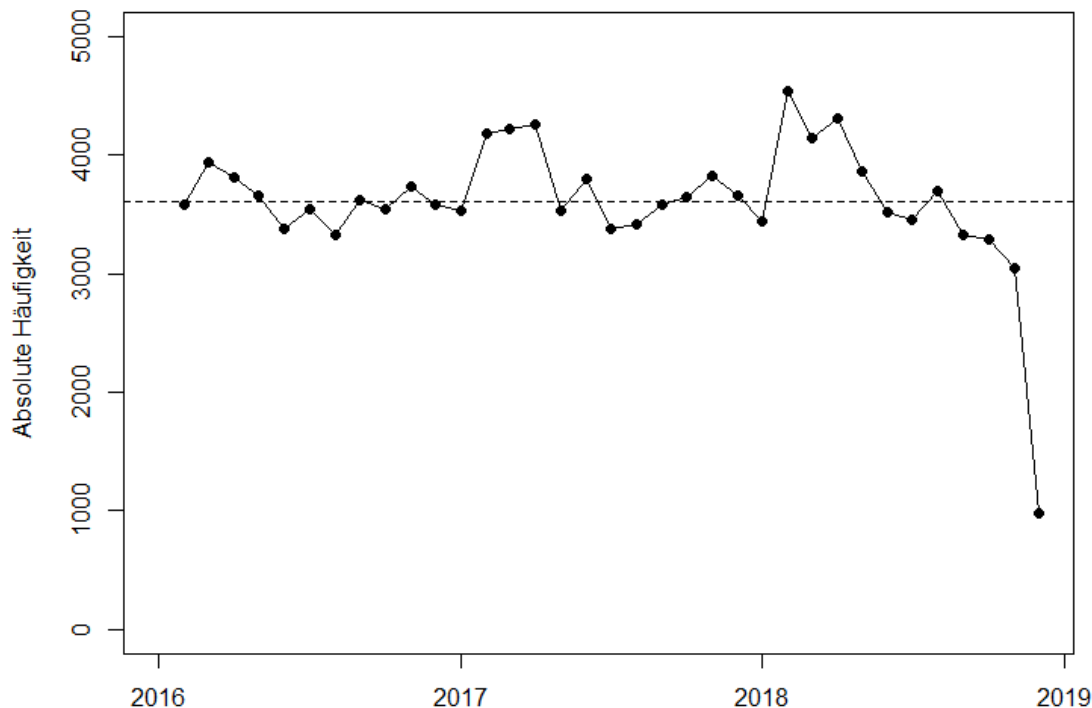
Als "Kontrollgruppe" kommen Personen in Frage, welche zwischen 2016 und 2018 keinen Pflegegeldbezug aufweisen. Aus dieser Gruppe wurde dann – nach Wohnregion, Geschlecht und 5-Jahres-Altersgruppe stratifiziert – eine anonymisierte Zufallsstichprobe aus dem Datenbestand des Dachverbands der Sozialversicherungsträger gezogen. Um eine ausreichende Variabilität innerhalb der erklärenden Variablen sicherzustellen, umfasst die Kontrollgruppe innerhalb aller Strata etwa doppelt so viele Personen wie die Anzahl der Pflegegeldbezieherinnen und -bezieher im gleichen Stratum ("Case-Control-Sampling"). Aufgrund der hohen Anzahl an Pflegegeldbezieherinnen und -bezieher innerhalb der Gruppe der 81- bis 85-Jährigen wurde für diese Altersgruppe eine Kontrollgruppe gezogen, welche in etwa der Anzahl der Pflegegeldbezieherinnen und -bezieher entspricht. Da eine Stichprobenziehung für die Kontrollgruppe in Größe der Anzahl der Pflegegeldbezieherinnen und -bezieher für Über-85-Jährige einer Vollerhebung gleichgekommen wäre, werden in diesem Modul ausschließlich Personen zwischen 60 und 85 Jahren betrachtet.

Aufgrund der Stichprobenziehung wird in den empirischen Modellen dieses Moduls der Eventzeitraum für den erstmaligen Pflegegeldbezug in das Jahr 2018 gelegt. Der Grund dafür ist, dass es ansonsten zu einem Selektionseffekt kommen könnte.¹³⁾

Insgesamt stehen aus dieser Altersgruppe für die Auswertung Daten für 435.332 Personen der Kontrollgruppe sowie für 126.400 Personen mit erstmaligem Pflegegeldbezug zwischen 2016 und 2018 zur Verfügung. In den Jahren 2016 – 2018 verzeichnen die Daten für die 60- bis 85-Jährigen pro Monat im Durchschnitt 3.611 Neueintritte in den Pflegegeldbezug. Lediglich im Dezember 2018 wird mit 982 eine deutlich geringere Zahl an erstmaligen Pflegegeldbezieherinnen und -bezieher ausgewiesen (Abbildung 5), was möglicherweise darauf zurückzuführen ist, dass nachträgliche Meldungen von Erstbezügen im Dezember 2018 in den vorhandenen Daten noch nicht berücksichtigt wurden.

¹³⁾ Die Kontrollgruppe wurde aus dem Bestand jener Personen gezogen, die in den Jahren 2016–2018 keinen erstmaligen Pflegegeldbezug aufwiesen und bis Ende 2018 im aktiven Datenbestand waren. Dies bedeutet, dass für Personen in der Kontrollgruppe im Zeitraum 2016–2018 kaum Todesfälle verzeichnet werden. *Ceteris paribus* weist diese Gruppe demnach in den Jahren 2016–2018 einen günstigeren Gesundheitsverlauf auf, sodass diese Gruppe positiv selektiert ist. Da die Entwicklung nach 2018 die Stichprobenauswahl nicht beeinflusst, sollte der Selektionseffekt am Ende der Sampleperiode (2018) deutlich geringer sein als zu Beginn (2016).

Abbildung 5: **Erstmalige Pflegegeldbezüge pro Monat für die Jahre 2016—2018, Altersgruppe 60 – 85 Jahre**



Q: DSVS, WIFO.

Zusätzlich zu diesen Daten, aus denen die binäre abhängige Variable $PG_{i,[t,t+b)}$ (= 1, wenn zwischen t und $t+b$ erstmalig Pflegegeld bezogen wird; sonst 0) abgeleitet wird und welche neben dem Pflegegeldbezug auch Auskunft über persönliche Faktoren wie Alter, Geschlecht und Wohnort geben, wurden Seitens des Dachverbands der Sozialversicherungsträger für alle dort beinhalteten Personen auch Daten über die in Anspruch genommenen Gesundheitsleistungen in den Jahren 2015 – 2018 für alle Personen zur Verfügung gestellt.

Die Daten zu den Gesundheitsleistungen gliedern sich in drei Gruppen. Die erste Gruppe umfasst Daten aus dem niedergelassenen Bereich und beinhaltet Informationen über die Kontakthäufigkeit zu (Fach-) Ärzten und Ärztinnen und Vertragspartnern unterteilt nach 58 Fachgebieten. Die höchste Kontakthäufigkeit wird im gesamten Datenbestand im Zeitraum 2015 – 2018 zu Ärzten und Ärztinnen für Allgemeinmedizin verzeichnet (Fachgebiets-Code 1, ca. 70 Mio. Kontakte), gefolgt von Rettungsdiensten und Beförderungsgewerbe (Code 65, ca. 7 Mio. Kontakte), Fachärzten und Fachärztinnen für medizinische und chemische Labordiagnostik (Code 50, ca. 4,5 Mio. Kontakte), Fachärzten und Fachärztinnen für Inneres (Code 7, ca. 3,6 Mio. Kontakte) sowie Bandagisten und Bandagistinnen (Code 61, 3,4 Mio. Kontakte). Die zweite Gruppe umfasst abgegebene Heilmittel (ohne magistrale Zubereitungen, keine Registrierung von Medikamenten unter den Kosten der Rezeptgebühr) auf der 5. Ebene der

Anatomisch-Therapeutisch-Chemischen Klassifikation (ATC). Da über 20% dieser Codes im vierjährigen Zeitraum 2015 – 2018 weniger als 100 Abgaben aufweisen, werden die Heilmittelabgaben auf ATC-Ebene 4 aggregiert (483 unterschiedliche Codes). Zu den am häufigsten abgegebenen Mitteln zählen Protonenpumpen-Hemmer (ATC-Code A02BC, 530.515 Abgaben; A02B: Mittel bei peptischem Ulkus und gastrooesophagealer Refluxkrankheit), HMG-CoA-Reduktasehemmer (ATC-Code C10AA, 404.469 Abgaben; C10A: Mittel, die den Lipidstoffwechsel beeinflussen), Kombinationen von Penicillinen, inkl. Beta-Lactamase-Inhibitoren (ATC-Code J01CR, 382.077 Abgaben; J01C: Betalactam-Antibiotika, Penicilline), Propionsäure-Derivate (ATC-Code M01AE, 324.267 Abgaben; M01A: Nichtsteroidale Antiphlogistika und Antirheumatika) sowie Pyrazolone (ATC-Code N02BB, 323.521 Abgaben; N02B: Andere Analgetika und Antipyretika).

Die dritte Gruppe beinhaltet Daten aus dem Krankenanstaltenbereich und informiert über Krankenhausaufenthalte (Beginn- und Enddatum), die gestellten Haupt- und Zusatzdiagnosen auf Ebene der vierstelligen Systematik der Internationalen statistischen Klassifikation der Krankheiten und Gesundheitsprobleme (ICD-10) sowie die erbrachten medizinischen Einzelleistungen. Von den 2.832.282 Krankenhausaufenthalten enden ca. 22,1% am selben Tag. Für diese Aufenthalte werden die gestellten Haupt- und Zusatzdiagnosen auf Ebene der dreistelligen ICD-Systematik (1.483 unterschiedliche Codes) sowie die medizinischen Einzelleistungen (1.553 unterschiedliche Codes) in der Untersuchung verwendet, wobei in der Auswertung nicht unterschieden wird, ob eine Diagnose als Hauptdiagnose oder als Zusatzdiagnose gestellt wurde. Für alle anderen Aufenthalte wird zudem die Anzahl der Nächte in stationärem Aufenthalt berücksichtigt. Die am häufigsten gestellten Diagnosen im Rahmen stationärer Aufenthalte sind essentielle (primäre) Hypertonie (ICD-Code I10, 801.317 Mal gestellt), Vorhofflimmern und Vorhofflattern (ICD-Code I48, 341.880 Mal gestellt), chronische ischämische Herzkrankheit (ICD-Code I25, 313.280 Mal gestellt), Diabetes Typ 2 (ICD-Code E11, 297.683 Mal gestellt) sowie chronische Nierenkrankheit (ICD-Code N18, 226.667 Mal gestellt). Zu den am häufigsten erbrachten medizinischen Einzelleistungen zählen: Physiotherapie im Rahmen eines stationären Aufenthaltes (MEL-Code PE010, 683.846 Mal), Computertomographie von Kopf und Hals (MEL-Code ZA010, 341.798 Mal), Computertomographie von Abdomen und Becken (MEL-Code ZC010, 223.532 Mal), extrakapsuläre Kataraktoperation mit Linsenimplantation (MEL-Code BF020, 172.579 Mal) sowie Computertomographie des Thorax (MEL-Code ZB010, 165.835 Mal).

2.3 Schätzungen auf Basis aggregierter Gesundheitsvariablen

Um eine Übersicht über das Prognosepotential dieser persönlichen Eigenschaften und konsumierten Gesundheitsleistungen für den erstmaligen Pflegegeldbezug zu gewinnen wird in einem ersten Schritt die logistische Regression auf den erstmaligen Pflegegeldbezug im Jahr 2018 – Eventzeitraum: [01/2018, 01/2019) – als abhängige Variable angewandt, welche durch persönliche Eigenschaften sowie die im Jahr 2017 erbrachten Gesundheitsleistungen – Beobachtungszeitraum: [01/2017, 01/2018) – erklärt werden sollen. Der Gesamtdatenbestand

an Personen, für welche vor Jänner 2018 kein Pflegegeldbezug beobachtet wird (473.982 Personen in Pflegegeld- und Kontrollgruppe) wird dazu in Trainingsdaten (355.486 Beobachtungen bzw. 75% des Gesamtdatenbestands) und Testdaten (118.496 Beobachtungen bzw. 25%) geteilt. Während erstere zur Schätzung herangezogen werden, dienen letztere der Beurteilung der Prognosekraft des Modells.

Übersicht 1: **Logit-Schätzung für erstmaligen Pflegegeldbezug im Eventzeitraum [01/2018, 01/2019) auf Basis der Trainingsdaten (75% des Gesamtdatenbestands)**

Variable	Abhängige Variable Erstmaliger Pflegegeldbezug		Variable	Abhängige Variable Erstmaliger Pflegegeldbezug	
	Koeffizient	Std		Koeffizient	Std
Alter	0,085***	(0,001)	NUTS3: AT225	-0,324***	(0,041)
Weiblich (= 1)	-0,027**	(0,013)	NUTS3: AT226	-0,434***	(0,055)
Kontakthäufigkeit FG 01 (AllgemeinmedizinerInnen)	0,031***	(0,0005)	NUTS3: AT311	-0,194***	(0,039)
Dauer stationärer Aufenthalte	0,051***	(0,001)	NUTS3: AT312	-0,215***	(0,031)
NUTS3: AT111	0,006	(0,073)	NUTS3: AT313	-0,234***	(0,047)
NUTS3: AT112	-0,101**	(0,044)	NUTS3: AT314	-0,300***	(0,055)
NUTS3: AT113	-0,033	(0,048)	NUTS3: AT315	-0,185***	(0,044)
NUTS3: AT121	-0,263***	(0,039)	NUTS3: AT321	-0,089	(0,115)
NUTS3: AT122	-0,235***	(0,037)	NUTS3: AT322	-0,194***	(0,050)
NUTS3: AT123	-0,248***	(0,048)	NUTS3: AT323	-0,088**	(0,036)
NUTS3: AT124	-0,217***	(0,037)	NUTS3: AT331	-0,242**	(0,115)
NUTS3: AT125	-0,163***	(0,047)	NUTS3: AT332	-0,065*	(0,039)
NUTS3: AT126	-0,055	(0,034)	NUTS3: AT333	-0,132	(0,086)
NUTS3: AT127	-0,058*	(0,035)	NUTS3: AT334	-0,233***	(0,066)
NUTS3: AT211	0,029	(0,034)	NUTS3: AT335	-0,209***	(0,045)
NUTS3: AT212	-0,051	(0,051)	NUTS3: AT341	-0,155**	(0,074)
NUTS3: AT213	-0,089**	(0,043)	NUTS3: AT342	-0,089**	(0,040)
NUTS3: AT221	-0,046	(0,031)	Konstante	-9,201***	(0,091)
NUTS3: AT222	-0,435***	(0,067)			
NUTS3: AT223	-0,242***	(0,040)			
NUTS3: AT224	-0,227***	(0,037)			
			Beobachtungen	355.486	
			Log Likelihood	-93.910,150	

Anmerkung: *p<0,1; **p<0,05; ***p<0,01.

Die Logit-Regression in Übersicht 1 berücksichtigt (neben persönlichen Eigenschaften wie Alter, Geschlecht und Wohnort auf NUTS-3-Ebene) in einem ersten Schritt lediglich die Anzahl der Kontakte zu Ärztinnen und Ärzte für Allgemeinmedizin (Fachgebiets-Code 01) sowie die Dauer stationärer Aufenthalte (gemessen durch die Anzahl der Nächte in Krankenhausaufenthalten) als Gesundheitsvariablen. Die Regressionsergebnisse für die Trainingsdaten zeigen, dass – wie erwartet – ein erstmaliger Pflegegeldbezug mit fortschreitendem Alter immer wahrscheinlicher wird: Im Durchschnitt erhöht ein zusätzliches Lebensjahr die Wahrscheinlichkeit,

im Zeitraum Jänner – Dezember 2018 erstmalig Pflegegeld zu beziehen, um 0,6 Prozentpunkte.¹⁴⁾ Da es sich hierbei um eine multivariate Regression handelt sind diese Effekte *ceteris paribus* zu interpretieren, d.h. gegeben die Werte aller anderen Variablen im Modell.

Übersicht 2: **Modellmetriken für Logit-Regression in Übersicht 1**

Daten	Trainingsdaten	Testdaten
Beobachtungen	355.486	118.496
Eventzeitraum	[01/2018, 01/2019]	
Beobachtungszeitraum	[01/2017, 01/2018]	
Relative Häufigkeit von $PG_{i,t,t+b} = 1$	0,086	0,085
Falschklassifikationsrate (bei $\tau = 0,5$)	0,086	0,085
Sensitivität (bei $\tau = 0,5$)	0,041	0,042
Spezifität (bei $\tau = 0,5$)	0,996	0,996
Ausfallsrate (bei $\tau = 0,5$)	0,004	0,004
AUC (Fläche unter ROC-Kurve)	0,730	0,730
Ausfallsrate bei Sensitivität von 0,8	0,516	0,518

Q: DSVS, WIFO.

Gegeben das Alter und die anderen im Modell enthaltenen Variablen ist die Wahrscheinlichkeit eines erstmaligen Pflegegeldbezugs im Eventzeitraum für Frauen um etwa 0,2 Prozentpunkte geringer als für Männer. Aber auch stationäre Aufenthalte (+0,4 Prozentpunkte pro Nacht) als auch die Häufigkeit von Arztbesuchen (+0,2 Prozentpunkte pro Kontakt) gehen mit einer höheren Wahrscheinlichkeit eines erstmaligen Pflegegeldbezugs einher. Überraschenderweise deuten die Koeffizienten für einige NUTS-3-Regionen zudem auf eine – verglichen mit der Basiskategorie Wien (NUTS-3 Code AT130) – signifikant geringere Wahrscheinlichkeit hin, erstmalig Pflegegeld zu beziehen. Gegeben das Alter, Geschlecht, die Kontakthäufigkeit zu Allgemeinmedizinerinnen und -mediziner sowie die Dauer stationärer Aufenthalte im Jahr 2017 weisen die Bewohner und Bewohnerinnen vieler Regionen eine signifikant geringere Wahrscheinlichkeit auf, im Jahr 2018 erstmalig Pflegegeld zu beziehen.

Würde man zur Klassifikation einen Schwellenwert von 0,5 heranziehen, sodass Beobachtungen ab einer prognostizierten Wahrscheinlichkeit von 50% ein "positiv" prognostizierter Wert (d.h. ein erstmaliger Pflegegeldbezug) zugeteilt wird (vgl. Abschnitt 2.1), ergibt sich in den Testdaten eine Falschklassifikationsrate von 8,5% (siehe Übersicht 2). Dies erscheint zwar gering, entspricht aber exakt der relativen Häufigkeit, mit der ein erstmaliger Pflegegeldbezug in den Testdaten beobachtet wird. Bei einem Schwellenwert von $\tau = 0,5$ entspricht die Prognosegüte demnach jener eines "naiven" Klassifikators, der für alle Beobachtungen den Wert $\widehat{PG}_{i,t,t+b} = 0$ prognostizieren würde. Zwar wäre die Spezifität (Richtig-Negativ-Rate) bei

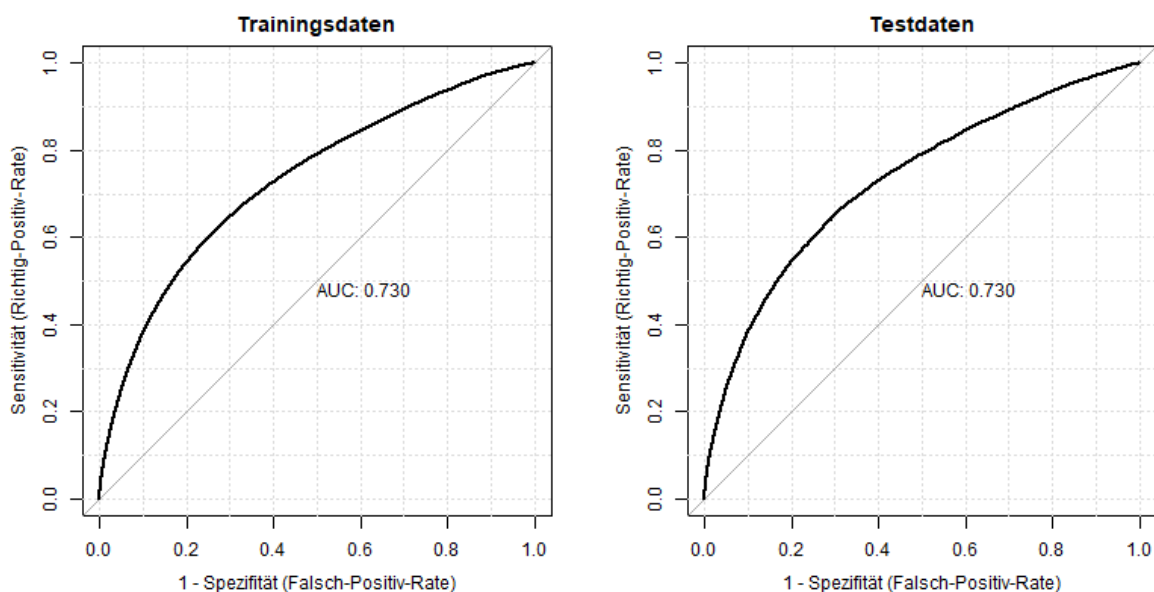
¹⁴⁾ Hier ist zu beachten, dass die Logit-Koeffizienten in Übersicht 1 nicht als Effekte auf die Wahrscheinlichkeit interpretiert werden können. Die hier genannten Veränderungen der Wahrscheinlichkeit basieren auf Berechnungen des durchschnittlichen marginalen Effekts, also der durchschnittlichen Veränderung der Wahrscheinlichkeit eines erstmaligen Pflegegeldbezugs wenn sich das Alter für alle Personen um ein Jahr erhöhen würde.

diesem Schwellenwert 99,6%, allerdings bei einer sehr geringen Sensitivität (Richtig-Positiv-Rate) von 4,2%.

Berücksichtigt man die gesamte Bandbreite an möglichen Schwellenwerte, weist das Modell mit einer Fläche unter der ROC-Kurve von $AUC = 0,73$ jedoch ein durchaus akzeptables Unterscheidungsvermögen zwischen erstmaligem Pflegegeldbezug und keinem erstmaligen Pflegegeldbezug auf (Hosmer – Lemeshow, 2000, S. 162). Jedoch würde z.B. bei einer Sensitivität von 0,8 die Ausfallsrate knapp über 50% betragen (siehe Abbildung 6): Setzt man den Schwellenwert so, dass 80% der tatsächlichen erstmaligen Pflegegeldbezieherinnen und -bezieher korrekt klassifiziert werden (bei $\tau = 0,064$), würde für 51,8% jener Beobachtungen, für welche kein erstmaliger Pflegegeldbezug registriert wird, das Modell fälschlicherweise einen erstmaligen Pflegegeldbezug prognostizieren. Berücksichtigt man, dass im Jahr 2018 für nur rund 8,6% der Personen im Datenbestand ein erstmaliger Pflegegeldbezug registriert wird, ergibt sich daraus eine doch erhebliche Fehlerquelle.

Overfitting scheint hingegen in diesem Modell noch keine Rolle zu spielen: Die Metriken des Modells unterscheiden sich zwischen Trainings- und Testdaten nur geringfügig. Jedoch kann erwartet werden, dass mit steigender Anzahl an erklärenden Variablen die Überanpassung an die Trainingsdaten zunimmt, sodass eine Regularisierung notwendig wird.

Abbildung 6: ROC-Kurven für Logit-Regression in Übersicht 1



Q: DSV, WIFO.

Übersicht 3: **Logit-Schätzung für erstmaligen Pflegegeldbezug im Zeitraum [01/2018, 01/2019) auf Basis der Trainingsdaten (75% des Gesamtdatenbestands)**

Abhängige Variable Erstmaliger Pflegegeldbezug			Abhängige Variable Erstmaliger Pflegegeldbezug		
Variable	Koeffizient	Std	Variable	Koeffizient	Std
Dauer stationärer Aufenthalte		(0,001)	Kontakte FG 17	-0,055***	(0,004)
MEL Kap. 12	0,024***	(0,006)	Kontakte FG 18	0,010	(0,025)
MEL Kap. NA	-0,064***	(0,023)	Kontakte FG 19	0,081***	(0,006)
MEL Kap. 15	0,016	(0,012)	Kontakte FG 20	0,043***	(0,008)
MEL Kap. 7	-0,035**	(0,017)	Kontakte FG 21	-0,071	(0,095)
MEL Kap. 10	-0,115***	(0,019)	Kontakte FG 24	-1,121	(0,996)
MEL Kap. 5	-0,035*	(0,020)	Kontakte FG 27	-0,047***	(0,006)
MEL Kap. 1	-0,024	(0,021)	Kontakte FG 29	-0,098	(0,079)
MEL Kap. 9	-0,099***	(0,023)	Kontakte FG 40	0,016	(0,018)
MEL Kap. 3	0,032	(0,032)	Kontakte FG 42	0,003*	(0,001)
MEL Kap. 18	0,061***	(0,009)	Kontakte FG 43	0,079	(0,236)
MEL Kap. 6	-0,354***	(0,127)	Kontakte FG 50	-0,009**	(0,004)
MEL Kap. 8	-0,123***	(0,022)	Kontakte FG 52	-0,185**	(0,087)
MEL Kap. 21	0,038***	(0,004)	Kontakte FG 53	-0,168***	(0,019)
MEL Kap. 16	-0,072***	(0,013)	Kontakte FG 55	-0,045	(0,050)
MEL Kap. 14	0,024	(0,027)	Kontakte FG 61	0,074***	(0,004)
MEL Kap. 2	0,047***	(0,012)	Kontakte FG 62	-0,047	(0,104)
MEL Kap. 4	-0,124**	(0,056)	Kontakte FG 63	0,008**	(0,004)
MEL Kap. 17	0,079**	(0,033)	Kontakte FG 64	0,017	(0,026)
MEL Kap. 20	-0,090**	(0,039)	Kontakte FG 65	0,015***	(0,001)
MEL Kap. 22	0,864***	(0,308)	Kontakte FG 66	0,022	(0,013)
MEL Kap. 11	-0,925***	(0,298)	Kontakte FG 67	-0,011	(0,042)
MEL Kap. 13	-0,009	(0,016)	Kontakte FG 68	-0,039	(0,027)
ICD Kap. 1	0,279*	(0,144)	Kontakte FG 69	0,081***	(0,013)
ICD Kap. 9	-0,043	(0,069)	Kontakte FG 72	0,033	(0,027)
ICD Kap. a	0,112*	(0,059)	Kontakte FG 73	0,484***	(0,087)
ICD Kap. b	-0,025	(0,067)	Kontakte FG 75	0,021	(0,014)
ICD Kap. c	0,725***	(0,033)	Kontakte FG 80	-0,007	(0,006)
ICD Kap. d	0,179***	(0,034)	Kontakte FG 81	0,124***	(0,027)
ICD Kap. e	0,077***	(0,027)	Kontakte FG 84	0,100***	(0,011)
ICD Kap. f	0,559***	(0,036)	Kontakte FG 85	-0,007***	(0,002)
ICD Kap. g	0,118***	(0,033)	Kontakte FG 86	-0,002	(0,007)
ICD Kap. h	-0,030	(0,034)	Kontakte FG 87	0,064***	(0,016)
ICD Kap. i	0,044*	(0,025)	Kontakte FG 91	-0,008***	(0,002)
ICD Kap. j	0,284***	(0,032)	Kontakte FG 92	0,012**	(0,005)
ICD Kap. k	-0,102***	(0,029)	Kontakte FG 99	0,020***	(0,004)
ICD Kap. l	0,295***	(0,056)	Weiblich (= 1)	0,0005	(0,015)
ICD Kap. m	-0,278***	(0,027)	NUTS3: AT111	0,067	(0,076)
ICD Kap. n	0,118***	(0,030)	NUTS3: AT112	-0,089*	(0,046)
ICD Kap. q	-0,065	(0,106)	NUTS3: AT113	-0,036	(0,050)
ICD Kap. r	0,120***	(0,030)	NUTS3: AT121	-0,172***	(0,041)
ICD Kap. s	0,399***	(0,071)	NUTS3: AT122	-0,046	(0,038)
ICD Kap. t	0,038	(0,051)	NUTS3: AT123	-0,110**	(0,050)
ICD Kap. u	0,211	(0,876)	NUTS3: AT124	-0,146***	(0,039)
ICD Kap. z	-0,027	(0,035)	NUTS3: AT125	-0,171***	(0,049)
ATC Gr. A	0,012***	(0,001)	NUTS3: AT126	0,041	(0,036)
ATC Gr. B	0,013***	(0,002)	NUTS3: AT127	0,057	(0,036)
ATC Gr. C	-0,0004	(0,001)	NUTS3: AT211	0,059*	(0,035)
ATC Gr. D	-0,001	(0,004)	NUTS3: AT212	0,021	(0,052)
ATC Gr. G	-0,001	(0,002)	NUTS3: AT213	-0,039	(0,045)
ATC Gr. H	0,009**	(0,004)	NUTS3: AT221	-0,069**	(0,032)
ATC Gr. J	0,001	(0,004)	NUTS3: AT222	-0,403***	(0,069)

ATC Gr. L	0,031***	(0,003)	NUTS3: AT223	-0,136***	(0,042)
ATC Gr. M	0,016***	(0,002)	NUTS3: AT224	-0,190***	(0,038)
ATC Gr. N	0,026***	(0,001)	NUTS3: AT225	-0,278***	(0,043)
ATC Gr. P	-0,079*	(0,040)	NUTS3: AT226	-0,360***	(0,057)
ATC Gr. R	0,023***	(0,001)	NUTS3: AT311	-0,251***	(0,042)
ATC Gr. S	0,0001	(0,002)	NUTS3: AT312	-0,249***	(0,033)
ATC Gr. V	0,041***	(0,014)	NUTS3: AT313	-0,283***	(0,049)
Kontakte FG 01	0,015***	(0,001)	NUTS3: AT314	-0,315***	(0,058)
Kontakte FG 02	0,129	(0,379)	NUTS3: AT315	-0,176***	(0,045)
Kontakte FG 03	-0,023***	(0,004)	NUTS3: AT321	-0,094	(0,118)
Kontakte FG 04	-0,015**	(0,007)	NUTS3: AT322	-0,116**	(0,053)
Kontakte FG 05	-0,017***	(0,004)	NUTS3: AT323	-0,037	(0,038)
Kontakte FG 06	-0,128***	(0,015)	NUTS3: AT331	-0,260**	(0,119)
Kontakte FG 07	0,002	(0,002)	NUTS3: AT332	-0,031	(0,041)
Kontakte FG 08	-0,215	(0,473)	NUTS3: AT333	-0,115	(0,088)
Kontakte FG 09	-0,055***	(0,007)	NUTS3: AT334	-0,254***	(0,069)
Kontakte FG 10	0,014**	(0,006)	NUTS3: AT335	-0,112**	(0,047)
Kontakte FG 11	0,046***	(0,008)	NUTS3: AT341	-0,084	(0,076)
Kontakte FG 12	-0,002	(0,002)	NUTS3: AT342	-0,085**	(0,042)
Kontakte FG 13	-0,009*	(0,005)	Alter	0,091***	(0,001)
Kontakte FG 14	-0,112***	(0,008)	Konstante	-9,664***	(0,096)
Kontakte FG 15	-0,074***	(0,021)			
Kontakte FG 16	-0,040***	(0,005)			
			Beobachtungen	355.486	
			LogLikelihood	-89.109,000	

Anmerkung: *p<0,1; **p<0,05; ***p<0,01.

Übersicht 4: Modellmetriken für Logit-Regression in Übersicht 3

Daten	Trainingsdaten	Testdaten
Beobachtungen	355.486	118.496
Eventzeitraum	[01/2018, 01/2018)	
Beobachtungszeitraum	[01/2017, 01/2018)	
Relative Häufigkeit von $PG_{i,t,t+b} = 1$	0,086	0,085
Falschklassifikationsrate (bei $\tau = 0,5$)	0,085	0,084
Sensitivität (bei $\tau = 0,5$)	0,082	0,083
Spezifität (bei $\tau = 0,5$)	0,993	0,993
Ausfallsrate (bei $\tau = 0,5$)	0,007	0,007
AUC (Fläche unter ROC-Kurve)	0,778	0,776
Ausfallsrate bei Sensitivität von 0,8	0,412	0,409

Q: DSVS, WIFO.

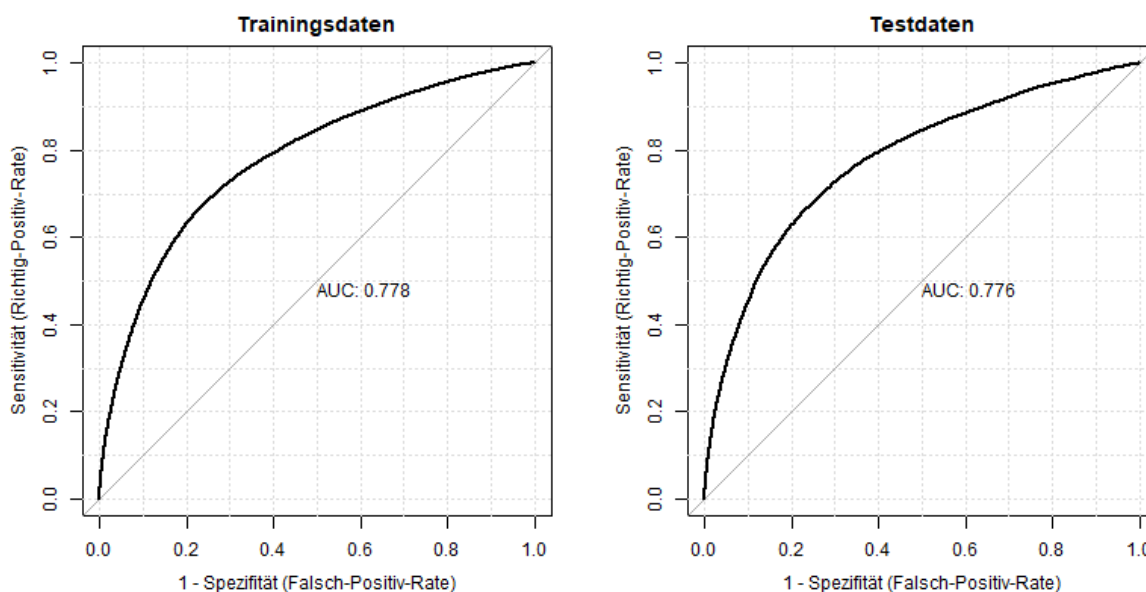
Aus diesem Grund werden, wenn das Modell im nächsten Schritt um weitere Variablen für die Gesundheitsleistungen im Jahr vor dem Eventzeitraum erweitert wird, die detailreichen Merkmale über Heilmittel, Diagnosen und medizinische Einzelleistungen vorerst in aggregierter Form verwendet: Abgaben von Heilmitteln werden auf der höchsten Ebene der Anatomisch-Therapeutisch-Chemischen Klassifikation (Zuordnung zu einer anatomischen Gruppe, 14 Variablen), Diagnosen auf Ebene der ICD-Krankheitskapitel (23 Variablen) und medizinische Einzelleistungen (MEL) auf Ebene der Kapitel im Leistungskatalog des BMASGK (23 Variablen) aggregiert. Zudem wird das Modell um Kontakte zu anderen Fachärzten und Fachärztinnen sowie anderen Vertragspartnern ergänzt (58 Variablen).¹⁵⁾

Die Berücksichtigung der Heilmittel, Diagnosen und medizinischen Einzelleistungen im Jahr 2017 erhöht die Klassifikationsfähigkeit des Modells deutlich: Auch wenn die Falschklassifikationsrate nur marginal abnimmt, verdoppelt sich die Sensitivität beim Schwellenwert $\tau = 0,5$ gegenüber dem vorigen Modell, in dem (neben den persönlichen Eigenschaften) nur die Tage in stationärem Aufenthalt sowie die Kontakte zu Ärztinnen und Ärzte für Allgemeinmedizin berücksichtigt wurden. Die Ausfallsrate bei einer Sensitivität von 0,8 verringert sich um ca. 10 Prozentpunkte auf 40,9% in den Testdaten (siehe auch Abbildung 7). Auch die Fläche unter der AUC-Kurve verbessert sich und liegt sowohl in den Trainings- als auch in den Testdaten bei ca. 0,78. Dies könnte erneut darauf hinweisen, dass Overfitting bei diesem aggregierten Datenbestand noch kein gravierendes Problem darstellt.

Erneut haben das Alter, die Anzahl der Kontakte zu Ärztinnen und Ärzte für Allgemeinmedizin sowie die Anzahl der Nächte in stationärem Aufenthalt signifikant positive Effekte auf die Wahrscheinlichkeit, erstmalig Pflegegeld zu beziehen. Weiterhin sind auch viele der regionalen Dummyvariablen für den Wohnort statistisch signifikant, was auf einen Unterschied relativ zur Basiskategorie (Wien) hinweist. Nach Kontrolle für die im Beobachtungszeitraum konsumierten Gesundheitsleistungen hat das Geschlecht hingegen keinen signifikanten Effekt mehr. Auch viele der Diagnosen und medizinischen Einzelleistungen sowie Kontakte zu anderen Fachärztinnen und -ärzten stehen signifikant mit einem erhöhten Risiko für erstmaligen Pflegegeldbezug in Verbindung.

¹⁵⁾ In der Schätzung können Variablen nicht berücksichtigt werden, welche keine Variation innerhalb der Gruppe der Personen, die im Eventzeitraum erstmalig Pflegegeld beziehen, und/oder innerhalb der Gruppe der Personen, die im Eventzeitraum nicht erstmalig Pflegegeld beziehen, aufweisen. Die Gesamtzahl der Variablen im Modell (147) ist damit etwas geringer als die Gesamtzahl der Variablen insgesamt (156).

Abbildung 7: ROC-Kurven für Logit-Regression in Übersicht 3



Q: DSV, WIFO.

Um die relative Bedeutung der Variablen einzuordnen, können sogenannte Beta-Koeffizienten herangezogen werden. Diese werden berechnet, indem der geschätzte Koeffizient einer Variable mit der Stichprobenstandardabweichung dieser Variable (als Schätzer für die Standardabweichung in der Grundgesamtheit) multipliziert wird.¹⁶⁾ Beta-Koeffizienten berücksichtigen damit, welche Streuung die Variablen annehmen, was Rückschlüsse darauf zulässt, welche erklärenden Variablen die abhängige Variable besonders stark beeinflussen: Bei gegebener Größe des geschätzten Koeffizienten beeinflusst eine Variable mit hoher Streuung in Summe die abhängige Variable mehr als eine Variable mit nur geringer Streuung.

Der größte Beta-Koeffizient kann für das Alter (0,57) beobachtet werden, das somit die stärkste Korrelation mit dem erstmaligen Pflegegeldbezug aufweist. Danach folgen die Abgaben von Heilmitteln der ATC-Gruppe "N" (Nervensystem, Beta-Koeffizient: 0,21) sowie die Anzahl der Kontakte zu Ärztinnen und Ärzten für Allgemeinmedizin und die Dauer stationärer Aufenthalte (je 0,17). Auch ob im Beobachtungszeitraum eine Diagnose aus einer ICD-Gruppe beginnend mit "C" (bösartige Neubildungen) gestellt wurde, hat eine deutliche Korrelation mit den erstmaligen Pflegegeldbezug (Beta-Koeffizient: 0,12). Ebenfalls relativ stark, aber negativ, korreliert der erstmalige Pflegegeldbezug mit der Anzahl an Kontakten zu

¹⁶⁾ Da in einer Logit-Regression die Varianz der zugrundeliegenden latenten abhängigen Variable unbekannt ist, handelt es sich dabei nicht um voll standardisierte Koeffizienten. Beta-Koeffizienten sind daher – wie Koeffizienten im Logit-Modell allgemein – nicht direkt interpretierbar, können jedoch in ihrem Verhältnis relativ zueinander interpretiert werden.

Fachärztinnen und -ärzten für Zahn-, Mund- und Kieferheilkunde (Fachgebiets-Code 17) sowie zu Fachärztinnen und -ärzten für Radiologie (Beta-Koeffizienten -0,12 bzw. -0,11).

Übersicht 5: **Fläche unter der ROC-Kurve (AUC) auf Basis der Testdaten für Modelle mit unterschiedlichen Betrachtungs- und Eventzeiträumen (t = 01/2018)**

Beobachtungszeiträume	$[t - a, t) / [t, t + b)$	Eventzeiträume			
		[01/2018, 02/2018)	[01/2018, 04/2018)	[01/2018, 07/2018)	[01/2018, 01/2019)
[10/2017, 01/2018) [07/2017, 01/2018) [01/2017, 01/2018) [01/2016, 01/2018)	$[t - a, t) / [t, t + b)$	0,883	0,830	0,799	0,772
		0,877	0,829	0,800	0,775
		0,862	0,822	0,797	0,776
		0,847	0,812	0,791	0,776
[10/2017, 12/2017) [07/2017, 12/2017) [01/2017, 12/2017) [01/2016, 12/2017)	$[t - a, t) / [t, t + b)$	0,818	0,787	0,773	0,754
		0,817	0,791	0,777	0,760
		0,806	0,787	0,777	0,763
		0,797	0,782	0,774	0,765

Q: DVS, WIFO.

Um zu zeigen, wie sich eine Veränderung des Event- und/oder Beobachtungszeitraums auswirkt, wird das Modell im nächsten Schritt für Beobachtungszeiträume von 3, 6, 12 und 24 Monaten Länge sowie für Eventzeiträume von 1, 3, 6 und 12 Monaten Länge geschätzt, wobei wie vorhin Jänner 2018 als Beginn des Eventzeitraums herangezogen wird.

Übersicht 5 zeigt die AUC aus den Testdaten für alle Kombinationen dieser Beobachtungs- und Eventzeiträume. Der obere Teil der Übersicht zeigt, dass die Schätzung des erstmaligen Pflegegeldbezugs vor allem in sehr kurzfristigen Zeiträumen eine gute Prognosekraft aufweist. Verwendet man die Gesundheitsleistungen der vorhergegangenen drei Monate, um den erstmaligen Pflegegeldbezug im nächsten Monat zu prognostizieren, ergibt sich eine Fläche unter der ROC-Kurve (AUC) von 0,883, was einem exzellenten Unterscheidungsvermögen entspricht (Hosmer – Lemeshow, 2000). Zudem zeigt Abbildung 8, dass das Modell bei einer Sensitivität von 80% eine Spezifität von über 80% erreicht.¹⁷⁾

Eine Verlängerung des Betrachtungszeitraums auf ein halbes oder ganzes Jahr bzw. zwei Jahre führt hingegen zu keiner Verbesserung der AUC. Dies spricht dafür, dass die Ereignisse

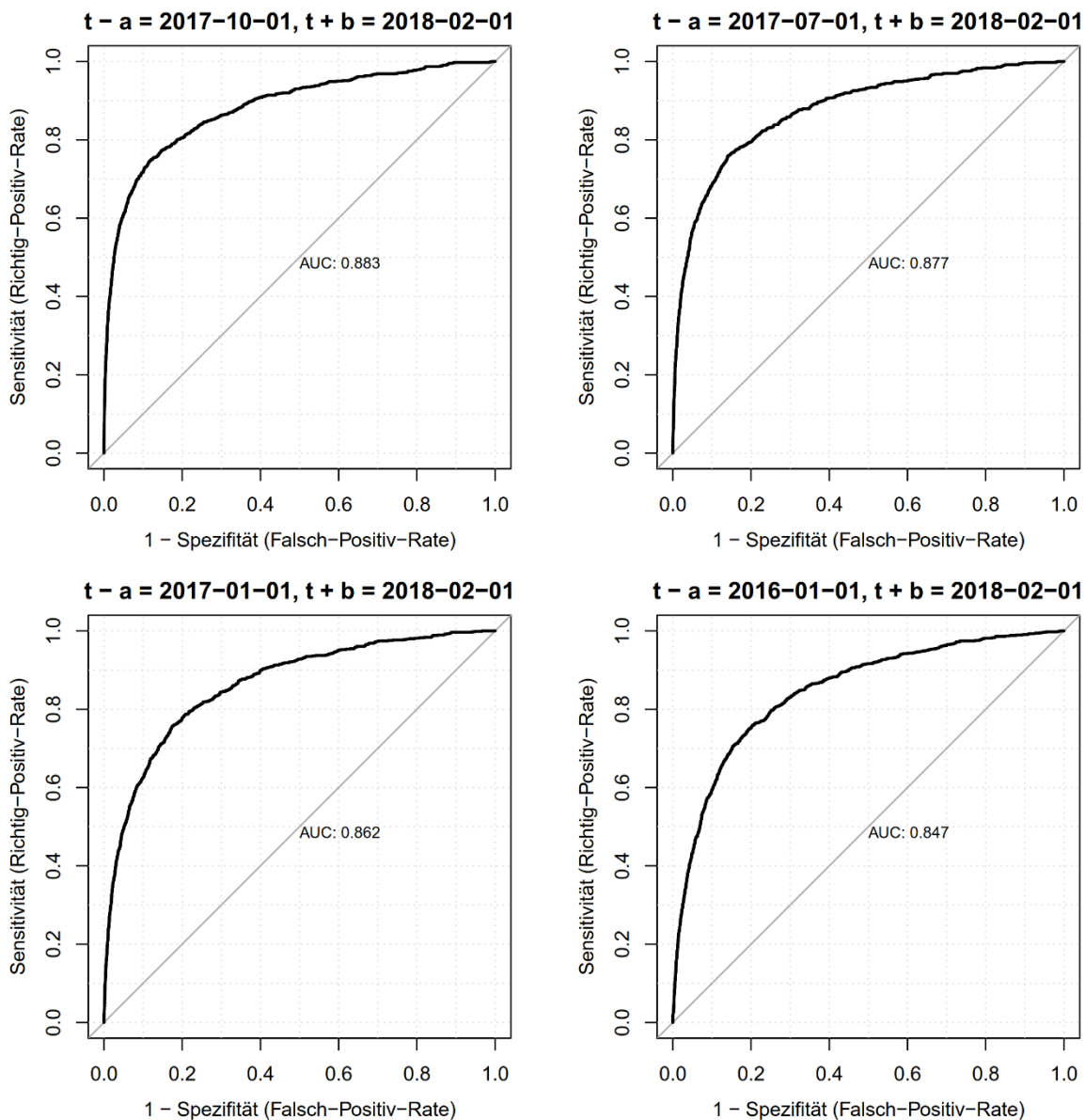
¹⁷⁾ Die ROC-Kurven der restlichen Modelle im oberen Teil von Übersicht 5 sind in Appendix 7.1 zu finden.

der unmittelbaren Vergangenheit für die Klassifizierung des erstmaligen Pflegegeldbezugs die höchste Bedeutung haben. Dehnt man hingegen den Beobachtungszeitraum aus, fließen weiter zurückliegende Informationen in das Modell ein, welche für den erstmaligen Pflegegeldbezug von geringerer Bedeutung sind. Dies führt insgesamt zu einer Reduktion der Modellgüte. Verlängert sich der Eventzeitraum, nimmt *ceteris paribus* die Prognosegüte (gemessen durch die AUC) ab. Die besten Ergebnisse werden im Allgemeinen für jene Modelle erzielt, in denen die Länge des Eventzeitraums dem Beobachtungszeitraum entspricht. Die AUC-Werte dieser Modelle sind in Übersicht 5 grau hinterlegt.

Da nicht ausgeschlossen werden kann, dass diese Resultate durch Ereignisse am unmittelbaren Ende des Beobachtungszeitraums getrieben werden, zeigt der untere Teil der Übersicht die Fläche unter der ROC-Kurve für Modelle, in denen der Monat unmittelbar vor Beginn des Eventzeitraums (Dezember 2017) im Beobachtungszeitraum ausgespart wird. Der Beobachtungszeitraum endet in diesen Modellen daher bereits im November 2017, während der Eventzeitraum weiterhin im Jänner 2018 beginnt. Diese Einschränkung reduziert die Fläche unter der ROC-Kurve vor allem in der kurzfristigen Prognose, jedoch bleibt die AUC mit 0,818 weiterhin relativ hoch, sodass dem Modell ein exzellentes Unterscheidungsvermögen attestiert werden kann, obwohl – wie aus Abbildung 9 ersichtlich wird – die Spezifität bei einer Sensitivität von 80% auf ca. 65% zurückgeht (bzw. die Falsch-Positiv-Rate auf 35 % ansteigt).¹⁸⁾

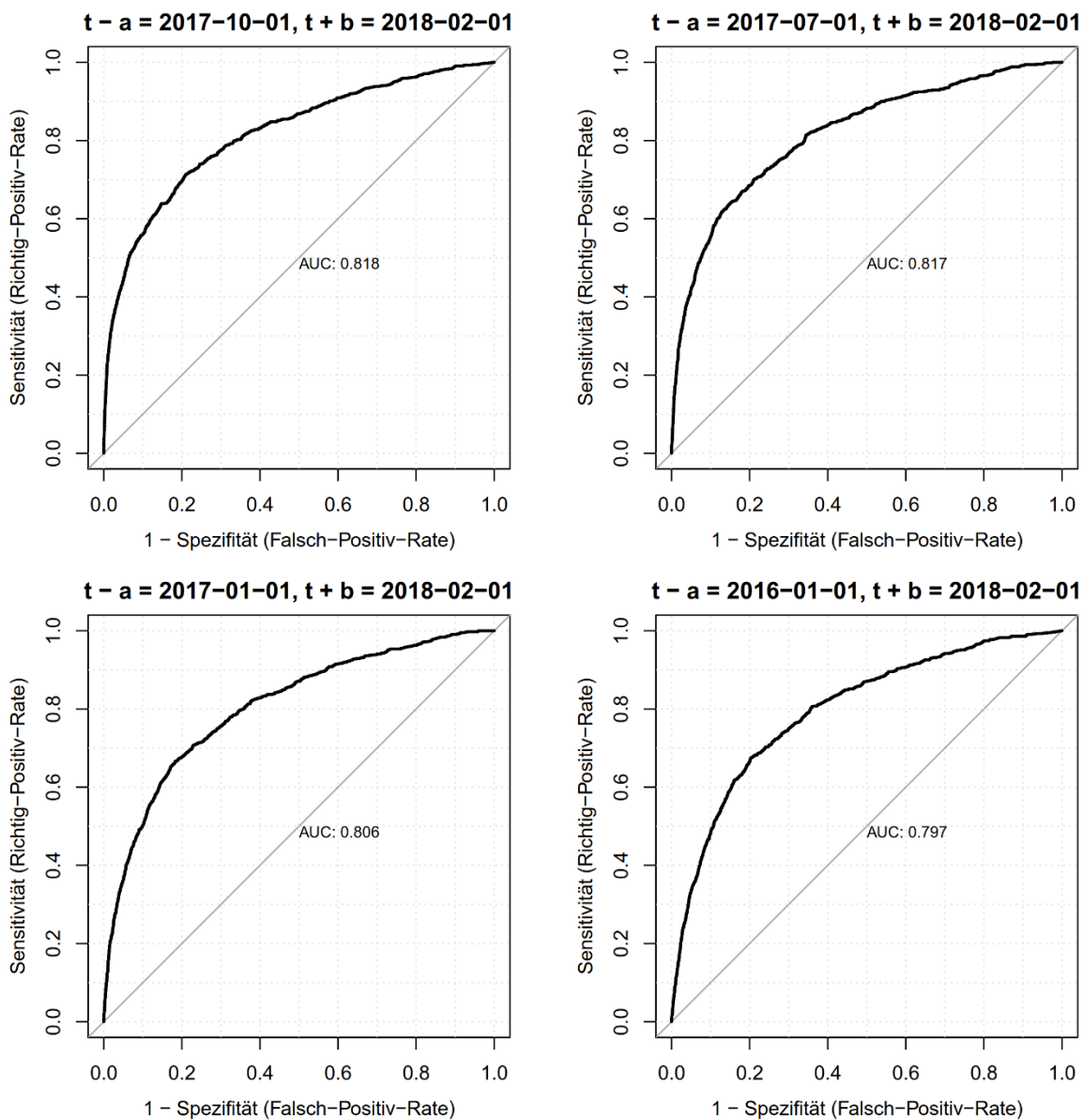
¹⁸⁾ Die ROC-Kurven für die restlichen Modelle im unteren Teil von Übersicht 5 finden sich in Appendix 7.2.

Abbildung 8: ROC-Kurven auf Basis der Testdaten für Modelle mit unterschiedlichem Beobachtungszeitraum, Eventzeitraum [01/2018, 02/2018)



Q: DSV, WIFO.

Abbildung 9: ROC-Kurven auf Basis der Testdaten für Modelle mit unterschiedlichem Beobachtungszeitraum, Eventzeitraum [01/2018, 02/2018), Beobachtungszeitraum ohne Dezember 2017



Q: DSV, WIFO.

2.4 Schätzungen auf Basis detaillierter Gesundheitsvariablen

Wie die Ergebnisse des vorigen Abschnitts zeigen, lässt sich der erstmalige Pflegegeldbezug am besten prognostizieren, wenn der Eventzeitraum relativ kurz ist und der Beobachtungszeitraum in etwa der Länge des Eventzeitraums entspricht. Aus diesem Grund werden in diesem Abschnitt Daten für den Beobachtungszeitraum [10/2017, 01/2018) und den Eventzeitraum [01/2018, 02/2018) verwendet, jedoch greifen die Modelle auf deutlich detailliertere Daten zurück: Neben den bereits im letzten Abschnitt verwendeten Variablen Alter, Geschlecht, Wohnort auf NUTS-3-Ebene (exkl. Wien, welches als Basiskategorie dient), Dauer des stationären Aufenthalts sowie Häufigkeit von Kontakten zu Allgemeinmedizinerinnen und -medizinern, Fachärztinnen und -ärzten und anderen Vertragspartnerinnen und -partnern werden die abgegebenen Heilmittel auf ATC-Level 4 (fünfstellig, 483 Variablen), Diagnosen auf Ebene der ICD-Kategorien (dreistellig, 1.483 Variablen) sowie medizinische Einzelleistungen in einer fünfstelligen Systematik (1.553 Variablen) verwendet. In Summe stehen damit 3.614 potenzielle Variablen zur Verfügung.

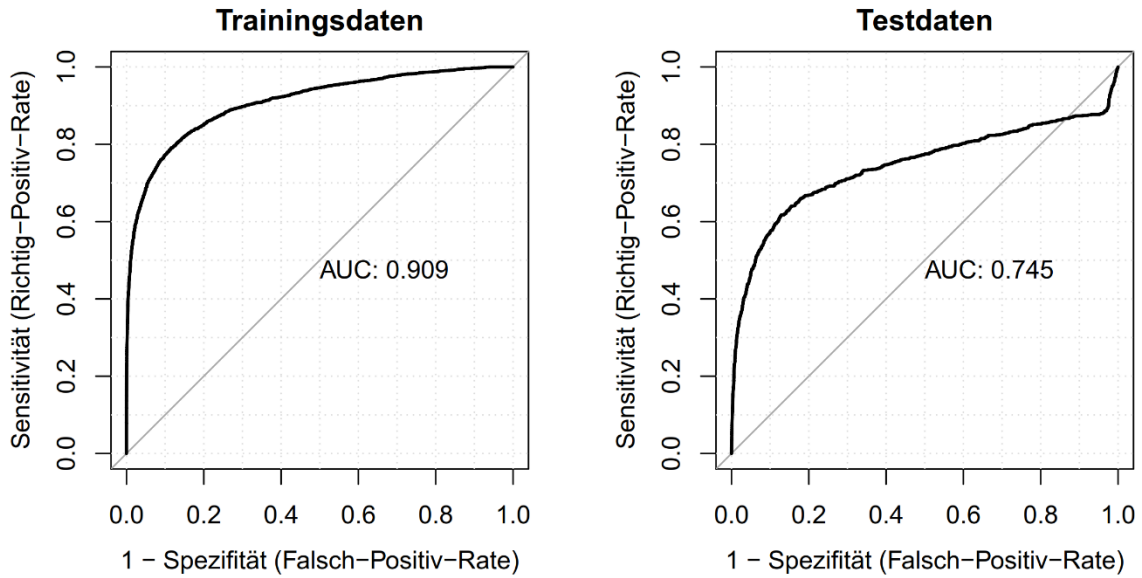
Die ersten beiden Spalten in Übersicht 6 zeigen Modellmetriken für eine Logit-Regression auf Basis dieses Datenbestands.¹⁹⁾ Verglichen mit den Ergebnissen für den gleichen Beobachtungs- und Eventzeitraum in Übersicht 4 weist das Logit-Modell zwar in den Trainingsdaten eine höhere Fläche unter der ROC-Kurve auf (AUC: 0,909), allerdings sinkt diese in den Testdaten deutlich (AUC: 0,745) und liegt sogar unter dem Wert für das weniger detailreiche Modell auf Basis der aggregierten Gesundheitsleistungen (AUC: 0,883, siehe Übersicht 5). Dies ist vermutlich durch ein deutliches "Overfitting" zu erklären (vgl. Abschnitt 2.1): Das detaillierte Modell kann zwar die Eigenheiten der Trainingsdaten gut erklären und weist in diesen ein herausragendes Unterscheidungsvermögen auf. Konfrontiert man das Modell jedoch mit neuen Daten, auf welche es nicht trainiert wurde, sinkt das Klassifizierungspotential deutlich.

Dies zeigt sich auch im Verlauf der ROC-Kurve des Logit-Modells in den Testdaten, wo sich die Zunahme der Sensitivität bei steigender Falsch-Positiv-Rate deutlich früher abflacht als in den Trainingsdaten. Ab einer Ausfallsrate von ca. 86% fällt die ROC-Kurve sogar unter die 45°-Linie. Genau genommen bedeutet dies, dass in diesem Bereich das Gegenteil des vom Modell prognostizierten Ereignisses eine höhere Auftretenswahrscheinlichkeit hat als das prognostizierte Ereignis, da die Sensitivität in diesem Bereich kaum noch zunimmt, obwohl die Ausfallsrate gegen Eins (die Spezifität gegen Null) geht.

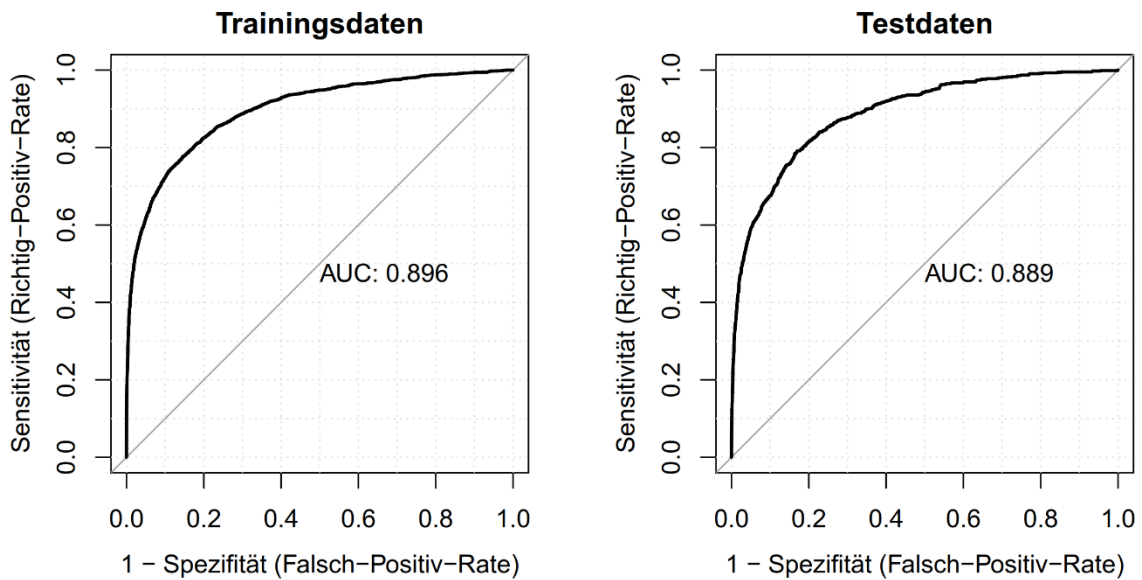
¹⁹⁾ Auf eine vollständige tabellarische Darstellung der Ergebnisse wird aufgrund des Modellumfangs verzichtet. Dass von insgesamt 3.614 Variablen nur 2.651 in der Logit-Regression enthalten sind liegt daran, dass viele ATC-, ICD- oder MEL-Codes entweder im Beobachtungszeitraum nicht oder nur so selten vorkamen, dass diese perfekt kollinear mit anderen Variablen waren oder keine Variation innerhalb der Gruppen der im Eventzeitraum (nicht) erstmalig Pflegegeldbeziehenden aufwiesen. Die Effekte solcher Variablen können in diesem empirischen Modell nicht geschätzt werden.

Abbildung 10: ROC-Kurven für Modelle in Übersicht 6

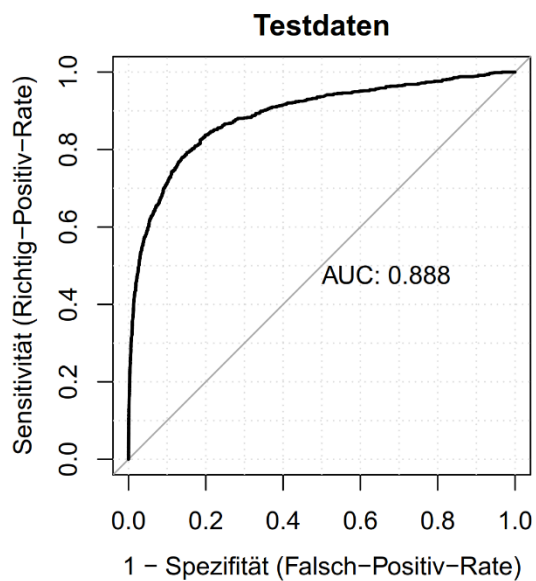
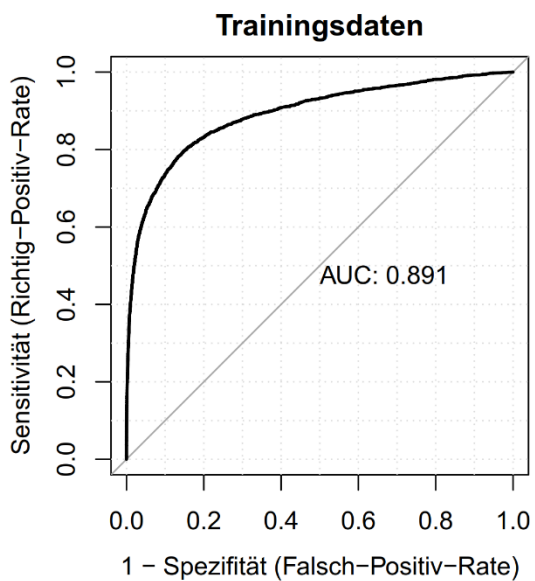
Logit:



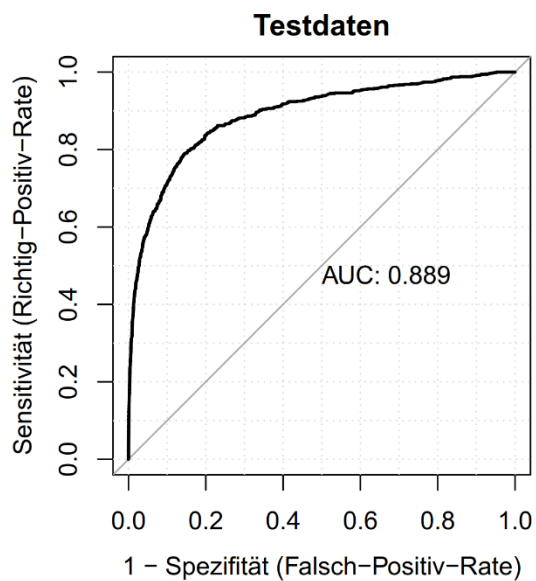
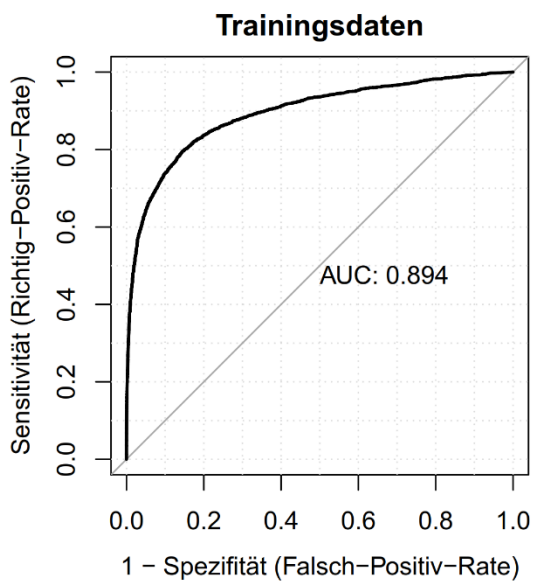
LASSO:



Ridge Regression:



Elastic Net:



Q: DSV, WIFO.

Übersicht 6: **Modellmetriken für Logit-, LASSO-, Ridge und Elastic Net Regressionen**

Modell	Logit		LASSO		Ridge Regression		Elastic Net	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Daten	Training	Test	Training	Test	Training	Test	Training	Test
Beobachtungen	349.223	116.126	349.223	116.126	349.223	116.126	349.223	116.126
Eventzeitraum	[01/2018, 02/2018)		[01/2018, 02/2018)		[01/2018, 02/2018)		[01/2018, 02/2018)	
Beobachtungszeitraum	[10/2017, 01/2018)		[10/2017, 01/2018)		[10/2017, 01/2018)		[10/2017, 01/2018)	
Anzahl aktiver Variablen	2.651		461		2.651		513	
Relative Häufigkeit von $PG_{i,t,t+b} = 1$	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074
Falschklassifikationsrate (bei $\tau = 0,5$)	0,0064	0,0090	0,0071	0,0075	0,0069	0,0074	0,0071	0,0075
Sensitivität (bei $\tau = 0,5$)	0,229	0,117	0,106	0,102	0,091	0,068	0,111	0,104
Spezifität (bei $\tau = 0,5$)	0,999	0,997	0,999	0,999	0,999	0,999	0,999	0,999
Ausfallsrate (bei $\tau = 0,5$)	0,001	0,003	0,001	0,001	0,001	0,001	0,001	0,001
AUC (Fläche unter ROC-Kurve)	0,909	0,745	0,891	0,888	0,896	0,889	0,894	0,889
Ausfallsrate bei Sensitivität von 0,8	0,128	0,594	0,154	0,164	0,173	0,185	0,151	0,161

Q: DSVV, WIFO.

Aus diesem Grund zeigt Übersicht 6 auch die Modellmetriken von Regularisierungsmethoden (siehe Abschnitt 2.1). Der optimale Wert des Strafterms λ wird mittels 5-facher Kreuzvalidierung aus den Trainingsdaten gewählt. Die Spalten 3 und 4 zeigen die Ergebnisse einer LASSO-Schätzung. Die LASSO-Regularisierung schrumpft die Koeffizienten für alle bis auf 461 Variablen auf Null. Verglichen mit dem Logit-Modell reduziert dies zwar die Fläche unter der ROC-Kurve in den Trainingsdaten, dafür erhöht sich die AUC in den Testdaten auf 0,888. Marginal höher ist die Fläche unter der ROC-Kurve aus einer Ridge Regression, wenngleich die ROC-Kurve einen etwas anderen Verlauf nimmt (siehe Abbildung 10). So beträgt die Ausfallsrate bei einer Sensitivität von 80% im LASSO-Modell nur 16,4%, in der Ridge Regression ist diese (trotz höherer AUC) hingegen mit 18,5% leicht schlechter.

Die letzten beiden Spalten in Übersicht 6 zeigen die Modellkennzahlen eines "Elastic Net"-Modells. Das "elastische Netz" verbindet LASSO und Ridge Regression, indem die Pönalterme beider Modelle mit einem Gewicht von je 50% in das Modell einfließen (siehe Appendix 7.3). Dadurch erhöht sich die Fläche unter der ROC-Kurve nochmals inkrementell: Mit einer AUC von 0,889 ist das Elastic-Net-Modell das Modell mit dem besten Unterscheidungsvermögen, wenngleich die Ergebnisse der drei Regularisierungsmethoden sehr ähnlich sind. Auch im Elastic-Net-Verfahren wird der Großteil Koeffizienten auf exakt Null geschrumpft, sodass lediglich 513 erklärende Variablen aktiv im Modell verbleiben. Vergleicht man dies mit den Ergebnissen für den gleichen Beobachtungs- und Eventzeitraum aus Übersicht 5, so stellt dieses deutlich komplexere Modell mit höherem Detailgrad, jedoch nur eine marginale Verbesserung der Prognosegüte (gemessen an der AUC) dar. Der höhere Detailgrad führt damit zu keiner

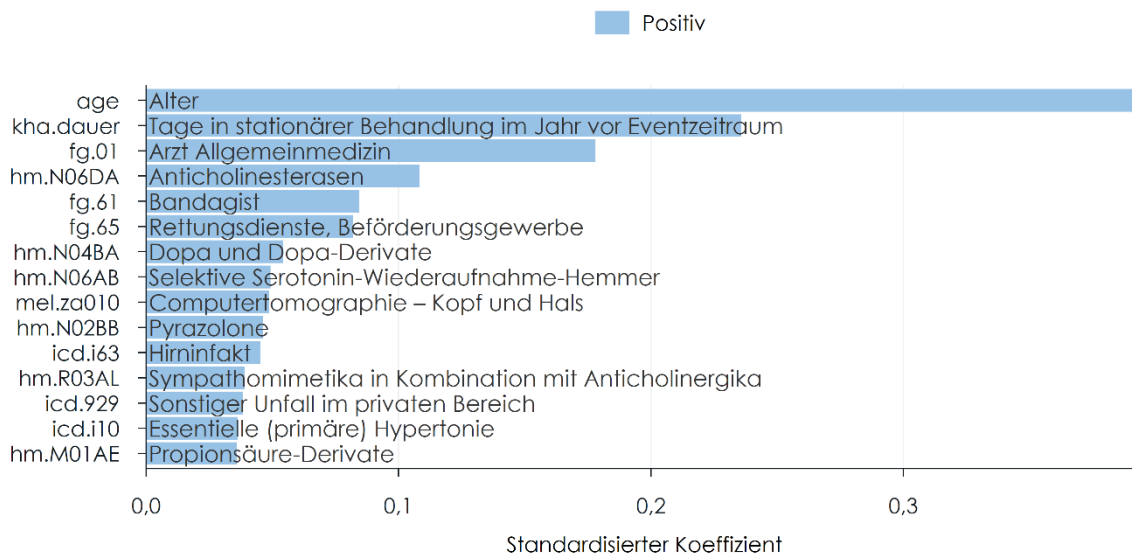
substanziellen Verbesserung gegenüber einem relativ einfachen Logit-Modell auf Basis aggregierter Gesundheitsleistungen.²⁰⁾

Um jene Variablen und Leistungsbezüge zu identifizieren, die in der kurzen Frist am stärksten mit dem erstmaligen Pflegegeldbezug korrelieren, werden aus dem Elastic-Net-Modell erneut standardisierte Beta-Koeffizienten berechnet. Wie im Modell auf Basis aggregierter Gesundheitsvariablen haben das Alter (Beta-Koeffizient: 0,39), die Anzahl der Nächte in stationärer Behandlung (0,24) sowie die Anzahl der Kontakte zu Allgemeinmedizinerinnen und -medizinern (0,18) die höchste Korrelation mit dem erstmaligen Pflegegeldbezug in der kurzen Frist (siehe Abbildung 11). Unter den zehn Variablen mit den größten Beta-Koeffizienten finden sich wiederum Abgaben von Heilmitteln der ATC-Gruppe "N" (Nervensystem), konkret die ATC-Codes N06DA (Cholinesterasehemmer, N06D: Antidementiva), N04BA (Dopa und Dopa-Derivate, N04B: Dopaminerge Mittel, N04: Antiparkinsonmittel), N06AB (Selektive Serotonin-Wiederaufnahmehemmer, N06A: Antidepressiva) sowie N02BB (Pyrazolone, N02B: Andere Analgetika und Antipyretika, N02: Analgetika). Weiters zählen Kontakte zu Bandagistinnen und Bandagisten (Fachgebiets-Code 61) und Rettungsdiensten (Fachgebiets-Code 65) sowie ob im Beobachtungszeitraum eine Computertomographie durchgeführt wurde (MEL-Code ZA010, siehe *BMSGPK*, 2020) zu den zehn Faktoren, welche die größte Korrelation mit dem erstmaligen Pflegegeldbezug in der kurzen Frist aufweisen.

Betrachtet man die einzelnen Variablenkategorien genauer (Abbildungen 12-15), so findet (neben den oben bereits erwähnten Allgemeinmedizinerinnen und -medizinern, Rettungsdiensten sowie Bandagistinnen und Bandagisten) auch die Gruppe der Orthopädietechnikerinnen und -techniker (Fachgebiets-Code 69), welche u.a. Orthesen und Prothesen herstellen und anpassen, unter den am stärksten korrelierten Fachgebieten (siehe Abbildung 12). Zu den wichtigsten Diagnosen, welche mit dem erstmaligen Pflegegeldbezug in der kurzen Frist korrelieren, zählen der Hirninfarkt (ICD-10-Code I63), sonstige Unfälle im privaten Bereich (Code 929), essentielle Hypertonie (I10) sowie die Fraktur des Femurs (Oberschenkelknochen, S72) und sekundäre bösartige Neubildung an sonstigen und nicht näher bezeichneten Lokalisationen (C79). Aber auch Demenzdiagnosen (vaskuläre Demenz, F01 und nicht näher bezeichnete Demenz, F03) scheinen eine bedeutende Korrelation mit dem erstmaligen Pflegegeldbezug in der kurzen Frist zu haben.

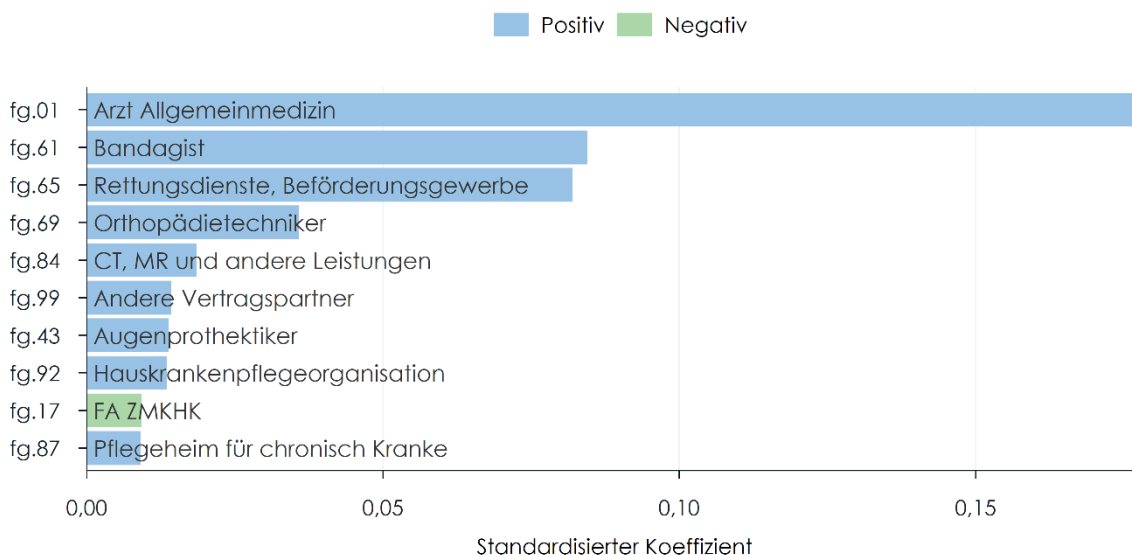
²⁰⁾ Als Alternative zu den in Übersicht 6 gezeigten Modellen wurde auch die Random-Forest-Methode (*Breimann*, 2001) angewandt, in welchen die Klassifizierung auf Basis mehrerer optimierter "Entscheidungsbäume" vorgenommen wird (siehe auch *James et al.*, 2013, S. 319ff oder *Hastie et al.*, 2009, S. 587ff). Jeder Entscheidungsbaum verwendet dabei nur eine zufällig ausgewählte Teilmenge aller verfügbaren Variablen zur Klassifizierung der abhängigen Variable, wodurch die Korrelation zwischen den Klassifikationen unterschiedlicher Bäume reduziert werden soll. Um eine neue Beobachtung zu klassifizieren, wird die von der Mehrheit der Entscheidungsbäume prognostizierte Klasse verwendet. Random Forest ergibt, auf die Testdaten angewandt, eine Fläche unter der ROC-Kurve von 0,899 und übertrifft damit den Punktschätzer für die AUC der Elastic-Net-Schätzung in Übersicht 6. Die AUC des Random-Forest-Modells liegt jedoch innerhalb des 95%-Konfidenzintervalls der Elastic-Net-Schätzung in den Testdaten (0,876 – 0,902).

Abbildung 11: **Beta-Koeffizienten der 15 am stärksten korrelierenden Variablen auf Basis der Elastic-Net-Regression in Übersicht 6**



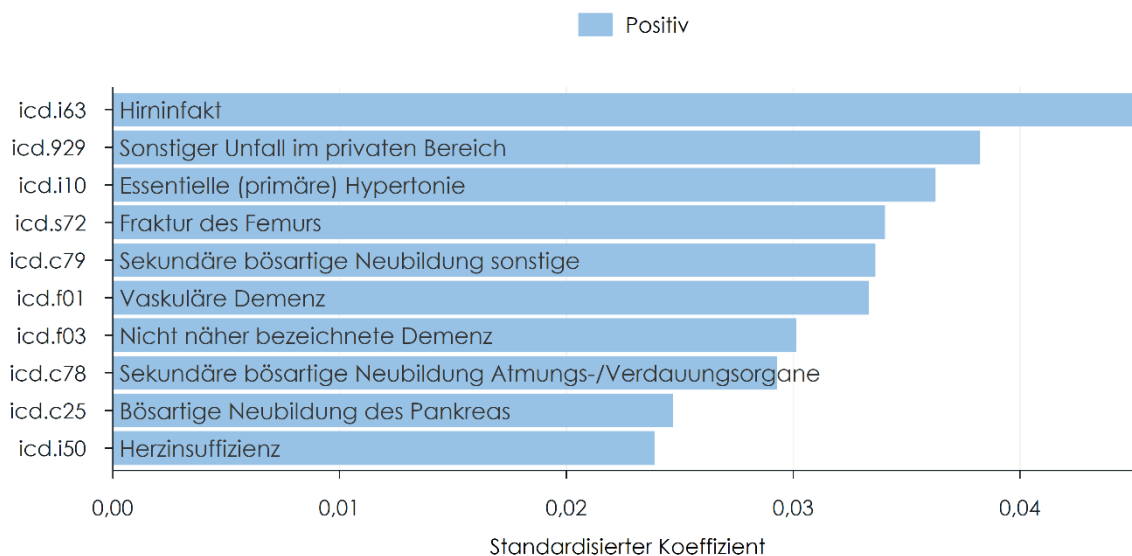
Q: DSVS, WIFO.

Abbildung 12: **Beta-Koeffizienten der 15 am stärksten korrelierenden Fachgebiets-Codes auf Basis der Elastic-Net-Regression in Übersicht 6**



Q: DSVS, WIFO.

Abbildung 13: **Beta-Koeffizienten der 15 am stärksten korrelierenden ICD-Diagnosen auf Basis der Elastic-Net-Regression in Übersicht 6**

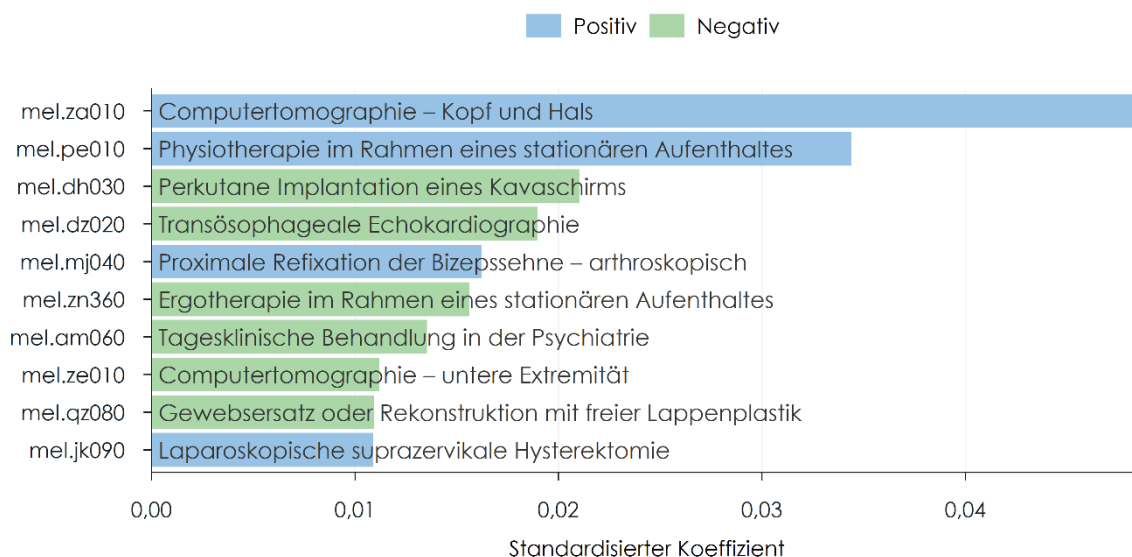


Q: DVS, WIFO.

Unter den medizinischen Einzelleistungen finden sich einige, welche negativ mit dem erstmaligen Pflegegeldbezug in der kurzen Frist korrelieren und somit mit einer geringeren Wahrscheinlichkeit eines erstmaligen Pflegegeldbezugs einhergehen (siehe Abbildung 14). An dieser Stelle muss jedoch nochmals darauf hingewiesen werden, dass es sich hierbei um keine Kausalzusammenhänge, sondern lediglich um Korrelationen handelt. Diese Eingriffe werden möglicherweise an Patienten vorgenommen, welche ceteris paribus einen stabilen Gesundheitsverlauf haben, sodass die kurzfristige Wahrscheinlichkeit eines erstmaligen Pflegegeldbezugs negativ mit diesen Leistungen korreliert.

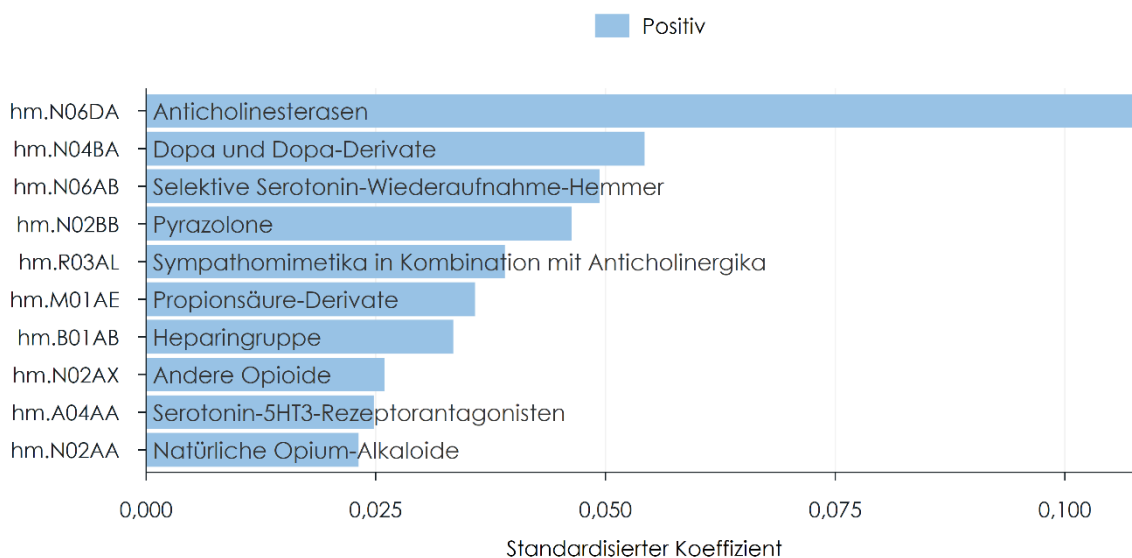
Unter den abgegebenen Heilmitteln finden sich (neben den oben bereits erwähnten Mitteln der ATC-Gruppe "N") zudem noch Sympathomimetika in Kombination mit Anticholinergika inkl. Dreifachkombinationen mit Corticosteroiden (ATC-Code: R03AL; R03: Mittel bei obstruktiven Atemwegserkrankungen), Propionsäure-Derivate (M01AE; M01: Antiphlogistika und Antirheumatika), Medikamente des ATC-Codes B01AB (Heparin-Gruppe; B01: Antithrombotische Mittel) sowie Serotonin-5HT₃-Rezeptorantagonisten (A04AA; A04: Antiemetika und Mittel gegen Übelkeit) unter den Variablen mit den größten Beta-Koeffizienten (siehe Abbildung 15).

Abbildung 14: **Beta-Koeffizienten der 15 am stärksten korrelierenden medizinischen Einzelleistungen auf Basis der Elastic-Net-Regression in Übersicht 6**



Q: DSVS, WIFO.

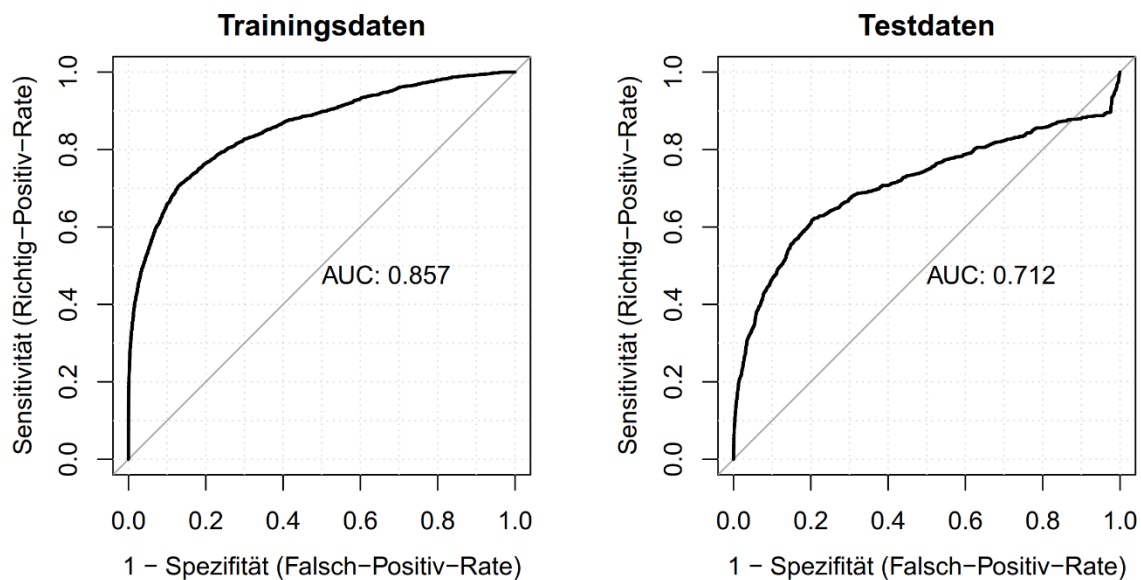
Abbildung 15: **Beta-Koeffizienten der 15 am stärksten korrelierenden ATC-Codes auf Basis der Elastic-Net-Regression in Übersicht 6**



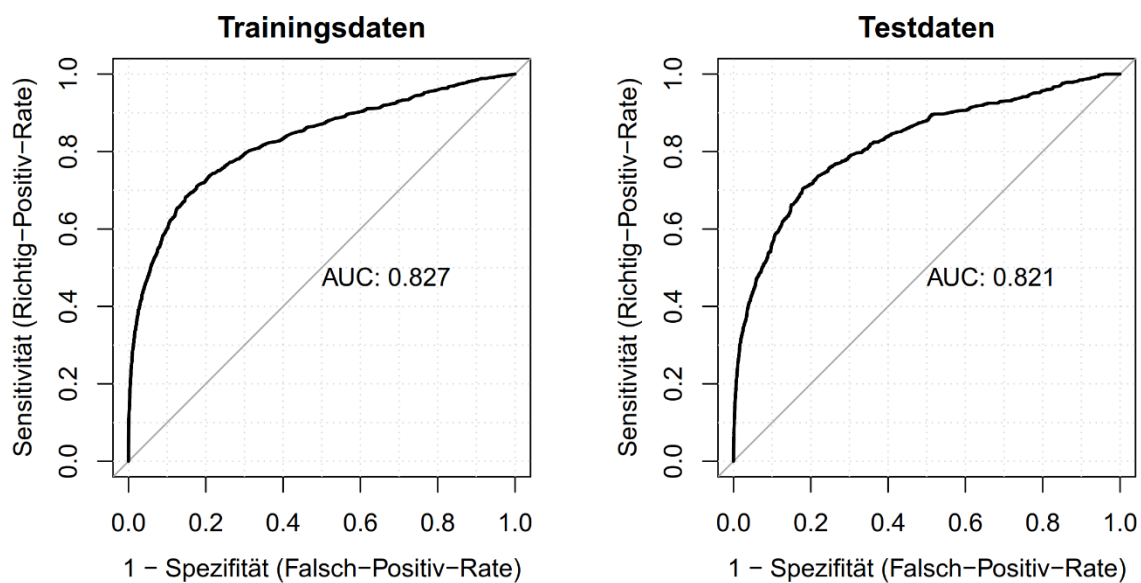
Q: DSVS, WIFO.

Abbildung 16: ROC-Kurven für Modelle in Übersicht 7, Beobachtungszeitraum [10/2017, 12/2017)

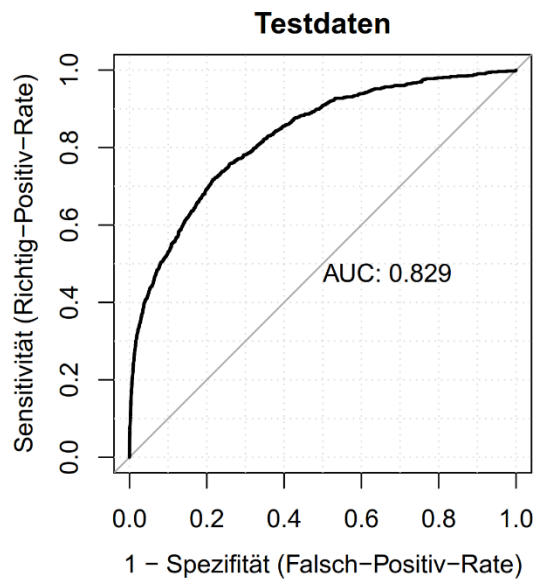
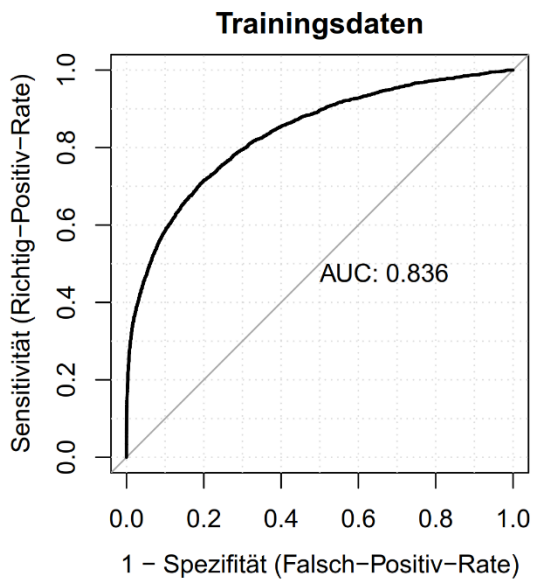
Logit:



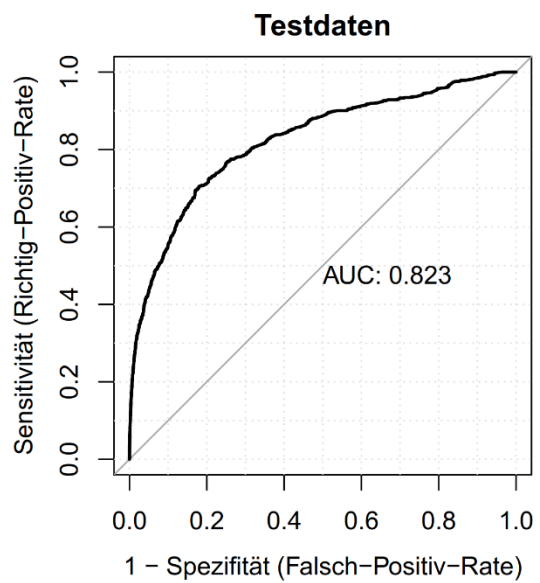
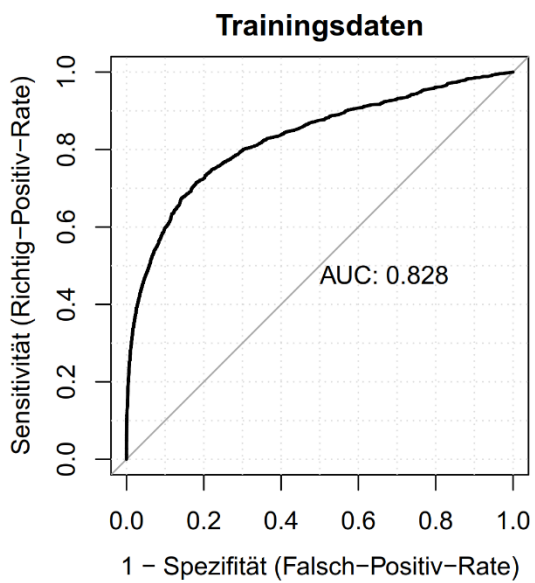
LASSO:



Ridge Regression:



Elastic Net:



Q: DSV, WIFO.

Übersicht 7: **Modellmetriken für Logit-, LASSO-, Ridge und Elastic Net Regressionen**

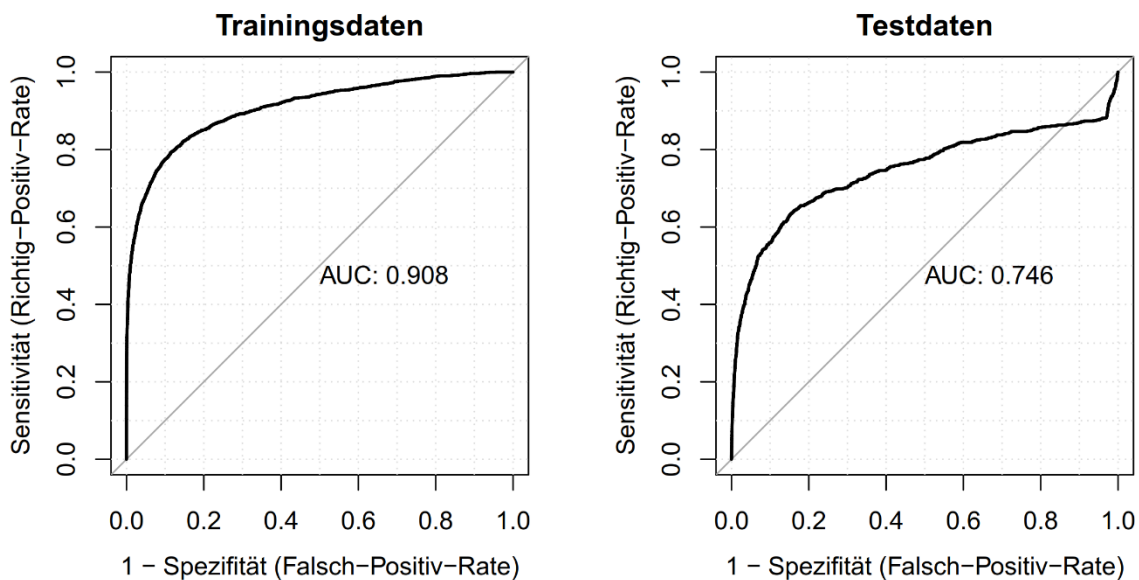
Modell	Logit		LASSO		Ridge Regression		Elastic Net	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Daten	Training	Test	Training	Test	Training	Test	Training	Test
Beobachtungen	349.223	116.126	349.223	116.126	349.223	116.126	349.223	116.126
Eventzeitraum	[01/2018, 02/2018]		[01/2018, 02/2018]		[01/2018, 02/2018]		[01/2018, 02/2018]	
Beobachtungszeitraum	[10/2017, 12/2017]		[10/2017, 12/2017]		[10/2017, 12/2017]		[10/2017, 12/2017]	
Anzahl aktiver Variablen	2.518		335		2.518		337	
Relative Häufigkeit von $PG_{i,t,t+b} = 1$	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074	0,0074
Falschklassifikationsrate (bei $\tau = 0,5$)	0,0066	0,0087	0,0072	0,0076	0,0071	0,0074	0,0072	0,0076
Sensitivität (bei $\tau = 0,5$)	0,151	0,077	0,057	0,037	0,057	0,039	0,060	0,037
Spezifität (bei $\tau = 0,5$)	0,999	0,998	0,999	0,999	0,999	0,999	0,999	0,999
Ausfallsrate (bei $\tau = 0,5$)	0,001	0,002	0,001	0,999	0,001	0,001	0,001	0,001
AUC (Fläche unter ROC-Kurve)	0,857	0,712	0,827	0,821	0,836	0,829	0,828	0,823
Ausfallsrate bei Sensitivität von 0,8	0,254	0,623	0,311	0,334	0,310	0,327	0,303	0,313

Q: DVSV, WIFO.

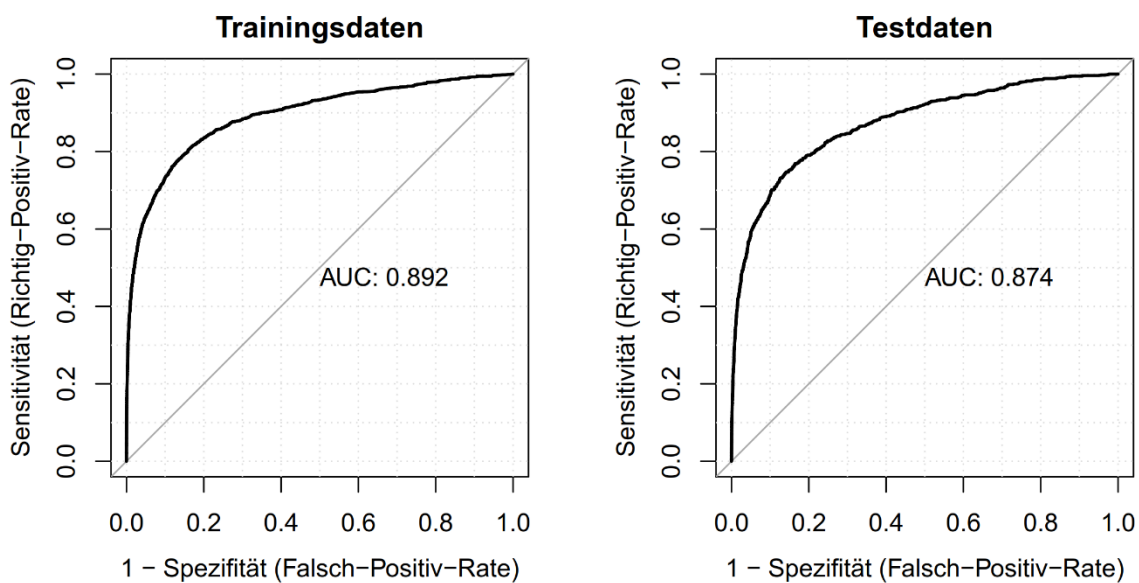
Wie im vorigen Abschnitt wird das Modell auch für Daten geschätzt, in denen der Monat vor Beginn des Eventzeitraums (Dezember 2017) im Beobachtungszeitraum ausgespart wird, so dass der Beobachtungszeitraum bereits im November 2017 endet. Dies dient dazu, abzuschätzen, wie robust die Methode ist, wenn Ereignisse am unmittelbaren Ende des Beobachtungszeitraums unberücksichtigt bleiben. Wie erwartet kommt es dadurch zu einem deutlichen Rückgang der Fläche unter der ROC-Kurve. Das beste Ergebnis kann hier mittels Ridge Regression erzielt werden (AUC: 0,829 in den Testdaten), wobei die Unterschiede zwischen den drei Regularisierungsmethoden vernachlässigbar sind. Deutlich verändert sich jedoch der Verlauf der ROC-Kurven (siehe Abbildung 16) im Vergleich zum Modell, welches auch den Monat unmittelbar vor dem Eventzeitraum berücksichtigt: Betrachtet man z.B. wieder die Ausfallsrate bei einer Sensitivität von 0,8, so kann in den Testdaten keines der Modelle eine Ausfallsrate von unter 30% erzielen. Berücksichtigt man also im Beobachtungszeitraum auch den Monat vor dem Eventzeitraum, lässt sich die Ausfallsrate bei gegebener Sensitivität von 80% beinahe halbieren. Dies spricht erneut dafür, dass der erstmalige Pflegegeldbezug von kurzfristigen Entwicklungen in der unmittelbaren Vergangenheit substantiell beeinflusst wird.

Abbildung 17: ROC-Kurven für Modelle in Übersicht 8, Eventzeitraum: [11/2018, 12/2018), Beobachtungszeitraum [08/2018, 11/2018)

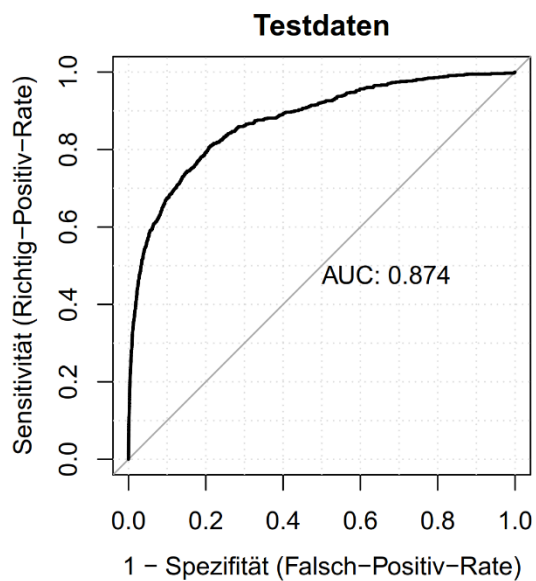
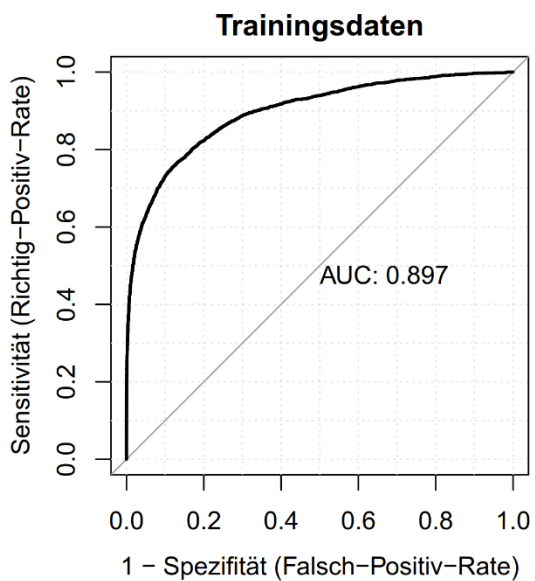
Logit:



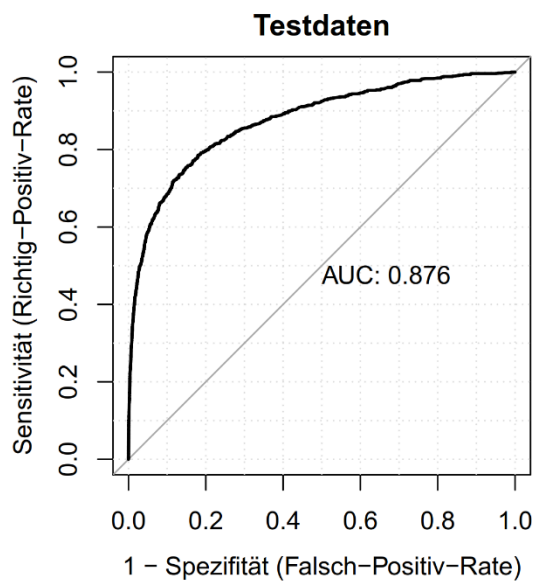
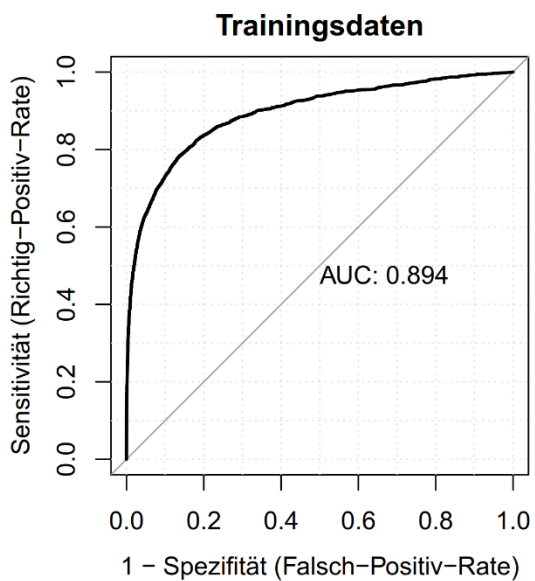
LASSO:



Ridge Regression:



Elastic Net:



Q: DSV, WIFO.

Übersicht 8: **Modellmetriken für Logit-, LASSO-, Ridge und Elastic Net Regressionen**

Modell	Logit		LASSO		Ridge Regression		Elastic Net	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Daten	Training	Test	Training	Test	Training	Test	Training	Test
Beobachtungen	328.236	109.274	328.236	109.274	328.236	109.274	328.236	109.274
Eventzeitraum	[11/2018, 12/2018)		[11/2018, 12/2018)		[11/2018, 12/2018)		[11/2018, 12/2018)	
Beobachtungszeitraum	[08/2018, 12/2018)		[08/2018, 12/2018)		[08/2018, 12/2018)		[08/2018, 12/2018)	
Anzahl aktiver Variablen	2.515		392		2.515		449	
Relative Häufigkeit von $PG_{i,t,t+b} = 1$	0,0069	0,0072	0,0069	0,0072	0,0069	0,0072	0,0069	0,0072
Falschklassifikationsrate (bei $\tau = 0,5$)	0,0057	0,0085	0,0066	0,0073	0,0062	0,0074	0,0066	0,0073
Sensitivität (bei $\tau = 0,5$)	0,254	0,105	0,115	0,078	0,142	0,077	0,118	0,078
Spezifität (bei $\tau = 0,5$)	0,999	0,997	0,999	0,999	0,999	0,999	0,999	0,999
Ausfallsrate (bei $\tau = 0,5$)	0,001	0,003	0,001	0,001	0,001	0,001	0,001	0,001
AUC (Fläche unter ROC-Kurve)	0,908	0,746	0,892	0,874	0,897	0,874	0,894	0,876
Ausfallsrate bei Sensitivität von 0,8	0,125	0,549	0,159	0,216	0,169	0,206	0,158	0,207

Q: DVSV, WIFO.

Ein weiterer Faktor, welcher die Prognosegüte beeinflussen kann, ist der Umstand, dass die Kontrollgruppe jener, die innerhalb des Zeitraums 2016 – 2018 nie Pflegegeld bezogen, so gewählt wurde, dass kaum Todesfälle in dieser Gruppe in diesem Zeitraum auftreten, während die Gruppe jener, die innerhalb des Zeitraums 2016 – 2018 (erstmalig) Pflegegeld bezogen, vollständig abgebildet wird. Dies birgt das Potential, dass Personen in der Kontrollgruppe über den Zeitraum 2016 – 2018 über diese Selektion einen besseren Gesundheitsverlauf hatten (vgl. Abschnitt 2.2). Um zu testen, ob die Prognosefähigkeit robust gegenüber diesem Selektionseffekt ist, verwenden die Modelle, deren Metriken in Übersicht 8 zusammengefasst sind, den Monat November 2018 als Eventzeitraum.²¹⁾ Da der Gesundheitsverlauf nach Dezember 2018 für die Datenauswahl der Kontrollgruppe irrelevant ist, sollte der Selektionseffekt in diesen Daten keine (oder nur eine geringfügige) Rolle mehr spielen. Als Beobachtungszeitraum werden erneut die drei Monate vor dem Eventzeitraum (August – Oktober 2018) herangezogen.

Die Ergebnisse ähneln sehr stark jenen von Übersicht 6, die geringfügigen Abweichungen sind vermutlich eher auf Zufallsschwankungen als auf Selektionseffekte zurückzuführen. Erneut weist die Elastic-Net-Regression die beste Prognosegüte gemessen an der AUC der Testdaten auf (0,876), während es in der Logit-Schätzung zu deutlichem Overfitting kommt. Die Regressionen zeigen zudem, dass die Verschiebung des Eventzeitraums zu keiner wesentlichen Änderung in der Aussagekraft der Modelle führt.²²⁾

²¹⁾ Der Monat November wurde anstelle des Monats Dezember gewählt, da der Anteil der erstmaligen Pflegegeldbezieherinnen und -bezieher in den Daten für Dezember 2018 ungewöhnlich gering ist, siehe Abschnitt 2.2.

²²⁾ Zusätzliche Schätzungen, welche Jänner 2017 als Eventzeitraum und Oktober – Dezember 2016 als Beobachtungszeitraum heranziehen, zeigen ebenfalls sehr ähnliche Werte für die Fläche unter der ROC-Kurve von 0,889 (LASSO) bis 0,891 (Ridge Regression), sodass weder die Selektion in der Kontrollgruppe noch die Auswahl des Eventzeitraums einen gravierenden Effekt auf die Modellgüte haben.

2.5 Fazit und Einschätzung des Potentials zur Entwicklung eines kurzfristigen Vorlaufindikators

Die Verbindung der abgerufenen Gesundheitsleistungen mit dem Pflegegeldbezug für die Altersgruppe 60 – 85 zeigt, dass der erstmalige Pflegegeldbezug relativ gut aus den Daten prognostiziert werden kann. Die geschätzten Modelle weisen ein durchwegs exzellentes Unterscheidungsvermögen auf und erzielen selbst in Daten, auf welche sie nicht trainiert wurden, Flächen unter der ROC-Kurve nahe an 0,9. Dabei zeigt sich, dass vor allem rezente Gesundheitsleistungen für den erstmaligen Pflegegeldbezug innerhalb eines Eventzeitraums von großer Bedeutung sind: Zieht man den Monat unmittelbar vor dem Eventzeitraum nicht in die Betrachtung mit ein, zeigt sich eine deutliche Verschlechterung der Modellgüte. In dieselbe Richtung weist auch der Umstand, dass Modelle, welche längere Eventzeiträume und/oder längere Beobachtungszeiträume nutzen, eine geringere Prognosefähigkeit aufweisen. Dennoch muss festgehalten werden, dass die Modelle selbst dann, wenn der unmittelbare Monat vor Beginn des Eventzeitraums ausgespart wird und die Beobachtungs- und Eventzeiträume auf ein ganzes Jahr verlängert werden, noch akzeptabel zwischen Personen mit und ohne erstmaligen Pflegegeldbezug im Eventzeitraum unterscheiden können. Ein höherer Detailgrad in der Messung der Gesundheitsleistungen führt hingegen nicht zwangsweise zu einer Verbesserung der Modellgüte: Der Zuwachs an Prognosefähigkeit durch eine tiefere Aufschlüsselung von abgegebenen Heilmitteln, Diagnosen oder medizinischen Einzelleistungen erhöht zwar die Modellgüte, jedoch nur marginal.

Zu den am stärksten mit dem erstmaligen Pflegegeldbezug in der kurzen Frist korrelierten Faktoren zählen laut den hier verwendeten multivariaten Modellen das Alter, stationäre Aufenthalte, sowie Kontakte zu Allgemeinmedizinerinnen und -medizinern, aber auch Abgaben von Heilmitteln der ATC-Gruppe "N" (Nervensystem). Aber auch durchgeführte Computertomographien oder Diagnosen wie "Hirnfarkt", Unfälle im privaten Bereich oder "essentielle Hypertonie" weisen eine deutliche Präzedenz für den erstmaligen Pflegegeldbezug in der kurzen Frist auf.

Betreffend die Möglichkeit, aus den Daten einen kurzfristigen Vorlaufindikator für den zukünftigen Pflegegeldbezug auf individueller Ebene zu entwickeln, muss trotz der hohen Unterscheidungsfähigkeit der Modelle (gemessen an der Fläche unter der ROC-Kurve) bedacht werden, dass die untersuchten Ereignisse in den zugrundeliegenden Daten nicht ausgewogen verteilt sind. Der Anteil der Personen, die im Zeitraum 2016 – 2018 innerhalb eines Monats erstmalig Pflegegeld beziehen, liegt in den zur Verfügung stehenden Daten relativ stabil bei rund 0,75% pro Monat.

Welches Problem sich daraus ergibt, lässt sich anhand eines einfachen Rechenbeispiels verdeutlichen: Setzt man den Schwellenwert τ so, dass eine Sensitivität (Richtig-Positiv-Rate) von 80% erreicht wird, ergibt das beste Modell auf Basis der kurzfristigen Daten eine Spezifität (Richtig-Negativ-Rate) von 83,9% bzw. eine Ausfallrate (Falsch-Positiv-Rate) von 16,1%. Zieht man nun eine hypothetische Gruppe von 100.000 Personen heran, welche vor einem Stichtag t kein Pflegegeld beziehen, dann kann man erwarten, dass im folgenden Monat

$100.000 \times 0,51\% = 510$ Personen erstmalig Pflegegeld beziehen werden. Das Modell würde 80% dieser Personen richtig klassifizieren, und somit 600 tatsächliche Pflegegeldbezieherinnen und -bezieher korrekt prognostizieren. Unter den 99.490 Personen, von denen man keinen erstmaligen Pflegegeldbezug im nächsten Monat erwarten kann, würde das Modell bei der gegebenen Ausfallrate jedoch $99.250 \times 16,1\% = 15.979$ Personen fälschlicherweise ebenfalls der Gruppe der erstmaligen Pflegegeldbezieherinnen und -bezieher zuteilen. Von jenen Personen, denen das Modell insgesamt einen erstmaligen Pflegegeldbezug unterstellt, werden demnach nur $600 / (600 + 15.979) = 3,6\%$ oder knapp jede(r) 28. tatsächlich im Folgemonat erstmalig Pflegegeld beziehen.

Um auf einen Anteil von zumindest 50% an tatsächlichen Pflegegeldbezieherinnen und -bezieher unter allen Beobachtungen mit prognostiziertem Erstbezug zu kommen, müsste die Spezifität bei einer Sensitivität von 80% bei 99,4% liegen bzw. die Ausfallrate bei nur 0,6%. Eine derart geringe Ausfallrate kann von den hier verwendeten Klassifizierungsmodellen bei gegebener Sensitivität nicht erreicht werden.²³⁾ Dieses Problem ist in der zugrundeliegenden Grundgesamtheit aller 60- bis 85-jährigen noch stärker ausgeprägt, da dort der Anteil der erstmaligen Pflegegeldbezieherinnen und -bezieher pro Eventzeitraum (und damit der Anteil jener Personen, welche Falsch-Positiv klassifiziert werden könnten) nochmals geringer ist als in den zur Verfügung stehenden Daten. Eine Anwendung als Vorlaufindikator auf individueller Ebene scheint damit trotz der guten Prognoseleistung auf Basis dieser Daten und Methoden verfrüht. Weitere Forschung (wie z.B. die Evaluierung alternativer Klassifizierungsalgorithmen) wäre noch notwendig, um alternative Modelle und Ansätze zu evaluieren welche möglicherweise die Modellgüte weiter verbessern.

²³⁾ Eine Verlängerung des Eventzeitraums, sodass eine etwas ausgewogene Verteilung jener Personen erreicht wird, welche erstmalig Pflegegeld beziehen, kann dieses Problem nicht vollständig lösen: Zwar ist z.B. im 12-monatigen Eventzeitraum [01/2018, 01/2019) der Anteil jener, die erstmalig Pflegegeld beziehen, mit 8,6% deutlich höher, allerdings weist z.B. ein Elastic-Net-Modell für den Eventzeitraum [01/2018, 01/2019) und den Beobachtungszeitraum [01/2017, 01/2018) auf Basis der detaillierten Gesundheitsvariablen (AUC: 0,791) bei einer Sensitivität von 80% eine Ausfallrate von 38,2% (Spezifität: 61,8%) in den Testdaten auf. Demnach würden erneut nur 16,5% bzw. nur jede(r) Sechste jener, für die das Modell einen erstmaligen Pflegegeldbezug prognostiziert, im Folgejahr tatsächlich erstmalig Pflegegeld beziehen. Um bei einer Sensitivität von 80% eine Parität zwischen Richtig-Positiv- und Falsch-Positiv-Klassifizierten zu erreichen, müsste die Spezifität selbst in dieser Anwendung bei 92,5% liegen.

3. Modul 2: Gesundheitliche Gründe für einen Übergang in eine höhere Pflegegeldstufe

3.1 Ziel des Moduls 2

Modul 2 analysiert die gesundheitlichen Gründe für einen Übergang von pflegegeldbeziehenden Personen in eine höhere Pflegegeldstufe. Es wird eruiert, welche Gesundheitsleistungen laut Gesundheitsleistungsdaten den Wechsel in eine höhere Pflegegeldstufe bestimmen. Ziel ist es, zu überprüfen, ob aus den Ergebnissen relevante Präventionsfelder abgeleitet werden können.

Der Wechsel in eine höhere Pflegegeldstufe ist durch eine Zunahme des Pflegeaufwandes begründet. Ein erhöhter Pflegeaufwand wird in der Regel durch die generelle Verschlechterung des Gesundheitszustands (z. B. durch Multimorbidität) oder durch ein plötzliches gesundheitliches Ereignis (z. B. Stürze oder bösartige Neubildungen) verursacht. Ziel ist es, auf Basis von Gesundheitsleistungsdaten (erstellte Diagnosen, verordnete Medikamente, erbrachte medizinische Einzelleistungen, erbrachte Leistungen nach Fachgebieten) und unter Berücksichtigung von persönlichen Merkmalen die wichtigsten Determinanten für den Übergang in eine höhere Pflegegeldstufe in einem multivariaten Modell zu analysieren.

Die Ergebnisse sind aus gesundheitspolitischer Perspektive wichtig, da sie zeigen, welche gesundheitliche Entwicklung und Charakteristika von älteren Personen zu einer Höhereinstufung im Pflegegeldsystem führen und – im Idealfall – in künftige Präventionskonzepte einfließen können.

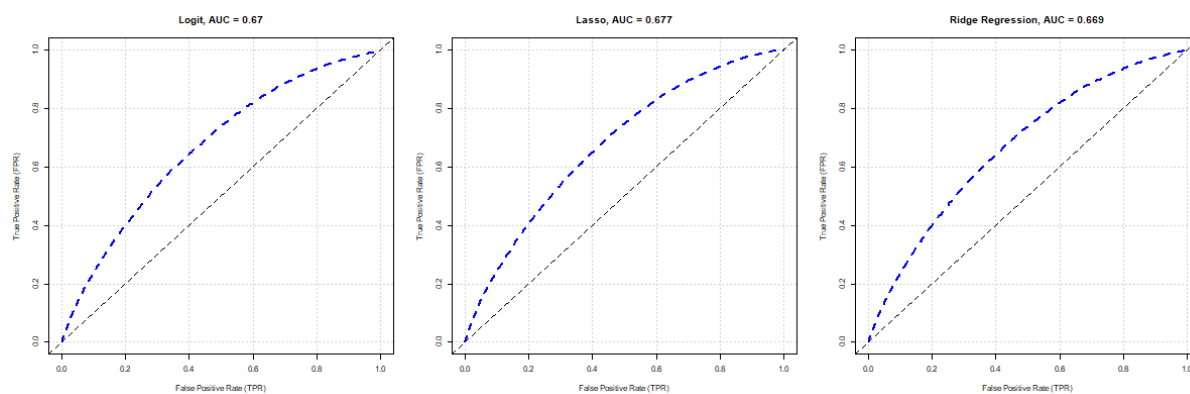
3.2 Methodik des Moduls 2 und Datenbeschreibung

Die Methodik der Analyse des Übergangs in eine höhere Pflegegeldstufe orientiert sich an der methodischen Vorgangsweise in Modul 1. Während jedoch dort (aufgrund der Notwendigkeit einer Kontrollgruppe) nur Personen zwischen 60 und 85 Jahren berücksichtigt werden konnten, bezieht sich Modul 2 (sowie die folgenden Module) ausschließlich auf Personen mit Pflegegeldbezug. In der empirischen Untersuchung werden daher alle Pflegegeldbeziehenden ab 60 Jahren berücksichtigt. Als Eventzeitraum für den Übergang in eine höhere Pflegegeldstufe wird der Zeitraum [01/2017, 01/2018) definiert. Die abhängige Variable ist damit erneut eine binäre Variable, welche den Wert 1 annimmt, wenn für eine Person im Eventzeitraum eine Erhöhung der Pflegegeldstufe registriert wird (sonst 0). Innerhalb des Eventzeitraums wird für 18,2% der Pflegegeldbeziehenden ein Übergang in eine höhere Stufe registriert.

Die Wahrscheinlichkeit, in eine höhere Pflegegeldstufe zu wechseln, wird erneut im Rahmen einer logistischen Regression als Funktion von persönlichen Eigenschaften (wie Alter, Wohnort, Geschlecht) sowie von den im Beobachtungszeitraum [01/2016, 01/2017) in Anspruch genommenen Gesundheitsleistungen (Heilmittel, Arztbesuche, Dauer stationärer Aufenthalte, medizinische Einzelleistungen) und gestellten Diagnosen (im Rahmen stationärer Aufenthalte) modelliert. Als zusätzlicher Faktor wird die Pflegegeldstufe im letzten Monat vor Beginn des Eventzeitraums (Dezember 2016) herangezogen.

Auch wenn die Entwicklung eines Vorlaufindikators in diesem Modul nicht im Zentrum steht, werden die Einflussfaktoren anhand ihres Potentials zur Klassifizierung der abhängigen Variable (Übergang in eine höhere Pflegegeldstufe/kein Übergang in eine höhere Pflegegeldstufe) in Daten, auf denen das Modell nicht trainiert wurde, eingeordnet. Dazu wird erneut auch auf Regularisierungsschätzer zurückgegriffen, um zwischen den 3.429 erklärenden Variablen zu diskriminieren. Die Wahl der optimalen Größe des Strafterms λ bei LASSO oder Ridge Regression erfolgt wiederum anhand der minimalen Deviance und wird mit 5-facher Kreuzvalidierung aus den Trainingsdaten (283.061 bzw. 75% der insgesamt 377.414 Beobachtungen) ermittelt.²⁴⁾ Zur Evaluierung der unterschiedlichen Modelle wird die Fläche unter der ROC-Kurve (AUC) in den Testdaten (94.353 bzw. 25% aller Beobachtungen) herangezogen.

Abbildung 18: **ROC-Kurven von Logit-, LASSO- und Ridge Regression-Modellen für den Übertritt in eine höhere Pflegegeldstufe auf Basis der Testdaten**



Q: WIFO, DVSV.

Abbildung 18 zeigt, dass die ROC-Kurve für alle drei Modelle – Logit, LASSO und Ridge Regression – sehr ähnlich verläuft. Die Fläche unter der Kurve ist dabei im LASSO-Modell am höchsten (AUC: 0,677), weshalb im Folgenden auf die Ergebnisse dieses Modells zurückgegriffen wird. Im Allgemeinen sind die Koeffizienten der drei Modelle relativ ähnlich, insbesondere die Parameter der regularisierten Schätzer weisen eine relativ hohe Korrelation auf ($\rho = 0,939$), aber auch die Korrelation zwischen den Parametern des Logit- und LASSO-Modells ($\rho = 0,607$) als auch des Logit- und Ridge-Regression-Modells ($\rho = 0,734$) weisen darauf hin, dass Variablen, welche im Logit-Modell einen stärkeren Einfluss haben, tendenziell auch in den beiden

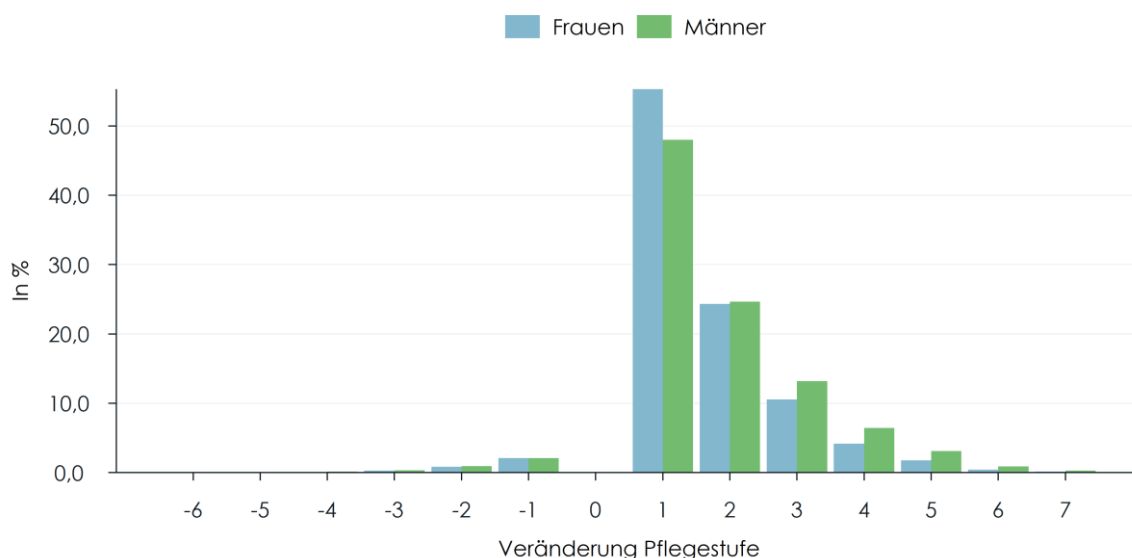
²⁴⁾ 6.112 der 377.414 PflegegeldbezieherInnen befanden sich im Dezember 2017 in der höchsten Pflegegeldstufe 7. Diese Personen können per Definition nicht in eine höhere Pflegegeldstufe übergehen. Aufgrund der geringen Häufigkeit beeinflusst dies die Ergebnisse der Schätzungen jedoch nicht, wie ein Vergleich mit Regressionen zeigt, in denen Personen in Pflegegeldstufe 7 aus dem Datenbestand entfernt wurden.

anderen Modellen einen stärkeren Einfluss auf den Übergang in eine höhere Pflegegeldstufe aufweisen.²⁵⁾

Wie in Modul 1 beschrieben, wurden für die vorliegende Studie Individualdaten von 550.960 Personen, die 60 Jahre oder älter waren und im Zeitraum 2016 bis 2018 Pflegegeld bezogen, analysiert. Davon bezogen 388.480 bereits im Jänner 2016 (erstmalig oder wiederholt) Pflegegeld und 162.480 Personen bezogen zwischen Februar 2016 und Dezember 2018 erstmalig Pflegegeld.

Der Neueintritt in das Pflegegeldsystem erfolgt in den meisten Fällen in Pflegegeldstufe 1. Die Hälfte der Eintritte erfolgte in Pflegegeldstufe 1, ein Fünftel in Pflegegeldstufe 2, und ein Achtel der Neueintretenden fanden sich in Pflegegeldstufe 3 wieder und ein Zwölftel in Pflegegeldstufe 4. Der Eintritt in eine höhere Pflegegeldstufe als 4 war selten der Fall. Abbildung 21 zeigt neben den Neueintritten in das Pflegegeldsystem auch die Übertritte in andere Pflegegeldstufen. Dabei wird ersichtlich, dass, erstens, die meisten Übertritte in höheren Pflegegeldstufen anlangen und zweitens, dass die häufigsten Übertritte jene in die nächsthöhere Pflegegeldstufe sind.

Abbildung 19: **Veränderung der Pflegegeldstufen nach Geschlecht**



Q: WIFO, DVSV.

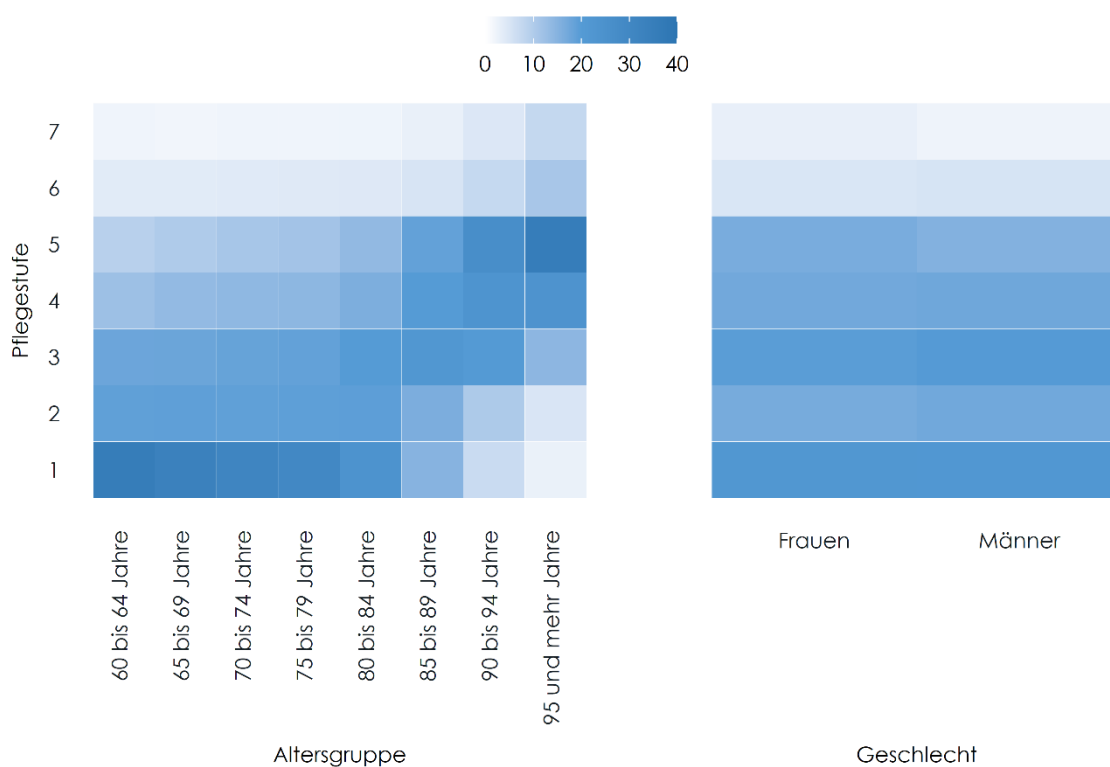
Übergänge in eine andere Pflegegeldstufe sind in der überwiegenden Mehrzahl Übergänge in eine höhere Pflegegeldstufe (Abbildung 19). Rund die Hälfte der Übergänge sind

²⁵⁾ Die etwas geringere Korrelation mit dem LASSO-Modell kommt daher, dass im LASSO-Modell Parameter auf exakt Null geschrumpft werden können. Im LASSO-Modell werden alle bis auf 940 erklärende Variablen auf Null geschrumpft.

Höhereinstufungen um eine Pflegegeldstufe, ein knappes Viertel der Übergänge sind Höhereinstufungen um zwei Pflegegeldstufen und gut ein Zehntel sind Höhereinstufungen um drei Pflegegeldstufen.

Die häufigsten Übertritte in eine andere Pflegegeldstufe finden einerseits beim Eintritt in das Pflegegeldsystem statt (von Pflegegeldstufe 0 in Pflegegeldstufe 1), aber auch Übertritte in höhere Pflegegeldstufen, insbesondere von Personen über 85 Jahren. Die Übertritte in andere Pflegegeldstufen sind zwischen den Geschlechtern relativ ähnlich verteilt (siehe Abbildung 20).

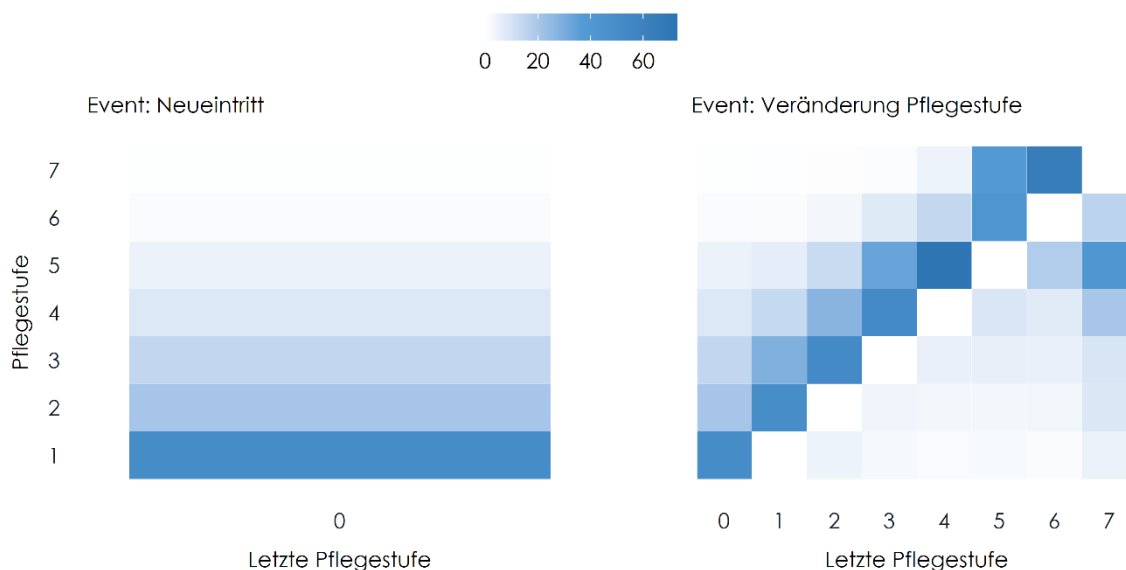
Abbildung 20: **Darstellung der Zielpflegegeldstufe bei Veränderung der Pflegegeldstufe, nach Alter und Geschlecht**



Q: WIFO, DVSV. Diese Abbildung zeigt, in welche Pflegegeldstufen die jeweiligen Altersgruppen (links) bzw. Geschlechter (rechts) in welchem Ausmaß wechseln (Lesebeispiel: die Gruppe der 60 bis 64-Jährigen wechseln zumeist in Pflegegeldstufe 1, am seltensten in Pflegegeldstufe 7, während die 95-Jährigen und älteren hauptsächlich in Pflegegeldstufe 5 wechseln).

Abbildung 21 zeigt neben den Neueintritten in das Pflegegeldsystem auch die Übertritte in andere Pflegegeldstufen. Dabei wird ersichtlich, dass, erstens, die meisten Übertritte in höheren Pflegegeldstufen anlangen und zweitens, dass die häufigsten Übertritte jene in die nächsthöhere Pflegegeldstufe sind.

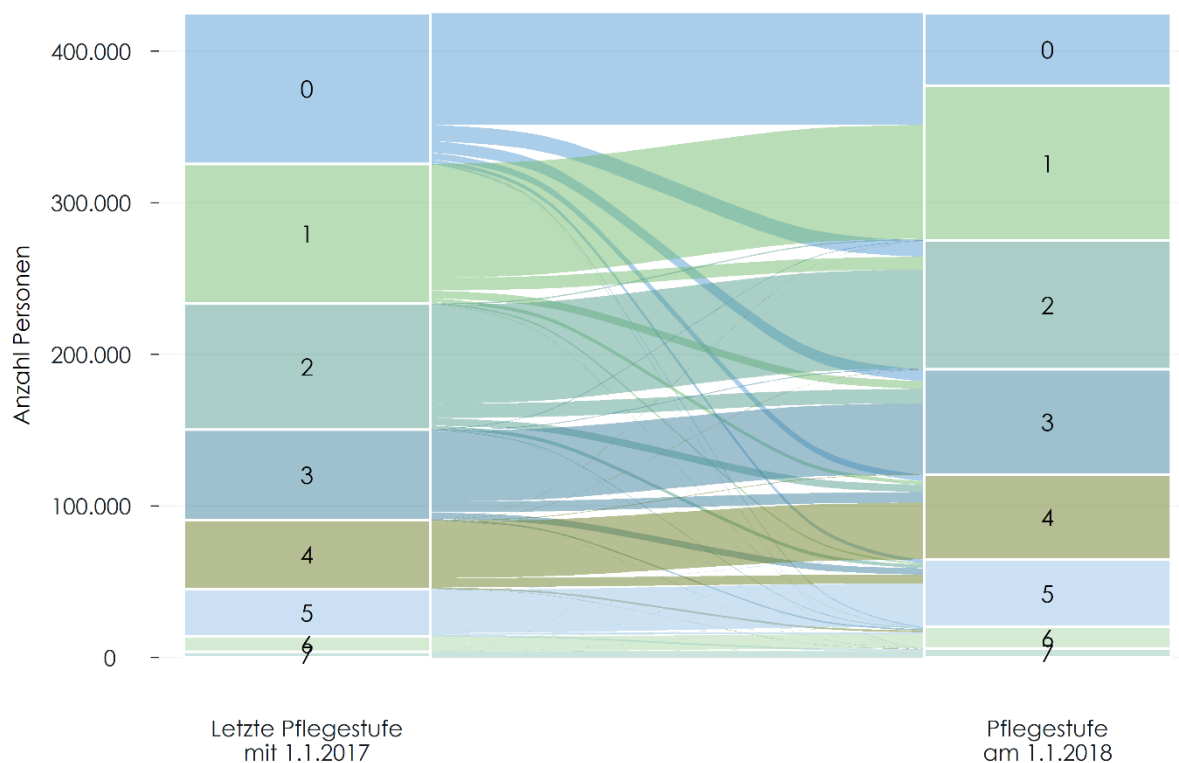
Abbildung 21: **Übertritte der Pflegegeldstufen**



Q: WIFO, DVSV.

Im folgenden Flow-Diagramm (Abbildung 22) sind alle Personen des Datensatzes über 60 Jahren und ihre Pflegegeldstufen zu zwei Zeitpunkten (Ende des Beobachtungszeitraums: 1.1.2017 sowie Ende des Eventzeitraums: 1.1.2018) dargestellt. Zwischen diesen beiden Zeitpunkten zeigt sich, dass der Großteil der Personen in derselben Pflegegeldstufe bleibt (über 80%). Weiters ist ersichtlich, dass die meisten Änderungen Höhereinstufungen sind, meist um eine Pflegegeldstufe, oft auch um zwei.

Abbildung 22: **Flow-Diagramm aller Übertritte und Verbleibe von Pflegegeldbeziehenden am Ende des Beobachtungszeitraumes und am Ende des Eventzeitraumes**



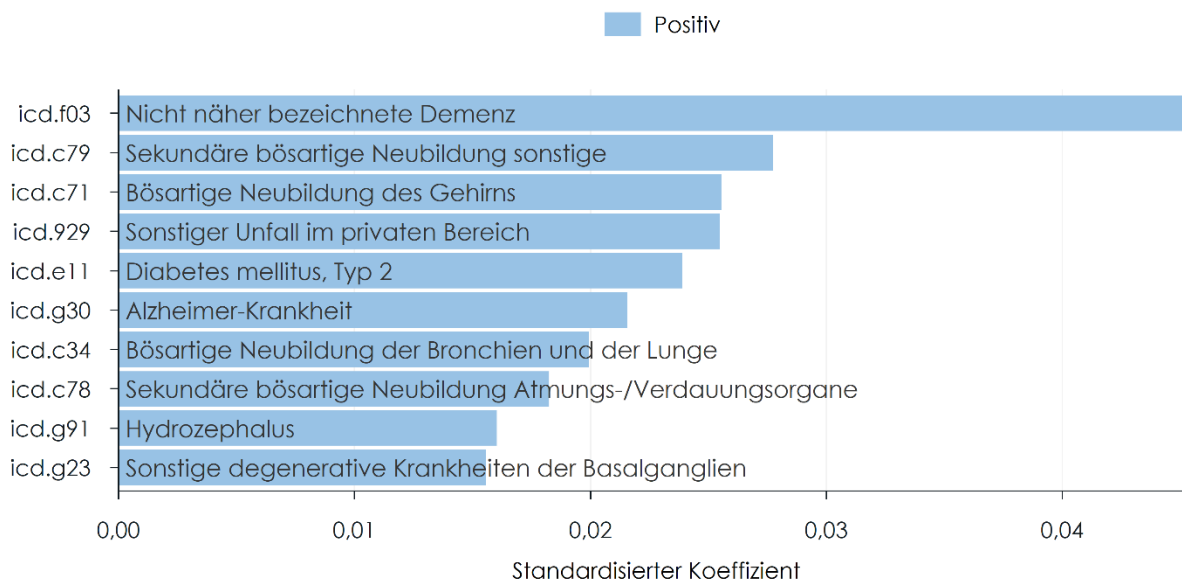
Q: WIFO, DVSV.

3.3 Analyse der gesundheitlichen Gründe für einen Übergang in eine höhere Pflegegeldstufe

Dieser Abschnitt stellt die zehn einflussreichsten Merkmale im Bereich der Diagnosen, der medizinischen Einzelleistungen, der Fachgebiete bei ärztlichen Leistungen, der Heilmittel und Medikamente sowie der sonstigen persönlichen Merkmale eines Übergangs in eine höhere Pflegegeldstufe dar. Betrachtet werden ausschließlich die Neuzugänge sowie die Übergänge in höhere Pflegegeldstufen - also keine Rückstufungen.

Die zehn wesentlichsten Diagnosen, die während stationären Krankenhausaufenthalten gestellt wurden und mit einer Höhereinstufen der Pflegegeldstufe in Verbindung stehen, werden in Abbildung 23 dargestellt. Diagnosen sind hier gestellte Haupt- und Zusatzdiagnosen auf Ebene der vierstelligen Systematik der Internationalen statistischen Klassifikation der Krankheiten und Gesundheitsprobleme (ICD-10). Die Ergebnisse zeigen die hohe Relevanz von dementiellen Erkrankungen. Ebenso wichtig sind bösartige Neubildungen in unterschiedlichen Körperregionen.

Abbildung 23: **Die zehn wichtigsten Diagnosen, die mit einer Höhereinstufung der Pflegegeldstufe in Verbindung stehen**

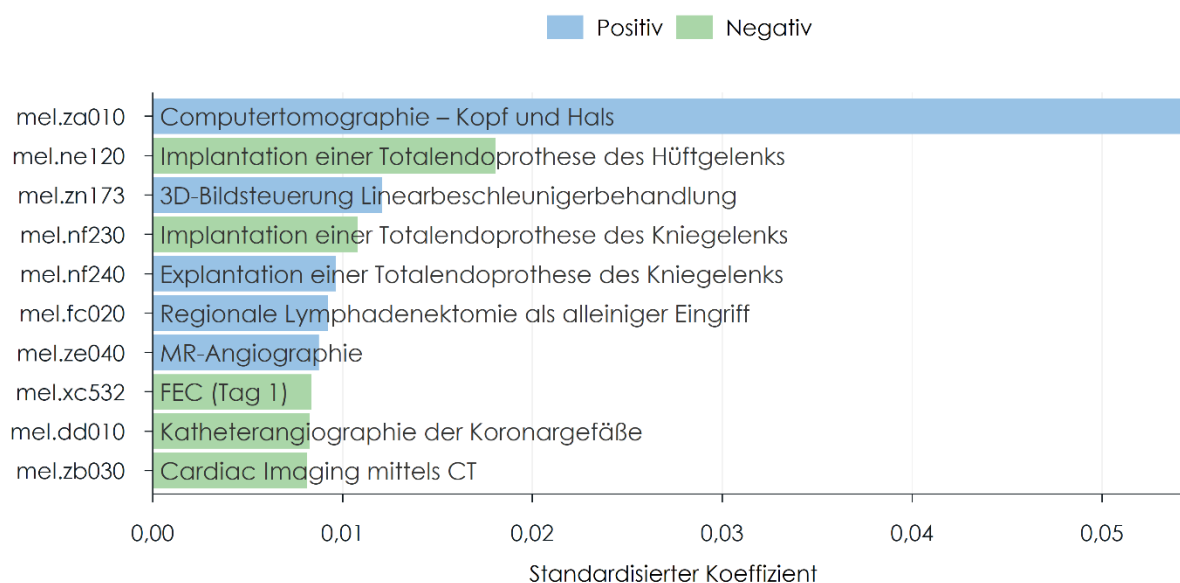


Q: WIFO, DVSV.

Im Bereich der im Beobachtungszeitraum bezogenen medizinischen Einzelleistungen vor einer Höhereinstufung der Pflegegeldstufe bzw. eines Neuzugangs zeigt die Leistung einer Computertomographie im Kopf- und Hals-Bereich (mel.za010) die größte positive Erklärungskraft (Abbildung 24). Einen ebenso positiven Effekt, wenn auch im Ausmaß kleiner, zeigen die Leistungen einer Bestrahlung (mel.zn173), die Explantation eines künstlichen Kniegelenks (mel.nf240), die operative Entfernung von Lymphknoten (mel.fc020) sowie eine diagnostische Darstellung von Blutgefäßen mit den Methoden der Magnetresonanztomographie (mel.ze040).

Die Implantation eines künstlichen Hüftgelenks und die Implantation eines künstlichen Kniegelenks (mel.nf230) zeigen eine negative Korrelation mit einer Höhereinstufung der Pflegegeldstufe – d.h. die Inanspruchnahme dieser medizinischen Leistungen reduzieren, ceteris paribus, die Wahrscheinlichkeit in eine höhere Pflegegeldstufe zu wechseln. Diese Eingriffe werden in der Regel nur bei Personen gemacht, die eine gute Aussicht haben, für den schwierigen Genesungsverlauf fit genug zu sein. Ebenso negative Korrelationen zeigen Chemotherapien (mel.xc532), Kathedereingriffe in die Herzkranzgefäße (mel.dd010) sowie Herz-CTs (mel.zb030). Wichtig hinzufügen ist, dass es sich hier nicht um kausale Interpretationen handelt, sondern um statistische Korrelationen.

Abbildung 24: **Medizinische Einzelleistungen und Übergang in eine höhere Pflegegeldstufe**



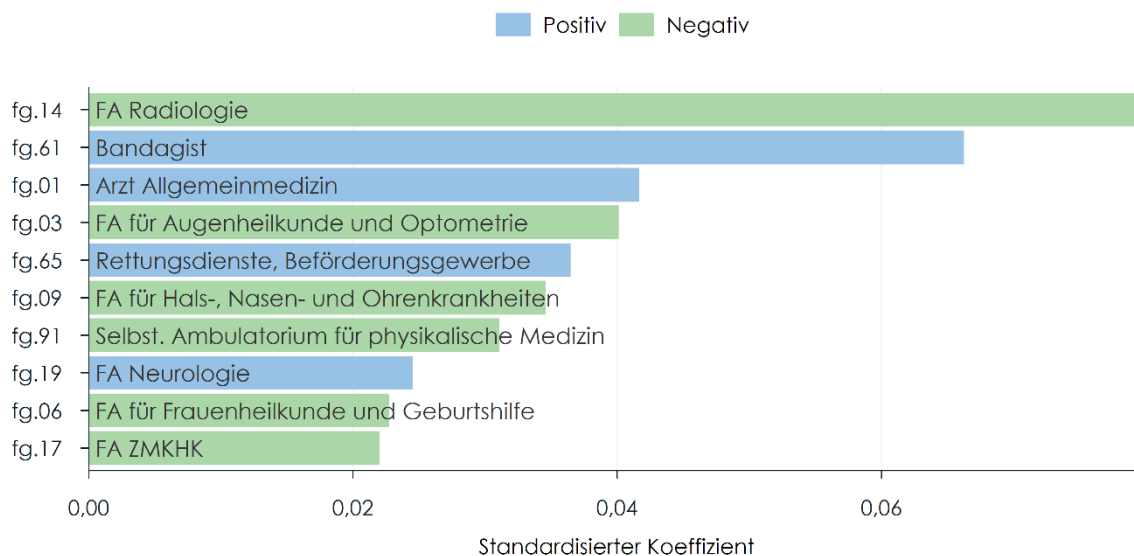
Q: WIFO, DVSV.

Die Fachgebiete der bezogenen ärztlichen Leistungen im niedergelassenen Bereich vor einer Erhöhung der Pflegegeldstufe zeigen heterogene Effekte (Abbildung 25). Der Besuch bei einer Fachärztin bzw. eines Facharztes für Radiologie korreliert negativ mit dem Wechsel in eine höhere Pflegegeldstufe. Ebenso – aber in einem geringeren Ausmaß – negative Effekte hat der Bezug von Leistungen von Augenärztinnen und -ärzten, HNO-Ärztinnen und -ärzten, eines Ambulatoriums für physikalische Medizin, von Gynäkologinnen und Gynäkologen sowie von Zahnärztinnen und -ärzten. Wiederum können die Besuche bei niedergelassenen Fachärztinnen und Fachärzten dieser Gruppen Hinweise darauf sein, dass die Personen entsprechend mobil sind bzw., ceteris paribus, einen besseren Gesundheitszustand aufweisen als Personen, die diese Facharztleistungen nicht in Anspruch nehmen (können). Den höchsten positiven Effekt hat der Bezug von Leistungen einer Bandagistin bzw. eines Bandagisten sowie – zu einem geringeren Ausmaß – von Allgemeinmedizinerinnen und -mediziner, Rettungsdiensten und Neurologinnen und Neurologen.

Die wichtigsten Heilmittel bzw. Medikamente, die mit einer Höhereinstufung des Pflegegeldes in Verbindung stehen, sind Substanzen, die in der Regel bei unterschiedlichen dementiellen Erkrankungen verabreicht werden (Anticholinesterasen, andere Medikamente gegen Demenz). So werden Anticholinesterasen, das sind Enzyme, die den Abbau von Neurotransmittern²⁶⁾ hemmen, vorwiegend bei Alzheimer-Erkrankungen eingesetzt. Medikamentös zugeführte Dopamine gleichen den Dopaminmangel z.B. bei Morbus Parkinson aus.

²⁶⁾ Botenstoff, der im Nervensystem Informationen von einer Nervenzelle zur anderen weitergibt.

Abbildung 25: **Fachgebiete ärztlicher Leistungen und Übergang in eine höhere Pflegegeldstufe**

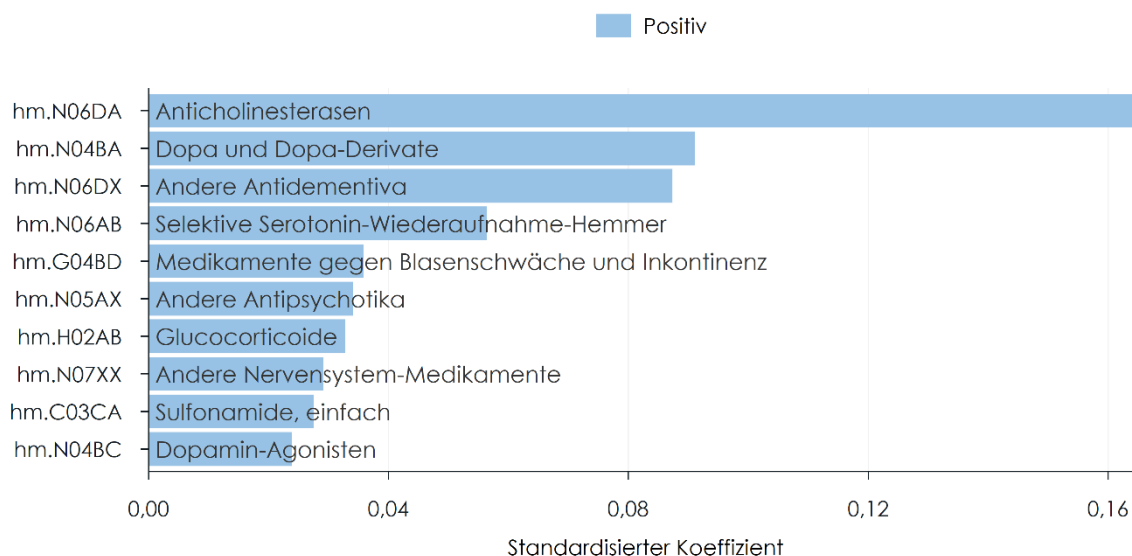


Q: WIFO, DVSV.

Ebenso in Verbindung mit einer Höhereinstufung des Pflegegelds stehen Serotonin (ebenfalls ein Neurotransmitter, oft eingesetzt gegen Depressionen), Anti-Psychotika und andere Nervenmedikamente. Eine weitere Gruppe sind Medikamente, die oft mit Harnwegserkrankungen (Sulfonamide)²⁷⁾ und Inkontinenz in Verbindung stehen. Glucocorticoide, die auch eine Rolle spielen, sind ein Cortison, wirken also entzündungshemmend und regulieren den Elektrolyt-haushalt. Aufgrund der Wirkung, dass sie das Immunsystem "herunterfahren", werden sie z.B. vor Transplantationen, aber auch bei starken Formen von Rheuma oder Arthritis eingesetzt. Dopaminagonisten werden zur Dopaminreduktion z.B. bei Psychosen, Epilepsie, manische Zustände, Entzugsbehandlungen, aber auch bei Alzheimer eingesetzt (v. a., wenn aggressives Verhalten ausgelöst wird).

²⁷⁾ Antibiotikum z.B. bei chronischen Harnwegsinfekten .

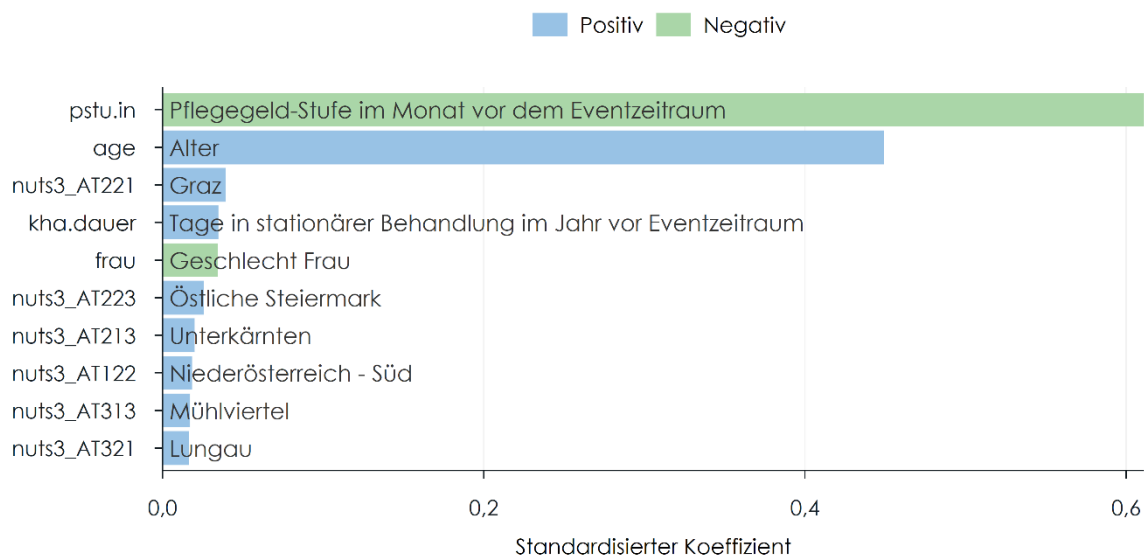
Abbildung 26: **Medikamente/Heilmittel und der Übergang in eine höhere Pflegegeldstufe**



Q: WIFO, DVSV.

Andere persönliche Merkmale – abseits von Diagnosen, Medikamenten/Heilmitteln, medizinischen Einzelleistungen und Fachgebieten medizinischer Leistungen – die mit einer Höhereinstufung der Pflegegeldstufe in Verbindung stehen, zeigt Abbildung 27. Eine höhere Pflegegeldstufe hat einen negativen Effekt auf eine Höhereinstufung der Pflegegeldstufe: Je höher die Pflegegeldstufe, desto höher die Wahrscheinlichkeit, nicht in eine höhere Pflegegeldstufe überzutreten. Einen ebenfalls negativen Effekt auf die Wahrscheinlichkeit in eine höhere Pflegegeldstufe zu wechseln, wenn auch in einem deutlich schwächeren Ausmaß, haben Frauen. Dies bedeutet, dass Frauen – im Vergleich zu Männern – bei gleichen gesundheitlichen Determinanten eine geringere Wahrscheinlichkeit haben, in eine höhere Pflegegeldstufe überzutreten. Einen stark positiven Effekt hat das Alter: Je älter die Personen, desto höher ist die Wahrscheinlichkeit, in eine höhere Pflegegeldstufe zu kommen. Ebenso positiven Einfluss, wenn auch deutlich schwächer, haben gewisse NUTS3-Regionen (Graz, östl. Steiermark, Unterkärnten, Niederösterreich-Süd, Mühlviertel und Lungau) sowie die Anzahl der Nächte in stationärer Behandlung in einem Krankenhaus im Beobachtungszeitraum.

Abbildung 27: **Sonstige persönliche Merkmale und der Übergang in eine höhere Pflegegeldstufe**



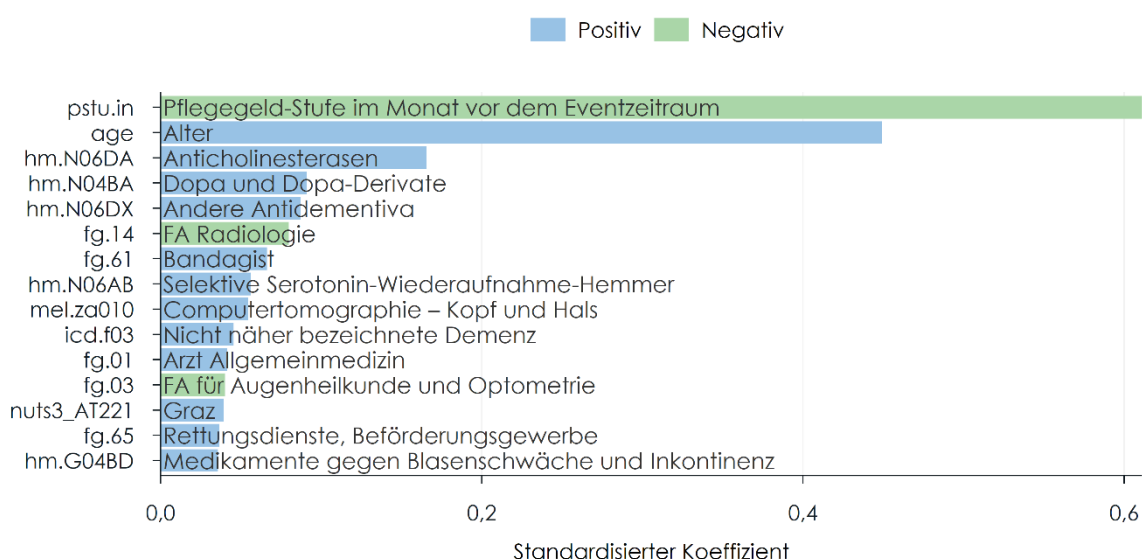
Q: WIFO, DVSV.

Die 15 einflussreichsten Merkmale auf eine Höhereinstufung der Pflegegeldstufe zeigt Abbildung 28. Hier sind alle Merkmalskategorien zusammengefasst: Diagnosen, Medikamente/Heilmittel, medizinische Einzelleistungen, Fachgebiete medizinischer Leistungen und sonstige persönliche Charakteristika. Den stärksten negativen Effekt hat die Höhe der Pflegegeldstufe im Monat vor dem Eventzeitraum: Eine bereits höhere Pflegegeldstufe im Ausgangszeitpunkt verringert die Wahrscheinlichkeit für eine Höhereinstufung deutlich. Der Grund dafür liegt wohl auch in der Definition der Anspruchsvoraussetzungen für die einzelnen Pflegegeldstufen. Die Pflegegeldstufen 5 bis 7 haben die gleiche Voraussetzung bezüglich des Pflegeaufwands (nämlich mehr als 180 Stunden im Monat), nur das qualitative Ausmaß des Pflegebedarfsdivergieren (Bundespflegegeldgesetz 1993 in der letztgültigen Fassung).²⁸⁾ Zusätzlich sind in den oberen Pflegegeldstufen relativ "stabiler" Zustände möglich (z.B. Wachkomapatientinnen und -patienten), wo sich der Zustand bis zum Tod nicht mehr ändert. Schwach negative Effekte zeigen die Faktoren der Inanspruchnahme von Leistungen von Radiologinnen und Radiologen sowie von Augenärztinnen und -ärzten. Die Analyse zeigt weiters, nicht überraschend, einen positiven Zusammenhang zwischen höherem Alter und einer Höhereinstufung der Pflegegeldstufe. Ebenso positive Effekte, wenn auch schwächer, zeigen mit Demenz verbundene Medikamente sowie Stimmungsaufheller (Serotonin). Die einzigen medizinischen

²⁸⁾ Pflegegeldstufe 5: Mehr als 180 Stunden und Vorliegen von außergewöhnlichem Pflegeaufwand. Pflegegeldstufe 6: Mehr als 180 Stunden und Erfordernis zeitlich unkoordinierter Betreuungsmaßnahmen oder Erfordernis dauernder Beaufsichtigung aufgrund Eigen- oder Fremdgefährdung. Pflegegeldstufe 7: Mehr als 180 Stunden und wenn keine zielgerichteten Bewegungen der vier Extremitäten mit funktioneller Umsetzung möglich sind oder ein gleichzuachtender Zustand vorliegt.

Einzelleistung, die in die Analyse der 15 wichtigsten Merkmale für eine Höhereinstufung der Pflegegeldstufe eingeht, ist jene der Computertomographie im Kopf- und Halsbereich. Als einzige Diagnose geht jene der (nicht näher bezeichneten) Demenz ein. Von den weiteren Merkmalen zeigt die NUTS-3-Region Graz einen positiven Zusammenhang. Pflegebedürftige in Graz haben – unter Kontrolle der vorhandenen Merkmale – eine höhere Wahrscheinlichkeit in eine höhere Pflegegeldstufe zu kommen als Personen in anderen Regionen. Der Grund dafür ist unklar.

Abbildung 28: **Analyse des Übergangs in eine höhere Pflegegeldstufe: Die 15 wichtigsten Merkmale**



Q: WIFO, DVSV.

3.4 Main Findings und Fazit

Das Modul 2 hat durch die Zusammenführung der Gesundheitsleistungsdaten mit Daten der Pflegegeldinformation die Zusammenhänge zwischen der gesundheitlichen Entwicklung und den Übertritten in höhere Pflegegeldstufen von pflegegeldbeziehenden Personen ab 60 Jahren analysiert. Als Beobachtungszeitraum wurde das Jahr 2016 und als Eventzeitraum das Jahr 2017 definiert. Im Jahr 2017 sind 18,2% der pflegegeldbeziehenden Personen ab 60 Jahren in eine höhere Pflegegeldstufe gewechselt. Übergänge in eine andere Pflegegeldstufe sind in der überwiegenden Mehrzahl Übergänge in eine höhere Pflegegeldstufe, meist in die nächsthöhere Pflegegeldstufe.

Die in Modul 2 vorgestellten Ergebnisse zeigen die Wahrscheinlichkeit des Übertritts in eine höhere Pflegegeldstufe unter Berücksichtigung persönlicher Eigenschaften, den im Beobachtungszeitraum in Anspruch genommenen Gesundheitsleistungen und gestellten Diagnosen sowie der Pflegegeldstufe am Ende des Beobachtungszeitraums (Dezember 2016). Die

deskriptiven Ergebnisse zeigen, dass die häufigsten Übertritte in eine höhere Pflegegeldstufe beim Eintritt in das Pflegegeldsystem (von Pflegegeldstufe 0 auf Pflegegeldstufe 1), sowie von Personen über 85 Jahren stattfinden.

Die Analyse der Modellwahl hat für das LASSO-Regressionsmodell die größte Erklärungskraft ergeben, weshalb die Ergebnisse dieses Modells ausgewiesen wurden. Die LASSO-Regressionen zeigen, dass – unter Berücksichtigung aller vorgestellten Merkmale – die stärkste Erklärungskraft für den Übergang in eine höhere Pflegegeldstufe im Jahr 2017 die Pflegegeldstufe im Dezember 2016 hat: eine bereits höhere Pflegegeldeinstufung hat einen negativen Effekt auf die Wahrscheinlichkeit einer Höhereinstufung der Pflegegeldstufe. Mögliche Erklärungen hierfür sind die Anspruchsvoraussetzungen der Pflegegeldstufen 5 bis 7, die sich nicht im Ausmaß des Pflegeaufwandes unterscheiden, sondern durch das qualitative Ausmaß des Pflegebedarfs. Ebenso können obere Pflegegeldstufen mit relativ stabileren Gesundheitszuständen verbunden sein, die sich vor dem Tod nicht mehr ändern.

Ein höheres Alter erhöht die Wahrscheinlichkeit, *ceteris paribus*, in eine höhere Pflegegeldstufe überzutreten, ebenso wie die Verschreibung von mit Demenz verbundenen Medikamenten. Die Analyse der Diagnosen hebt die Relevanz von Demenz und bösartigen Neubildungen für den Übertritt in höhere Pflegegeldstufen hervor. Die wesentlichste medizinische Einzelleistung, die einen positiven Zusammenhang mit dem Übergang in eine höhere Pflegegeldstufe hat, ist jene der Computertomographie des Kopf- und Halsbereichs.²⁹⁾ Die Implantation eines künstlichen Hüftgelenks – ebenso wie die Implantation eines künstlichen Kniegelenks – haben einen negativen Zusammenhang mit einer Höhereinstufung: Diese Eingriffe reduzieren, *ceteris paribus*, die Wahrscheinlichkeit eines Wechsels in eine höhere Pflegegeldstufe. Eine mögliche Begründung könnte sein, dass diese Eingriffe nur an vergleichsweise "fit-ten" Personen durchgeführt werden. Höhere Wahrscheinlichkeiten resultieren jedoch aus Leistungen einer Bestrahlung (wohl idR als Folge einer bösartigen Neubildung). Die Medikamente bzw. Heilmittel, die am stärksten mit dem Übergang in eine höhere Pflegegeldstufe korrelieren, sind Substanzen, die mit unterschiedlichen dementiellen Erkrankungen in Verbindung stehen. Diese Ergebnisse zeigen die Wichtigkeit von dementiellen Erkrankungen bei der Projektion eines Wechsels in eine höhere Pflegegeldstufe.

Die Analyse der Rolle der persönlichen Merkmale zeigt – neben dem negativen Zusammenhang der Pflegegeldstufe und dem positiven Zusammenhang des Alters mit einer Höhereinstufung der Pflegegeldstufe – einen positiven Zusammenhang zwischen dem Übergang in eine höhere Pflegegeldstufe und manchen Regionen.

²⁹⁾ Mögliche Zusammenhänge könnten in der medizinischen Abklärung von Demenz, Schlaganfällen oder nach Unfällen zu finden sein.

4. Modul 3: Gesundheitliche Gründe für den Eintritt in stationäre Pflege

4.1 Ziel des Moduls 3

Da die Pflegegeldinformationsdaten auch eine Identifikation des Eintritts in stationäre Pflege erlauben, können mithilfe der Gesundheitsleistungsdaten die gesundheitlichen Zusammenhänge mit dem Eintritt in stationäre Pflege ermittelt werden. Dafür wird der Effekt des Bezuges unterschiedlicher Gesundheitsleistungen im Jahr vor dem Eintritt in stationäre Pflege analysiert. Ziel ist die Identifikation der wesentlichen gesundheitlichen Faktoren für den Eintritt in stationäre Pflege.

Die Struktur der Personen in stationärer Pflege hat sich über die Jahre hinweg gewandelt. In allen Bundesländern besteht seit einigen Jahren die Tendenz, verstärkt Personen in höheren Pflegegeldstufen in Pflegeheimen aufzunehmen (*Firgo – Famira-Mühlberger, 2019*). Mit der Änderung des Pflegefondsgesetzes im Jahr 2017 wurde für alle Bundesländer einheitlich festgelegt, dass Eintritte nur ab Pflegegeldstufe 4 zu erfolgen haben. Dies bedeutet, dass Personen, die in stationäre Pflege eintreten, tendenziell älter und pflegebedürftiger sind als dies noch vor einigen Jahren der Fall war.

Ein Übertritt in stationäre Pflege ist in der Regel mit einer Verschlechterung des Gesundheitszustands (z. B. durch Multimorbidität) oder durch ein plötzliches gesundheitliches Ereignis (z. B. Stürze oder bösartige Neubildungen) verursacht.³⁰⁾ Ziel ist es auf Basis von Gesundheitsleistungsdaten (erstellte Diagnosen, verordnete Medikamente, erbrachte medizinische Einzelleistungen, erbrachte Leistungen nach Fachgebieten) und unter Berücksichtigung von persönlichen Merkmalen die wichtigsten Determinanten für den Übertritt in stationäre Pflege in einem multivariaten Modell zu analysieren.

Die Ergebnisse sind aus pflegepolitischer Perspektive wichtig, da sie zeigen, welche gesundheitliche Entwicklung und Charakteristika von älteren Personen zu einem Übertritt in stationäre Pflege führen. Die Ergebnisse können – im Idealfall – in künftige Projektionen über die Nachfrage nach stationärer Pflege einfließen.

4.2 Empirisches Modell und Beschreibung der Daten

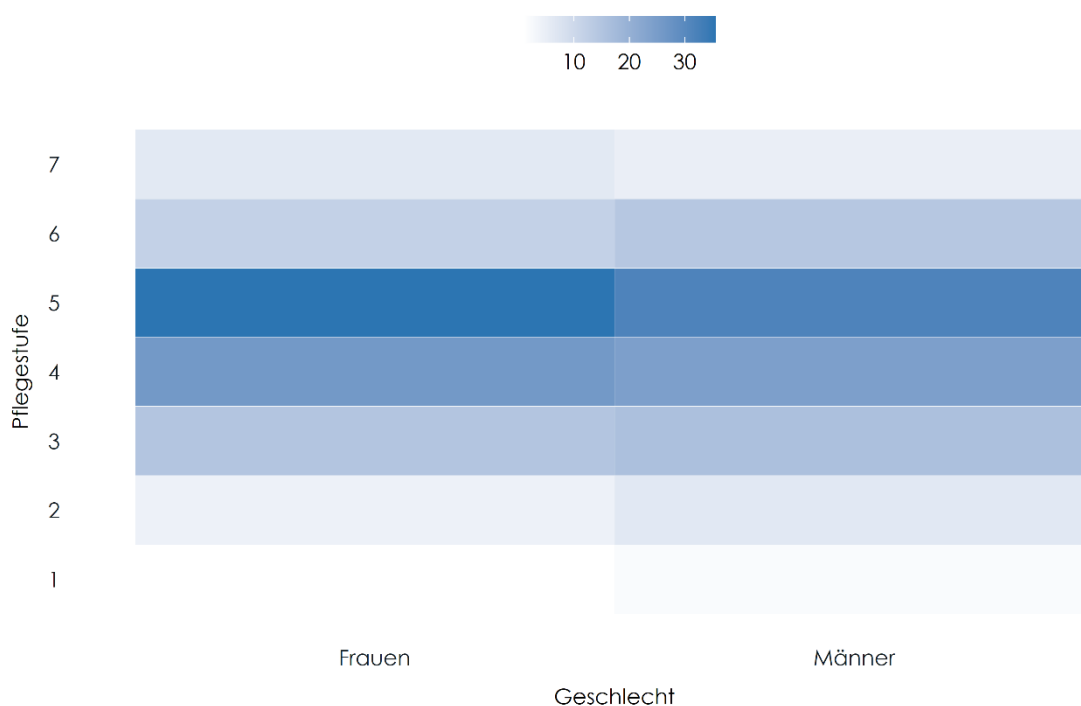
Zur Analyse des Eintritts in stationäre Pflege werden die gleichen Daten (Pflegegeldbeziehenden und -bezieher ab 60 Jahren) wie in Modul 2 verwendet. Als Eventzeitraum wird wiederum das Jahr 2017 herangezogen. Die binäre abhängige Variable nimmt für eine Person den Wert 1 an, wenn innerhalb des Eventzeitraums ein Eintritt in stationäre Pflege verzeichnet wird (sonst 0). Dies traf im Eventzeitraum [01/2017, 01/2018] auf ca. 7,7% der 377.414 Pflegegeldbeziehenden in den vorliegenden Daten zu. Die Wahrscheinlichkeit dieses Ereignisses wird erneut als Funktion persönlicher Eigenschaften, den im Beobachtungszeitraum [01/2016, 01/2017] in Anspruch genommenen Gesundheitsleistungen und gestellten Diagnosen sowie

³⁰⁾ Auch Veränderungen im Familiengefüge können ausschlaggebend sein.

der Pflegegeldstufe im Monat Dezember 2016 modelliert. Als Eintritt in stationäre Pflege werden alle Eintritte in stationäre Pflege subsummiert, auch jene in Kurzzeit- oder Übergangspflege. Der Übertritt in ein Pflegeheim wird in dieser Analyse dann als Übertritt bewertet, wenn in der Pflegegeldinformation vermerkt ist, dass das Pflegegeld auf ein Pflegeheim übertragen wurde ("Übergangsgrund (Pflegeheim J/N)").

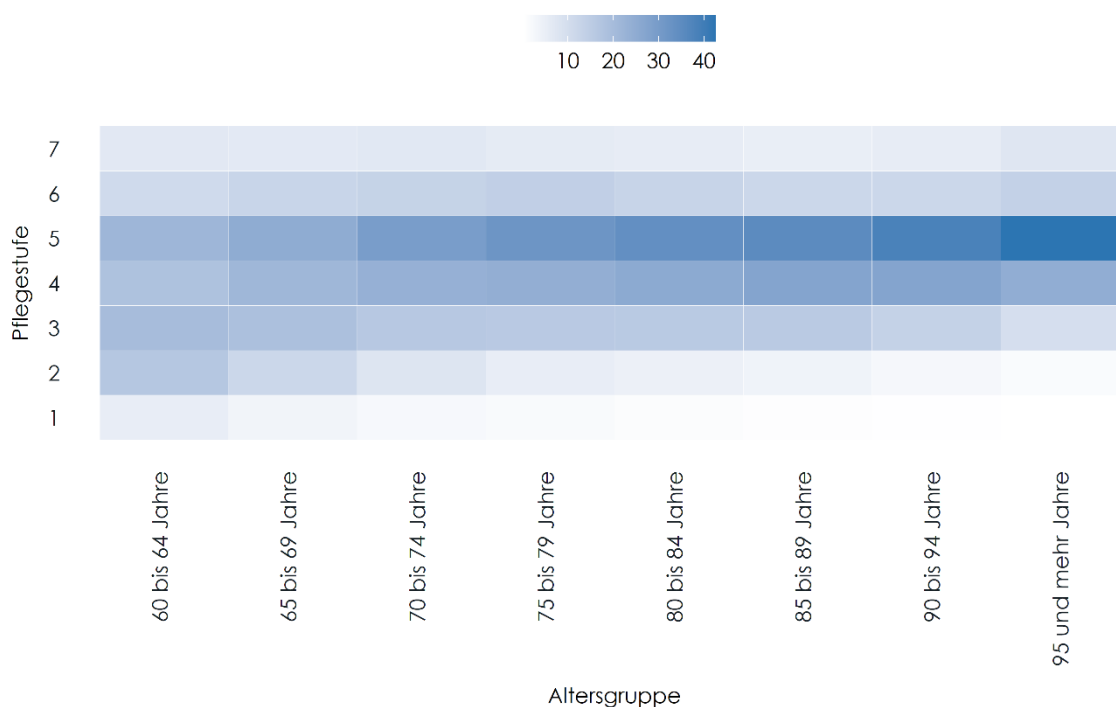
Die deskriptiven Statistiken zeigen, dass Frauen, die in stationäre Pflege übertreten, sich tendenziell in höheren Pflegegeldstufen befinden als Männer (Abbildung 29). Die verkreuzte Darstellung zwischen Altersgruppen und Pflegegeldstufen zeigt, dass es – wenig überraschend – vermehrt Ältere und Personen in höheren Pflegegeldstufen sind, die in stationäre Pflege übertreten (Abbildung 30).

Abbildung 29: **Eintritt in stationäre Pflege nach Geschlecht**



Q: WIFO, DVSV.

Abbildung 30: **Eintritt in stationäre Pflege nach Pflegegeldstufe und Alter**

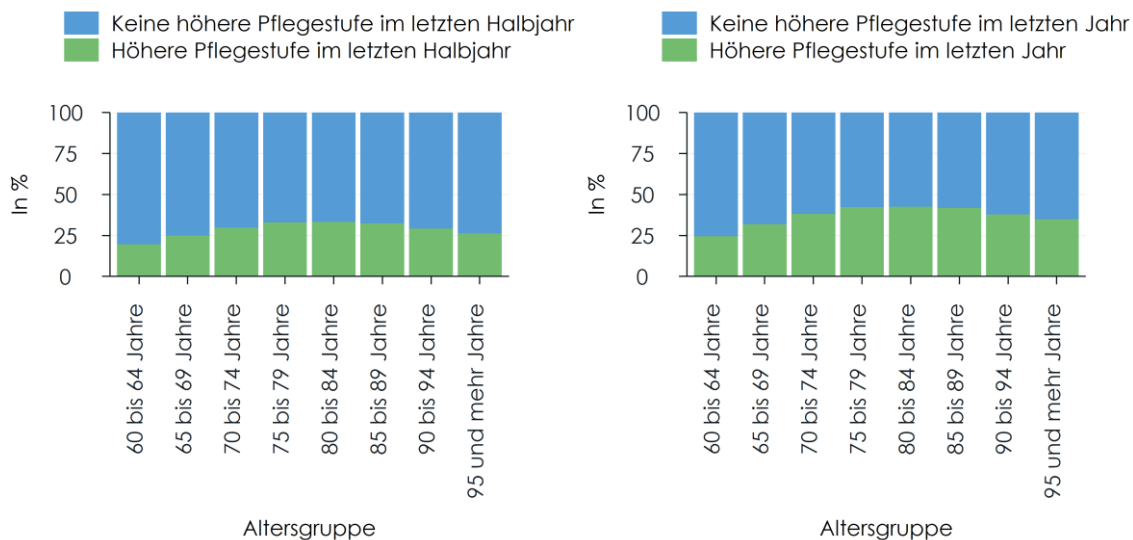


Q: WIFO, DVSV

Abbildung 31 stellt die Anteile der Personen, die vor dem Eintritt in stationäre Pflege in eine höhere Pflegegeldstufe gewechselt sind. Mindestens ein Viertel der Personen in allen Altersgruppen ist im Jahr vor dem Eintritt in stationäre Pflege auch in eine höhere Pflegegeldstufe übergetreten. Betrachtet man hingegen nur die letzten sechs Monate vor dem Eintritt in stationäre Pflege zeigt sich ein deutlich geringerer Anteil.

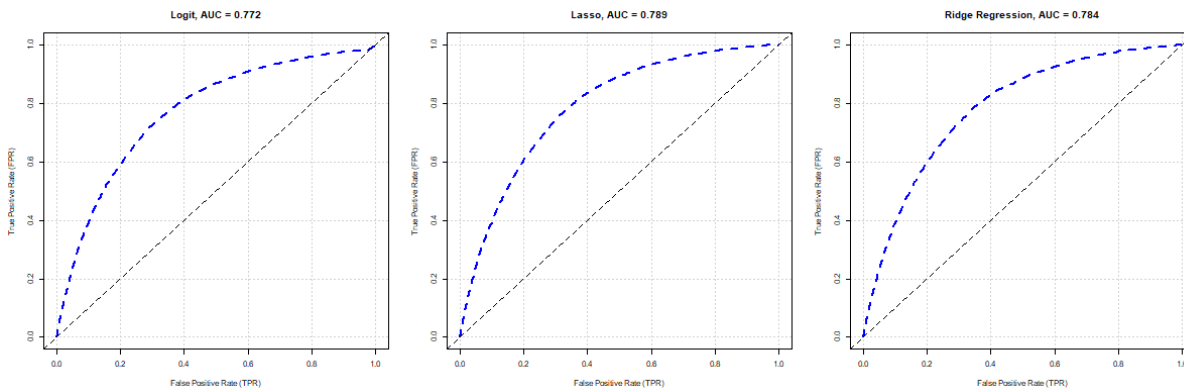
Die Einordnung der erklärenden Variablen erfolgt erneut anhand ihres Potentials zur Klassifizierung der abhängigen Variable in den Testdaten, auf denen das Modell nicht trainiert wurde (25% aller Beobachtungen). Geschätzt werden Logit, LASSO und Ridge Regression wiederum auf Trainingsdaten, welche 75% des Gesamtdatenbestands umfassen. Die letzten beiden Modelle verwenden 5-fache Kreuzvalidierung in den Trainingsdaten um den "Strafterm" λ zu optimieren.

Abbildung 31: **Eintritt in stationäre Pflege und Höhereinstufung der Pflegegeldstufe ein halbes Jahr und ein Jahr vor Eintritt**



Q: WIFO, DVSV.

Abbildung 32: **ROC-Kurven von Logit-, LASSO- und Ridge Regression-Modellen für den Eintritt in eine höhere Pflegegeldstufe auf Basis der Testdaten**



Q: WIFO, DVSV.

Obwohl im LASSO-Modell die Parameter der meisten erklärenden Variablen auf Null geschrumpft werden und nur 826 Variablen aktiv im Modell verbleiben weist es die größte Fläche unter der ROC-Kurve auf (AUC: 0,79, siehe Abbildung 32). Aus diesem Grund wird auch in Modul 3 vorrangig auf die Ergebnisse dieses Modells zurückgegriffen. Erneut sind die Parameter der regularisierten Schätzer relativ hoch korreliert ($\rho = 0,923$), die Korrelation mit den geschätzten Koeffizienten des Logit-Modells kann jedoch – aufgrund der ausgeprägten

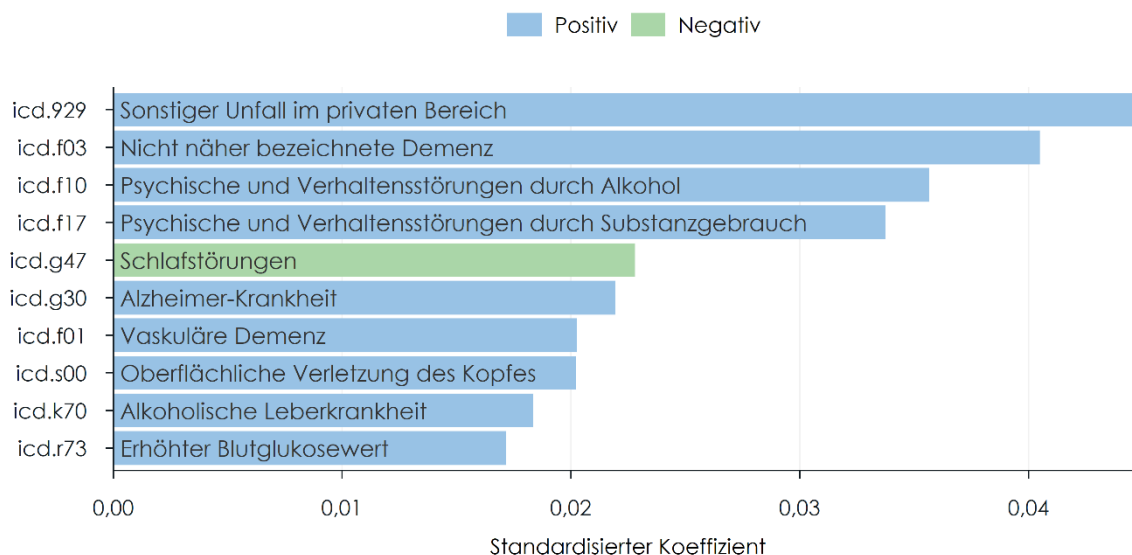
Regularisierung – nur noch als mittelstark bezeichnet werden (LASSO: $\rho = 0,280$; Ridge Regression: $\rho = 0,401$).

4.3 Analyse des Effekts von Gesundheitsleistungen auf den Eintritt in stationäre Pflege

Der folgende Abschnitt präsentiert die Ergebnisse von LASSO-Regressionen, die analog zu den übrigen Modulen den Zusammenhang von Diagnosen, medizinischen Einzelleistungen, Fachgebiete der medizinischen Leistungen, Heilmittel/Medikamente sowie sonstigen persönlichen Charakteristika von pflegegeldbeziehenden Personen und deren Eintritt in stationäre Pflege analysieren. In die jeweiligen LASSO-Regressionen, deren Ergebnisse in diesem Abschnitt dargestellt sind, sind jeweils alle verfügbaren Variablen aller beschriebenen Kategorien in die Analyse eingegangen. Im Folgenden werden die Ergebnisse getrennt nach Kategorien dargestellt – jeweils die 10 wichtigsten Variablen je Kategorie – und im Anschluss die 15 insgesamt wichtigsten Variablen.

Abbildung 33 zeigt die Effekte der zehn wichtigsten Diagnosen auf einen Eintritt in stationäre Pflege. Unfälle im privaten Bereich haben den größten positiven Effekt auf die Wahrscheinlichkeit für einen Eintritt in die stationäre Pflege. Einen ebenso positiven Effekt, aber in einem geringeren Ausmaß, haben unterschiedliche dementielle Erkrankungen sowie psychische und Verhaltensstörungen durch Alkohol oder Substanzgebrauch. Schlafstörungen hingegen zeigen einen negativen Effekt: die Diagnose Schlafstörungen reduziert, ceteris paribus, die Wahrscheinlichkeit in stationäre Pflege überzutreten.

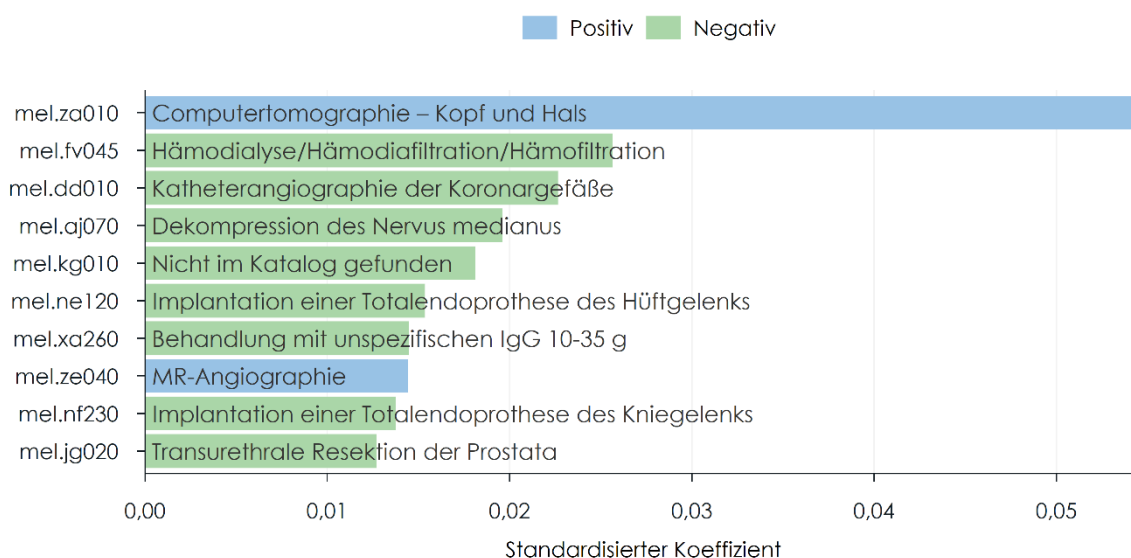
Abbildung 33: **Effekte von unterschiedlichen Diagnosen auf Eintritte in stationäre Pflege**



Q: WIFO, DVSV.

Die medizinische Einzelleistung, die den stärksten positiven Effekt auf die Wahrscheinlichkeit für den Eintritt in stationäre Pflege hat, ist eine Computertomographie des Kopf- und Halsbereiches. Ebenso einen positiven Effekt hat – wenn auch in einem geringeren Ausmaß – die Magnetresonanztomographie³¹⁾. Die restlichen in Abbildung 34 dargestellten medizinischen Einzelleistungen haben negative Effekte auf die Wahrscheinlichkeit des Eintritts in stationäre Pflege. Allen voran Nierenersatztherapien (Hämodialyse, -filtration und diafiltration), Kathetereingriffe der Herzkranzgefäße, Dekompression des Mittelarmnervs oder eine Implantation eines künstlichen Hüftgelenks. Ein mögliche Erklärung dafür könnte sein, dass es sich dabei zum Teil um Eingriffe handelt, die ein gewisses Maß an gesundheitlicher Stabilität erfordern, sodass diese Eingriffe bei Personen mit sehr labilem Gesundheitszustand nicht durchgeführt werden können.

Abbildung 34: **Effekte von medizinischen Einzelleistungen auf Eintritte in stationäre Pflege**



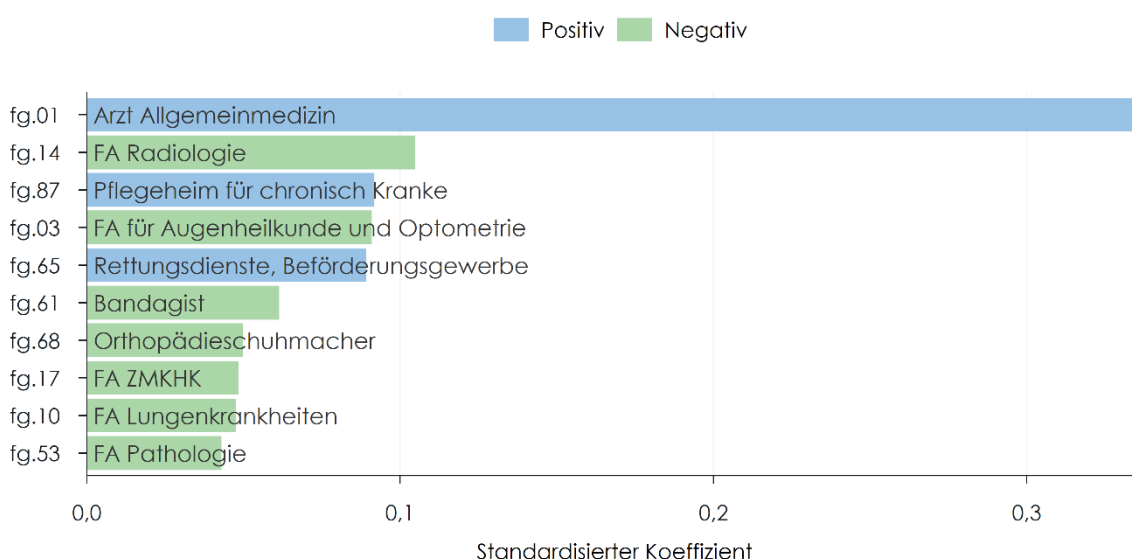
Q: WIFO, DVSV.

Bei den zehn wichtigsten Fachgebieten ärztlicher Leistungen, die mit einem Eintritt in stationäre Pflege in Verbindung stehen, zeigt der Besuch mit Allgemeinmedizinerinnen und -mediziner den größten positiven Effekt in Bezug auf die Wahrscheinlichkeit für den Eintritt in die stationäre Pflege (Abbildung 35). Ebenso positive Effekte, wenn auch im deutlich geringeren Ausmaß, zeigen der Aufenthalt in einem Pflegeheim für chronisch Kranke und die Inanspruchnahme einer Krankenbeförderung. Letztere könnte ein Hinweis auf das Fehlen von Angehörigen sein, welche informelle Pflege leisten und damit den Eintritt in die stationäre Pflege verhindern bzw. verzögern könnten. Die restlichen in Abbildung 35 gelisteten Fachgebiete

³¹⁾ Bildgebendes Verfahren zur diagnostischen Darstellung von Blutgefäßen.

ärztlicher Leistungen zeigen negative Effekte auf die Wahrscheinlichkeit für den Übertritt in stationäre Pflege. Die drei wesentlichsten darunter sind Leistungen von Radiologinnen und Radiologen, Augenärztinnen und -ärzte sowie Bandagistinnen und Bandagisten. Auch hier könnte eine Interpretationsmöglichkeit sein, dass die Besuche bei diesen niedergelassenen Fachärztinnen und Fachärzte nur möglich sind, wenn die Personen entsprechend mobil sind bzw., ceteris paribus, einen besseren Gesundheitszustand aufweisen als Personen, die diese Facharztleistungen nicht in Anspruch nehmen (können).

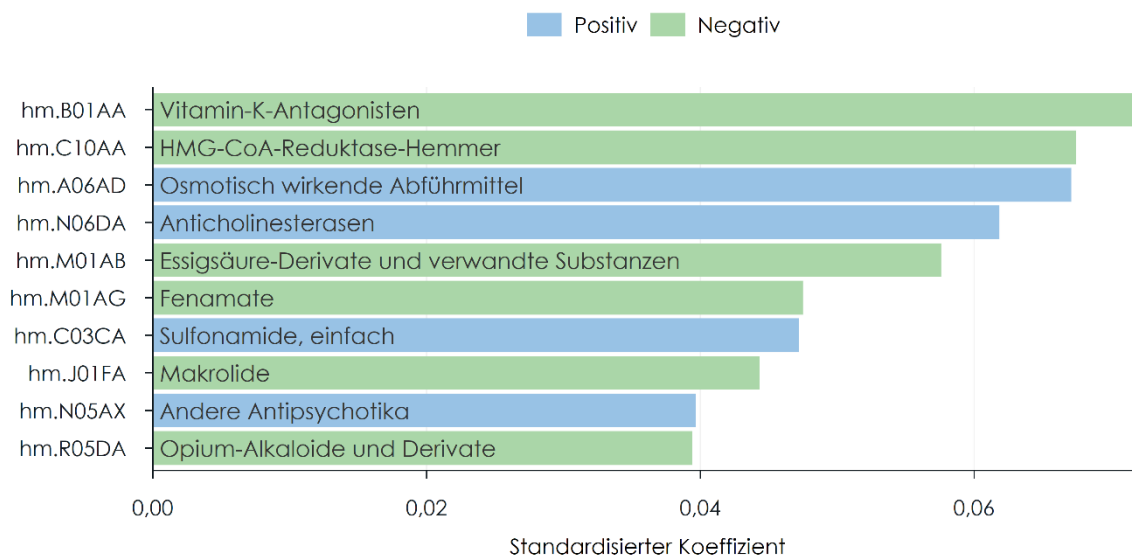
Abbildung 35: **Effekte der Inanspruchnahme von Gesundheitsanbietern nach Fachgebieten auf Eintritte in stationäre Pflege**



Q: WIFO, DVSV.

Die beiden Medikamentengruppen mit dem stärksten negativen Zusammenhang mit dem Eintritt in stationäre Pflege sind gerinnungshemmende Wirkstoffe (Vitamin-K-Antagonisten) und Cholesterinsenker (HMG-CoA-Reduktase) (Abbildung 36). Ebenso negative Zusammenhänge zeigen unterschiedliche Gruppen von Schmerzmitteln (Essigsäure-Derivate, Fenamates), Antibiotika (Makrolide) und schmerzlindernde Opiate. Positive Effekte auf die Wahrscheinlichkeit für den Eintritt in stationäre Pflege zeigen Medikamente im Bereich von Abführmittel (osmotische Laxantien), Anticholinesterasen (vorwiegend bei Alzheimer-Erkrankungen eingesetzt), Antibiotika (Sulfonamide, z.B. bei chronischen Harnwegsinfekten bzw. bei Inkontinenz, wenn in Verbindung mit Harnwegseffekten) sowie Anti-Psychotika. Die Verschreibung dieser Medikamente zeigt einen positiven Zusammenhang auf die Wahrscheinlichkeit in stationäre Pflege überzutreten und verdeutlicht damit die hohe Bedeutung der jeweils damit verbundenen Grunderkrankungen. Wichtig ist hinzuzufügen, dass diese Medikamente bzw. Heilmittel nicht per se in Verbindung mit der Wahrscheinlichkeit des Eintritts in stationäre Pflege stehen, sondern die dahinterliegenden gesundheitlichen Probleme.

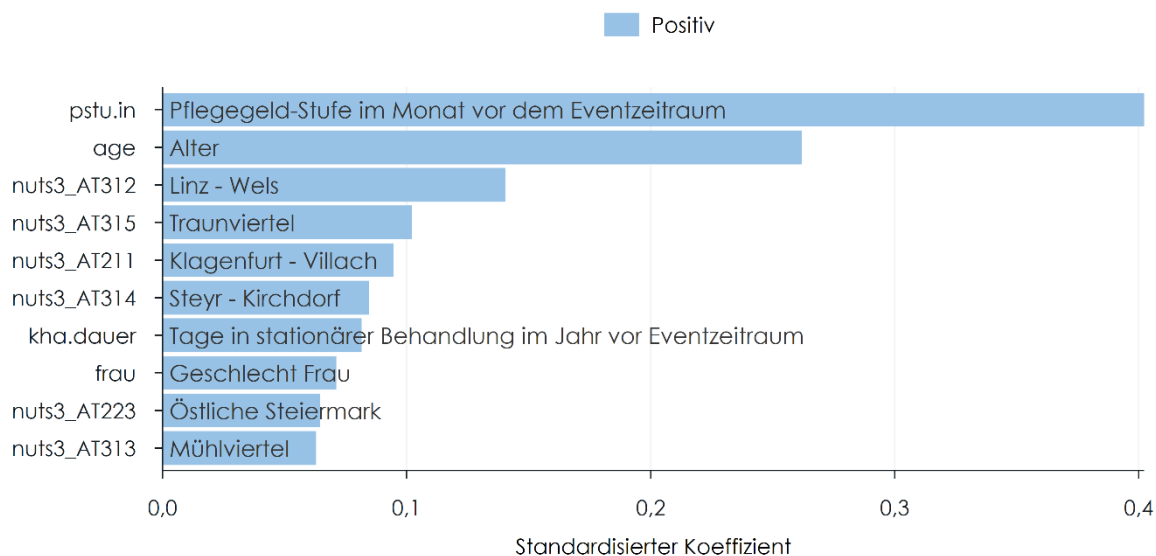
Abbildung 36: **Zusammenhänge von bezogenen Medikamenten/Heilmitteln auf Eintritte in stationäre Pflege**



Q: WIFO, DVSV.

Die zehn wesentlichsten sonstigen Merkmale, die mit einem Eintritt in stationäre Pflege in Verbindung stehen, zeigt Abbildung 37. Die Pflegegeldstufe vor Eintritt steht positiv in Verbindung mit dem Eintritt in ein Pflegeheim: Je höher die Pflegegeldstufe, desto wahrscheinlicher wird der Eintritt in ein Pflegeheim. Ebenso einen positiven Effekt hat das Alter der Übertretenden. Darüber hinaus gibt es – in schwächerem Ausmaß – positive Effekte von manchen NUTS3-Regionen: Linz-Wels, Traunviertel, Klagenfurt-Villach, Steyr-Kirchdorf, die östliche Steiermark und das Mühlviertel. In all diesen Regionen zeigt sich ein positiver Effekt auf einen Eintritt in stationäre Pflege. Ebenso positive Effekte zeigen sich bei der Anzahl der Nächte in stationärer Spitalsbehandlung: eine höhere Anzahl der Nächte stationärer Aufenthalte im Krankenhaus ist mit einer höheren Wahrscheinlichkeit in stationäre Pflege überzutreten, verbunden.

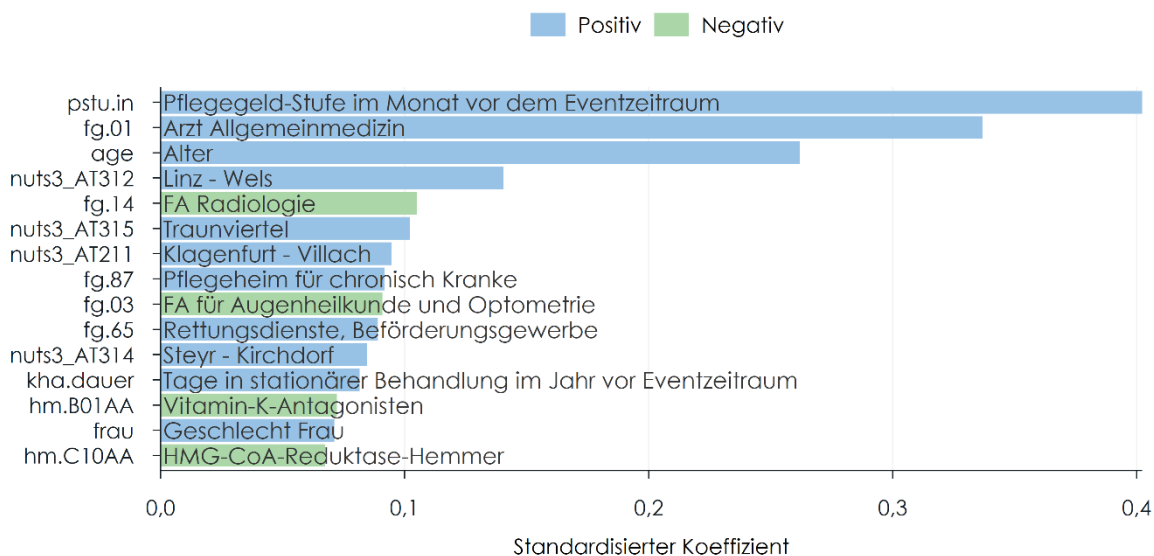
Abbildung 37: **Effekte von sonstigen persönlichen Charakteristika auf Eintritte in stationäre Pflege**



Q: WIFO, DVSV.

Die Ergebnisse aller Variablenkategorien zeigt Abbildung 38, die die 15 wesentlichsten Variablen für den Eintritt in stationäre Pflege darstellt. Den stärksten positiven Effekt auf Eintritte in stationäre Pflege hat die Pflegegeldstufe im Monat vor dem Eventzeitraum. Personen in höheren Pflegegeldstufen haben eine höhere Wahrscheinlichkeit in stationäre Pflege überzutreten. Ebenso, jedoch in einem geringeren Ausmaß, Personen, die im Beobachtungszeitraum eine Ärztin bzw. einen Arzt für Allgemeinmedizin konsultiert haben. Auch Personen in einem höheren Alter haben eine größere Wahrscheinlichkeit, in stationäre Pflege überzutreten. Weiters zeigen sich positive Effekte in gewissen NUTS3-Regionen (Linz-Wels, Traunviertel, Klagenfurt-Villach und Steyr-Kirchdorf), durch den Aufenthalt in Pflegeheimen für chronisch Kranke, die Inanspruchnahme von Rettungsdiensten und durch die Tage der stationären Behandlung in Krankenhäusern. Frauen haben – im Gegensatz zu Männern – eine höhere Wahrscheinlichkeit in stationäre Pflege überzutreten. Negative Effekte zeigen die Inanspruchnahme von Gesundheitsleistungen von Radiologinnen und Radiologen, Augenärztinnen und -ärzte, gerinnungshemmende Medikamente (Vitamin-K-Antagonisten) und Cholesterinsenker (HMG-CoA-Reduktase).

Abbildung 38: **Die 15 wichtigsten Merkmale, die einen Effekt auf Übertritte in stationäre Pflege haben**



Q: WIFO, DVSV.

4.4 Main Findings und Fazit

Das Modul 3 hat durch die Zusammenführung der Gesundheitsleistungsdaten mit Daten der Pflegegeldinformation die Zusammenhänge zwischen der gesundheitlichen Entwicklung und den Übertritten in stationäre Pflege von pflegegeldbeziehenden Personen ab 60 Jahren analysiert. Als Beobachtungszeitraum wurde das Jahr 2016 und als Eventzeitraum das Jahr 2017 definiert. Im Jahr 2017 sind 7,7% der 377.414 erfassten Pflegegeldbeziehenden ab 60 Jahren in stationäre Pflege übergetreten. Die in Modul 3 diskutierten Ergebnisse zeigen die Wahrscheinlichkeit des Eintritts in stationäre Pflege unter Berücksichtigung persönlicher Eigenschaften, den im Beobachtungszeitraum in Anspruch genommenen Gesundheitsleistungen und gestellten Diagnosen sowie der Pflegegeldstufe am Ende des Beobachtungszeitraums (Dezember 2016). Die deskriptiven Ergebnisse zeigen, dass Frauen, die in stationäre Pflege übertreten, tendenziell in höheren Pflegegeldstufen sind und dass vermehrt ältere Personen, jenseits der 80 Jahre, und Personen in höheren Pflegegeldstufen in stationäre Pflege wechseln.

Die Analyse der Modellwahl hat für das LASSO-Regressionsmodell die größte Erklärungskraft ergeben, weshalb die Ergebnisse dieses Modells ausgewiesen wurden. Die LASSO-Regressionen zeigen, dass – unter Berücksichtigung aller vorgestellten Merkmale – die stärkste Erklärungskraft für die Übertritte in stationäre Pflege im Jahr 2017 die Pflegegeldstufe im Dezember 2016 hat. Dabei haben höhere Pflegegeldstufen eine höhere Wahrscheinlichkeit des Eintritts. Ebenso starke Erklärungskraft haben die ärztlichen Leistungen von Allgemeinmedizinerinnen und -mediziner – in diesem Zusammenhang wohl ein Indiz für einen schlechten Gesundheitszustand – sowie ein höheres Alter. Die Analyse der Diagnosen hebt die Relevanz von

Unfällen im privaten Bereich (z.B. Stürze), Demenz und psychische Störungen als Ursache für den Wechsel in stationäre Pflege hervor. Die wesentlichste medizinische Einzelleistung, die mit dem Eintritt in stationäre Pflege in Zusammenhang steht, ist eine Computertomographie des Kopf- und Halsbereichs. PatientInnen mit Nierenersatztherapien und Kathetereingriffen der Herzkranzgefäße haben, *ceteris paribus*, eine geringere Wahrscheinlichkeit in stationäre Pflege überzutreten. Es zeigt sich, dass manche kurative Eingriffe (auch z.B. künstliche Hüftgelenke oder Kniegelenke) die Wahrscheinlichkeit eines Wechsels in stationäre Pflege reduzieren. Eine mögliche Erklärung wäre, dass es sich dabei zum Teil um Eingriffe handelt, die nicht bei Personen mit labilem Gesundheitszustand durchgeführt werden können. Ebenso einen negativen Effekt auf die Wahrscheinlichkeit des Eintritts in stationäre Pflege haben, *ceteris paribus*, gerinnungshemmende Medikamente und Cholesterinsenker. Die Analyse der persönlichen Merkmale zeigt – neben der Pflegegeldstufe und dem Alter – positive Zusammenhänge zwischen dem Eintritt in stationäre Pflege und manchen Regionen in Österreich, wobei Regionen in Oberösterreich überproportional betroffen sind.

5. Modul 4: Häufigkeit von Aufenthaltstagen im Krankenhaus von Pflegegeldbeziehenden

5.1 Ziel des Moduls 4

Krankenhausaufenthalte zeigen sich in den bisherigen Analysen zum Eintritt in den Pflegegeldbezug (Modul 1) wie auch beim Übergang in eine höhere Pflegegeldstufe (Modul 2) als zentrale Merkmale. Im vorliegenden Modul stehen die Krankenhausaufenthalte der Pflegegeldbeziehenden im Vordergrund sowie die Frage, welche gesundheitlichen Faktoren zu einem vermehrten Krankenhausaufenthalt führen. Auf der Grundlage der Haupt- und Zusatzdiagnosen aus dem Krankenanstaltenbereich (Internationale statistische Klassifikation der Krankheiten und Gesundheitsprobleme (ICD-10)) und medizinischer Einzelleistungen sind jene Gesundheitsdienstleistungen eruiert, die besonders häufig im Zusammenhang mit den Krankenhausaufenthalten von pflegegeldbeziehenden Personen auftreten.

Die Quantifizierung der wesentlichsten Gesundheits(dienst)leistungen jener Pflegegeldbeziehenden, die auch Krankenhausaufenthalte haben, ermöglicht die Ableitung von Präventionsfeldern. Mit anderen Worten weisen die ermittelten wichtigsten Gesundheitsdiagnosen, medizinischen Einzelleistungen und Heilmittel darauf hin, dass eine intensive Betreuung der betroffenen Pflegegeldbeziehenden die Krankenhausaufenthalte verringern könnten.

5.2 Methodische Vorgangsweise: Empirisches Modell

Zur Untersuchung der Häufigkeit von Krankenhausaufenthalten wird als abhängige Variable die Gesamtdauer stationärer Aufenthalte in Tagen im Eventzeitraum (1/2017 bis 1/2018) herangezogen. Diese Definition wurde einer Messung durch die reine Anzahl von Aufenthalten innerhalb des Eventzeitraums vorgezogen, da es bei Verwendung letzterer zu einer Verzerrung kommen kann: So würden etwa bei Verwendung der reinen Häufigkeit mehrere kurze Aufenthalte eine höhere Intensität suggerieren als ein langer Krankenhausaufenthalt. Als stationärer Aufenthalt gelten dabei Krankenhauskontakte, deren Entlassungsdatum nicht dem Aufnahmedatum entspricht³²⁾. Liegen innerhalb des Eventzeitraums mehrere dieser Aufenthalte, so werden diese addiert. Im Durchschnitt verbrachten Pflegegeldbeziehende im Eventzeitraum 8,1 Nächte in stationärer Behandlung, wenngleich die Verteilung deutlich linkssteil ist.

Da es sich bei der Gesamtdauer stationärer Aufenthalte um ein diskretes Merkmal handelt, das entweder den Wert null oder eine ganzzahlig positive Zahl annimmt, werden deren Einflussfaktoren – im Unterschied zu den vorhergehenden Modellen – nicht mittels eines binären Klassifikationsmodells, sondern mit Hilfe einer Poisson-Regression geschätzt. Hierbei handelt es sich um ein Regressionsmodell für Zähldaten, welche entweder null oder ganzzahlig positiv sein können. Der Zusammenhang zwischen der Dauer stationärer Aufenthalte $S_{i,[t,t+b]}$ im

³²⁾ Von den 2.832.282 Krankenhausaufenthalten der Pflegegeldbeziehenden enden immerhin 22,1% am selben Tag (vgl. Abschnitt 2.2).

Eventzeitraum als abhängiger Variable, den persönlichen Eigenschaften $X_{i,t}$ und den im Beobachtungszeitraum $[t - a, t)$ konsumierten Gesundheitsleistungen $G_{i,[t-a,t)}$ wird dabei folgendermaßen modelliert (siehe Long, 2008, S. 221):

$$S_{i,[t,t+b)} = \exp(X_{i,t}\beta + G_{i,[t-a,t)}\gamma)$$

Die Parameter des Modells β und γ werden wiederum mit der Maximum-Likelihood-Methode aus den Trainingsdaten (75% des Gesamtdatenbestands) geschätzt, in der folgende Log-Likelihood-Funktion maximiert wird:

$$LL(\beta, \gamma) = \frac{1}{N} \sum_{i=1}^N (S_{i,[t,t+b)}(X_{i,t}\beta + G_{i,[t-a,t)}\gamma) - \exp(X_{i,t}\beta + G_{i,[t-a,t)}\gamma) - \ln(S_{i,[t,t+b)}!))$$

Um ein "Overfitting" der Parameter an die Trainingsdaten zu verhindern, kann die zu maximierende Log-Likelihood – analog zu den logistischen Modellen in den Modulen 1 bis 3 – wiederum um einen Pönalterm ergänzt werden. Je nachdem, ob dieser Pönalterm die Summe der quadrierten Parameter β und γ oder die Summe der Absolutwerte dieser Parameter beinhaltet, ergibt sich eine Ridge-Regression oder eine LASSO-Schätzung (vgl. detaillierte Beschreibung zu Modul 1) des Poisson-Modells. Zum Vergleich der Modelle wird anstelle der Fläche unter der ROC-Kurve nun allerdings auf den "durchschnittlichen absoluten Fehler" (MAE: "mean absolute error") in den Testdaten (25% des Gesamtdatenbestands) zurückgegriffen. Der MAE misst, wie groß die absolute Differenz zwischen der (aus dem geschätzten Modell) prognostizierten Häufigkeit von Krankenhausaufenthalten $\hat{S}_{i,[t,t+b)}$ und der tatsächlich beobachteten Häufigkeit $S_{i,[t,t+b)}$ im Durchschnitt ist:

$$MAE = \frac{1}{N} \sum_{i=1}^N |S_{i,[t,t+b)} - \hat{S}_{i,[t,t+b)}|$$

Modelle mit kleinerem MAE sind dabei zu bevorzugen.

Während ein Standard-Poisson-Modell in den Trainingsdaten einen MAE von 5,83 erreicht (d. h. im Durchschnitt weicht die beobachtete Häufigkeit um 5,83 Tage von der prognostizierten Häufigkeit ab), beträgt der durchschnittliche absolute Fehler in den Testdaten 90,57. Dies könnte auf eine deutliche Überanpassung des Modells an die Trainingsdaten hinweisen. Die Verwendung der LASSO-Regularisierung führt zu einer relativ starken Reduktion der Anzahl an Parametern: Nur 63 Variablen verbleiben nach optimaler Wahl der Regularisierungsstärke mittels fünffacher Kreuzvalidierung aktiv im Modell. Von den Gruppen der verwendeten Gesundheitsvariablen (ärztliche Leistungen, Heilmittel, medizinische Einzelleistungen) fallen die ärztlichen Leistungen, die Heilmittel und sonstige Charakteristika gänzlich weg bzw. werden im Modell auf null geschrumpft. Die geringere Anzahl der verbleibenden Variablen führt zu

einem deutlich geringeren Schätzfehler: Die durchschnittliche absolute Abweichung der prognostizierten von den beobachteten Werten beträgt sowohl in den Trainings- als auch in den Testdaten nur mehr 6,80. Auch die Ridge-Regression weist einen geringeren Fehler in den Testdaten auf als das Poisson-Modell ohne Regularisierung (MAE: 8,27 in den Testdaten, 6,47 in den Trainingsdaten). Im Folgenden wird daher auf die Ergebnisse des Poisson-Modells mit LASSO-Regularisierung zurückgegriffen³³).

5.3 Krankenhaushäufigkeit und Aufenthaltsdauer von Pflegegeldbeziehenden

Eine zentrale Größe für die Nutzung des Gesundheitssystems ist die Zahl der Krankenhausaufenthalte; hier kann zwischen Nächtingungen und Tagesaufenthalten unterschieden werden. Erste Einblicke gibt hier die österreichische Gesundheitsbefragung (ATHIS) aus dem Jahr 2019 mit einer Stichprobengröße von knapp 16.000 über-15-jährigen Personen. Mit steigendem Lebensalter steigt die Häufigkeit wie auch die Dauer von Krankenhausaufenthalten kontinuierlich an. Ein Aufenthalt der jungen Bevölkerung (20- bis 29-Jährige) dauerte gemäß Österreichischer Gesundheitsbefragung 2019 im Schnitt 4,6 Tage bei Männern und 5,4 Tage bei Frauen; in der Gruppe der über-75-Jährigen waren es im Schnitt 13 (Männer) bzw. 14,8 (Frauen) Tage (vgl. Übersicht 9). Unabhängig vom Pflegegeldbezug zeigt sich mit steigendem Lebensalter sowohl eine Zunahme des Bevölkerungsanteils mit Krankenhausaufenthalten wie auch eine Zunahme der Dauer der jeweiligen Aufenthalte (Galama – Van Kippersluis, 2018, BMASK, 2018).

Übersicht 9: **Durchschnittliche Inanspruchnahme der Gesundheitsdienstleistung Krankenhausaufenthalt nach Alter und Geschlecht, 2019**

Alter in Jahren	Anteil In % der Bevölkerung	Männer		Frauen		
		Dauer In Tagen	Tagespatient In % der Bevölkerung	Anteil In % der Bevölkerung	Dauer In Tagen	Tagespatientin In % der Bevölkerung
15 bis 29	7,1	4,6	16,8	13,2	5,4	14,1
30 bis 44	9,4	7,2	15,0	13,7	8,2	14,6
45 bis 59	15,0	9,2	16,1	15,0	8,6	19,1
60 bis 74	22,9	10,6	18,6	20,5	9,5	15,5
75+	34,5	13,0	24,8	31,7	14,8	24,5

Q: Österreichische Gesundheitsbefragung 2019 (ATHIS), Krankenhausnächtingungen bzw. -aufenthalte in den letzten 12 Monaten.

³³) Da der Poisson-Verteilung eine Äquidispersionsannahme (Mittelwert ist gleich Varianz) zugrunde liegt, welche in den gegebenen Daten möglicherweise verletzt ist, wurde alternativ auch ein Negativ-Binomial-Modell mit und ohne Regularisierung geschätzt. Dieses wies jedoch in den Testdaten mit einem minimalen MAE von 7,85 (bei LASSO-Regularisierung) einen höheren Fehler auf als das Poisson-Modell.

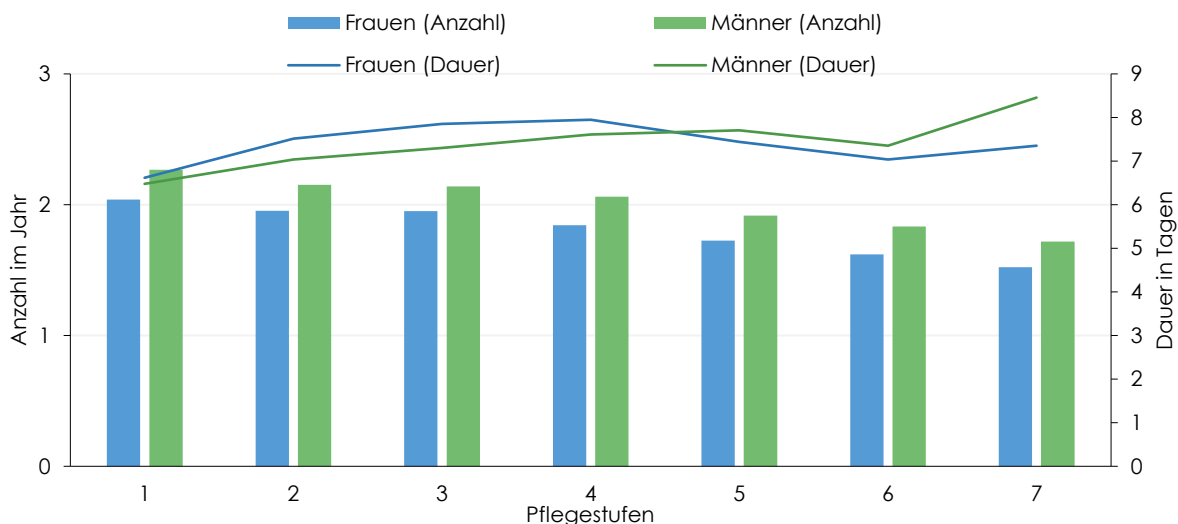
Die Anteile aller Pflegegeldbeziehenden mit Krankenhausaufenthalten im Alter von 60 Jahren und älter im Zeitraum 1/2017 bis 1/2018 zeigt eine sinkende Anzahl von Aufenthalten mit zunehmendem Alter (vgl. Abbildung 39): Während die Altersgruppe der 60- bis 64-Jährigen im Schnitt 2,6 (Frauen) bzw. 2,5 (Männer) Aufenthalte hatte, sind es bei den 80- bis 84-Jährigen 1,9 (Frauen) und 2,0 (Männer) Aufenthalte. Die sinkende Zahl an Aufenthalten könnte im Zusammenhang mit einer verstärkten Betreuung daheim, mit einer längeren Krankenhausaufenthalte von Pflegegeldbeziehenden oder auch mit einer Übersterblichkeit der Personen mit schlechter Gesundheit stehen (Klotz 2018, Psihoda, 2019) stehen.

Ein unerwarteter Zusammenhang zeigt sich zwischen den Pflegegeldstufen und Krankenhausaufenthalten: Einerseits steigen die im Krankenhaus verbrachten Tage pro Aufenthalt in höheren Pflegegeldstufen an, andererseits sinkt die Anzahl mit der Höhe der Pflegegeldstufen. Während in der Stufe 1 die Zahl an Krankenhausaufenthalten der Frauen und Männer bei zwei- bzw. 2,3-mal jährlich liegt, sinkt sie in Stufe 7 auf 1,5 bzw. 1,7 ab. Hier gibt es einen weiterführenden Forschungsbedarf bei den Pflegegeldbeziehenden zum Zusammenhang ihrer Krankenhausaufenthalte und ihrer stationären Pflege (vgl. Modul 3) insgesamt.

Der Vergleich der Pflegegeldbeziehenden mit der altersgleichen Gesamtbevölkerung (Übersicht 9) bei der durchschnittlichen Krankenhausaufenthaltsdauer zeigt deutliche Unterschiede: Während 2018 die Aufenthaltsdauer der 60- bis 74-jährigen Männer in den letzten 12 Monaten in Summe bei 10,6 Tagen liegt, hatten pflegegeldbeziehende Männer dieser Altersgruppe im Schnitt 18,9 Krankenhaustage. Auf pflegegeldbeziehende Frauen dieser Altersgruppe entfallen mit 18,1 Tagen um 8,6 Aufenthaltstage mehr als auf alle Frauen dieser Altersgruppe. Bei den über-74-Jährigen beträgt die Differenz 3 (Männer) bzw. 0,5 (Frauen) Tage.

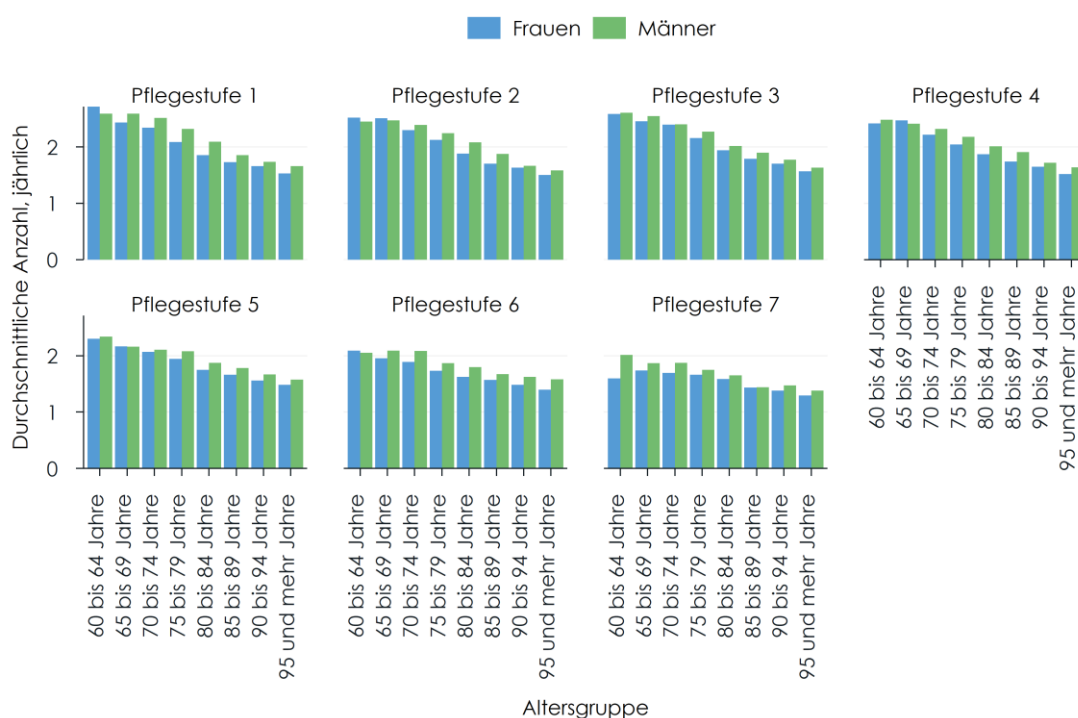
Die mit steigendem Lebensalter rückläufige Anzahl von Krankenhausaufenthalten zeigt sich über alle Pflegegeldstufen hinweg (Abbildung 40). Die zunehmende Dauer der Aufenthalte im Zusammenhang mit dem Lebensalter zeigt sich hingegen nur bis zur Pflegestufe 3. In höheren Stufen nimmt die Summe der Aufenthaltstage tendenziell wieder ab (Abbildung 41).

Abbildung 39: Anzahl (linke Achse) und Dauer (rechte Achse) von Krankenhausaufenthalten der Pflegegeldbeziehenden, 2018



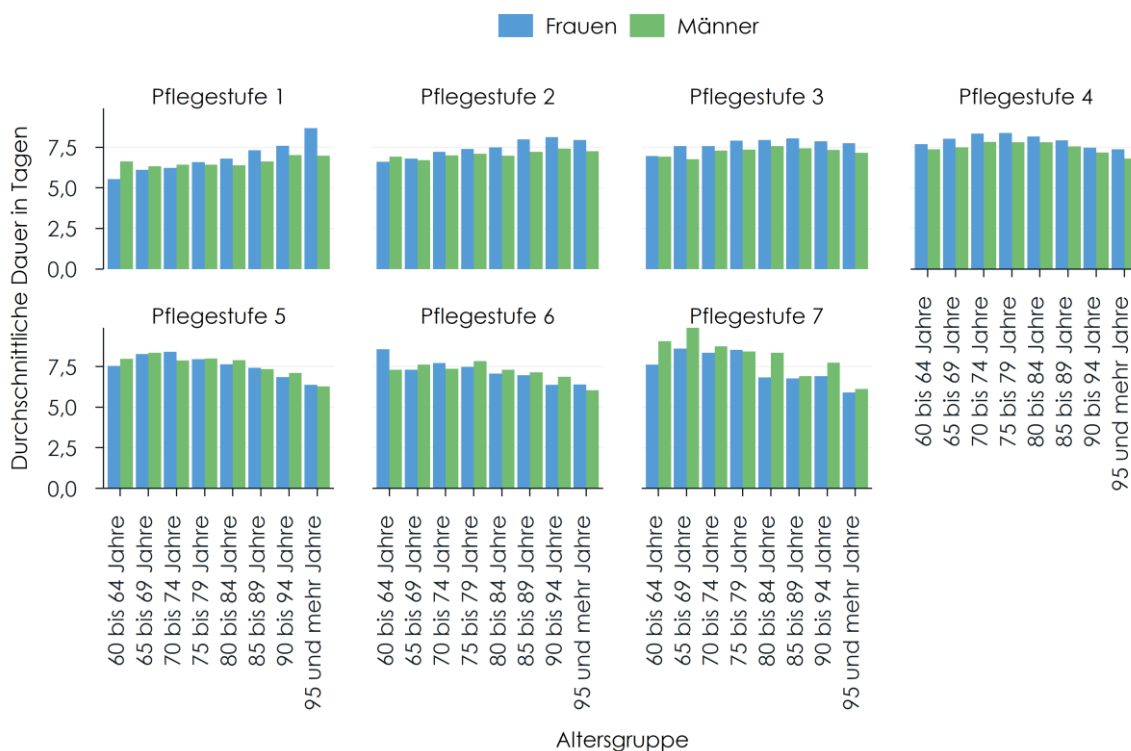
Q: DVSV, WIFO.

Abbildung 40: Anzahl von Krankenhausaufenthalten nach Pflegegeldstufen und Alter, 2018



Q: DVSV, WIFO.

Abbildung 41: **Dauer von Krankenhausaufenthalten nach Pflegegeldstufen und Alter, 2018**



Q: DSVS, WIFO.

5.4 Effekte der Diagnosen medizinischen Leistungen auf die Krankenhausaufenthaltsdauer

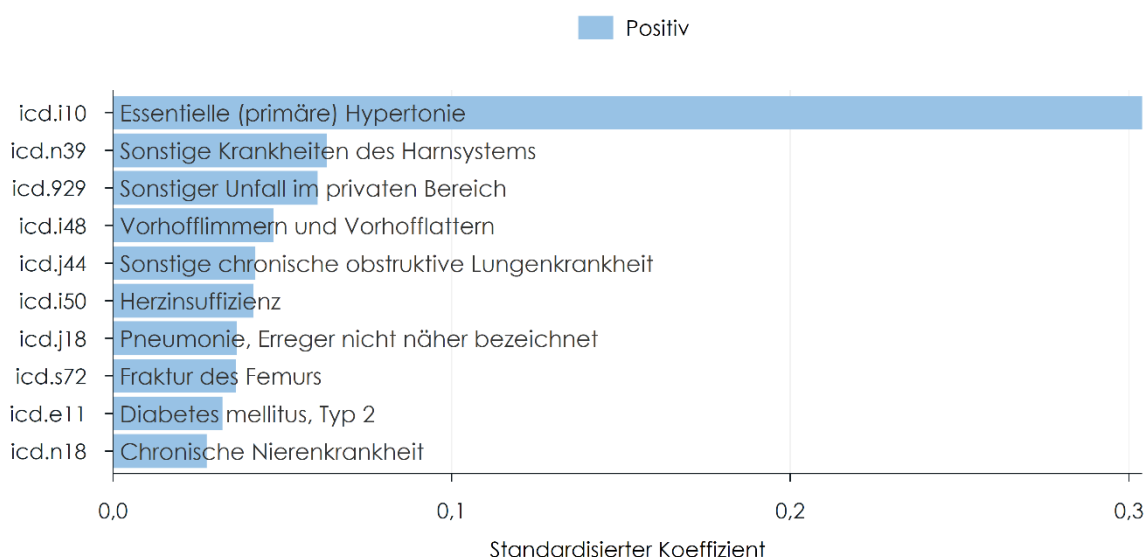
Bislang stand die Gesundheitsdienstleistung Krankenhausaufenthalt beim Eintritt in einen Pflegegeldbezug (Modul 2) und beim Übertritt in eine höhere Pflegegeldstufe (Modul 3) im Vordergrund. Nachfolgend geht es um die Frage, welche Diagnosen und Heilmittel die Pflegegeldbeziehenden im Beobachtungszeitraum 1/2016 bis 1/2017 hatten, wenn sie im Eventzeitraum 1/2017 bis 1/2018 einen Krankenhausaufenthalt verzeichneten.

Im Eventzeitraum 1/2017 bis 1/2018 waren von den 2.832.282 Krankenhausaufenthalten der Pflegegeldbeziehenden knapp 78% mit Nächtingungen verbunden. Im Schnitt verbrachten Pflegegeldbeziehende 8,1 Nächtinge pro Jahr im Krankenhaus, mit einer heterogenen Verteilung über die Pflegegeldstufen. Nachfolgend sind die Haupt- und Zusatzdiagnosen, die während der stationären Krankenhausaufenthalte erstellt wurden, auf Ebene der dreistelligen Systematik der Internationalen statistischen Klassifikation der Krankheiten und Gesundheitsprobleme (ICD-10, insgesamt 1.483 unterschiedliche Codes) näher beleuchtet.

Den mit Abstand stärksten Effekt im Zusammenhang mit stationären Krankenhausaufenthalten von Pflegegeldbeziehenden übt die essentielle primäre Hypertonie (ICD.i10), also Bluthochdruck, der nicht durch andere Erkrankungen verursacht ist (Abbildung 42), aus.

Nach den sonstigen Krankheiten des Harnsystems (ICD.39) stehen sonstige Unfälle im privaten Bereich (ICD.929) an dritter Stelle im Zusammenhang mit stationären Aufnahmen. Davon zu unterscheiden ist die Fraktur des Oberschenkelknochens (ICD.s72), die ebenfalls Aufenthalte erhöht. Vorhofflimmern und Vorhofflattern (ICD.i48), sonstige chronische obstruktive Lungenkrankheiten (ICD.j44), Herzinsuffizienz (ICD.i50), Pneumonie (ICD.j18) sowie Diabetes mellitus Typ 2 (ICD.e11) und chronische Nierenkrankheiten (ICD.n18) vervollständigen die zehn häufigsten Hauptdiagnosen, die bei Pflegegeldbeziehenden im Zusammenhang mit stationären Krankenhausaufenthalten auftreten. Die zehn häufigsten Diagnosen sind für 26% aller Krankenhaustage mitverantwortlich.

Abbildung 42: **Die zehn wichtigsten medizinischen Diagnosen von Pflegegeldbeziehenden mit Effekten auf ihre Krankenhausaufenthaltsdauer**



Q: DSVS, WIFO.

Die zehn wesentlichsten Diagnosen, die im Jahr vor den stationären Krankenhausaufenthalten des Eventzeitraums erfolgen, unterschieden sich grundsätzlich sowohl von jenen, die zu einer Höhereinstufen im Pflegegeld (vgl. Modul 2, Abschnitt 3) als auch zur Aufnahme in eine stationäre Pflegeeinrichtung (vgl. Modul 3, Abschnitt 4) führen, einzig sonstige Unfälle im privaten Bereich (ICD.929) sind in allen drei Bereichen ein gleichermaßen einflussreicher Faktor.

5.5 Medizinische Leistungen und Krankenhausaufenthalte

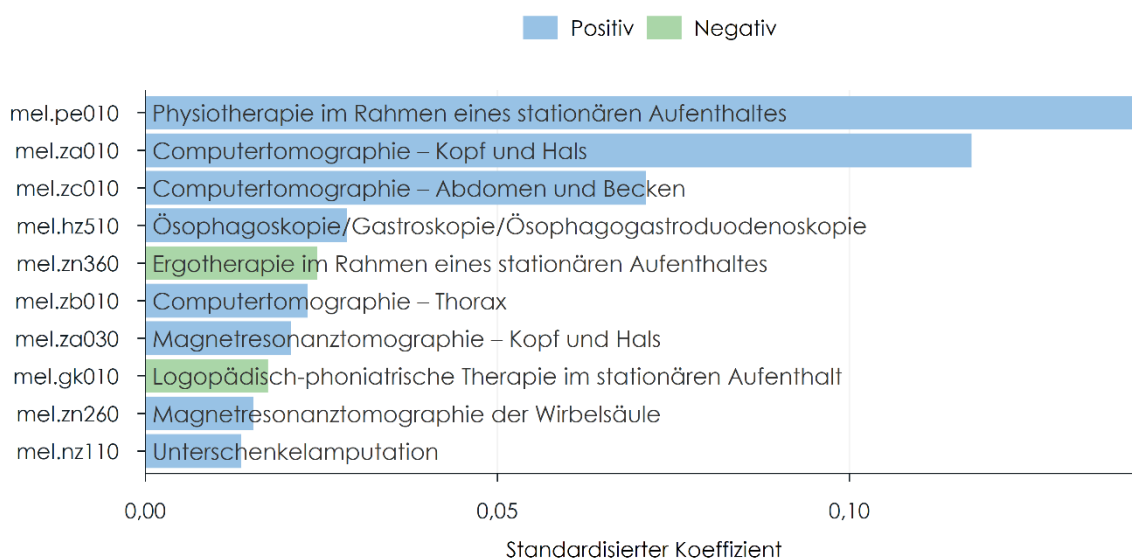
Das zugrundeliegende Berechnungsmodell der medizinischen Leistungen im Beobachtungszeitraum (1/2016 bis 1/2017) greift auf detaillierte Daten medizinischer Einzelleistungen in der fünfstelligen Systematik mit 1.553 Variablen zurück. Von diesen verbleiben wenige aktiv im Modell und die zehn wichtigsten finden sich in Abbildung 43. Ergotherapie (mel.zn360) und

logopädisch-phoniatrische Therapie (mel.gk010) bei Pflegegeldbeziehenden haben einen negativen, also verkürzenden Effekt auf die Krankenhausaufenthalte.

Den gegenteiligen Effekt zeigen hingegen Physiotherapie (mel.pe10), Computertomographie von Kopf und Hals (mel.za010) und Abdomen (mel.zc010), endoskopische Untersuchung von Speiseröhre/Magen/Zwölffingerdarm (mel.hz510), Computertomographien (mel.zb010) und Magnetresonanztomographie von Hals und Kopf (mel.za030) sowie der Wirbelsäule (mel.zn260). Diese medizinischen Einzelleistungen im Beobachtungszeitraum gehen Hand in Hand mit einem längeren stationären Aufenthalt der Pflegegeldbeziehenden im Eventzeitraum.

Ein hoher Effekt der medizinischen Einzelleistung Computertomographien von Kopf und Hals zeigt sich sowohl im Zusammenhang mit stationären Krankenhausaufenthalten als auch beim Übertritt in eine höhere Pflegegeldstufe (Modul 2, Abschnitt 3) und beim Übertritt in eine stationäre Pflege (Modul 3, Abschnitt 4). Alle anderen der zehn häufigsten medizinischen Einzelleistungen im Zusammenhang mit stationären Krankenhausaufenthalten von Pflegegeldbeziehenden treten bei Pflegegeldstufenübertritten bzw. Übertritten in stationäre Pflege nicht auf.

Abbildung 43: **Die zehn wichtigsten medizinischen Einzelleistungen von Pflegegeldbeziehenden mit Effekten auf ihre Krankenhausaufenthalte**



Q: DSVS, WIFO.

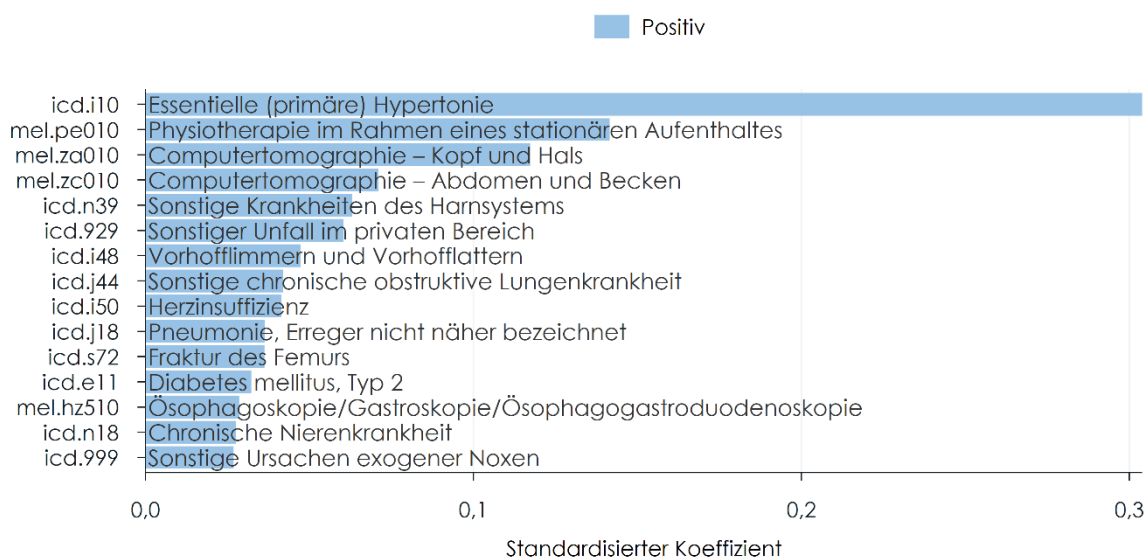
Die integrierte Betrachtung der Krankheitsdiagnosen und medizinischen, Einzelleistungen, die die größten Effekte auf die Krankenhausaufenthalte von Pflegegeldbeziehenden haben, zeigt folgendes Bild: Von den 15 Merkmalen wirken elf Krankheitsdiagnosen und vier medizinische Einzelleistungen auf die Wahrscheinlichkeiten der Krankenhausaufenthalte im Eventzeitraum (vgl.

Abbildung 38). Neben den bereits in Abbildung 42 enthaltenen und diskutierten zehn wichtigsten ICD-Diagnosen stellt ICD-999, sonstige exogene Noxen (gefährdende Substanzen), ein weiteres wichtiges Charakteristikum dar, das die Wahrscheinlichkeit der Dauer von stationären Aufenthalte erhöht. Von den medizinischen Einzelleistungen verbleiben vier Charakteristika (Physiotherapie, CT von Kopf/Hals bzw. Abdomen/Becken und endoskopische Untersuchung von Speiseröhre/Magen/Zwölffingerdarm) unter den Hauptmerkmalen.

Pflegegeldbeziehende mit Diagnosen der essentiellen (primäre) Hypertonie oder auch Physiotherapie im Rahmen eines stationären Aufenthaltes zeigen im Folgejahr längere Krankenhausaufenthalte. Bei den medizinischen Einzelleistungen zeigt sich ein Zusammenhang zwischen Computertomographien von Kopf und Hals bzw. Abdomen und Becken und der Aufenthaltsdauer im Folgejahr.

Die abschließende Frage nach den potenziellen Präventionsfeldern, die die Dauer der Krankenhausaufenthalte zu verkürzen imstande sind, sollten in erster Linie bei den berechneten Haupteinflussfaktoren der Gesundheitsdiagnosen entwickelt werden. Da medizinische Einzelleistungen Antworten auf sehr konkrete Gesundheitslagen sind, würden Präventivmaßnahmen in den elf berechneten Gesundheitsdiagnosen, allen voran die essentielle Hypertonie, am ehesten Optionen enthalten, die die stationäre Krankenhausaufenthaltsdauern zu verkürzen vermögen.

Abbildung 44: **Analyse der Krankenhausaufenthaltsdauern: Die 15 wichtigsten Merkmale**



Q: DVS, WIFO.

5.6 Main Findings – Kurzzusammenfassung

Das Modul 4 zeigt in der Zusammenführung der Gesundheitsleistungsdaten mit Daten der Pflegegeldinformation die Effekte unterschiedlicher medizinischer Leistungen auf die Dauer

stationärer Krankenhausaufenthalte von Pflegegeldbeziehenden. Es wird der Frage nachgegangen, wie die Summe der Krankenhaustage im Jahr 2017 von welchen der konsumierten medizinischen Leistungen des Jahres 2016 beeinflusst ist.

Im angewendeten LASSO-Regressionsmodell verblieben die Gesundheitsdiagnosen und die medizinischen Einzelleistungen als erklärende Charakteristika, währende ärztliche Leistungen, Heilmittel wie auch sonstige Charakteristika auf null schrumpften, also wegfallen.

Im Durchschnitt verbrachten Pflegegeldbeziehende im Jahr 2017 8,1 Nächte in stationärer Behandlung. Diese durchschnittliche Dauer eines Krankenhausaufenthaltes liegt bei Pflegegeldbeziehenden deutlich höher als in der altersvergleichbaren Gesamtbevölkerung.

Die Pflegegeldbeziehenden haben mit höheren Pflegegeldstufen sinkende Krankenhausaufenthalte: Während in der Stufe 1 die Zahl an Krankenhausaufenthalten der Frauen und Männer bei zwei- bzw. 2,3-mal jährlich liegt, sinkt sie in Stufe 7 auf 1,5- bzw. 1,7-mal jährlich. Pflegegeldbeziehende Männer haben über alle Pflegegeldstufen hinweg eine höhere Anzahl an stationären Krankenhausaufenthalten im Vergleich zu pflegegeldbeziehenden Frauen. Allerdings erhöht sich mit steigendem Lebensalter der Pflegegeldbeziehenden, sowohl bei Männern als auch bei Frauen, die Aufenthaltsdauer bis zur Pflegegeldstufe 3 deutlich. In den anschließenden höheren Stufen nimmt die Länge der Krankenhausaufenthalte tendenziell wieder ab.

Gemäß den Modellberechnungen haben Gesundheitsdiagnosen eine größere Erklärungskraft für die Krankenhausaufenthaltsdauern im Vergleich zu den medizinischen Einzelleistungen. Von den zehn wichtigsten Diagnosen im Beobachtungszeitraum ist die essentielle (primäre) Hypertonie die bedeutendste. Die Effekte auf die Dauer der stationären Krankenhausaufenthalte im Eventzeitraum sind auch bei den sonstigen Krankheiten des Harnsystems, sonstigen Unfällen im privaten Bereich, Vorhofflimmern, sonstigen chronischen obstruktiven Lungenerkrankungen, Herzinsuffizienz, Pneumonie, Fraktur des Femurs, Diabetes mellitus Typ 2 und chronischen Nierenerkrankungen gegeben. Allerdings wirken sie deutlich schwächer auf die Krankenhaustage als die Diagnose Hypertonie.

Pflegegeldbeziehende mit den medizinischen Einzelleistungen Physiotherapie im Rahmen stationärer Aufenthalte, Computertomographie von Kopf/Hals und Abdomen/Becken sowie Speiseröhren- und Magenspiegelungen haben im Folgejahr längere Krankenhausaufenthalte im Vergleich zu jenen, die diese Leistungen nicht bezogen.

Die zehn Diagnosen mit den größten Effekten auf die stationären Krankenhausaufenthaltsdauern der Pflegegeldbeziehenden im Zeitraum 1/2017 bis 12/2018 zeigen wenig Übereinstimmung mit jenen Diagnosen, die für den Übertritt in eine höhere Pflegegeldstufe (Modul 2, Abschnitt 3) oder den Übertritt in stationäre Pflege (Modul 3, Abschnitt 4) hauptbestimmend sind. Einzig die essentielle primäre Hypertonie findet sich in allen Bereichen als eines der maßgeblichsten Erklärungsmerkmale wieder.

Die Entwicklung und Intensivierung von Präventivmaßnahmen würde vor allem bei der Diagnose Hypertonie den größten Effekt auf die Reduktion der Krankenhaustage liefern können.

Allerdings sind gerade hier die Ursachen und Auslöser des Krankheitsbildes sehr vielfältig, weshalb es hier sehr vielschichtige Konzepte bräuchte.

6. Zusammenfassung und Schlussfolgerungen

Die demographische Entwicklung der kommenden Jahrzehnte lässt eine stark wachsende Zahl an alten und hochaltrigen Menschen erwarten. Damit wird auch die Zahl an pflegebedürftigen Menschen steigen. Mit der steigenden Nachfrage sind – ceteris paribus – höhere Kosten für die öffentliche Hand verbunden (*Famira-Mühlberger, 2020; Klien et al., 2020*). Die Perspektive einer zunehmenden Bedeutung des Pflegesystems für den öffentlichen Haushalt hat den Auftraggeber dazu veranlasst, durch das WIFO überprüfen zu lassen, welche Zusammenhänge zwischen den abgerufenen Gesundheitsleistungen und

- dem Eintritt in das Pflegegeldsystem (Modul 1),
- dem Übertritt in eine höhere Pflegegeldstufe (Modul 2),
- dem Eintritt in die stationäre Pflege (Modul 3), sowie
- den gesundheitlichen Faktoren für die Krankenhausaufenthalte von pflegegeldbeziehenden Personen (Modul 4)

bestehen. Die Ziele der Studie sind:

- die Überprüfung der Möglichkeit einer systematischen Vorhersehbarkeit von Pflegegeldbezug ("Vorlaufindikator") (Modul 1),
- die Analyse der wichtigsten Determinanten für den Übertritt in eine höhere Pflegegeldstufe (Modul 2),
- die Analyse der wichtigsten Determinanten für den Eintritt in die stationäre Pflege (Modul 3), sowie
- die Analyse der wichtigsten Determinanten für die Krankenhausaufenthalte von pflegegeldbeziehenden Personen (Modul 4).

Die Verbindung der abgerufenen Gesundheitsleistungen mit dem Pflegegeldbezug für die Altersgruppe 60 bis 85 Jahre zeigt, dass der erstmalige Pflegegeldbezug relativ gut aus den Daten prognostiziert werden kann. Die geschätzten empirischen Modelle weisen eine hohe Prognosefähigkeit auf. Dabei zeigt sich, dass vor allem rezente Gesundheitsleistungen für den erstmaligen Pflegegeldbezug innerhalb eines nachgelagerten Eventzeitraums von Bedeutung sind: Wird der Monat unmittelbar vor dem Eventzeitraum nicht in die Betrachtung miteinbezogen, zeigt sich eine deutliche Verschlechterung der Modellgüte. In dieselbe Richtung weist auch der Umstand, dass Modelle, welche längere Eventzeiträume und/oder längere Beobachtungszeiträume nutzen, eine geringere Prognosefähigkeit aufweisen. Dennoch muss festgehalten werden, dass die Modelle selbst dann, wenn der unmittelbare Monat vor Beginn des Eventzeitraums ausgespart wird und die Beobachtungs- und Eventzeiträume auf ein ganzes Jahr verlängert werden, noch akzeptabel zwischen Personen mit und ohne erstmaligen Pflegegeldbezug im Eventzeitraum unterscheiden können.

Ein höherer Detailgrad in der Messung der Gesundheitsleistungen führt hingegen nicht zwangsweise zu einer Verbesserung der Modellgüte: Der Zuwachs an Prognosefähigkeit durch eine tiefere Aufschlüsselung der abgegebenen Heilmittel, Diagnosen oder medizinischen Einzelleistungen erhöht die Modellgüte zwar, jedoch nur marginal.

Zu den am stärksten mit dem erstmaligen Pflegegeldbezug korrelierten Faktoren in der kurzen Frist zählen laut den hier verwendeten multivariaten Modellen das Alter, stationäre Aufenthalte sowie Kontakte zu Allgemeinmedizinerinnen und -medizinern, aber auch Abgaben von Heilmitteln der ATC-Gruppe "N" (Nervensystem). Auch durchgeführte Computertomographien oder Diagnosen wie "Hirnfarkt", Unfälle im privaten Bereich oder "essentielle Hypertonie" weisen eine deutliche Präzedenz für den erstmaligen Pflegegeldbezug in der kurzen Frist auf.

Betreffend die Möglichkeit, aus den Daten einen kurzfristigen Vorlaufindikator für den zukünftigen Pflegegeldbezug auf individueller Ebene zu entwickeln, muss trotz der hohen Prognosefähigkeit bedacht werden, dass die untersuchten Ereignisse in den vorhandenen Daten und der zugrundeliegenden Grundgesamtheit nicht ausgewogen verteilt sind. Daraus ergibt sich das Problem einer zu hohen Falsch-Positiv-Klassifizierung, d. h. das Modell würde zu viele Personen fälschlicherweise der Gruppe der erstmaligen Pflegegeldbezieherinnen und -bezieher zuteilen. Eine Anwendung als Vorlaufindikator auf individueller Ebene scheint damit trotz der bereits guten Prognoseleistung der hier verwendeten Methoden weitere Forschung vorauszusetzen, welche alternative fortgeschrittene Klassifikationsmethoden evaluiert.

Modul 2 analysiert die statistischen Zusammenhänge zwischen den abgerufenen Gesundheitsleistungen (im Beobachtungszeitraum 2016) und den Übertritten in höhere Pflegegeldstufen von pflegegeldbeziehenden Personen ab 60 Jahren (im Eventzeitraum 2017). Im Jahr 2017 wechselten 18,2% der pflegegeldbeziehenden Personen ab 60 Jahren in eine höhere Pflegegeldstufe.

Die multivariate Analyse auf Basis von LASSO-Regressionen zeigt die Wahrscheinlichkeit des Übertritts in eine höhere Pflegegeldstufe unter Berücksichtigung persönlicher Eigenschaften, den im Beobachtungszeitraum abgerufenen Gesundheitsleistungen und gestellten Diagnosen sowie der Pflegegeldstufe am Ende des Beobachtungszeitraums (Dezember 2016). Die LASSO-Regressionen zeigen, dass die Pflegegeldstufe im Dezember 2016 die stärkste Korrelation mit dem Übergang in eine höhere Pflegegeldstufe im Jahr 2017 hat: Eine bereits höhere Pflegegeldeinstufung hat – ceteris paribus – einen negativen statistischen Effekt auf die Wahrscheinlichkeit einer Höhereinstufung der Pflegegeldstufe. Mögliche Erklärungen hierfür sind die Anspruchsvoraussetzungen der Pflegegeldstufen 5 bis 7, die sich nicht im Ausmaß des Pflegeaufwandes unterscheiden, sondern durch die gesundheitlichen Voraussetzungen. Ebenso können die oberen Pflegegeldstufen mit relativ "stabileren" Gesundheitszuständen verbunden sein, die sich vor dem Tod nicht mehr ändern.

Ein höheres Alter zeigt – ceteris paribus – eine positive Korrelation mit dem Übertritt in eine höhere Pflegegeldstufe, ebenso wie die Verschreibung von mit Demenz verbundenen Medikamenten. Die Analyse der Diagnosen hebt die Relevanz von Demenz und bösartigen Neubildungen für den Übertritt in höhere Pflegegeldstufen hervor. Die wesentlichste medizinische Einzelleistung, die einen positiven statistischen Zusammenhang mit dem Übergang in eine höhere Pflegegeldstufe hat, ist jene der Computertomographie des Kopf- und Halsbereichs. Die Implantation eines künstlichen Hüftgelenks hat, ebenso wie die Implantation eines künstlichen

Kniegelenks, einen negativen Zusammenhang mit einer Höhereinstufung: Diese Eingriffe korrelieren – *ceteris paribus* – negativ mit der Wahrscheinlichkeit eines Wechsels in eine höhere Pflegegeldstufe. Eine mögliche Begründung könnte sein, dass diese Eingriffe nur an vergleichsweise "fitten" Personen durchgeführt werden. Leistungen einer Bestrahlung (wohl in der Regel als Folge einer bösartigen Neubildung) sind verbunden mit höheren Wahrscheinlichkeiten des Übergangs in eine höhere Pflegegeldstufe. Die Medikamente bzw. Heilmittel, die am stärksten einem Übergang in eine höhere Pflegegeldstufe vorausgehen, sind Substanzen, die mit unterschiedlichen dementiellen Erkrankungen in Verbindung stehen. Die Ergebnisse des Moduls 2 zeigen die Wichtigkeit von dementiellen Erkrankungen bei der Projektion eines Wechsels in eine höhere Pflegegeldstufe.

Die Analyse der Rolle der persönlichen Merkmale zeigt den oben bereits beschriebenen negativen Zusammenhang mit der Pflegegeldstufe im Monat vor dem Übergang in eine höhere Pflegegeldstufe sowie den positiven Zusammenhang mit dem Alter der Personen. Diese Analyse zeigt zusätzlich einen positiven statistischen Zusammenhang zwischen dem Übergang in eine höhere Pflegegeldstufe und manchen Regionen.

Modul 3 analysiert die Zusammenhänge zwischen den abgerufenen Gesundheitsleistungen (im Beobachtungszeitraum 2016) und den Übertritten in die stationäre Pflege von pflegegeldbeziehenden Personen ab 60 Jahren (im Eventzeitraum 2017). Im Jahr 2017 sind 7,7% der Pflegegeldbeziehenden ab 60 Jahren in die stationäre Pflege übergetreten. Die Ergebnisse zeigen die Wahrscheinlichkeit des Eintritts in die stationäre Pflege unter Berücksichtigung persönlicher Eigenschaften, den im Beobachtungszeitraum abgerufenen Gesundheitsleistungen und gestellten Diagnosen sowie der Pflegegeldstufe am Ende des Beobachtungszeitraums (Dezember 2016).

Die multivariate Analyse der LASSO-Regressionen zeigt, dass die Pflegegeldstufe im Dezember 2016 die stärkste Korrelation mit einem Eintritt in die stationäre Pflege im Jahr 2017 hat: Je höher die Pflegegeldstufe, desto höher ist – *ceteris paribus* – die Wahrscheinlichkeit des Eintritts. Ebenso stark korrelieren die ärztlichen Leistungen von Allgemeinmedizinerinnen und -medizinern – in diesem Zusammenhang wohl ein Indiz für einen schlechten Gesundheitszustand – sowie ein höheres Alter. Die Analyse der Diagnosen hebt die Relevanz von Unfällen im privaten Bereich (z. B. Stürze), Demenz und psychischen Störungen als Einflussfaktoren für den Wechsel in die stationäre Pflege hervor. Die wesentlichste medizinische Einzelleistung, die die Wahrscheinlichkeit des Eintritts in die stationäre Pflege erhöht, ist eine Computertomographie des Kopf- und Halsbereichs. Patientinnen und Patienten mit Nierenersatztherapien und Kathetereingriffen der Herzkranzgefäße haben – *ceteris paribus* – eine geringere Wahrscheinlichkeit, in die stationäre Pflege überzutreten. Es zeigt sich, dass manche kurativen Eingriffe (auch z. B. künstliche Hüft- oder Kniegelenke) die Wahrscheinlichkeit eines Wechsels in die stationäre Pflege reduzieren. Eine mögliche Erklärung wäre, dass es sich dabei zum Teil um Eingriffe handelt, die nicht bei Personen mit labilem Gesundheitszustand durchgeführt werden können. Ebenso einen negativen Effekt auf die Wahrscheinlichkeit des Eintritts in die stationäre Pflege haben – *ceteris paribus* – gerinnungshemmende Medikamente und Cholesterinsenker. Die

Analyse der persönlichen Merkmale zeigt – neben der Pflegegeldstufe und dem Alter – positive Zusammenhänge zwischen dem Eintritt in die stationäre Pflege und manchen Regionen in Österreich. Eine Erklärung dafür könnte ein Angebotseffekt in diesen Regionen sein, nämlich, dass in diesen Regionen zusätzliche Kapazitäten geschaffen wurden oder mehr Kapazitäten vorhanden sind als in anderen Regionen.

Modul 4 analysiert die Zusammenhänge zwischen den medizinischen Einzelleistungen und den Gesundheitsdiagnosen der Pflegegeldbeziehenden im Beobachtungszeitraum 2016 und ihren Tagen an stationären Krankenhausaufenthalten im Eventzeitraum 2017. Im Schnitt verbrachten Pflegegeldbeziehende 8,1 Nächte im Krankenhaus.

Die multivariate Analyse der LASSO-Regressionen zeigt einen stärkeren Effekt der Gesundheitsdiagnosen gegenüber den medizinischen Einzelleistungen auf die Dauer der Krankenhausaufenthalte. Die mit Abstand stärkste Korrelation mit der Krankenhausedauer weist die essentielle (primäre) Hypertonie auf, gefolgt von Krankheiten des Harnsystems und sonstigen Unfällen im privaten Bereich.

Diese Studie hat erstmals die Zusammenhänge zwischen der gesundheitlichen Entwicklung von älteren Personen (gemessen an den bezogenen Gesundheitsleistungen) und dem Pflegegeldsystem für Österreich untersucht. Die wesentlichen Schlussfolgerungen sind folgende:

- Der erstmalige Pflegegeldbezug lässt sich relativ gut prognostizieren.
- Sehr rezente Gesundheitsleistungen sind für den Eintritt ins Pflegegeldsystem von großer Bedeutung.
- Zu den am stärksten mit dem erstmaligen Pflegegeldbezug in der kurzen Frist korrelierten Faktoren zählen
 - das Alter,
 - stationäre Aufenthalte in Krankenanstalten,
 - Kontakte zu Allgemeinmedizinerinnen und -medizinern, sowie
 - Heilmittel, die mit dem Nervensystem in Verbindung stehen.
- Die Etablierung eines Vorlaufindikators auf individueller Ebene scheint trotz der guten Prognoseleistung auf Basis dieser Daten und Methoden verfrüht. Weitere Forschung wäre notwendig, um alternative Modelle und Ansätze zu evaluieren, welche möglicherweise die Modellgüte weiter verbessern.
- Die höchste Präzedenz für den Übertritt in eine höhere Pflegegeldstufe weisen
 - die Pflegegeldstufe am Ende des Beobachtungszeitraums,
 - das Alter, sowie
 - Heilmittel, die mit dem Nervensystem in Verbindung stehen auf.
- Zu den wichtigsten Faktoren, die mit dem Eintritt in die stationäre Pflege korrelieren, zählen
 - die Pflegegeldstufe am Ende des Beobachtungszeitraums,

- Kontakte zu Allgemeinmedizinerinnen und -medizinern, sowie
- das Alter.
- Zu den Faktoren mit der höchsten Korrelation mit der Dauer der Krankenhausaufenthalte zählen
 - die Gesundheitsdiagnose der essentiellen (primären) Hypertonie, sowie
 - Computertomographien von Kopf und Hals sowie Abdomen und Becken.
 - Einzig Ergotherapien und logopädisch-phoniatische Therapien stehen in einem Zusammenhang mit einer kürzeren Krankenhausaufenthaltsdauer.

Folgende gesundheitspolitische Schlussfolgerungen können aus den Ergebnissen der Studie gezogen werden:

- Die Studie zeigt die Relevanz des Alters: Personen, die gleiche Charakteristika und dieselben bezogenen Gesundheitsleistungen aufweisen (= *ceteris paribus*) und sich nur durch ihr Alter von Vergleichspersonen unterscheiden, haben eine höhere Wahrscheinlichkeit, ins Pflegesystem einzutreten, in höhere Pflegegeldstufen überzutreten und in die stationäre Pflege zu wechseln. Dies könnte bedeuten, dass das Alter für die Evaluierung der Pflegebedürftigkeit eine überproportionale Rolle spielt und Personen aufgrund ihres Alters, aber nicht primär aufgrund ihres Gesundheitszustandes eingestuft werden, wobei ein höheres Alter oftmals mit einem Funktionsverlust für die Aktivitäten des täglichen Lebens einhergeht.
- Weiters hebt die Studie die Bedeutung von Demenz hervor: Sowohl beim Eintritt in das Pflegegeldsystem als auch bei einer Höhereinstufung spielen dementielle Erkrankungen eine wesentliche Rolle. Hier sollten die Präventionsfelder ausgebaut werden.
- Allgemeinmedizinerinnen und -mediziner spielen sowohl beim Eintritt in das Pflegegeldsystem als auch beim Eintritt in die stationäre Pflege eine wesentliche Rolle. Häufige Besuche bedeuten eine höhere Wahrscheinlichkeit, in das Pflegegeldsystem bzw. in die stationäre Pflege einzutreten. Die Erklärung dafür ist vermutlich, dass sich der allgemeine Gesundheitszustand verschlechtert und Besuche bei Allgemeinmedizinerinnen und -medizinern notwendig macht. Dies bedeutet, dass Allgemeinmedizinerinnen und -mediziner für die Einschätzung der Entwicklung des Pflege(geld)systems zentral sind und diese Informationen durch verstärkte Kooperation in die weitere Gestaltung des Pflege(geld)systems einfließen sollten (siehe auch *Famira-Mühlberger et al., 2019*).
- Zwischen manchen Regionen Österreichs unterscheiden sich – *ceteris paribus* – sowohl die Wahrscheinlichkeiten, erstmalig Pflegegeld zu beziehen, in eine höhere Pflegegeldstufe zu wechseln, oder in die stationäre Pflege überzutreten. Die Gründe dafür müssten genauer evaluiert werden.
- Manche kurativen Eingriffe könnten das Pflegegeldsystem entlasten, da die betroffenen Personen durch die Eingriffe länger fit bleiben können und nicht auf Pflege angewiesen sind (z. B. Hüft- oder Kniegelenksimplantationen oder kurative Herzeingriffe). In einer Kosten-Nutzen-Analyse müsste das berücksichtigt werden.

- Im Bereich der medizinischen Einzelleistungen korrelieren Ergotherapien im Rahmen eines stationären Aufenthalts im Beobachtungszeitraum negativ mit den Krankenhaustagen im Eventzeitraum. Eine mögliche Erklärung hierfür ist, dass Ergotherapien pflegegeldbeziehende Personen beim Umgang mit neuen Situationen unterstützen (z.B. durch Hilfsmittelanpassung).
- Die Diagnose der essentiellen (primären) Hypertonie³⁴⁾ steht in einem statistischen Zusammenhang mit einem längeren Krankenhausaufenthalt der Pflegegeldbeziehenden. Hier gilt es, die Wirkung von Präventivmaßnahmen zu evaluieren.
- Die medizinischen Einzelleistungen der Ergo- und logopädisch-phoniatriischen Therapien zeigten in den Analysen aufenthaltssenkende Effekte bei stationären Krankenhausaufenthalten. Hier tritt offenbar neben dem kurativen auch ein präventiver Effekt zutage.

³⁴⁾ Bluthochdruck, der nicht durch andere Erkrankungen verursacht ist.

Literatur

- BMASGK (2018). *Eingliederungsindikatoren 2017. Kennzahlen für soziale Inklusion in Österreich*. Wien.
- BMSGPK (2020). *Leistungskatalog BMASGK 2020*. Wien: Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz.
- Breimann, L. (2001). *Random Forests*. *Machine Learning* 45, S. 5-32.
- Elo, I. T. (2009). Social Class Differentials in Health and Mortality: Patterns and Explanations in Comparative Perspective. *Annual Review of Sociology*, 35, S. 553-572.
- Famira-Mühlberger, U. (2020). *Pflegevorsorge in Gemeinden*. WIFO-Studie, Wien.
- Famira-Mühlberger, U., Firgo, M., Streicher, G. (2019). *Geriatrische Versorgung in Wien im Kontext des demographischen Wandels*. WIFO Monographien.
- Friedman, J., Hastie, T., Tibshirani, R. (2010). *Regularization Paths for Generalized Linear Models via Coordinate Descent*. *Journal of Statistical Software* 33(1), S. 1-22.
- Galama, T. J., Van Kippersluis, H. (2018). *A Theory of Socio-Economic Disparities in Health over the Life Cycle*. *The Economic Journal*, 129(617), S. 338-374.
- Hastie, T., Tibshirani, R., Friedman, J. (2009). *The Elements of Statistical Learning. Data Mining, Inference and Prediction* (2. Auflage). Springer Series in Statistics, New York: Springer.
- Hoerl, A.E., Kennard, R. (1970). *Ridge Regression: Biased Estimation for Nonorthogonal Problems*. *Technometrics* 12, S. 55-67.
- Hosmer, D.W., Lemeshow, S. (2000). *Applied Logistic Regression* (2. Ausgabe). Wiley Series in Probability and Statistics, New York: John Wiley & Sons.
- James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R*. Springer Texts in Statistics, New York: Springer.
- Klien, M., Pitlik, H., Firgo, M., Famira-Mühlberger, U. (2020). *Ein Modell für einen strukturierten vertikalen Finanzausgleich in Österreich*. WIFO-Studie, Wien.
- Klotz, J. (2010). *Convergence or Divergence of Educational Disparities in Mortality and Morbidity? The Evolution of Life Expectancy and Health Expectancy by Educational Attainment in Austria in 1981-2006*. *Vienna Yearbook of Population Research*, 8, S. 139-174.
- Long, J. S. (2008). *Regression Models for Categorical and Limited Dependent Variables* (11. Auflage). Thousand Oaks: SAGE Publications.
- Nykodym, T., Kraljevic, T., Wang, A., Wong, W. (2021). *Generalized Linear Modeling with H2O* (7. Ausgabe). Mountain View, CA: H2o.ai, Inc.
- Tibshirani, R. (1996). *Regression shrinkage and selection via the Lasso*. *Journal of the Royal Statistical Society* 58, S. 267-288.
- Winkelmann, R., Boes, S. (2009). *Analysis of Microdata* (2. Ausgabe). Heidelberg: Springer.

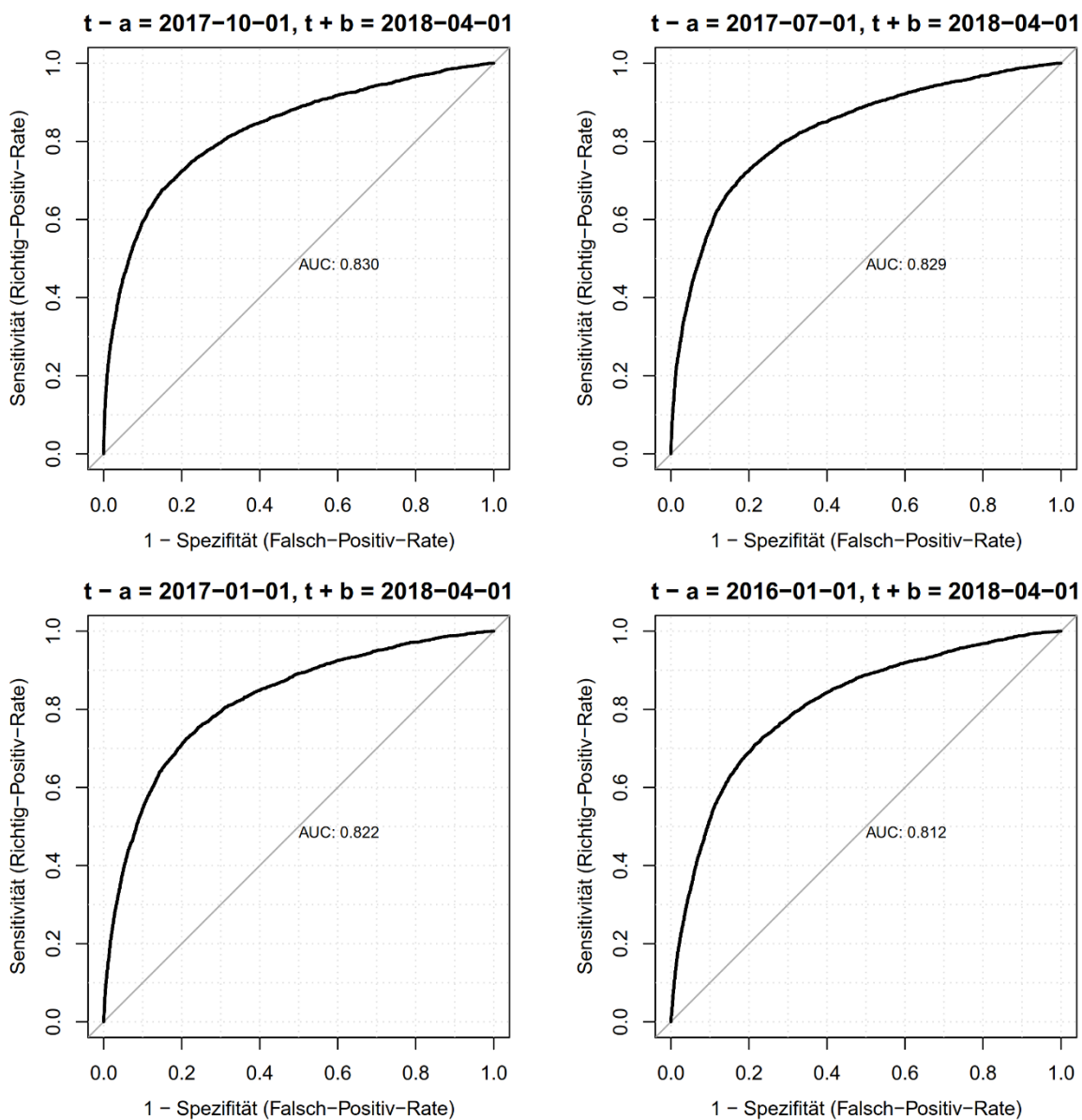
7. Appendix

7.1 ROC-Kurven bei unterschiedlichen Beobachtungs- und Eventzeiträumen

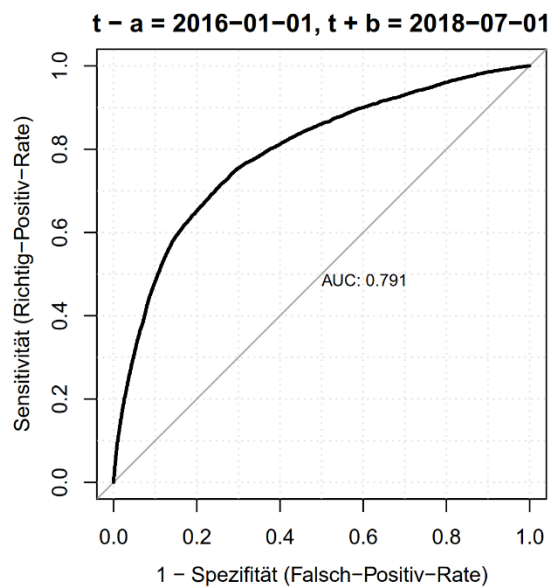
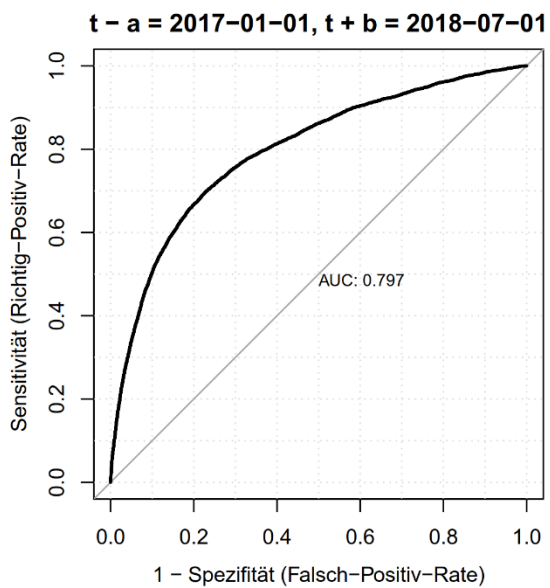
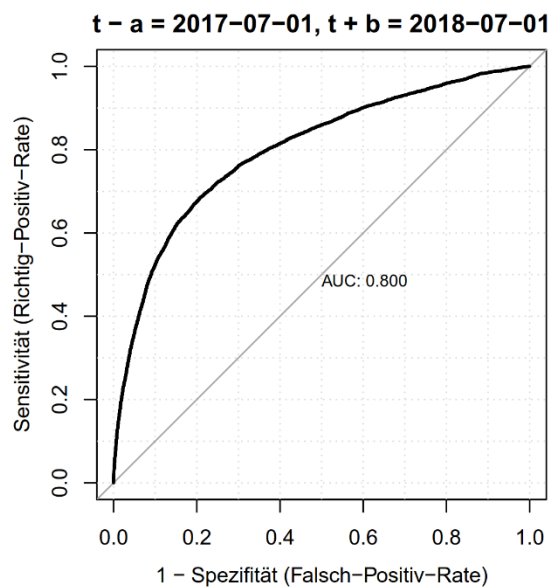
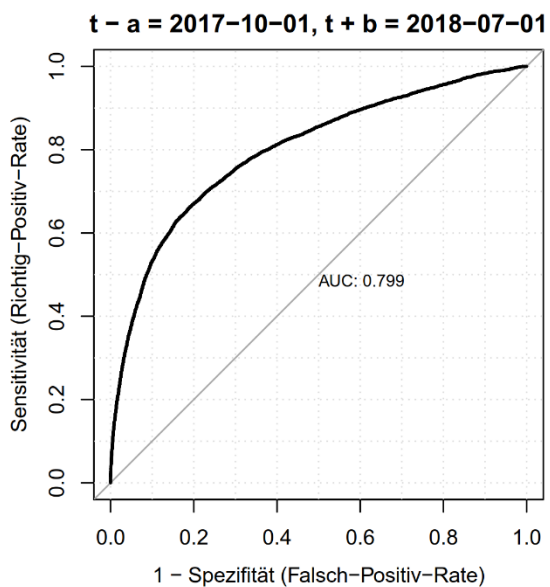
Dieser Anhang zeigt die ROC-Kurven (auf Basis der Testdaten) der Modelle im oberen Teil von Übersicht 5 mit aggregierten Gesundheitsvariablen für unterschiedliche Beobachtungs- und Eventzeiträume

Abbildung A 1: ROC-Kurven auf Basis der Testdaten für Modelle mit unterschiedlichem Beobachtungszeitraum

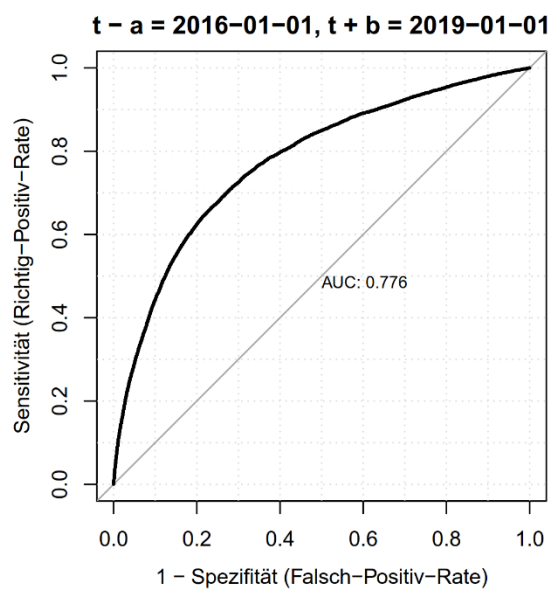
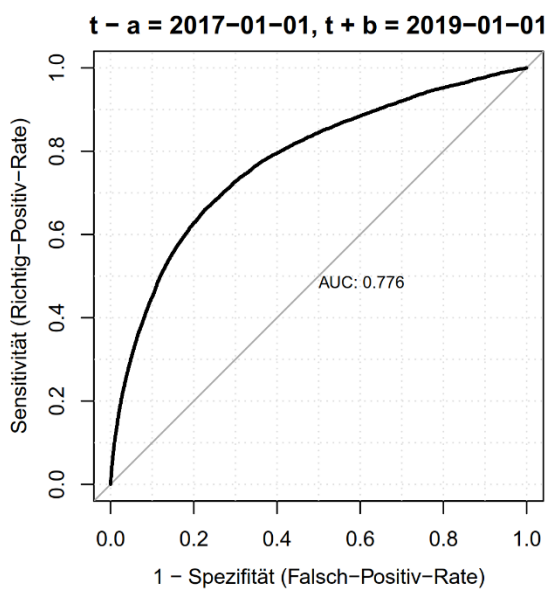
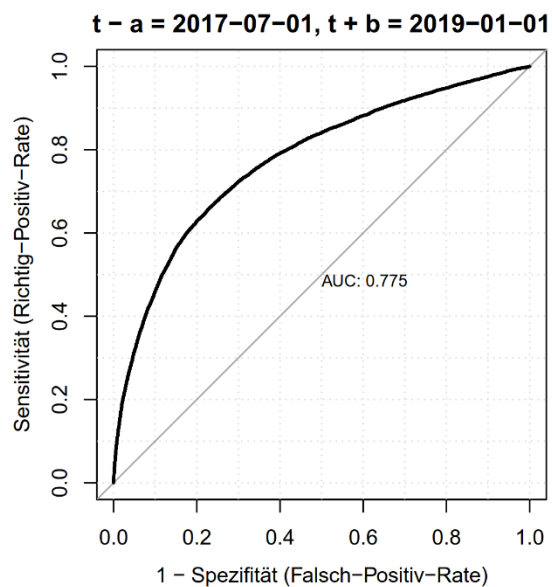
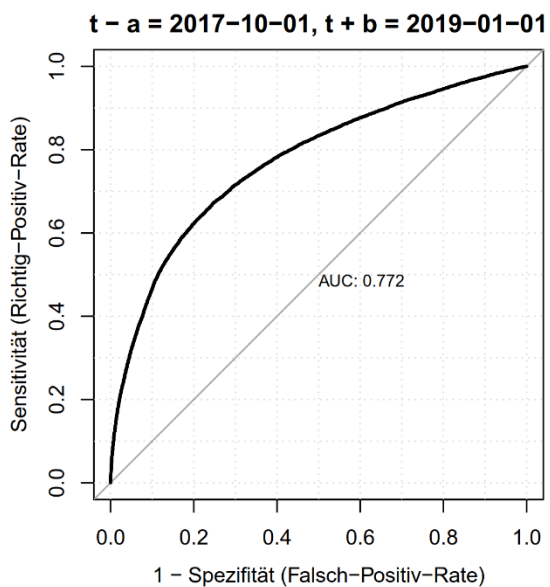
Eventzeitraum: [01/2018, 04/2018)



Eventzeitraum: [01/2018, 07/2018)



Eventzeitraum: [01/2018, 01/2019)

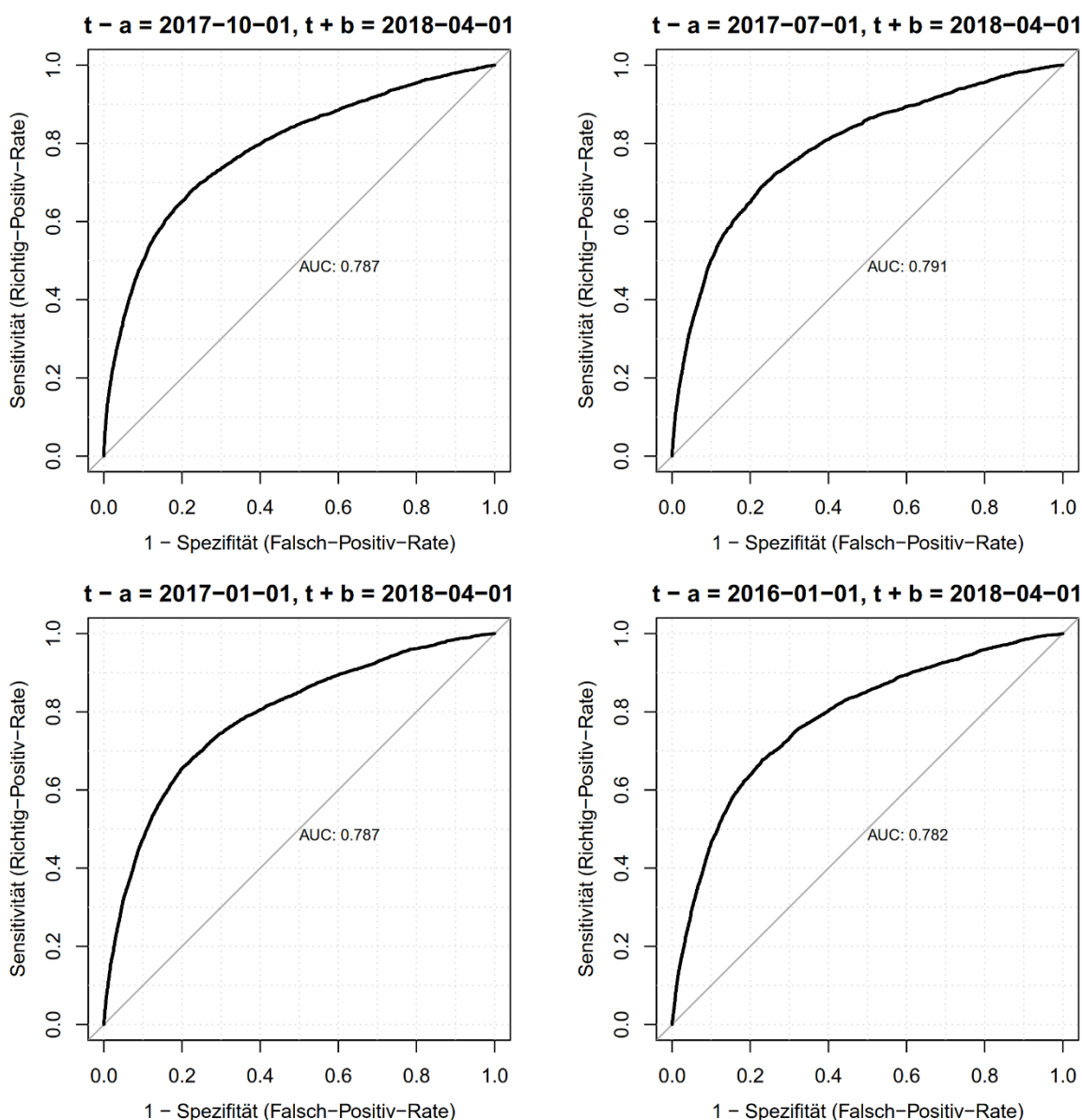


7.2 ROC-Kurven bei unterschiedlichen Beobachtungs- und Eventzeiträumen, ohne unmittelbarem Vormonat

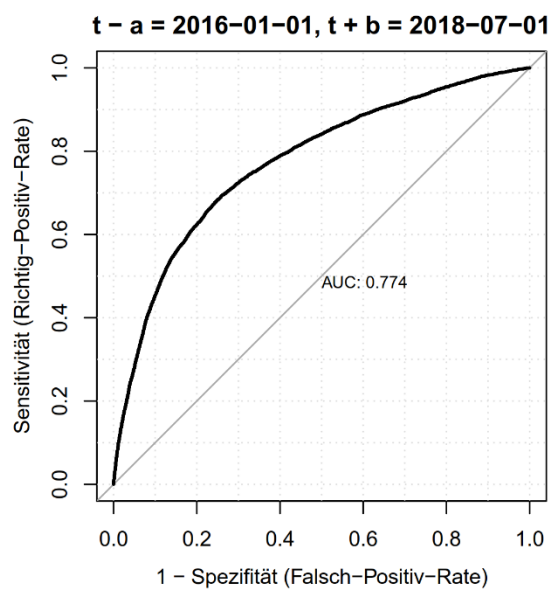
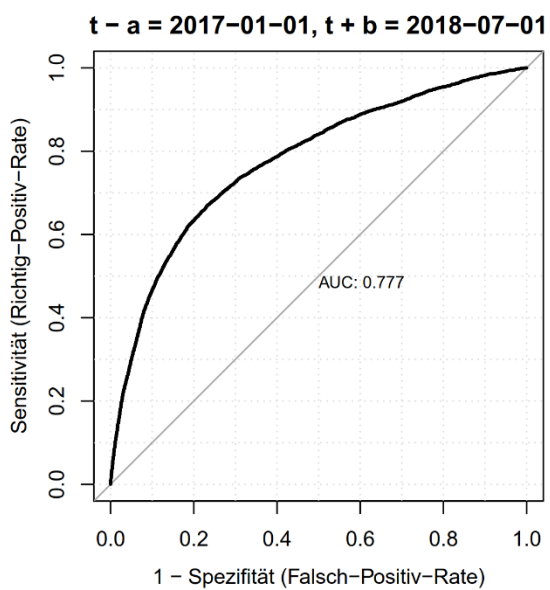
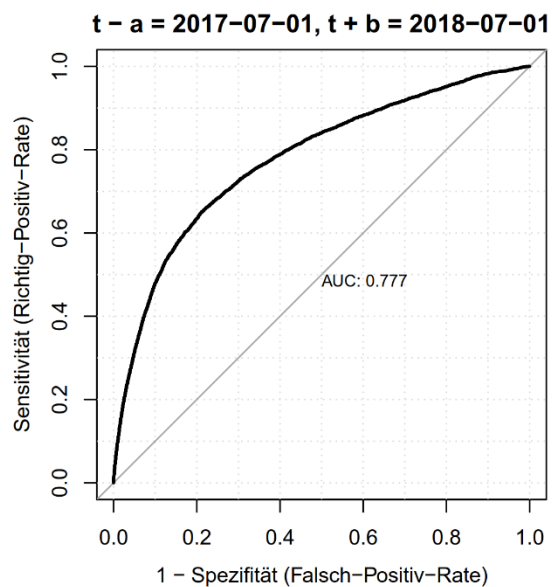
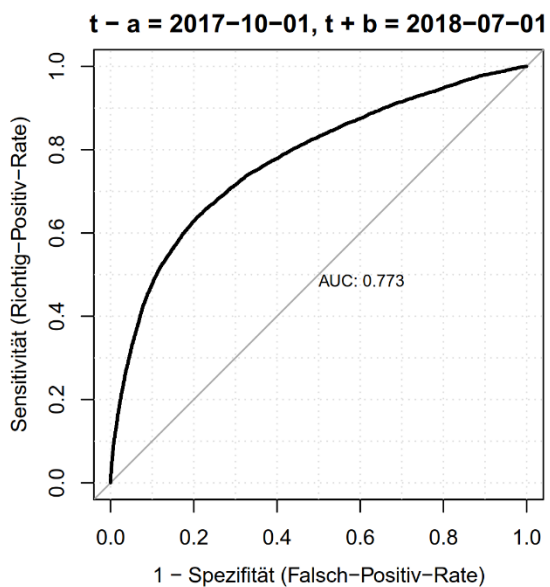
Dieser Anhang zeigt die ROC-Kurven (auf Basis der Testdaten) der Modelle um unteren Teil von Übersicht 5 mit aggregierten Gesundheitsvariablen für unterschiedliche Beobachtungs- und Eventzeiträume; Beobachtungszeitraum ohne Dezember 2017.

Abbildung A 2: ROC-Kurven auf Basis der Testdaten für Modelle mit unterschiedlichem Beobachtungszeitraum

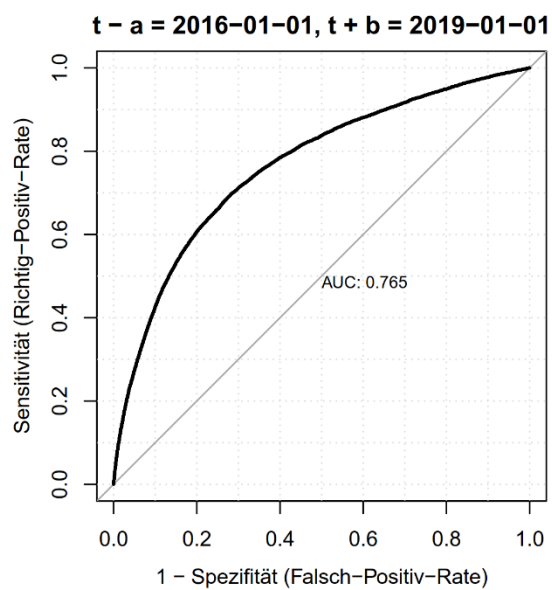
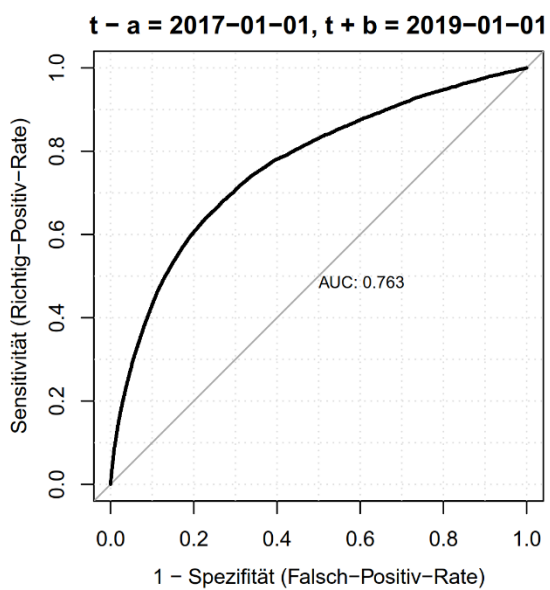
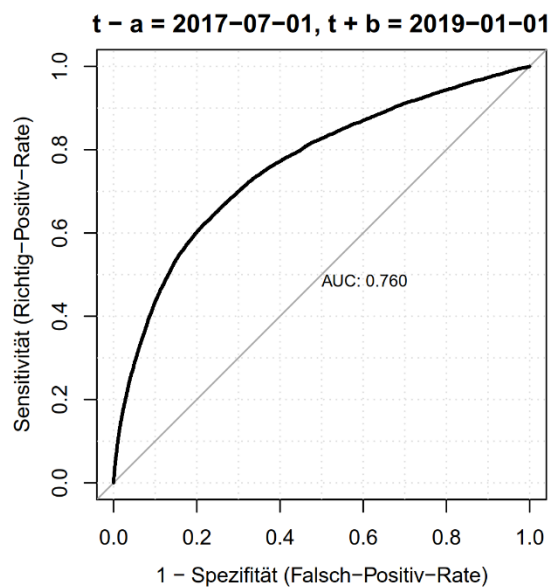
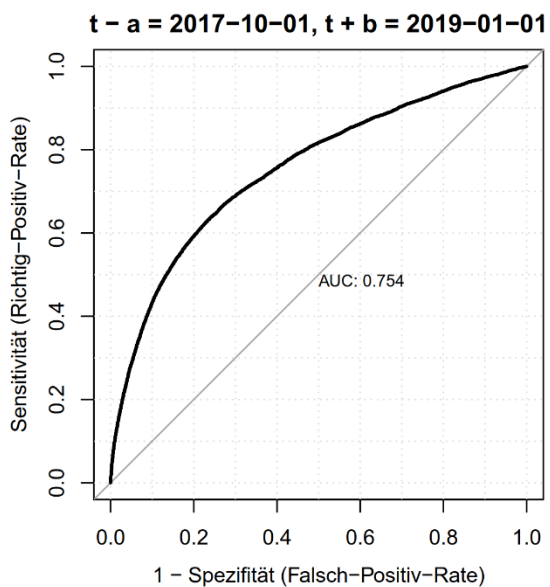
Eventzeitraum: [01/2018, 04/2018)



Eventzeitraum: [01/2018, 07/2018)



Eventzeitraum: [01/2018, 01/2019)



7.3 Ausführliche Methodenbeschreibung

7.3.1 Grundmodell

Das Ziel der empirischen Analyse in Modul 1 ist, auf Basis statistischer Modelle zu ermitteln, welche Wahrscheinlichkeit eines erstmaligen Bezugs von Pflegegeld innerhalb eines Eventzeitraums $[t, t + b)$ sich aus den persönlichen Eigenschaften einer Person i , welche mit X_i bezeichnet werden, sowie aus den von Person i im vorgelagerten Beobachtungszeitraum $[t - a, t)$ konsumierten Gesundheitsleistungen G_i ableiten lässt.³⁵⁾

Abbildung A 3: **Schematische Darstellung der Zeiträume in der empirischen Untersuchung**



Q: Eigene Darstellung.

Weist man einer Indikatorvariable $PG_{i,[t,t+b)}$ den Wert 1 zu, wenn die Person i im Zeitraum $[t, t + b)$ erstmalig Pflegegeld bezieht und den Wert 0, wenn dies nicht der Fall ist, dann stellt die Untersuchung des Zusammenhangs zwischen der Wahrscheinlichkeit eines erstmaligen Pflegegeldbezugs als abhängiger und den persönlichen Charakteristiken und konsumierten Gesundheitsleistungen als erklärenden Variablen ein binäres Klassifikationsproblem dar (*Hastie et al., 2009; James et al., 2013*), bei dem diese Wahrscheinlichkeit als Funktion der erklärenden Variablen modelliert wird:

$$\Pr(PG_{i,[t,t+b)} = 1) = f(X_{i,t}, G_{i,[t-a,t)}) + \varepsilon_{i,t}$$

Formell besteht das Ziel also darin, die binäre Indikatorvariable PG als Funktion der erklärenden Variablen X und G sowie eines Fehlerterms ε zu modellieren.³⁶⁾

7.3.2 Logistische Regression

Eine der gängigsten Methoden zur Modellierung solcher Zusammenhänge ist die logistische Regression (Logit-Regression). Die logistische Regression modelliert den Logarithmus der "Odds" (darunter versteht man das Verhältnis der Wahrscheinlichkeit, dass ein Ereignis eintritt, zur Wahrscheinlichkeit, dass das Ereignis nicht eintritt) als lineare Funktion erklärender Variablen:

$$\log\left(\frac{\Pr(PG_{i,[t,t+b)} = 1)}{\Pr(PG_{i,[t,t+b)} = 0)}\right) = X_{i,t}\beta + G_{i,[t-a,t)}\gamma$$

³⁵⁾ Diese Intervalle sind inklusive dem Startzeitpunkt aber exklusiv dem Endzeitpunkt zu verstehen. Der Zeitpunkt t ist daher z.B. nur im Eventzeitraum enthalten, nicht jedoch im Beobachtungszeitraum.

³⁶⁾ Der Fehlerterm $\varepsilon_{i,t}$ beinhaltet rein zufällige, nicht-systematische Einflüsse, aber auch z.B. Variablen, welche mit dem Ereignis PG_i in Zusammenhang stehen, in den erklärenden Variablen jedoch nicht enthalten sind.

wobei β und γ die Parameter(vektoren) darstellen, welche den Einfluss der erklärenden Variablen auf die Log-Odds messen.³⁷⁾

Die Parametervektoren β und γ werden in einer Logit-Regression mittels der Maximum-Likelihood-Methode geschätzt. Dabei werden diese in einem iterativen Verfahren so gewählt, dass sie die Wahrscheinlichkeit maximieren, die vorliegenden Daten zu beobachten. Grundlage dieser Maximierung ist die sog. Log-Likelihood-Funktion:

$$LL(\beta, \gamma) = \frac{1}{N} \sum_{i=1}^N \left(PG_i(X_{i,t}\beta + G_{i,[t-a,t]}\gamma) - \log(1 + \exp(X_{i,t}\beta + G_{i,[t-a,t]}\gamma)) \right)$$

Die aus einer Stichprobe geschätzten Parameter dieses statistischen Modells können in Folge dazu genutzt werden, den zukünftigen Status der abhängigen Variable "erstmaliger Pflegegeldbezug" auch für Beobachtungen zu prognostizieren, die sich nicht innerhalb der Stichprobe befinden, indem die aus dem Modell prognostizierte Wahrscheinlichkeit für $PG_{i,[t+b]} = 1$ als Vorlaufindikator für den erstmaligen Pflegegeldbezug in der (un-)mittelbaren Zukunft herangezogen wird.

7.3.3 Trainings- und Testdaten zur Evaluierung der Prognosegüte

Um die Prognosegüte des Modells zu evaluieren werden die vorhandenen Beobachtungen in Trainings- und Testdaten geteilt (vgl. *Hastie et al.*, 2009, S. 222f). Die Trainingsdaten werden dazu genutzt, das statistische Modell zu optimieren, d.h. die Modellparameter werden so gewählt, dass sie optimal an die in den Trainingsdaten enthaltenen Beobachtungen angepasst sind und die abhängige Variable in den Trainingsdaten bestmöglich erklären. Dies erlaubt jedoch nur begrenzte Rückschlüsse darauf, wie gut das Modell die abhängige Variable in Daten prognostizieren würde, auf die es nicht "trainiert" ist.

Die Prognosegüte wird daher nicht anhand der Trainingsdaten, sondern anhand der Testdaten evaluiert. Dazu werden das aus den Trainingsdaten geschätzte Modell und die in den Testdaten vorhandenen Werte der erklärenden Variablen verwendet, um die abhängige Variable zu prognostizieren. Der Vergleich zwischen der in den Testdaten prognostizierten Wahrscheinlichkeit $\widehat{Pr}(PG_{i,[t,t+b]} = 1)$ und den dort tatsächlich beobachteten Werten von $PG_{i,[t+b]}$ kann dann als Grundlage für die Evaluierung der Prognosegüte verwendet werden.³⁸⁾

Eine Möglichkeit, die Prognosegüte zu evaluieren ist, die Beobachtungen anhand der prognostizierten Wahrscheinlichkeiten zu klassifizieren, indem Beobachtungen, deren prognostizierte Wahrscheinlichkeit einen Schwellenwert τ übersteigen, ein "positiver" prognostizierter Wert $\widehat{PG}_{i,[t,t+b]} = 1$ zugeteilt wird und Beobachtungen, deren prognostizierte

³⁷⁾ Üblicherweise nimmt der erste Eintrag im Vektor der erklärenden Variablen $X_{i,t}$ den Wert 1 an, sodass der Koeffizientenvektor β auch eine Konstante umfasst.

³⁸⁾ Die Teilung der Daten in Trainings- und Testdaten erfolgt anhand einer Zufallsauswahl. Für die Wahl der relativen Größe der beiden Datensätze gibt es nur Faustregeln, welche üblicherweise Werte zwischen 20% und 30% der Gesamtstichprobe für die Testdaten empfehlen. In dieser Studie werden durchgehend 25% der Beobachtungen zufällig den Testdaten zugeteilt, die restlichen 75% den Trainingsdaten.

Wahrscheinlichkeit τ nicht überschreitet, ein "negativer" prognostizierter Wert $\widehat{PG}_{i,[t,t+b]} = 0$. Naheliegender wäre z.B. eine Klassifizierung nach dem wahrscheinlichsten Ergebnis mit einem Schwellenwert von $\tau = 0,5$, sodass ab $\widehat{\Pr}(PG_{i,[t,t+b]} = 1) > 0,5$ ein erstmaliger Pflegegeldbezug prognostiziert wird.

7.3.4 Wahrheitsmatrix, Sensitivität, Spezifität und ROC-Kurve

Die Gegenüberstellung der prognostizierten und tatsächlich beobachteten Werte von $PG_{i,[t,t+b]}$ ergibt eine "Wahrheitsmatrix" ("Confusion Matrix", siehe James et al., 2013, S. 145). Die Wahrheitsmatrix zeigt die Anzahl der "richtig positiv" (RP) klassifizierten Beobachtungen, für welche $\widehat{PG}_{i,[t,t+b]} = PG_{i,[t,t+b]} = 1$ gilt, die Anzahl der "richtig negativ" (RN) klassifizierten Beobachtungen mit $\widehat{PG}_{i,[t,t+b]} = PG_{i,[t,t+b]} = 0$, sowie die Anzahl der "falsch negativen" (FN) Beobachtungen mit $\widehat{PG}_{i,[t,t+b]} = 0$ und $PG_{i,[t,t+b]} = 1$ und die Anzahl der "falsch positiven" (FP) Beobachtungen mit $\widehat{PG}_{i,[t,t+b]} = 1$ und $PG_{i,[t,t+b]} = 0$ (vgl. Übersicht A 1). Aus diesen Einträgen lassen sich Kennzahlen für die Prognosegüte eines Modells berechnen.

Übersicht A 1: **Schematische Darstellung einer Wahrheitsmatrix**

Prognostiziert	Beobachtet		Summe
	$PG_{i,[t,t+b]} = 0$	$PG_{i,[t,t+b]} = 1$	
$\widehat{PG}_{i,[t,t+b]} = 0$	RN	FN	\widehat{TN}
$\widehat{PG}_{i,[t,t+b]} = 1$	FP	RP	\widehat{TP}
Summe	TN	TP	N

Q: Eigene Darstellung

So wird z.B. die Falschklassifikationsrate ("Error Rate") wird aus dem Verhältnis der falsch positiv oder falsch negativ klassifizierten Beobachtungen zu allen Beobachtungen berechnet:

$$\text{Falschklassifikationsrate} = \frac{FN + FP}{N}$$

Sie gibt damit an, wie hoch der relative Anteil der falsch klassifizierten Beobachtungen insgesamt ist. Zur Beurteilung der Prognosegüte ist die Falschklassifikationsrate vor allem dann geeignet, wenn die beobachteten Ereignisse $PG_{i,[t,t+b]} = 1$ und $PG_{i,[t,t+b]} = 0$ relativ ausgewogen verteilt sind. Ist hingegen der Anteil der Beobachtungen mit einer bestimmten Ausprägung der abhängigen Variable sehr hoch, ist die Falschklassifikationsrate weniger informativ. Be trägt z.B. der Anteil der Beobachtungen mit $PG_{i,[t,t+b]} = 0$ an die 90%, dann würde selbst ein "naiver" Klassifikator, der immer dem Wert $\widehat{PG}_{i,[t,t+b]} = 0$ prognostiziert, eine (vermeintlich gute) Falschklassifikationsrate von lediglich 10% aufweisen.

Als Alternative bieten sich daher die "Sensitivität" und "Spezifität" an. Als Sensitivität oder Richtig-Positiv-Rate bezeichnet man den Anteil der "richtig positiv" klassifizierten Beobachtungen an allen tatsächlich "positiven" Beobachtungen. Dies entspricht dem Anteil der Personen, denen vom Modell korrekterweise ein erstmaliger Pflegegeldbezug prognostiziert wird, an allen Personen, die tatsächlich erstmalig Pflegegeld beziehen. Die Spezifität oder Richtig-Negativ-Rate hingegen gibt den Anteil der "richtig negativ" prognostizierten Werte an allen tatsächlich

"negativen" Beobachtungen an, d.h. den Anteil der Personen, denen vom Modell korrekterweise kein erstmaliger Pflegegeldbezug vorhergesagt wird, an allen Personen, die tatsächlich nicht erstmalig Pflegegeld beziehen:

$$\text{Sensitivität} = \frac{RP}{RP + FN}$$

$$\text{Spezifität} = \frac{RN}{RN + FP}$$

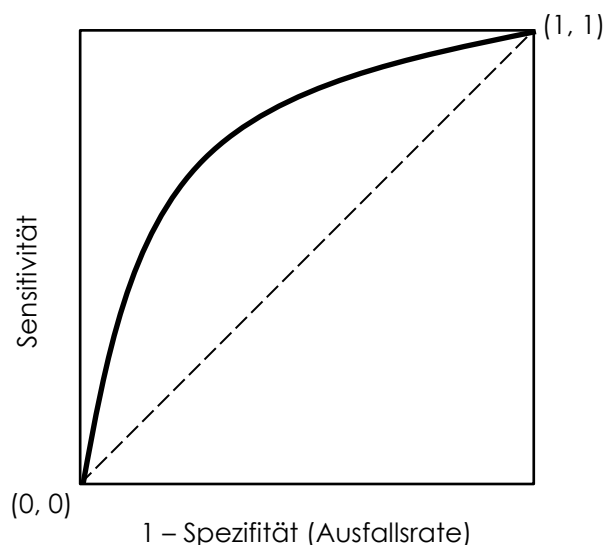
$$\text{Ausfallsrate} = 1 - \text{Spezifität} = \frac{RN}{RN + FP} = \frac{FP}{RN + FP}$$

Ein gutes Modell sollte bei gegebener Sensitivität eine möglichst hohe Spezifität aufweisen und *vice versa*. Häufig wird anstelle der Spezifität auch die "Ausfallsrate" oder Falsch-Positiv-Rate verwendet. Diese gibt an, wie hoch der Anteil der fälschlicherweise als positiv klassifizierten Beobachtungen an allen tatsächlich negativen Beobachtungen ist.

Dabei besteht eine Wechselbeziehung zwischen Sensitivität und Spezifität (bzw. Ausfallsrate), die sich durch Variation des Schwellenwerts τ leicht illustrieren lässt. Setzt man z.B. den Schwellenwert $\tau = 0$, sodass alle Beobachtungen ab einer prognostizierten Wahrscheinlichkeit von $\widehat{\text{Pr}}(PG_{i,[t,t+b)} = 1) = 0$ bereits als "positiv" $\widehat{PG}_{i,[t,t+b)} = 1$ klassifiziert werden, dann wird die Sensitivität ihren Maximalwert von 1 annehmen: Alle tatsächlich positiven Werte werden auch als positiv prognostiziert. Die Spezifität beträgt hingegen 0 (die Ausfallsrate gleich 1), da keine Beobachtung richtig negativ klassifiziert wird. Setzt man hingegen den Schwellenwert auf $\tau = 1$, sodass Beobachtungen erst ab einer prognostizierten Wahrscheinlichkeit $\widehat{\text{Pr}}(PG_{i,[t,t+b)} = 1) = 1$ als "positiv" $\widehat{PG}_{i,[t,t+b)} = 1$ klassifiziert würden, dann würde die Sensitivität 0 betragen, die Spezifität hingegen ihren Maximalwert von 1 (bzw. die Ausfallsrate den Wert 0) annehmen: Alle tatsächlich "negativen" Beobachtungen hätten auch einen negativ prognostizierten Wert. Für Werte von τ zwischen diesen Extremwerten ergeben sich Kombinationen der Sensitivität und Spezifität, bei denen üblicherweise eine hohe Sensitivität mit einer geringen Spezifität (einer hohen Ausfallsrate) einhergeht und *vice versa*.

Um diese Wechselwirkung grafisch darzustellen wird häufig auf die Darstellung der "Receiver Operating Characteristic" (ROC)-Kurve zurückgegriffen (siehe *Winkelmann – Boes, 2009, S. 127f* oder *James et al., 2013, S. 147f*). Diese stellt in einer zweidimensionalen Darstellung mit der Sensitivität auf der vertikalen und der Ausfallsrate ($1 - \text{Spezifität}$) auf der horizontalen Achse die Kombinationen von Sensitivität und $1 - \text{Spezifität}$ für alle Werte von τ im Intervall $[0, 1]$ dar (vgl. Abbildung A 4). Die sich daraus ergebende Kurve ist, wie oben ausgeführt, durch die Punkte (0, 0) und (1, 1) begrenzt und liegt üblicherweise über der diagonalen Linie zwischen diesen beiden Punkten. Diese Diagonale stellt die ROC-Kurve eines nichtinformativen "Modells" dar, das Beobachtungen rein zufällig in eine der beiden Klassen einteilt. In diesem Fall würde die Sensitivität auch immer der Ausfallsrate entsprechen.

Abbildung A 4: **Schematische Darstellung einer ROC-Kurve**



Q: Eigene Darstellung

Die Fläche unter dieser ROC-Kurve wird auch AUC ("Area Under the Curve") genannt und stellt eine Maßzahl dafür dar, wie gut Sensitivität und Spezifität eines empirischen Modells sind. Diese variiert zwischen 0,5 (der Fläche unter der diagonalen Linie, welche die ROC-Kurve eines nichtinformativen Modells darstellt) und 1 (der Fläche unter der ROC-Kurve eines optimalen Modells mit einer Sensitivität von 1 und einer Ausfallsrate von 0). Je näher die AUC an 1 liegt, desto besser das empirische Modell. Als Faustregeln für die Einordnung der Prognosegüte anhand der AUC bescheinigen Hosmer – Lemeshow (2000, S. 162) Modellen mit $0,7 \leq AUC < 0,8$ ein "akzeptables" Unterscheidungsvermögen ("acceptable discrimination"), Modellen mit $0,8 \leq AUC < 0,9$ ein "exzellentes" Unterscheidungsvermögen ("excellent discrimination") und Modellen mit $AUC \geq 0,9$ ein "überragendes" ("outstanding discrimination") Unterscheidungsvermögen.

7.3.5 "Overfitting", Regularisierung und Kreuzvalidierung

Die Anpassung des statistischen Modells an die Trainingsdaten birgt, insbesondere bei einer hohen Anzahl an erklärenden Variablen, die Gefahr des "Overfitting". Darunter versteht man eine Überanpassung des statistischen Modells an die Idiosynkrasien der Trainingsdaten: Ein komplexes statistisches Modell, welches auf die Trainingsdaten "trainiert" ist, kann die abhängige Variable in diesen Daten mitunter sehr gut erklären, in "neuen" Daten (wie den Testdaten) kann es jedoch möglicherweise einen größeren Prognosefehler aufweisen als weniger komplexe Modelle. Daraus ergibt sich ein Zielkonflikt zwischen der bestmöglichen Prognose innerhalb der vorhandenen Daten (geringer "Bias") und der bestmöglichen Prognose, wenn das Modell mit neuen Daten konfrontiert wird (geringe "Varianz", vgl. James et al., 2009, S. 34f bzw. S. 204).

Eine Möglichkeit, das Risiko des "Overfitting" zu reduzieren besteht in der Anwendung von Regularisierungsmethoden. Dabei werden die Parameter des Modells geschrumpft, sodass eine Überanpassung durch zu starke Einflüsse einzelner Variablen reduziert wird. Dies kann zwar den Prognosefehler ("Bias") in den Trainingsdaten erhöhen, die Varianz in den Testdaten aber senken. Zwei der am weitest verbreiteten Regularisierungsmethoden sind die "Ridge Regression" (Hoerl – Kennard, 1970) und der "Least Absolute Shrinkage and Selection Operator" (LASSO, Tibshirani, 1996). Bei beiden Methoden wird eine pönalisierte Version der Log-Likelihood-Funktion maximiert. Während Ridge Regression von der Log-Likelihood die Summe der quadrierten Parameter abzieht:³⁹⁾

$$LL(\beta, \gamma) - \lambda \left(\sum_{j=1}^J \beta_j^2 + \sum_{k=1}^K \gamma_k^2 \right),$$

reduziert LASSO die Log-Likelihood um die Summe der Absolutwerte der Parameter (siehe Friedman et al., 2010, oder Nykodym et al., 2021):

$$LL(\beta, \gamma) - \lambda \left(\sum_{j=1}^J |\beta_j| + \sum_{k=1}^K |\gamma_k| \right)$$

Der Unterschied zwischen Ridge Regression und LASSO besteht darin, dass die Regularisierung im LASSO dazu führen kann, dass einige Parameter auf exakt Null geschrumpft werden. In der Ridge Regression werden die Parameter zwar gegen, aber nie exakt auf Null geschrumpft. Der LASSO dient demnach auch der Variablenselektion, was insbesondere in Anwendungen mit vielen potenziellen Einflussvariablen von Vorteil ist.⁴⁰⁾ Verbindet man LASSO und Ridge Regression in einem Modell, indem die Log-Likelihood sowohl um die Summe der Absolutwerte der Parameter als auch um die Summe der quadrierten Parameter reduziert, spricht man von einer "Elastic Net"-Regression.

Da die Größe der Parameter auch von der Skalierung der erklärenden Variable abhängt, was die Berechnung der Pönalterme verzerren würde, werden im Regressionsmodell durchgehend standardisierte Variablen verwendet. Dazu wird von einer Variable der jeweilige Mittelwert abgezogen und das Ergebnis anschließend durch die Standardabweichung der Variable dividiert. Diese Normierung hat den zusätzlichen Vorteil, dass alle Variablen im Modell nach Standardisierung gleich skaliert sind, sodass sich der relative Einfluss auf die abhängige Variable anhand der relativen Größe der geschätzten Parameter messen lässt.

Sowohl in der Ridge Regression als auch beim LASSO hängt die Stärke der Regularisierung von der Größe des "Strafterms" λ ab. Ist dieser Null, entsprechen beide Methoden der einfachen logistischen Regression. Die optimale Größe des Strafterms wird auf Basis eines k -fach

³⁹⁾ J und K bezeichnen dabei die Gesamtzahl der zu schätzenden Parameter für die persönlichen Charakteristiken X bzw. die Gesundheitsleistungen G .

⁴⁰⁾ Eine alternative Interpretation dieser Regularisierungsmethoden ist, dass sie die Log-Likelihood unter der Nebenbedingung maximieren, dass die Summe der quadrierten Parameter (Ridge Regression) bzw. der Absolutwerte der Parameter (LASSO) ein gewisses "Budget" nicht übersteigt (James et al., 2013, S. 220f).

stratifizierten Kreuzvalidierungsverfahrens ("k-Fold Cross-Validation") ermittelt. Dabei werden die Trainingsdaten zunächst auf Zufallsbasis in k Gruppen gleicher Größe unterteilt. Danach wird eine dieser Gruppen zur Validierung des Modells verwendet: Während das Modell auf die verbleibende $k - 1$ Gruppen trainiert wird, dient die Validierungsgruppe der Abschätzung der Modellgüte anhand der "Deviance". Darunter versteht man die Differenz zwischen der maximierten Log-Likelihood des geschätzten Modells und der Log-Likelihood, die ein "saturiertes" Modell ergeben würde.⁴¹⁾

Dieser Vorgang wird für unterschiedliche Werte von λ und schließlich für alle k Gruppen wiederholt. Jener Wert des Strafterms λ , welcher über alle k Gruppen hinweg die geringste Deviance ausweist, wird schließlich als "optimaler" Wert festgelegt und abschließend anhand der Testdaten evaluiert (vgl. *James et al.*, 2013, S. 181). Typische Werte für k sind z.B. $k = 5$ oder $k = 10$ (*Hastie et al.*, 2009, S. 242).

⁴¹⁾ Unter einem saturierten Modell versteht man ein theoretisches Modell, das für jede Beobachtung im Datensatz einen Parameter beinhaltet. Da ein solches Modell perfekt an die vorhandenen Daten angepasst wäre, dient die Log-Likelihood dieses Modells als theoretischer Maßstab für das beste denkbare Modell.

7.4 Deskriptive Statistiken

Übersicht A 2: **Neueintritte in Pflegegeldstufen, nach Geschlecht**

Pflegegeldstufe	Frauen	Männer	Anteil Frauen	Anteil Männer
1	50.857	29.352	55,64	43,65
2	18.037	14.535	19,73	21,61
3	11.545	11.184	12,63	16,63
4	6.204	6.752	6,79	10,04
5	3.493	3.895	3,82	5,79
6	912	1.128	1,00	1,68
7	353	401	0,39	0,60

Q: WIFO, DVSV.

Übersicht A 3: **Alter bei Neueintritt in Pflegegeld, nach Geschlecht**

Altersgruppe	Frauen	Männer	Anteil Frauen	Anteil Männer
60 bis 64 Jahre	4.321	4.760	4,73	7,08
65 bis 69 Jahre	6.891	6.868	7,54	10,21
70 bis 74 Jahre	11.551	9.396	12,64	13,97
75 bis 79 Jahre	22.371	15.152	24,48	22,53
80 bis 84 Jahre	23.232	14.519	25,42	21,59
85 bis 89 Jahre	16.831	12.127	18,41	18,03
90 bis 94 Jahre	5.534	3.979	6,05	5,92
95 und mehr Jahre	670	446	0,73	0,66

Q: WIFO, DVSV.

Übersicht A 4: **Veränderung der Pflegegeldstufe, nach Geschlecht**

Veränderung Pflegestufe	Frauen	Männer	Anteil Frauen	Anteil Männer
-6	13	12	0,01	0,01
-5	43	45	0,02	0,03
-4	223	177	0,09	0,12
-3	709	500	0,29	0,35
-2	2.087	1.296	0,86	0,91
-1	5.141	2.978	2,11	2,10
1	134.821	68.023	55,29	47,98
2	59.297	34.942	24,32	24,65
3	25.686	18.675	10,53	13,17
4	10.138	9.108	4,16	6,42
5	4.292	4.381	1,76	3,09
6	1.045	1.227	0,43	0,87
7	353	401	0,14	0,28

Q: WIFO, DVSV.