

*Prof. Dr. Thomas P. Zahn, DCC-Risikoanalytik*

---

# **KI-basierte Vorhersage beeinflussbarer Gesundheits- risiken mit Leistungsdaten**

---

Fachgruppe KRANKEN 28.April 2026

# Agenda

Risk analysis with health insurance claims

Smart Casaplus for Hospitalization Risk

AI based Risk Prediction and Treatment Effect estimation

AI based Nursing Risk

## 20 Jahre Risikoprognose mit Abrechnungsdaten der Krankenversicherungen

- 2006-2008 Accenture - Analytic Care Management Services (u.a. KKH, TK, Barmer, APKV, Central, CH-Helsana)
- 2008-2010 DxCG Gesundheitsanalytik/Verisk – Morbi-RSA (u.a. BVA/BAS, 219 GKV, UK-NHS, F-CNAMTS, P-ACC)
- seit 2010 DCC Risikoanalytik – RiskSmartSuite (u.a. 32 GKV, 3 PKV, 4 Klinik, 2 Arztnetz, 4 DiGA, BfArM-FDZ, PWC-vfa)
- 2011-2013 Elsevier - Health Risk Analytics – CaseSmartSuite (8 Pharma, 5 Kliniken, 5 PKV)
- 2011-2013 spectrumK - Health Risk Institute (Ingef) – ContractSmartSuite / EVA (63 GKV, EMA, 9 Förderprojekte)
- 2013-2019 AOK Nordost – GF Gesundheitswissenschaftliches Institut – (u.a. 35 Studien, 19 Förderprojekte, 5 Produkte)
- 2016-2024 bbw Hochschule – Prof Health Data Science - fib Forschungsinstitut/Prorektor Forschung (11 Förderprojekte)
- seit 2019 DSI – Data Science Institute – Health Data Analyst & Scientist – (5 Unternehmen, 4 HS, 6 Pflege, 2 Arztnetze)



praktischer  
Nutzen durch  
statistische  
und KI-  
Methoden



*KI basierte Analyse von Gesundheitsrisiken auf Basis von Abrechnungs- und Behandlungsdaten seit 2010 in D,A,CH,F,S,P,UK,US,CA*



*Risk Smart Suite seit 2010 - Module enthalten u.a.:*

**Smart Predict** – Prognose individueller und populationsbezogener Gesundheitsrisiken -> beeinflussbare Versicherte


**Smart Effect** – Evaluation auf Basis riskodjustierter Kontrollgruppen

## Welche Daten nutzt die DCC Forschungsdatenbasis

### Übersicht der Stamm und Leistungsdaten

Tabellenname	Erläuterung
<b>VERS_STAMM</b>	Versicherten-Stammdaten mit PID (Patienten Pseudonym), Geburts- und Sterbedatum sowie Geschlecht
<b>VERS_TAGE</b>	Versicherungszeiten mit erstem und letzten Versicherungstag und der Dauer der Versicherung
<b>VERS_SOZDEM</b>	Versicherungsdaten mit Versichertenart,-status, Nationalität, Tätigkeit und Gültigkeit
<b>VERS_REGION</b>	Regionalzuordnung auf Wohnortbasis mit PLZ, KGS, Länderschlüsseln sowie Auslandskennzeichen und Gültigkeit
<b>LEISTUNGSKOSTEN</b>	Leistungskosten der 5 Hauptleistungsbereiche, Dialysekosten sowie Gesamtleistungskosten
<b>KH_FALL</b>	Krankenhaus Fallnachweise mit DRG, Fall ID, Entlass- und Aufnahmedaten, IK, Abteilungsangaben sowie Rechnungsdaten
<b>KH_DIAGNOSE</b>	Krankenhaus Diagnosen mit ICD, Fall ID, Entlassungsdatum und Angaben zur Diagnoseart und Primär Kennzeichen
<b>KH_OPS</b>	Krankenhaus Prozeduren mit OPS, Operationsdatum, DRG, Entlassungsdatum, IK Nummern sowie weiteren Angaben zur Prozedur
<b>ARZT_FALL</b>	Einzelfallnachweise mit Datumsangaben, EFN-ID, Bestriebsstätten Pseudonym, Arzt Pseudonym, Fachgruppe sowie Leistungskosten
<b>ARZT_DIAGNOSE</b>	Einzelfallnachweise mit Zeitangaben, ICD, Diagnosesicherheit, Lokalisation und Angaben zum behandelnden Arzt
<b>ARZT_GOP</b>	GOP mit Behandlungsdatum, EFN-ID, Multiplikator und Angaben zum behandelnden Arzt
<b>AM_EVO</b>	Arzneimittelverordnungen mit PZN, ATC, Menge, Verordnungsdatum, Abgabedatum, Apotheken-ID, Arzt-ID, Kosten
<b>HIMI_EVO</b>	Hilfsmittelverordnungen mit EFN, ICD, Produktangaben und Kosten sowie Angaben zum Datum und verordnendem Arzt
<b>HEMI_EVO</b>	Heilmittelverordnungen mit EFN, ICD, Leistungsangaben und Kosten sowie Angaben zum Datum und verordnendem Arzt
<b>AU_FALL</b>	AU Bescheinigungen mit Beginn/Ende AU und KG sowie Hauptdiagnose, KG Tage & KG Betrag, Ursache und Vers.Gruppe
<b>AU_DIAGNOSE</b>	Diagnosen der AU Bescheinigungen, mit Diagnoseart, Anzahl und Ursache
<b>PFLEGE_GRADE</b>	Pflege Grade mit Beginn und Ende sowie Kennzeichen für Pflegeheim und Betreuer

# Welche Möglichkeiten bietet das nationale Forschungsdatenzentrum FDZ- Gesundheit

Das FDZ Gesundheit sitzt am BfArM:  Bundesinstitut für Arzneimittel und Medizinprodukte



- Alle Links und Downloads
- Antragsvorbereitung >
- Antragsportal
- Daten am FDZ Gesundheit >
- Datenbereitstellungsformen >

<https://www.forschungsdatenzentrum-gesundheit.de/>


eröffnet im  
Oktober 2025

Abrechnungsdaten  
aller 72 Mill GKV Vers.  
2009-2023

<https://www.forschungsdatenzentrum-gesundheit.de/infoportal/externe-seiten-auf-einen-blick>

<https://www.risikoanalytik.de/produkte/fdz-support.html>

<p><b>Datensatzbeschreibung</b></p> <p>Nutzen Sie zur Vorbereitung Ihres Antrags die Datensatzbeschreibung. So erkennen Sie schnell, welche Variablen für Ihr Forschungsvorhaben relevant sind und können diese im Antrag gezielt und präzise benennen.</p> <p>ZUR DATENSATZBESCHREIBUNG</p>	<p><b>Antragsportal des FDZ Gesundheit</b></p> <p>Reichen Sie Ihren Antrag beim FDZ Gesundheit einfach und sicher über unser digitales Antragsportal ein. Dort führen wir Sie Schritt für Schritt durch den gesamten Antragsprozess.</p> <p>ZUM ANTRAGSPORTAL</p>	<p><b>Public Use File</b></p> <p>Das Public Use File (PUF) bildet die Struktur des Echtdatensatzes ab. Es ermöglicht Ihnen, bereits vor der Antragstellung erste Auswertungsskripte zu entwickeln und methodische Abläufe zu testen.</p> <p>ZUM PUBLIC USE FILE DOWNLOAD</p>	<p><b>Trainingsangebot</b></p> <p>Wenn Sie die FDZ Forschungsdatenbasis genauer kennen lernen oder sich darauf vorbereiten wollen, effiziente Abfragen zur Ihrer Forschungsfrage zu erstellen, können Sie hier spezifische Trainingsangebote, Jupyter Notebooks und einen SQL CoPilot der DCC Risikoanalytik nutzen.</p> <p>ZUM TRAININGSANGEBOT</p>
--	---	--	--

 Als Partner des FDZ Gesundheit bieten wir Ihnen hier folgende Support Angebote:

- Trainingsumgebung** A) Trainings zum detaillierten Kennenlernen der Tabellen und Felder des aktuellsten Datenmodells. (im Aufbau)
- Studienpopulation synthetisieren** B) Eine *Jupyter Hub* Umgebung mit dem FDZ *PublicUseFile* des Datenmodells 3 (PUF-DM3) und konfigurierbarem Notebook zur Synthetisierung einer Studienpopulation für Ihre Forschungsfrage aus den relevanten PUF-DM3 Tabellen.
- SQL Script erstellen** C) Einen *SQL-CoPilot* zur Erstellung von effizient auf den FDZ Tabellen lauffähigen SQL Scripts auf Basis Ihrer synthetisierten Studienpopulation, z. B. zur Extraktion Ihrer Kohorte aus den FDZ Echtdate, aber auch zur Beantwortung Ihrer Forschungsfragen auf Basis Ihrer Kohorte.
- Beratungstermin buchen** Bei Interesse wenden Sie sich bitte an: [service@risikoanalytik.de](mailto:service@risikoanalytik.de) oder buchen Sie einen kostenfreien Beratungstermin über die Taste Termin buchen

# Agenda

Risk analysis with health insurance claims

Smart Casaplus for Hospitalization Risk

AI based Risk Prediction and Treatment Effect estimation

AI based Nursing Risk

# Beispiel 1: Smart Casaplus Programm für risikobasierte Vermeidung von Krankenhaus Kosten

Wissenschaft

MVF 01/2024 17. Jahrgang 02.02.2024

Prof. Dr. rer. pol. Stephan Burger  
 Prof. Dr.-Ing. Thomas P. Zahn MSc  
 Markus M. Müller MA

## smartCasaplus – ein risikobasiertes geriatrisches Gesund- heitscoaching für multi- morbide Versicherte

Chronische Krankheiten bestimmen mehr und mehr das Krankheitsgeschehen einer alternden Gesellschaft und führen zu veränderten Versorgungsbedarfen in der Gesundheitsversorgung. Hierbei ist insbesondere eine stärker präventiv ausgerichtete Gesundheitsversorgung von Bedeutung, da die

### Zusammenfassung

Das geriatrische Schulungsprogramm Casaplus der MedicalContact AG bietet bereits seit 2007 telefonische Gesundheitscoachings für multimorbide geriatrische Versicherte, von dem bereits ca. 30.000 Versicherte profitiert haben. Die erfolgskritischen Faktoren für ein solches Programm sind die risikobasierte Identifikation der geeigneten Kandidat:innen, die Teilnehmergewinnung sowie die individuelle Betreuung durch speziell ausgebildete medizinische Fachkräfte. Seit 2017 wird das Casaplus-Programm gemeinsam mit der VIACTIV-Krankenkasse und der DCC Risikoanalytik GmbH um KI-basierte Komponenten sowohl bei der Prädiktion von beeinflussbaren Risiken im Rahmen der Identifikation als auch bei der Nutzung risikoadjustierter Propensity Scores im Rahmen der Evaluation zum smartCasaplus-Coaching weiterentwickelt. Die aktuellen Evaluationsergebnisse der VIACTIV-Krankenkasse zeigen, dass damit unter Berücksichtigung der Zuweisungseffekte und Programmkosten ein Return-on-Invest von 1,7 im Mittel über die ersten 4 Programmjahre erreicht wird.

### Schlüsselwörter

Gesundheitsdatennutzung, Multimorbidität, Gesundheitscoaching, Selbstmanagement, Risikoprognose, KI, Propensity Score Matching, Evaluation

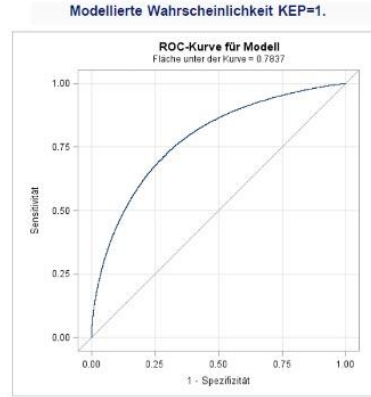
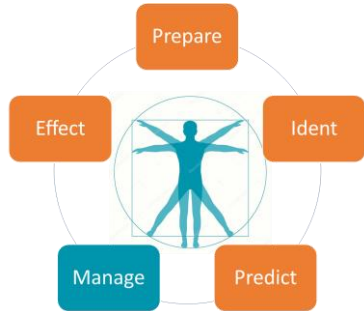
### Crossref/doi

<http://doi.org/10.24945/MVF.01.24.1866-0533.xxxx>

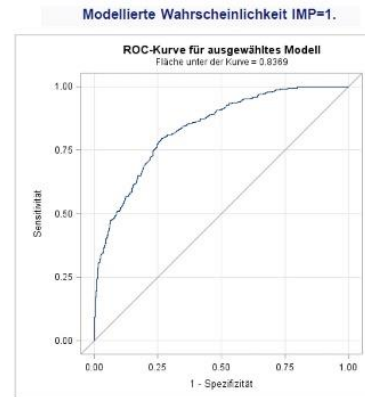
selten verursacht Multimedikation zusätzlich unspezifische Beschwerden wie Müdigkeit, Appetitlosigkeit, Schwindel, Verwirr-

## Beispiel 1: Smart Casaplus: Start 2007 → seit 2017 Fokussierung durch adaptive Risikoprognose

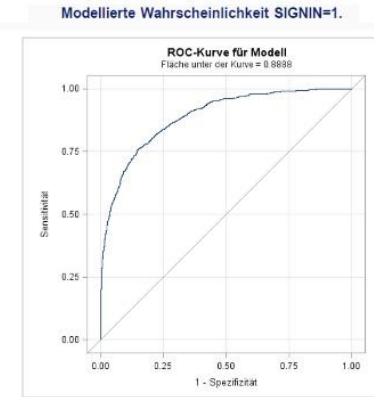
jährliche  
evaluationsbasierte  
Anpassung der 40-50  
beeinflussbaren Target  
ICDx Gruppen



kontinuierliche lernendes  
programmspezifisches KEP  
(Krankenhaus-Einweisungs-  
Prognosemodell)



auf Evaluations-  
ergebnissen lernendes  
adaptives IMP (Impact)  
Beflussbarkeitsmodell

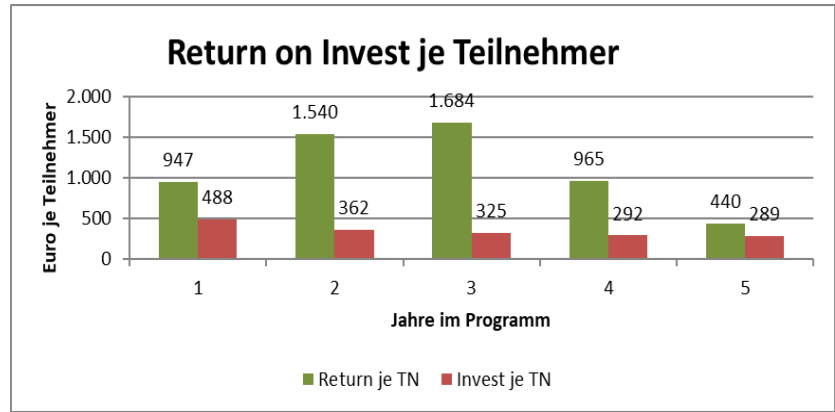
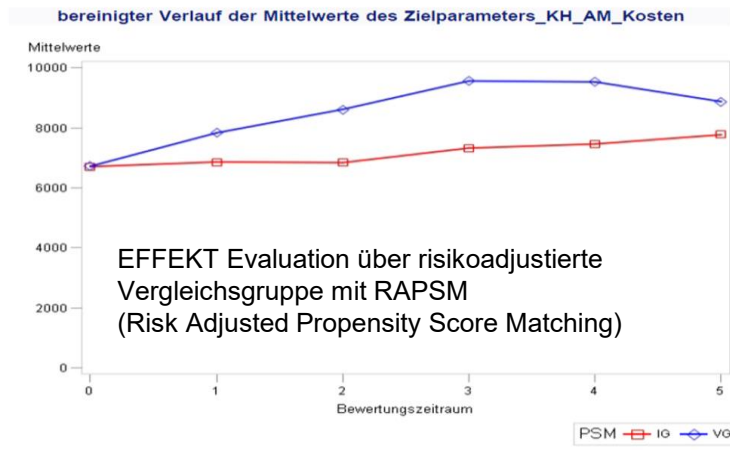


auf Gewinnungserfolg  
lernendes adaptives  
SIGNIN  
Einschreibungsmodell

PID individueller smartCasaplus Combi Score (60% KEP, 30% IMP, 10% SIGNIN)



# Beispiel 1: Smart Casaplus: weniger Invest und RoI von 3,2 in 5 Programmjahren



	Programmjahr	IGmembers	VGmembers	IG_Mean	VG_Mean	VGIgmeanDiff	pValue	Programm
Matching - 4 Quartale vor Interventionsbeginn	0	1635	1635	6.709	6.724	14	0,9514	-
1.-4. Quartal der Intervention	1	1627	1627	6.865	7.834	970	<b>0,0009</b>	-1.554.098
5.-8. Quartal der Intervention	2	1106	1106	6.843	8.618	1.775	<b>0,0000</b>	-1.947.133
9.-12. Quartal der Intervention	3	803	803	7.328	9.565	2.236	<b>0,0000</b>	-1.784.361
13.-16. Quartal der Intervention	4	555	555	7.463	9.532	2.069	<b>0,0003</b>	-1.140.481
17.-20. Quartal der Intervention	5	268	268	7.773	8.875	1.102	0,1674	-291.519
1. Programmjahr D2D					155	1.111	<b>955</b>	
2. Programmjahr D2D					134	1.894	<b>1.761</b>	
3. Programmjahr D2D					619	2.841	<b>2.222</b>	
4. Programmjahr D2D					754	2.809	<b>2.055</b>	
5. Programmjahr D2D					1.064	2.151	<b>1.088</b>	
1.-5. Programmjahr D2D							<b>8.081</b>	
Extremwertbereinigung								
paarweise kombiniert > 2,0 σ								

**smart Casaplus PQ Programmeffekte bezogen auf Gesamtleistungskosten und Zuweisungen**

nach der Compare to Control Messmethode

Programmjahr	TN	HMG Zuw je TN	GLK je TN	Return je TN	Invest je TN	Netto je TN	RoI	Return Summe
1.-4. Quartal der Intervention	1	1.758	-18	965	488	459	1,9	1.664.544
5.-8. Quartal der Intervention	2	1.198	-555	2.095	362	1.178	4,3	1.844.389
9.-12. Quartal der Intervention	3	851	-1.057	2.741	325	1.360	5,2	1.433.505
13.-16. Quartal der Intervention	4	606	-1.673	2.639	292	673	3,3	584.862
17.-20. Quartal der Intervention	5	294	-1.504	1.944	289	151	1,5	129.362
<b>total 1-5 Jahr</b>		<b>-</b>	<b>4.808</b>	<b>10.384</b>	<b>5.576</b>	<b>1.755</b>	<b>3,2</b>	<b>5.527.300</b>

\*In den Investitionskosten des 1. Programmjahres sind auch die Gewinnungskosten im Jahr 0 enthalten

## Beispiel 1: Smart Casaplus: Zusammenfassung

Die smart Casaplus Evaluation auf Basis der Quartalskohorten 1-21 zeigt erstmals für den **gesamten Programmverlauf von 5 Programmjahren**:

- eine signifikante Dämpfung der Krankenhauskosten um 6.279 € je Teilnehmer
- eine signifikante Dämpfung der Krankenhauseinweisungen um 25%
- eine signifikante Dämpfung der Gesamtleistungskosten um 10.384 € je Teilnehmer
- eine Reduktion der Morbiditätszuweisungen um 4.808 € je Teilnehmer
- so dass nach Abzug der Zuweisungseffekte einen Return von 5.576 € je Teilnehmer
- die Investitionen von 1.755 € je Teilnehmer überkompensiert
- und **ein Return on Invest von 3,2 entsteht**

Bezogen auf alle Teilnehmer erzeugt smartCasaplus einen **Netto Return von 5,53 Mill €**

## Beispiel 1: Smart Casaplus: Take Aways

Individuelle Kostenrisiken sind kurz- bis mittelfristig beeinflussbar wenn

- Die Risikoprognose auf beeinflussbare Risiken (Targets) fokussiert wird
- Die Risikoprognose mit Beeinflussbarkeit und Teilnahme-Wahrscheinlichkeit kombiniert wird
- Die Prognosemodelle sich kontinuierlich an die Datenlage und Umfeldbedingungen anpassen
- Die Evaluationsergebnisse kontinuierlich zur Fokussierung der Modelle genutzt werden
- Ausreichende GAPs zur Gewinnung und Intervention bei der Modellierung beachtet werden

# Agenda

Risk analysis with health insurance claims

Smart Casaplus for Hospitalization Risk

AI based Risk Prediction and Treatment Effect estimation

AI based Nursing Risk

## Beispiel 2: AI Based Risk Prediction and Treatment Effect Estimation – AIR\_PTE

Partner:



[www.ai-evidence.de](http://www.ai-evidence.de)



 Springer




Prävention und Gesundheitsförderung

Originalarbeit

Privat Gesundheitsf  
<https://doi.org/10.1007/s11553-023-01048-1>  
Eingegangen: 9. Januar 2023  
Angenommen: 8. April 2023

© Der/die Autor(en) 2023



Thomas P. Zahn<sup>1,2</sup>  · Eric Hesse<sup>3</sup> · Anamaria Herrera-Rodriguez<sup>1</sup> · Jochen Walker<sup>3</sup> · Dirk Enders<sup>3</sup>

<sup>1</sup>DCC Risikoanalytik GmbH Berlin, Berlin, Deutschland

<sup>2</sup>bbw Hochschule, Berlin, Deutschland

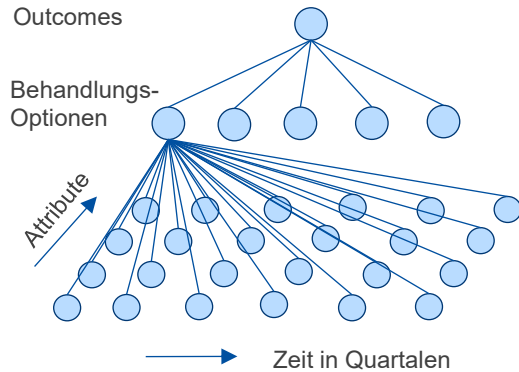
<sup>3</sup>InGeF – Institut für angewandte Gesundheitsforschung Berlin, Berlin, Deutschland

**KI basierter Rapid Evidence  
Generator mit risikoadjustiertem  
Propensity Score Matching**

Supported by:



## Beispiel 2: AIR\_PTE – Rapid Evidence Generator (REG)



 EVA: Self-Service Analytics



 SAHRA  
Smart Analysis Health Research Access

 macadamian  
healthconnect

Mit KI Methoden **Real World Evidenz** zur Wirkung von Behandlungsoptionen aus anonymisierten Abrechnungsdaten der nationalen Krankenversicherungen gewinnen und als **Rapid Evidence Generator** für Entscheider und Behandler praktisch nutzbar machen.

## Beispiel 2: AIR\_PTE – REG: Behandlungsoptionen für Blutungen bei Tiefer Venen Trombose (VTE)

### Rapid Evidence Generator (REG) Principle



- Identify de-personalized retrospective study population
- Identify Outcomes of interest in study population
- Identify Intervention groups with specific treatments
- Apply smart grouping algorithm to de-personalize study population
- Obtain all confounding attributes of PID in study population

### Venous Thrombo-Embolism (VTE) Example



- Patient Pseudonyms (PID) with incident VTE ICD10 Diagnostic Code
- PID with bleeding ICD 10 Code within 181 days after VTE Diagnose
- PID treated with Antiplet or Heparin within 14 days after VTE Diagnose
- all attributes with more than 20 occurrences in study population
- Morbidity (ICD-X), Drugs (ATC-X), Age, Sex within 365 days before VTE Diagnose

### REG\_BASE Results

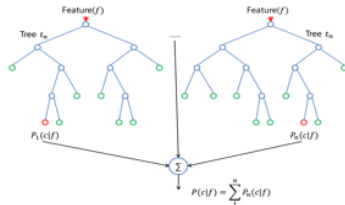
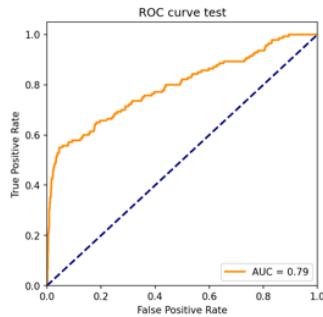


- 54.065 PID with incident VTE between 2011 and 2021
- 2.757 PID with bleeding in 181 days  
Prevalence 5 %
- 5 treatment options  
(DOAC, VKA, Heparin, Antiplet, None)  
1.153 PID with Antiplet, 5.254 with Heparin
- 25 Age-Sex-Groups,  
439 Outpatient Diagnostic Groups A-ICD-X,  
16 Inpatient Diagnostic Groups S-ICD-X,  
101 Drug Substance Groups - ATC-X  
in total 581 potential risk confounders

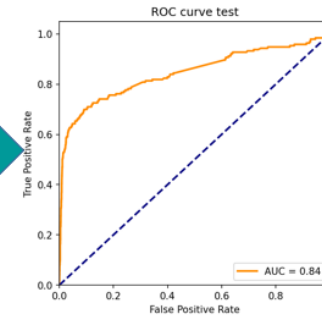
## Beispiel 2: AIR\_PTE – REG: VTE

2. create the best outcome risk model based on all high dimensional population attributes

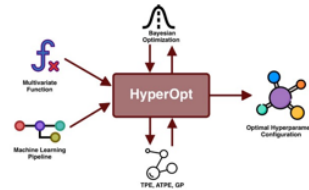
inital bleeding risk model



optimized bleeding risk model



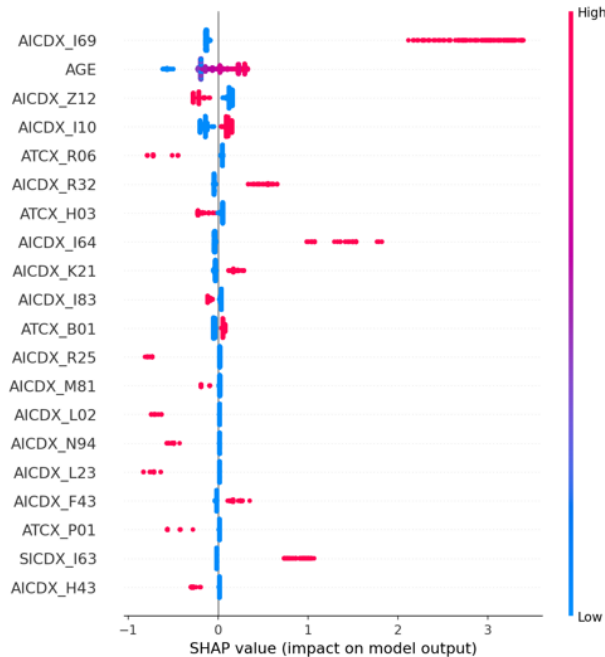
Bayesian Hyperparameter  
Optimization of Gradient  
Boost Model



Bleeding Risk Model		
metric	training	testing
auc	0,94	0,84
accuracy	96%	96%
sensitivity	59%	50%
specificity	98%	99%
ppv	67%	74%
prevalence	6%	6%

## Beispiel 2: AIR\_PTE – REG: VTE

### 3. identify all significant outcome risk confounders based on Shapley Analysis



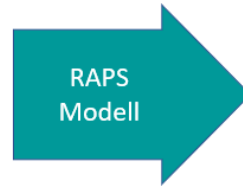
71 bleeding  
risk  
confounders  
identified

column_name	shap_importance
AICDX_I69	0.2388711442241687
AGE	0.19338454205999075
AICDX_Z12	0.17155987391562
AICDX_I10	0.13248673225124358
ATCX_R06	0.09237450590836623
AICDX_R32	0.08266798335167604
ATCX_H03	0.07738900803567948
AICDX_I64	0.07162208677442355
AICDX_K21	0.05388143585047584
AICDX_I83	0.050902568683450555
ATCX_B01	0.04926442130056033
AICDX_R25	0.03784622061028651
AICDX_M81	0.03509648205519373
AICDX_L02	0.03463299458671306
AICDX_N94	0.033640479905783206
AICDX_L23	0.032268623569867434
AICDX_F43	0.03167227790059113
ATCX_P01	0.031073176977821367
SICDX_I63	0.029837762868642162
AICDX_H43	0.029475877531911176
AICDX_J40	0.026999061753443576
SEX	0.02427926195962105
AICDX_M25	0.02383115819585383
ATCX_D06	0.02317838211491122
AICDX_G81	0.023102512750263197
ATCX_B03	0.02209185041675485
AICDX_R49	0.021344504797127732
AICDX_K59	0.01838081231480998

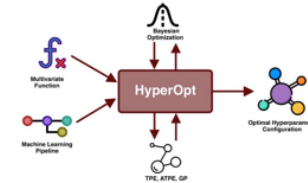
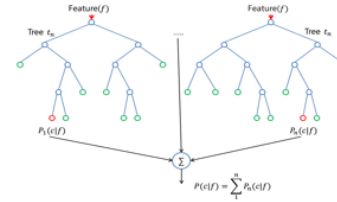
## Beispiel 2: AIR\_PTE – REG: VTE

### 4. find the best propensity model based on outcome risk confounders

column_name	shap_importance
AICDX_I69	0.2388711442241687
AGE	0.19338454205999075
AICDX_Z12	0.17155987391562
AICDX_I10	0.13248673225124358
ATCX_R06	0.09237450590836623
AICDX_R32	0.08266798335167604
ATCX_H03	0.07738900803567948
AICDX_I64	0.07162208677442355
AICDX_K21	0.05388143585047584
AICDX_I83	0.050902568683450555
ATCX_B01	0.04926442130056033
AICDX_R25	0.03784622061028651
AICDX_M81	0.03509648205519373
AICDX_L02	0.03463299458671306
AICDX_N94	0.033640479905783206
AICDX_L23	0.032268623569867434
AICDX_F43	0.03167227790059113
ATCX_P01	0.031073176977821367
SICDX_I63	0.029837762868642162
AICDX_H43	0.029475877531911176
AICDX_J40	0.026999061753443576
SEX	0.02427926195962105
AICDX_M25	0.02383115819585383
ATCX_D06	0.02317838211491122
AICDX_G81	0.023102512750263197
ATCX_B03	0.02209185041675485
AICDX_R49	0.021344504797127732
AICDX_K59	0.01838081231480998



71  
attributes for  
propensity  
model

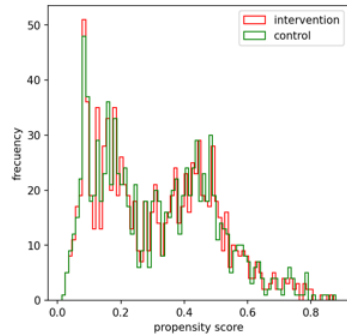


risk adjusted Propensity Model		
metric	training	testing
auc	0,85	0,81
accuracy	84%	83%
sensitivity	23%	21%
specificity	98%	97%
ppv	70%	58%
prevalence	18%	18%

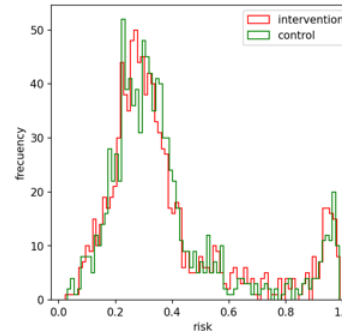
## Beispiel 2: AIR\_PTE – REG: VTE

### 6. after RAPS Matching - outcome risk adjusted control groups are available

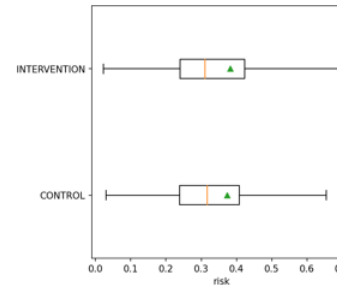
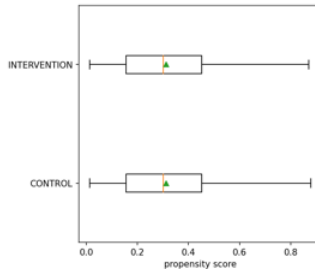
RA-Propensity Score Matching



outcome risk distribution

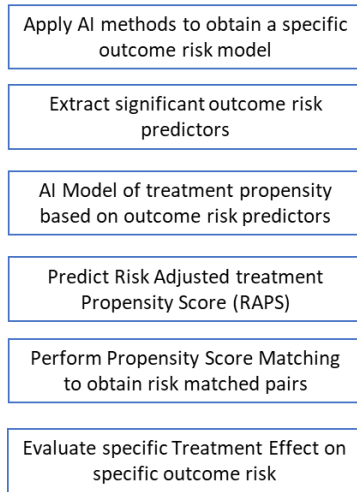


p-value 0.7537 > 0.05

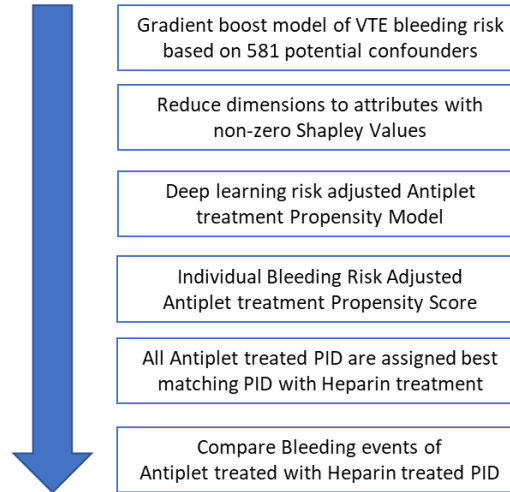


## Beispiel 2: AIR\_PTE – REG: VTE

### Rapid Evidence Generator (REG) Principle



### Venous Thrombo-Embolism (VTE) Example



### REG\_EFFECT Module Example Results (15 min)

Bleeding risk AUC 0.84 on 50% test sample  
Sensitivity 0.54 , Specificity 0.99 at cutoff 0.5

**Shapley analysis ->  
71 significant confounding attributes  
contributing to bleeding risk**

Antiplatelet Treatment Propensity  
AUC 0.81 on 50% test sample  
Sensitivity 0.69 , Specificity 0.97 at cutoff 0.5

SUMMARY	VALUE
number_of_pairs	532
mean_target_intervention	0,05
mean_target_control	0,10
average_treatment_effect	-0,05
p-value effect fisher test	0,003
p-value risk kolmogorov-test	0,754

**Treatment Effect of  
VTE Patients  
treated  
with Antiplatelet  
compared to  
Heparin  
with regard to  
bleeding  
within 6 Month  
after Treatment  
50% (0.05),  
p = 0.003 , n= 532**

## Beispiel 2: AIR\_PTE – Methods - applied to different VTE-Treatments on Research Database

Intervention	Control	#Patients Interv.	# Patients Control	# matched pairs	p-value risk adj. *	# target interv.	# target control	# no target interv.	# no target control	target risk Interv.	target risk control	ATT_EFFEKT	p-value ATT
VKA	NONE	9635	71274	7667	0,068	414	374	7253	7293	5,4%	4,9%	0,5%	0,1500
DOAC	NONE	30912	71274	25173	0,330	1244	1030	23929	24143	4,9%	4,1%	0,9%	0,0000
HEPARIN	NONE	21922	71274	19815	0,270	872	735	18943	19080	4,4%	3,7%	0,7%	0,0005
ANTIPILET	NONE	1958	71274	1191	0,064	84	120	1107	1071	7,1%	10,1%	-3,0%	0,0120
DOAC HEPARIN	NONE	2971	71274	2519	0,080	118	86	2401	2433	4,7%	3,4%	1,3%	0,0260
DOAC	VKA	30912	9635	8087	0,280	315	380	7772	7707	3,9%	4,7%	-0,8%	0,0130
HEPARIN	VKA	21922	9635	7592	0,025	306	339	7286	7253	4,0%	4,5%	-0,4%	0,1970
ANTIPILET	VKA	1958	9635	1263	0,980	85	119	1178	1144	6,7%	9,4%	-2,6%	0,0150
DOAC HEPARIN	VKA	2971	9635	2260	0,688	99	86	2161	2174	4,4%	3,8%	0,6%	0,3670
VKA	HEPARIN	9635	21922	7636	0,089	310	343	7326	7293	4,1%	4,5%	-0,4%	0,2005
DOAC	HEPARIN	30912	21922	18801	0,580	704	719	18097	18082	3,7%	3,8%	-0,1%	0,7051 Cancer
HEPARIN_VKA	HEPARIN	21922	21922	5338	0,809	165	169	5173	5169	3,1%	3,2%	-0,1%	0,8675
ANTIPILET	HEPARIN	1958	21922	1042	0,970	51	93	991	949	4,9%	8,9%	-4,0%	0,0004
DOAC HEPARIN	HEPARIN	2971	21922	2566	0,506	74	84	2492	2482	2,9%	3,3%	0,4%	0,4671

\* Note: For the risk-adjusted p-value: it is based on the Kolmogorov-Smirnov Test. If the p-value is high; we cannot reject the hypothesis that the distributions of the outcome risk between intervention and control samples are the same

**The study protocol used to identify the VTE population is based on the VTE study defined in the AIR\_PTE project to compare VKA (control) and DOAC (intervention)**

The VTE population of about 125.000 PID is identified on ingef\_FDB years 2012-2019 (Status 12/21)

**Target (Outcome): Bleeding Event between day 14 and 181 after initial VTE diagnosis**

Observational Period: 365 days before initial VTE diagnosis

**Method: Risk Adjusted Propensity Score Matching (RA\_PSM)**

The RA\_PSM Model is based on a GB Model using the significant impact factors of the initial optimized AI based Outcome Risk Model only

Nutzbar für Hinweise auf alternative Behandlungsoptionen bei Erstattung von Verordnungen

## Beispiel 2: AIR\_PTE: Take Aways

Der Effekt spezifischer Interventionen auf individuelle (Kosten) Risiken ist identifizierbar wenn

- Deren Wirkung auf historischen Abrechnungsdaten analysierbar ist
- Ein signifikanter Effekt auf den Outcome nachweisbar ist
- Die Effektbewertung im Vergleich zu risikoadjustierten Kontrollgruppen erfolgt
- Die Kontrollgruppen sowohl bezüglich Intervention als auch Outcome Risiko adjustiert sind
- Verschiedene Interventionsoptionen verglichen werden können

Nutzer:



# Agenda

Risk analysis with health insurance claims

Smart Casaplus for Hospitalization Risk

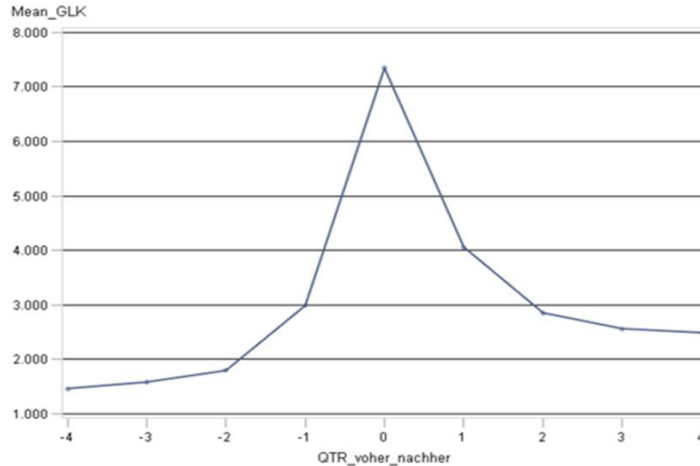
AI based Risk Prediction and Treatment Effect estimation

AI based Nursing Risk

## Beispiel 3: risikobasierte Pflegeberatung

Die Gesamtleistungskosten haben einen Peak bei Eintritt in die Pflege,

Verlauf der GLK in den 4 Quartalen vor und nach dem Erst Eintritt in die Pflege



QTR_voher_nachher	ANZ_PID	Sum_GLK	Mean_GLK
-4	63200	92.372.644	1.462
-3	65695	103.864.054	1.581
-2	67933	121.841.327	1.794
-1	70124	209.328.877	2.985
0	72432	532.220.715	7.348
1	73308	297.691.046	4.061
2	72768	207.595.235	2.853
3	71504	183.139.502	2.561
4	67085	166.950.434	2.489

Basis: 77.600 PID > 60 mit Ersteintritt in die Pflege zwischen 2014 und 2025

## Beispiel 3: risikobasierte Pflegeberatung

... dessen Treiber zu >40% vermeidbare KH Fälle wegen Sturz, Demenz und Immobilität sind.

ICDx_BEZEICHNUNG	ANZAHL_FALLE	SUM_KH_KOSTEN	Target	Legende
S72.- Fraktur des Femurs	1387	13.958.026	1	1 Sturz
I61.- Intrazerebrale Blutung	367	9.112.631	1	2 Demenz
R26.- Störungen des Ganges und der Mobilität	1038	7.604.482	3	3 Immobilität
S06.- Intrakranielle Verletzung	845	6.628.919	1	
S71.- Offene Wunde der Hüfte und des Oberschenkels	615	6.580.983	1	
S32.- Fraktur der Lendenwirbelsäule und des Beckens	721	5.284.348	1	
T84.- Komplikationen durch orthopädische Endoprothesen, Implantate oder Transplantate	321	4.242.215	3	
F33.- Rezidivierende depressive Störung	1081	3.938.265	2	
S42.- Fraktur im Bereich der Schulter und des Oberarmes	449	3.843.667	1	
M48.- Sonstige Spondylopathien	472	3.283.610	3	
M16.- Koxarthrose [Arthrose des Hüftgelenkes]	415	3.080.764	3	
F05.- Delir, nicht durch Alkohol oder andere psychotrope Substanzen bedingt	429	3.051.701	2	
T82.- Komplikationen durch Prothesen, Implantate oder Transplantate im Herzen und in den Gefäßen	252	2.587.973	3	
M17.- Gonarthrose [Arthrose des Kniegelenkes]	320	2.351.949	3	
G62.- Sonstige Polyneuropathien	173	2.323.343	3	
M80.- Osteoporose mit pathologischer Fraktur	274	2.079.418	1	
S82.- Fraktur des Unterschenkels, einschließlich des oberen Sprunggelenkes	262	2.024.736	1	
S22.- Fraktur der Rippe(n), des Sternums und der Brustwirbelsäule	339	1.999.425	1	
S41.- Offene Wunde der Schulter und des Oberarmes	220	1.935.630	1	
T85.- Komplikationen durch sonstige interne Prothesen, Implantate oder Transplantate	88	1.913.195	3	
S31.- Offene Wunde des Abdomens, der Lumbosakralgegend und des Beckens	215	1.906.702	1	
F32.- Depressive Episode	762	1.891.260	2	

## Beispiel 3: risikobasierte Pflegeberatung als Intervention zur proaktiven Kostendämpfung

### Rechtsgrundlagen SGB XI (GKV & PKV)

- § 7a SGB XI** Zentrale Anspruchsgrundlage für die **individuelle Pflegeberatung**: Beratung und Hilfestellung, Ermittlung des Hilfebedarfs, *Versorgungsplan*, Hinwirken auf Leistungen, *Begleitung/Anpassung*. Der Anspruch besteht auch schon bei Antragstellung, *wenn Hilfe- und Beratungsbedarf erkennbar* ist.
- § 7a Abs. 5** Speziell relevant für Private: Private Pflege-Pflichtversicherer können Pflegeberater der Pflegekassen nutzen oder untereinander Vereinbarungen zur Bereitstellung von Pflegeberatern treffen.
- § 7b SGB XI** **Pflicht zum Beratungsangebot**: Nach erstmaligem Antrag oder erklärtem Begutachtungsbedarf muss ein *Beratungstermin innerhalb von zwei Wochen angeboten oder ein Beratungsgutschein* ausgestellt werden. **Abs. 4 ordnet an, dass diese Regeln für private Pflege-Pflichtversicherer entsprechend gelten.**

### Rechtsgrundlagen SGB V (GKV)

- § 25b Datengestützte Erkennung individueller Gesundheitsrisiken durch die Kranken- und Pflegekassen**
- (1) Die Kranken- und Pflegekassen können zum Gesundheitsschutz eines Versicherten datengestützte Auswertungen vornehmen und den Versicherten auf die Ergebnisse dieser Auswertung hinweisen, soweit die Auswertungen den folgenden Zwecken dienen:
1. der Erkennung von seltenen Erkrankungen,
  2. der Erkennung von Krebserkrankungen,
  3. der Erkennung von schwerwiegenden Gesundheitsgefährdungen, die durch die Arzneimitteltherapie entstehen können,
  - 4. der Erkennung einer noch nicht festgestellten oder drohenden Pflegebedürftigkeit nach § 14 des Elften Buches,**
  5. der Erkennung ähnlich schwerwiegender Gesundheitsgefährdungen, soweit dies aus Sicht der Kranken- und Pflegekassen im überwiegenden Interesse der Versicherten ist, oder
  6. der Erkennung des Vorliegens von Impfindikationen für Schutzimpfungen, die von der Ständigen Impfkommision nach § 20 Absatz 2 des Infektionsschutzgesetzes empfohlen sind.
  - 7. Der Erkennung von schwerwiegenden Erkrankungen des Herz-Kreislauf-Systems oder eines hierfür erhöhten Risikos.**

\* Anpassungsvorschläge im GeDIG Referentenentwurf vom 1.4.2026

## Beispiel 3: risikobasierte Pflegeberatung

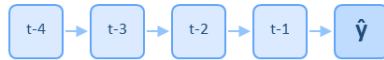
Wenn die beeinflussbaren KH Fälle bei 30% der Versicherten vermieden werden können, ergibt sich bei 300€ je Pflegeberatung und 500 € für Pflegeleistungen ein RoI von 1,6

<b>zusätzliche KV Kosten im Jahr des Eintritts in die Pflege je PID</b>	<b>10.072 €</b>
Anteil beeinflussbare KH Fälle & Sturzfolgen (Typ 1-3)	42%
<b>beeinflussbare KV Kosten je PID im Jahr des Eintritts in die Pflege</b>	<b>4.183 €</b>
Einflussquote (PPV) des Prognosemodells	30%
<b>Einsparpotential je Hochrisiko PID mit Intervention</b>	<b>1.267 €</b>
angebotene Interventionskosten für Pflegeberatung je PID	300 €
geschätzte Interventionskosten für Pflegeleistungen je PID	500 €
Netto Effekt je PID mit Intervention	467 €
<b>Return on Invest je PID</b>	<b>1,6</b>



## Beispiel 3: risikobasierte Pflegeberatung : Pflegerisiko – LSTM vs. Regression Trees

### LSTM bevorzugen, wenn ...

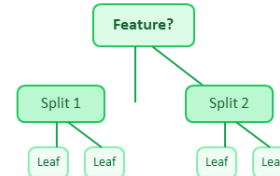


Sequenzverarbeitung

#### Typische Vorteile

- Zeitliche Abhängigkeiten, Verzögerungen und Kontext werden explizit mit modelliert.
- Kann komplexe nichtlineare Muster aus langen Sequenzen lernen.
- Geeignet für Mehrschritt-Prognosen, Anomalieerkennung und NLP-nahe Aufgaben.

### Regression Trees, wenn ...



#### Typische Vorteile

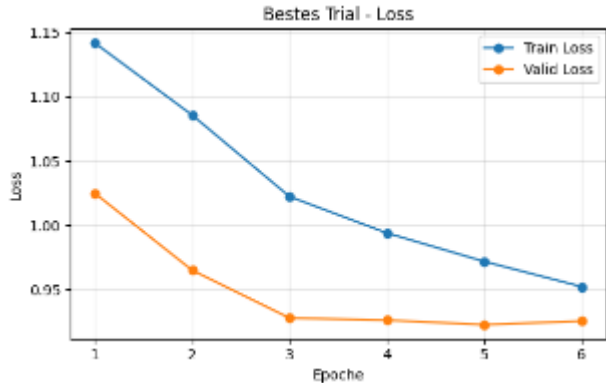
- Funktioniert stark auf strukturierten Tabellen mit numerischen und kategorialen Merkmalen.
- Regeln, Schwellenwerte und Feature-Effekte sind für Fachbereiche gut nachvollziehbar.
- Schnell zu trainieren und oft schon mit kleineren Datensätzen nützlich.



# Beispiel 3: risikobasierte Pflegeberatung : Pflegerisiko – LSTM

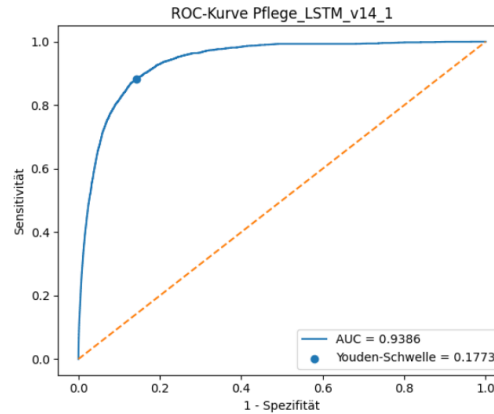
## Hyperparameter Optimierung mit Valid

Best params: {'hidden\_size': 512, 'num\_layers': 3, 'dropout': 0.1, 'dense\_hidden\_size': 128, 'batch\_size': 128, 'eval\_batch\_size': 64, 'learning\_rate': 0.001, 'weight\_decay': 1e-05, 'max\_epochs': 8, 'patience': 3}  
 Best validation AUC: 0.8260835281480672



## Prognosegüte mit Test

Test AUC: 0.93859401209284  
 Youden threshold: 0.1772514432668686  
 Sensitivität: 0.883545069570478  
 Spezifität: 0.8580949489328837



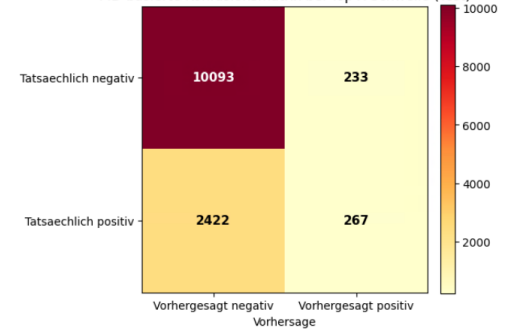
## Präzision bei TOP 500 mit Test

Top-X threshold (from test split):

0.8731461763381958

"precision" 0.534 "prevalence" 0.012

PID-basierte Konfusionsmatrix bei Top-X-Schwelle (Test)



rank	pid	score	target
0	024602795	0,9398	0
1	006288415	0,9292	1
2	127928651	0,9249	0
3	032564604	0,9249	1
4	003393644	0,9241	1
5	008394404	0,9216	1
6	035030090	0,9207	1
7	000454217	0,9203	0
8	011653391	0,9197	0
9	027800319	0,9196	1

## Beispiel 3: risikobasierte Pflegeberatung : VIVO Pilot 2026

### Versicherten Information online und als Aushang

<https://www.viactiv.de/leistungen/gesundheits-und-medizin/innovative-medizin/viactiv-vorsprung>

#### **GUT VORBEREITET AUF DEN PFLEGEFALL**

Wussten Sie, dass es typische Ereignisse gibt, die mit einer beginnenden Pflegebedürftigkeit zusammenhängen können? Stürze gehören zum Beispiel dazu. Die möchten wir natürlich vermeiden. Genau wie den Fall, dass Sie unvorbereitet vor der Herausforderung Pflege stehen. Deshalb ermitteln wir Ihr persönliches Risiko, pflegebedürftig zu werden. Und bieten die Möglichkeit, sich frühzeitig beraten zu lassen. Außerdem helfen wir Ihnen etwa dabei, einen Pflegegrad zu beantragen – oder unterstützen Sie mit vorsorgenden Pflegeleistungen.



**GESUNDHEITSDATEN GEZIELT NUTZEN.  
GESUNDHEITSRISIKEN EFFEKTIV VORBEUGEN.**

#### **VIACTIV Vorsprung**

Der § 25b SGB V erlaubt es uns als Ihrer Kranken- und Pflegekasse, gesundheitsbezogene Daten auszuwerten. Dafür haben wir das Konzept VIACTIV Vorsprung entwickelt.

Darum geht es konkret: Wir nutzen Ihre Gesundheitsdaten, um individuelle Gesundheitsrisiken frühzeitig zu erkennen. Und Ihnen in Zukunft Leistungen anzubieten, die optimal zu Ihnen und Ihrer Lebenssituation passen. Die Datenauswertung wird sich dabei auf die frühzeitige Erkennung einer noch nicht festgestellten Pflegebedürftigkeit beziehen.

##### **WELCHE DATEN WIR NUTZEN**

- Krankenhaus- und Reha-Aufenthalte
- Arztrechnungsdaten
- Verordnungen von Arznei-, Heil- und Hilfsmitteln
- Pflegegrad- und Leistungsdaten

##### **IHR GUTES RECHT**

Der Nutzung Ihrer Daten können Sie jederzeit widersprechen. Natürlich entstehen Ihnen dadurch in Zukunft keine Nachteile.

Füllen Sie einfach unser Widerspruchsformular aus unter [viactiv.de/viactiv-vorsprung](https://www.viactiv.de/viactiv-vorsprung). Oder richten Sie Ihren Widerspruch formlos mit Ihrem vollständigen Namen und Ihrer Versichertennummer an die

VIACTIV Kranken- oder Pflegekasse  
Suttner-Nobel-Allee 3-5  
44803 Bochum  
0800 2221211  
E-Mail: [service@viactiv.de](mailto:service@viactiv.de)

##### **ALLES ZUM DATENSCHUTZ**

Die Verarbeitung erfolgt gemäß DS-GVO und SGB X. Unter [viactiv.de/viactiv-vorsprung](https://www.viactiv.de/viactiv-vorsprung) finden Sie Details.

## Beispiel 3: risikobasierte Pflegeberatung : VIVO Pilot 2026

### Beratungsgutschein für TOP 500 Risiko PID

#### Beratungsgutschein – VIVO Pflegeberatung

für eine qualifizierte Beratung durch eine Pflegeberater\*in. Auf Wunsch gerne bei Ihnen zu Hause.

#### Inhalte der Beratung:

- Mobilität und Wohnumfeld
- Alltagsaktivitäten und Versorgungssituation
- Soziale Beziehungen
- Ansprüche auf Sozialleistungen
- Gesunderhaltung und Risiken
- Praktische Tipps zu Alltagshilfen und haushaltsnahen Dienstleistungen
- ggf. Pflegegrad-Berechnung
- Unterstützung bei Antragstellungen
- Dokumentation und Maßnahmenfestlegung

#### Kostenlos einzulösen bei unserem Kooperationspartner spectrumK.

Ihr Ansprechpartner: Herr Michael Vögeding

Telefonische Beratung:

Mo. - Fr.: 08.00 - 18.00 Uhr

Hotline: 0800-7237267

Fax: 030 586 945 39434

per E-Mail an: [pflgeberatung@spectrumK.de](mailto:pflgeberatung@spectrumK.de).

Ich wünsche eine qualifizierte Pflegeberatung durch den Vertragspartner spectrumK GmbH und erkläre dies mit meiner Unterschrift:

Datum: \_\_\_\_\_ Unterschrift des Versicherten bzw. des Bevollmächtigten:

### Leitfragen für Unterstützungsbedarfe

<b><u>Mobilität &amp; Wohnumfeld</u></b>
Ist das Sturzrisiko erhöht? (Siehe Screening in der Anlage)
Wo bestehen Gefahrenquellen in der Wohnung? (Siehe Screening in der Anlage)
Welche Hilfsmittel wären geeignet, um das Wohnumfeld sicherer zu machen?
<b><u>Alltagsaktivitäten &amp; Versorgungssituation</u></b>
Wie kommen Sie im Alltag mit den hauswirtschaftlichen Aufgaben (Kochen, Putzen, Wäsche oder Einkaufen) zurecht?
Gibt es Tätigkeiten im Haushalt, die Ihnen zunehmend schwerfallen?
Haben Sie schon Unterstützung im Haushalt oder würden Sie sich dabei manchmal Hilfe wünschen?
Gibt es Angehörige/Hilfspersonen und sind diese ggf. belastet?
<b><u>Soziale Beziehungen</u></b>
Wie ist Ihr Kontakt zu den Menschen in Ihrer Nachbarschaft?
Gibt es Nachbarn, mit denen Sie regelmäßig sprechen oder sich gegenseitig unterstützen?
Gibt es Gruppen, Treffen oder Veranstaltungen, an denen Sie gerne teilnehmen?
Finden Sie derzeit Angebote, die für Sie interessant wären?
Haben Sie früher an Gruppen teilgenommen, was heute vielleicht schwieriger geworden ist?
Fühlen Sie sich im Alltag gut eingebunden und begleitet – oder gibt es Momente, in denen Sie sich allein fühlen?
Haben Sie Menschen, mit denen Sie über persönliche Sorgen oder wichtige Themen sprechen können?
Gibt es Angehörige oder nahestehende Personen, die Sie derzeit unterstützen oder pflegerisch versorgen?
Wenn ja, brauchen Sie für diese Unterstützung manchmal Entlastung?
<b><u>Ansprüche auf Sozialleistungen</u></b>
Haben Sie in letzter Zeit mehr Unterstützung im Alltag benötigt – zum Beispiel beim Waschen, Anziehen, im Haushalt oder bei der
Falls ja, wer unterstützt Sie aktuell, und reicht Ihnen diese Hilfe aus?
Möchten Sie gemeinsam prüfen, ob ein Pflegegrad für Sie infrage kommt?
Pflegegrad-Rechner inkl. Pflegeantrag: <a href="https://mein.pflegeabc.de/pflegegradrechner/">https://mein.pflegeabc.de/pflegegradrechner/</a>
Falls ein Pflegeantrag aufgenommen wurde: Wünschen Sie eine ausführliche Pflegeberatung nach § 7a SGB XI, um
Liegt eine Schwerbehinderung vor?
Bedarf an Hilfsmitteln (§ 33 SGB V):
• Gibt es Hilfsmittel, die Ihnen den Alltag erleichtern könnten – etwa beim Gehen, Duschen oder Aufstehen?
Präventionskurse (§ 20 SGB V) – Bewegung, Ernährung, Stressbewältigung, Suchtmittelkonsum:
• Haben Sie Interesse an Angeboten, die Ihre Gesundheit stärken – etwa Bewegungs-, Ernährungs- oder Entspannungskurse?
• Wären Kurse i. E. hilfreich, um bestimmte Gewohnheiten zu verbessern oder Stress abzubauen?

## Beispiel 3: risikobasierte Pflegeberatung: Take Aways

Eine risikobasierte Pflegeberatung wirkt auf die mit dem Pflegeeintritt assoziierte (Kosten) Risiken wenn

- Diese auf Versicherte mit beeinflussbaren Risiken fokussiert wird
- Das freiwillige Beratungsangebot auf die TOP x Risikopatienten begrenzt wird
- Die Risikomodellierung unter Berücksichtigung zeitlicher Diagnose- und Therapiemuster erfolgt (LSTM)
- Die Pflegeberatung inhaltlich auf die Vermeidung kostenintensiver Akutereignisse ausgerichtet ist
- Die Vorgaben des SGB XI §§ 7a,b &17, SGB V §25b (GKV) oder BDSG §22 (PKV) eingehalten werden

Nutzer:

**VI**ACTIV  
Krankenkasse

**spectrum**>k  
Gemeinsam Gesundheit gestalten.

**ME** : **CO**  
WWW.MEDICAL-CONTACT.DE

Vielen Dank für Ihre Aufmerksamkeit und Ihre Fragen

[www.risikoanalytik.de](http://www.risikoanalytik.de)



 **RISIKOANALYTIK GMBH**      WIR ÜBER UNS    PRODUKTE ▾    PROJEKTE ▾    KONTAKT ▾     ENGLISH

**PRAKTISCHER NUTZEN DURCH  
STATISTISCHE METHODEN**

Ihr Partner für Identifikation, Prädiktion, Evaluation und maschinelles Lernen

### Kontakt:

Prof. Dr.-Ing Thomas P. Zahn  
Prof. Health Data Science  
Geschäftsführer  
DCC Risikoanalytik GmbH und  
DSI - Data Science Institute

Tel. +49 30 65485298  
mobil. +49 171 7909941  
[thomas.zahn@risikoanalytik.de](mailto:thomas.zahn@risikoanalytik.de)

---

**Besuchen Sie  
unsere Webseite**

*[www.aktuar.de](http://www.aktuar.de)*

---