

Gesundheitsdaten sicher nutzen – und Versorgung besser verstehen

*Von der Idee zur Analyse: Sichere Datennutzung
im FDZ Gesundheit*

Alisa Hauser
Kordinatorin Antragsbearbeitung



Inhalt

- Ziele & Nutzen des FDZ Gesundheit
- Welche Daten gibt es im FDZ Gesundheit?
- Antragsstellung Schritt für Schritt

Was sind die Ziele des FDZ?

Die Ziele des Forschungsdatenzentrum Gesundheit

Wir möchten die **medizinische Versorgung** von Patient:innen in Deutschland und der EU maßgeblich verbessern. **Denn wir machen Forschung möglich**, die bislang noch nicht möglich war.

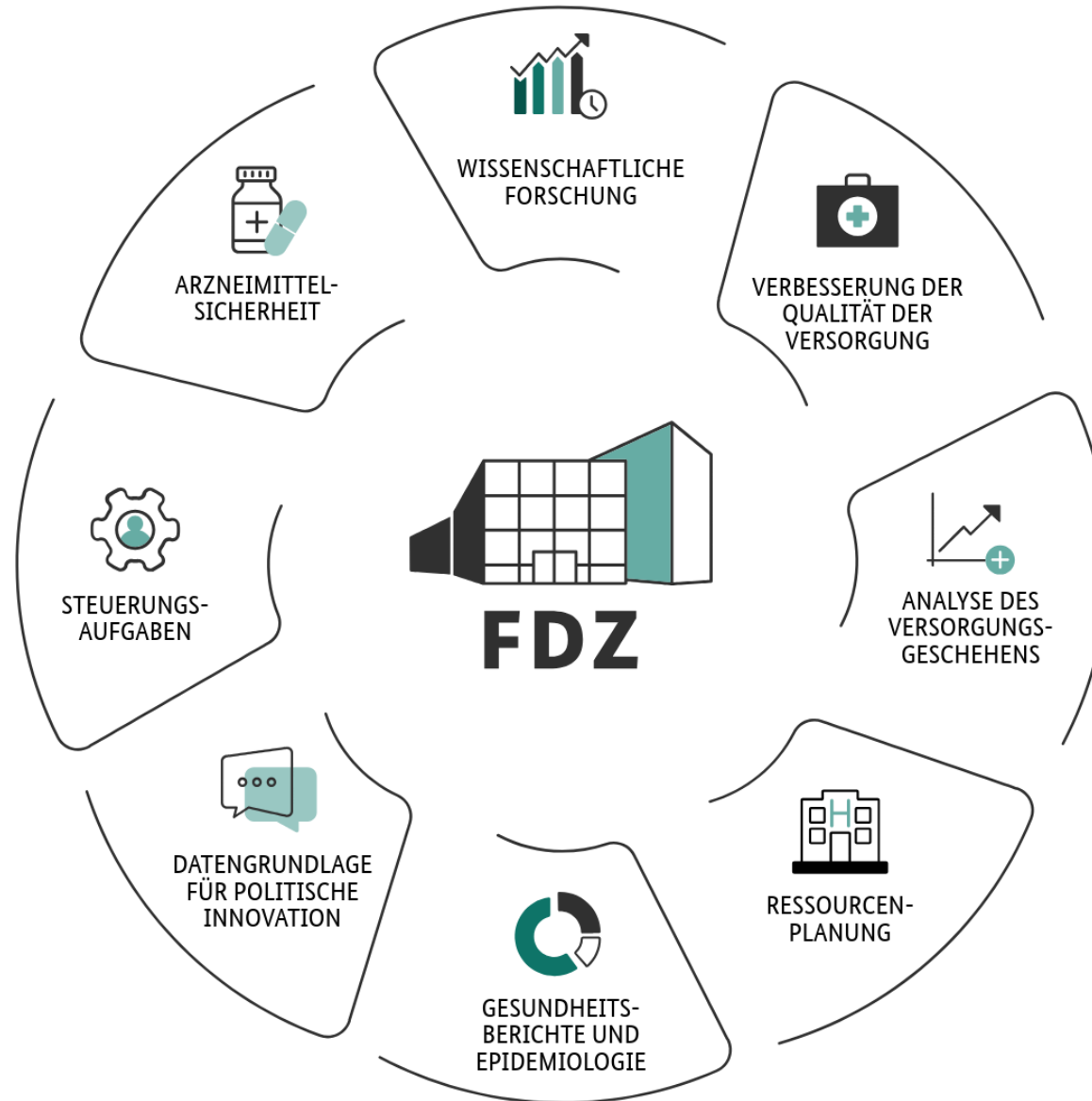
Forschungsorientierung

- **Zugang vereinfachen**
- **Iterative** und interaktive Forschungsprozesse ermöglichen
- **Transparenz und Planbarkeit** des Antragsprozesses
- Enge Abstimmung mit **Nutzungsberechtigten**

Sicherheit

- Berücksichtigung der **Sensibilität** der Gesundheitsdaten
- **Moderne Verfahren** der IT-Sicherheit
- Minimierung des **Reidentifikationsrisikos**
- Enge Abstimmung mit **BSI und BfDI**

Wofür der Aufwand?



Welches Potenzial haben diese Daten für Patient:innen?

Der Gewinn für alle PatientInnen...

Realitätsnahe Erkenntnisse



Klinische Studien zeigen umfassend, was unter **idealen** Bedingungen für eine **kleine Gruppe** funktioniert.

FDZ-Datenanalysen zeigen, was in der **alltäglichen** Versorgung **für alle** passiert.

Gerechtigkeit in der Versorgung



- Regionale Unterschiede sichtbar machen
- Unterschiede zwischen Altersgruppen erkennen (Erkrankungen und Versorgung)
- Unterschiede zwischen Geschlechtern differenzieren
- Vermeidbare Komplikationen entdecken und künftig vermeiden
- Grundlage für gesundheitspolitische Entscheidungen bieten

Systematische Verbesserungen



Ergebnisse von Analysen können helfen:

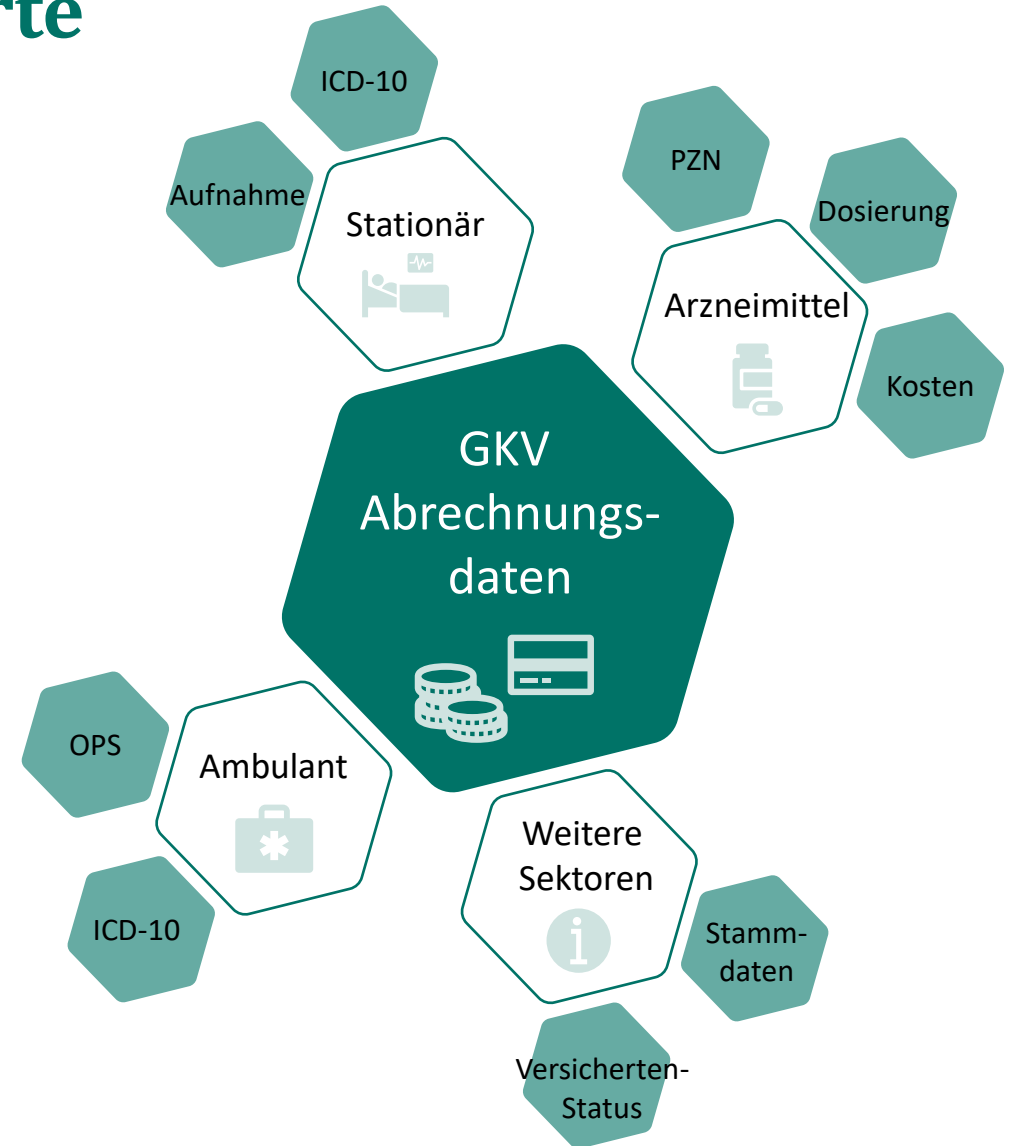
- Leitlinien und Nachsorgeprogramme zu verbessern
- Behandlungsabläufe zu verbessern
- Typische Komplikationen zu reduzieren
- Über-/Unterversorgung zu erkennen
- Systematische Unterschiede aufzudecken
- Versorgungsplanung an Bedarfe auszurichten

Welche Daten gibt es beim FDZ?

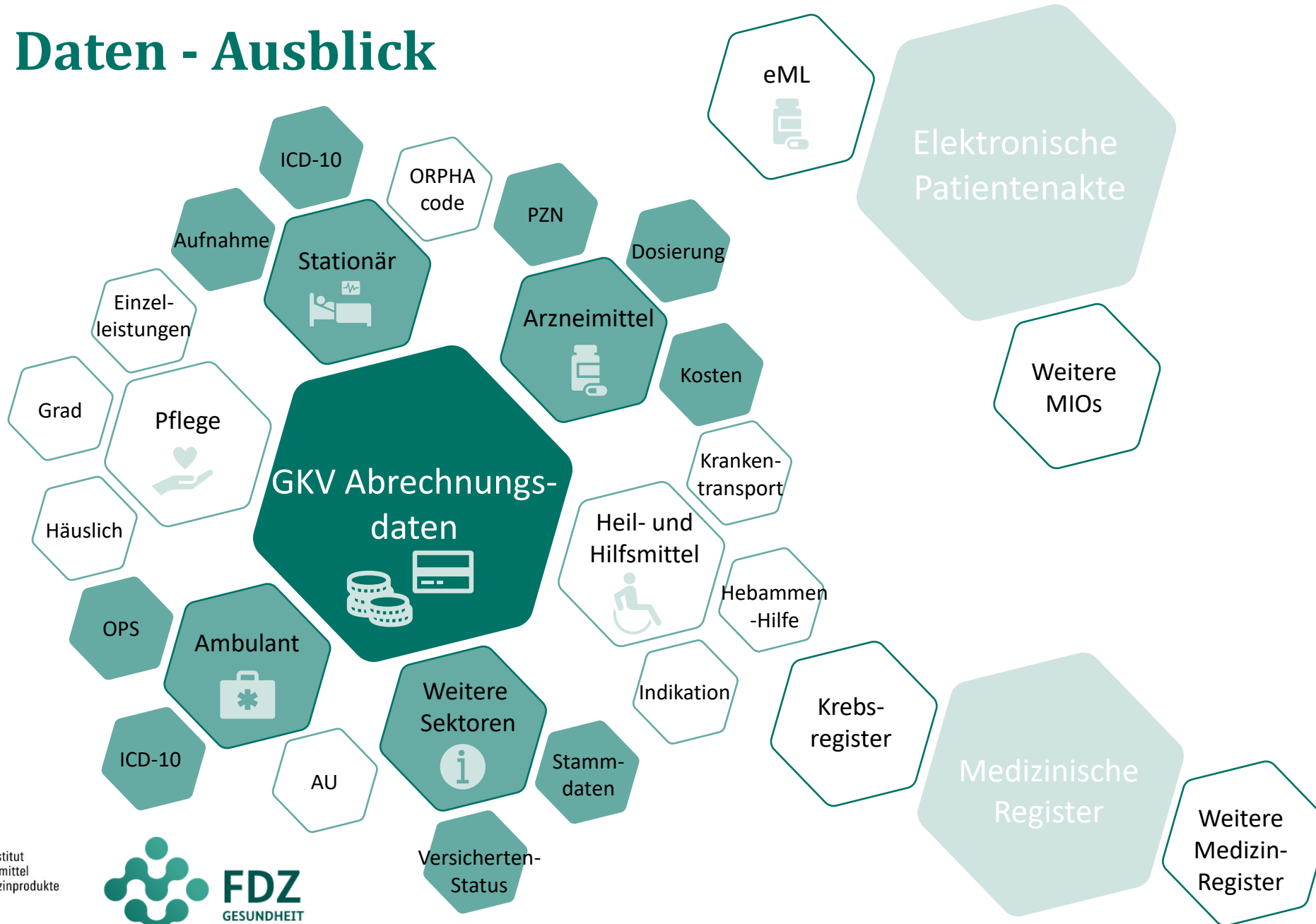
FDZ Daten – 74 Millionen Versicherte

Daten seit 2009, erste große Erweiterung 2019

- Aktuellste Daten: 4 zentrale Sektoren mit Fall-Inhalten in insgesamt 17 Detailtabellen
- Zwischen 10 Mio. und 3 Mrd. Zeilen pro Jahr pro Tabelle
- Durchschnittlich rund 650 Mio. Zeilen pro Jahr pro Tabelle und 11 Mrd. Zeilen pro Jahr

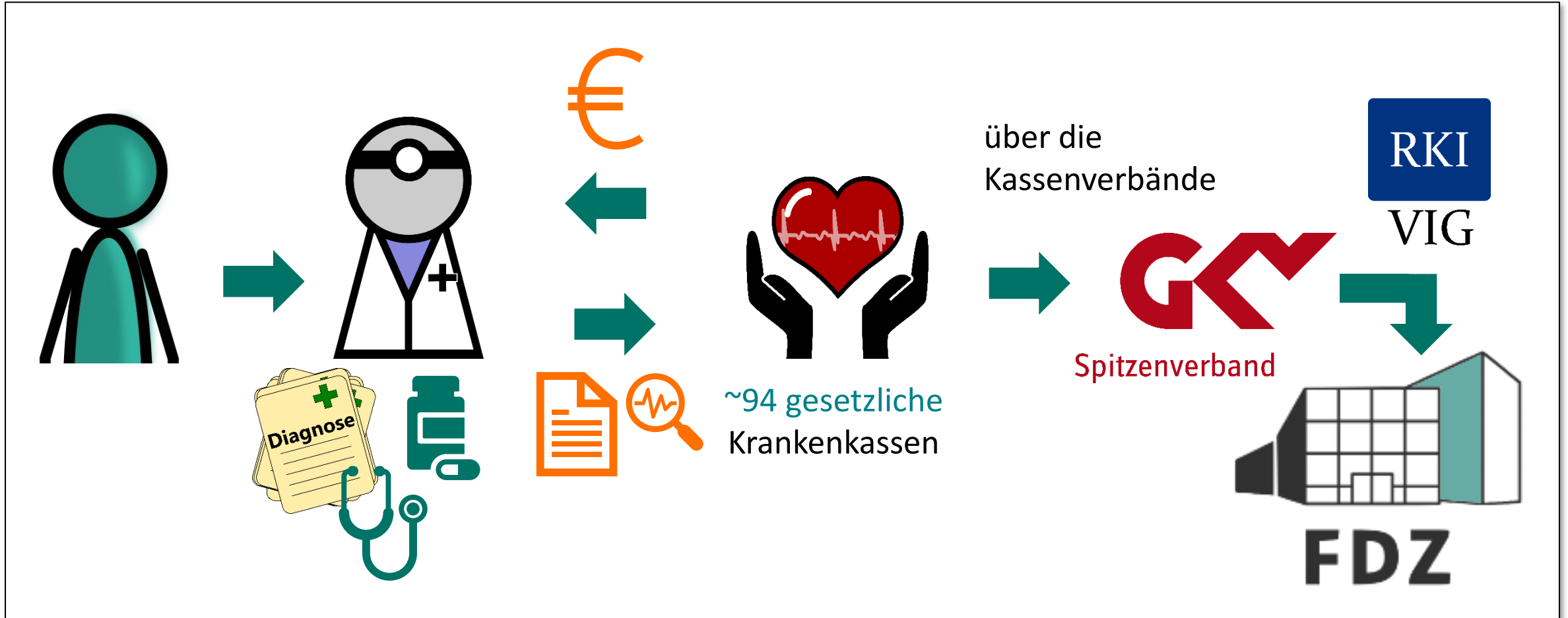


FDZ Daten - Ausblick



Woher kommen die Daten im FDZ?

Der Weg der Daten: Von den Versicherten bis ins FDZ...



Was beinhalten die Daten im FDZ? Und was nicht?

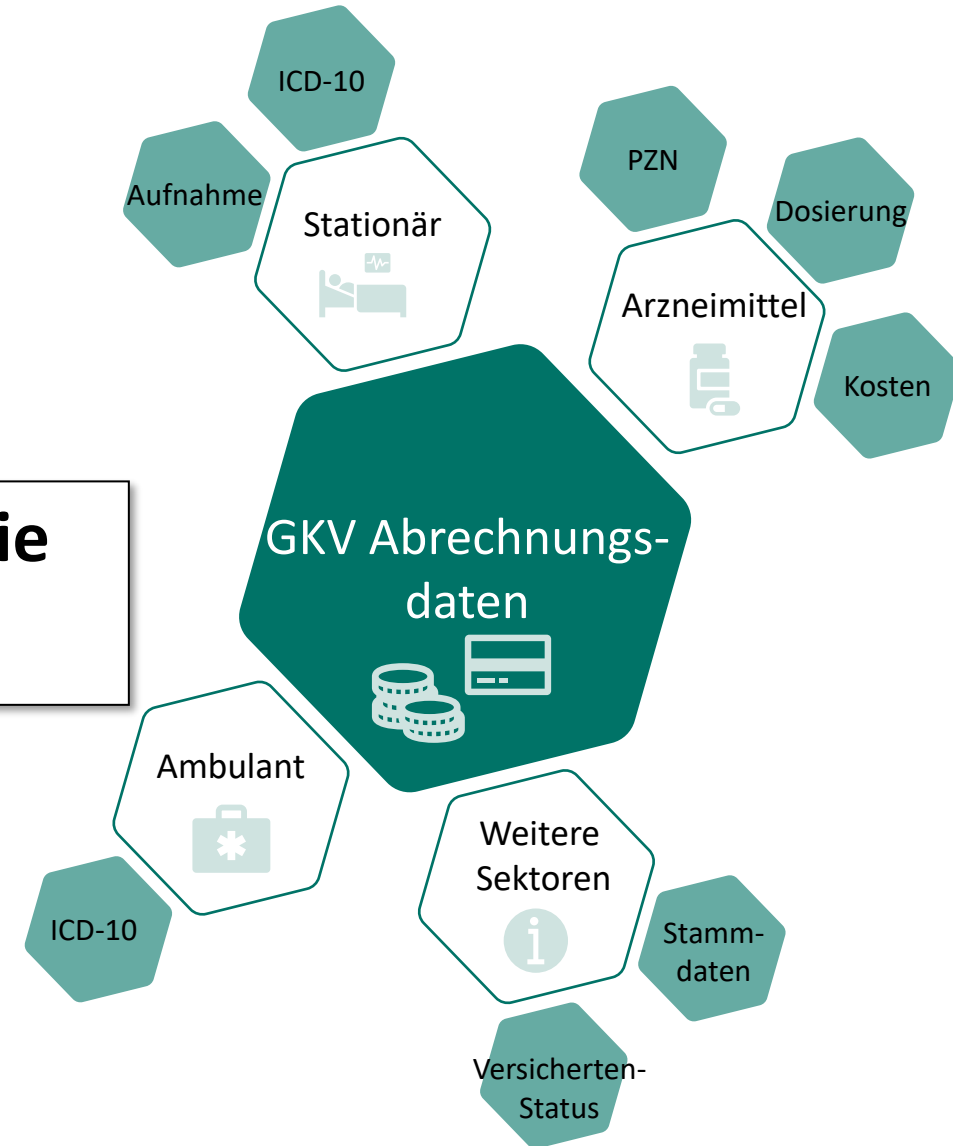
Dateninhalte und Pseudonymisierung

Diagnosen
Medikamente
Behandlungen (z.B. Hausarzt, Facharzt, Krankenhaus)
Durchgeführte Prozeduren (z.B. Operationen)
Zahnärztliche Behandlungen
Zeitliche Verläufe über viele Jahre
Versorgung in unterschiedli
...



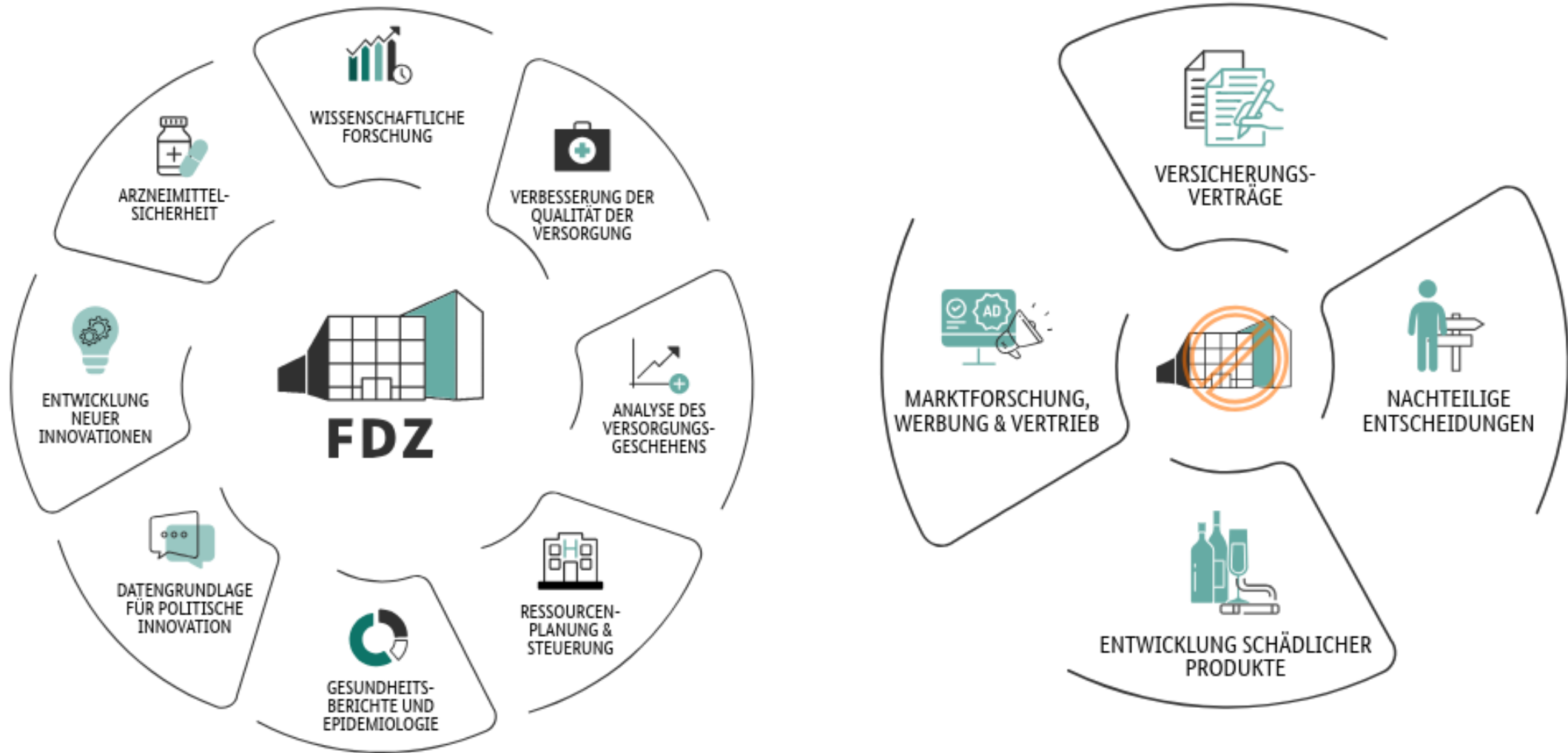
Details aller Inhalte über die Datensatzbeschreibung

Namen
Adressdaten
Konkrete Geburts-/Sterbedaten (nur Jahr)
Krankenversicherungsnummer / Lebenslange Arztnummer
Freitexte
Arztberichte
Gesprächsinhalte
...



Wie komme ich an die FDZ Daten?

Datenzugriff nach Nutzungszwecken



Technische und organisatorische Vorbereitungen

Personelle Überlegungen - Wer macht was?

Institutionsregistrierung:

- institutsbeauftragte Person
- Institutionsleitung

Rollen im Antragsprozess:

- Antragstellende Person
- Datennutzende Person
- Zeichnungsberechtigte Person



Technische und organisatorische Vorbereitungen

Informationsmaterialien auf der Webseite

- Musterantrag
- Gebührenverordnung
- Anleitung zur Anmeldung im Portal
- Nutzerhandbücher, uvm.
- **Regelmäßige Webinare**



Statistikportal



Forschungsdatenzentrum Gesundheit

Statistikportal [?]

Das Statistikportal ist ein ergänzendes Angebot des FDZ_Gesundheit und erlaubt die mengenmäßige Analyse von Diagnosedaten. Mit der Nutzung dieses Angebots stimmen Sie unseren [Nutzungsbedingungen](#) (Stand: August 2025) zu.

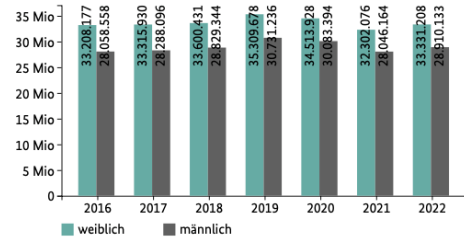
Alle dargestellten Zahlen sind aus den Abrechnungsdaten der gesetzlichen Krankenkassen in Deutschland berechnet. Bei den ambulanten Diagnosen werden ausschließlich gesicherte Diagnosen verwendet. Für eine genaue Beschreibung der zugrundeliegenden Daten siehe unsere [Datensatzbeschreibung](#).

Zur Wahrung der Anonymität der Patienten nutzen wir eine 95%-Stichprobe. Weiterhin sind kleine Fallzahlen und alle Rückrechenmöglichkeiten aus dem Datensatz entfernt worden. Nicht binäre Geschlechtsinformationen werden zufällig den Geschlechtern männlich/weiblich zugeordnet. Für weitere Informationen zur Anonymisierung siehe unser [Anonymisierungskonzept](#).

icd: any_diagnosis

Auswahl Berichtsjahr

► Erläuterung



Suche nach Einzeldiagnosen oder ICD-Gruppen

Anzahlen pro PLZ-Gebiet

► Erläuterung

Auflösung PLZ-Gebiete:

Grob (1-stellig)



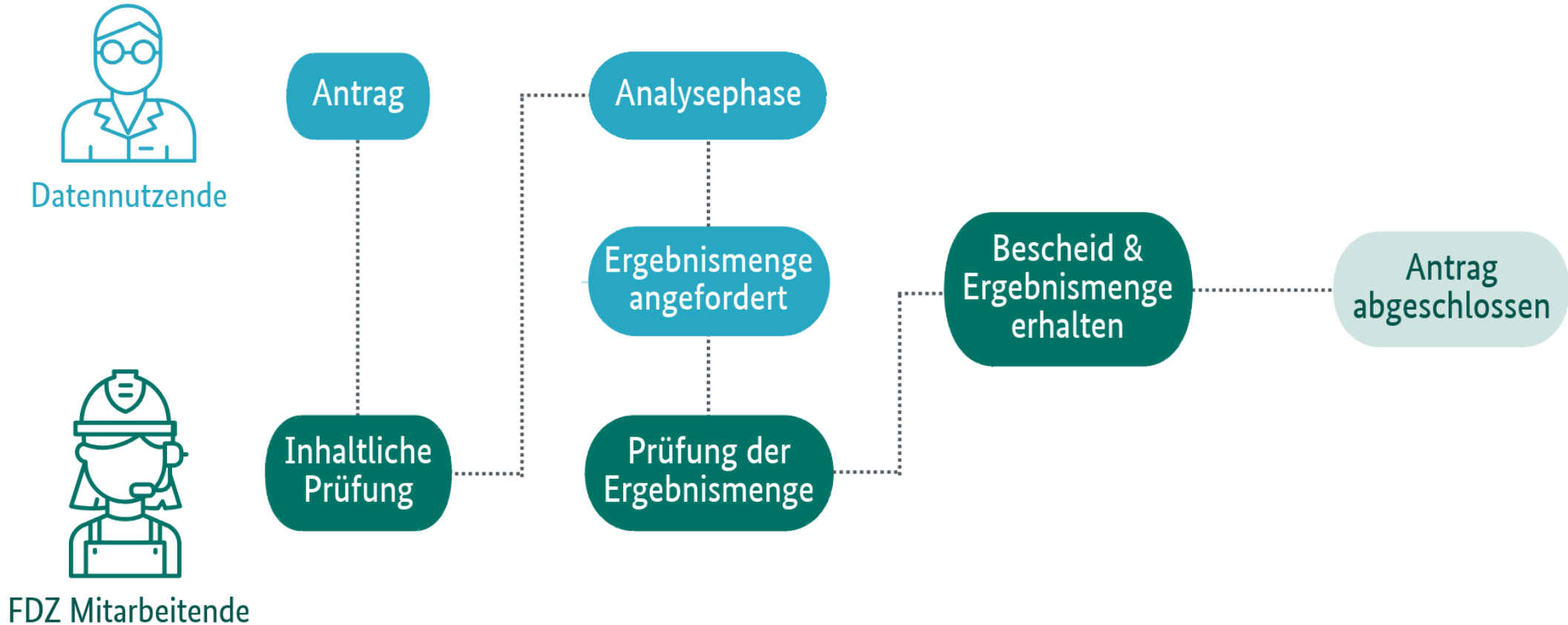
Häufigkeitsverteilung nach Alter und Geschlecht

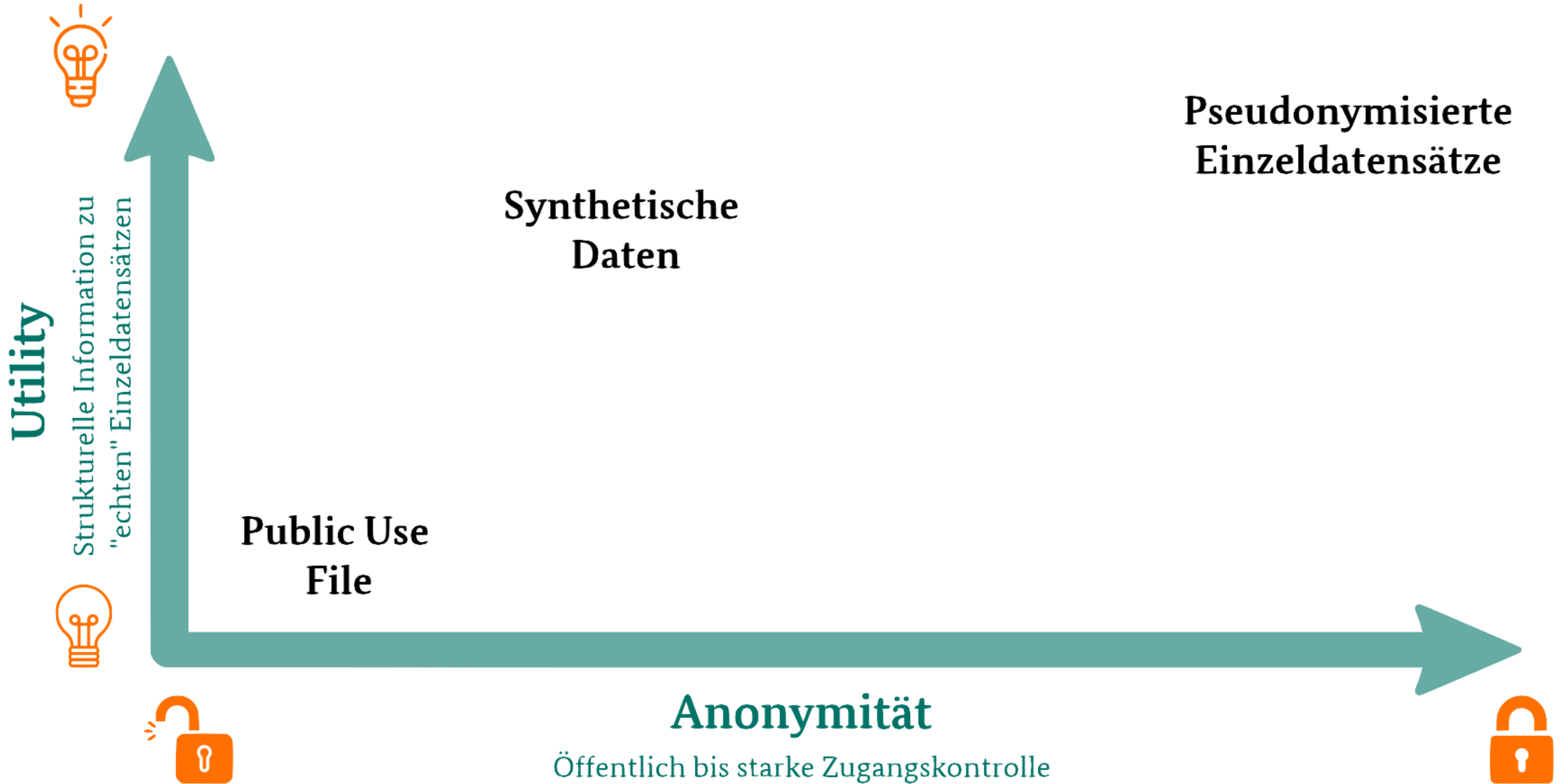
► Erläuterung

Altersgruppen:

Grob (3 Stufen)

Übersicht Antragsprozess





Posit Workbench

dm0000051 Sign Out

Sessions + New Session Suspend All Quit All

- JupyterLab Session 2** (lab icon) FINISHED Info Rename Quit
CREATED: 9/25/2023
- VS Code Session** (VS icon) ACTIVE Info Rename Quit
/home/dm0000051 CREATED: 9/14/2023
- RStudio Pro Session** (R icon) ACTIVE Info Rename Suspend Quit
R 4.2.3 (Home) CREATED: 9/6/2023 LAST USED: 9/27/2023
- JupyterLab Session** (lab icon) FINISHED Info Rename Quit
CREATED: 9/6/2023

Projects

- Untitled.ipynb (Home) OWNER: ME
LAST USED:
- analyseraum_002 ~/analyseraum_002 OWNER: ME
LAST USED:

Posit Workbench 2023.06.1+524.pro1, "Mountain Hydrangea" (7025e95a) © 2009 - 2023 Posit Software, PBC

Posit Workbench x analyseraum_003 - RStudio Pro x + DE German (Germany)

https://rstudio-workbench-3.ds.devqas-integrated.fdz.tir.bdr.de/s/cbc31b20e4e633c7d7789?launcher=1

File Edit Code View Plots Session Build Debug Profile Tools Help

performance_test_complex_survival.R

```

65 summary(outcomes[[1]])
66 head(outcomes[[2]])
67 summary(outcomes[[2]])
68 head(outcomes[[3]])
69 summary(outcomes[[3]])
70
71
72 ### Ergebnisse
73 ### Creating a Kaplan Meier curve and cox regression
74
75 parallel <- FALSE
76 if(parallel) {
77   library(foreach)
78   cl <- parallel::makeCluster(3)
79   doParallel::registerDoParallel(cl)
80 }
81
82 survivalResults <- list()
83 survivalResults <- foreach(i = 1:3
84   , .packages = c("tidyverse", "survminer", "survival")
85   ) %do% {
86
87   outcomesCombined <- riskFactors %>% left_join(outcomes[[i]], relationship = "many-to-many")
88
89   fit <- survfit(Surv(LQ, EVENT) ~ METFORMIN+GLUCOCORTICOSTEROIDS+DEMENTZ+AEMWEGE, data = outcomesCombined)
90
91   p1 <- ggsurvplot(fit, data = outcomesCombined,
92     ylim = c(0.95, 1),
93     size = 1,
94     pval = TRUE,
95     ggtheme = theme_bw()
96   )
97
98   p2 <- ggsurvplot(fit, data = outcomesCombined,
99     size = 1,
100    conf.int = TRUE,
101    pval = TRUE,
102    risk.table = TRUE,
103    risk.table.col = "strata",
104    risk.table.height = 1,
105    ggtheme = theme_bw()
106  )
107
108  modelFit <- coxph(Surv(LQ, EVENT) ~ METFORMIN+GLUCOCORTICOSTEROIDS+DEMENTZ+AEMWEGE, data = outcomesCombined)
109
116:26 [Top Level]

```

Console Terminal Background Jobs Workbench Jobs

R 4.2.3 --/analyseraum_003/

R version 4.2.3 (2023-03-15) -- "Shortstop Beagle"
 Copyright (C) 2023 The R Foundation for Statistical Computing
 Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
 You are welcome to redistribute it under certain conditions.
 Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
 Type 'contributors()' for more information and
 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
 'help.start()' for an HTML browser interface to help.
 Type 'q()' to quit R.

```

> CONN <- dbConnect(odbc(), "Analyseraum_003_INT")
> silent_check()

```

Environment History Connections Git Tutorial

Your branch is ahead of 'origin/main' by 1 commit.

Staged	Status	Path
<input type="checkbox"/>	<input checked="" type="checkbox"/>	performance_test_complex_survival.R
<input type="checkbox"/>	<input checked="" type="checkbox"/>	performance_test_complex_survival.ipynb
<input type="checkbox"/>	<input checked="" type="checkbox"/>	performance_test_simple_prevalence.ipynb
<input type="checkbox"/>	<input checked="" type="checkbox"/>	sql_debug_25092023.sql

Files Plots Packages Help Viewer Presentation

New Folder New Blank File Upload Delete Rename More

Home > analyseraum_003

Name	Size	Modified
.		
performance_test_simple_prevalence.ipynb	4.7 KB	Oct 17, 2023, 10:57 AM
performance_test_complex_survival.ipynb	10.1 KB	Oct 17, 2023, 10:57 AM
review_sqlalchemy.ipynb	5.6 KB	Oct 17, 2023, 10:53 AM
analyseraum_003.Rproj	205 B	Oct 17, 2023, 10:51 AM
testdate1.Rmd	815 B	Oct 17, 2023, 10:49 AM
performance_test_complex_survival.R	3.5 KB	Sep 25, 2023, 1:35 PM
sql_debug_25092023.sql	57 B	Sep 25, 2023, 10:59 AM
hane_survival.sql	64 KB	Sep 22, 2023, 10:37 AM
per_test_R_simple_prevalence.R	1.3 KB	Sep 22, 2023, 10:37 AM
performance_test_complex_pur.R	13.6 KB	Sep 22, 2023, 10:37 AM
performance_test_simple_medication.ipynb	6.7 KB	Sep 22, 2023, 10:37 AM
performance_test_simple_medication.R	1.6 KB	Sep 22, 2023, 10:37 AM
performance_test_simple_medication.sql	1.6 KB	Sep 22, 2023, 10:37 AM
performance_test_simple_prevalence.R	994 B	Sep 22, 2023, 10:37 AM
performance_test_simple_prevalence.sql	1.1 KB	Sep 22, 2023, 10:37 AM
sql_debug2.sql	241 B	Sep 22, 2023, 10:37 AM
sql_debug.sql	31 B	Sep 7, 2023, 4:15 PM
Rhistory	64 B	Sep 7, 2023, 3:10 PM
gitignore	72 B	Sep 7, 2023, 11:50 AM
DB_Connection.ipynb	1.1 KB	Sep 7, 2023, 11:50 AM
Python_test.ipynb	1.1 KB	Sep 7, 2023, 11:50 AM
README.md	6.1 KB	Sep 7, 2023, 11:50 AM
test.txt	27 B	Sep 7, 2023, 11:50 AM

There are more named users of Posit Workbench than your license supports. Please obtain an updated license to continue using the product.

Inhaltliche Vorbereitung: Webinare

Veranstaltungen des FDZ Gesundheit

Veranstaltung	Datum und Uhrzeit	Link zur Registrierung
Webinar: Antragstellung und sichere Verarbeitungsumgebung	Voraussichtlich: 23. Juni 2026	
Webinar: Datenbestand des FDZ Gesundheit	Voraussichtlich: 15. September 2026	
Webinar: Antragstellung und sichere Verarbeitungsumgebung	Voraussichtlich: 13. Oktober 2026	
Webinar: Datenbestand des FDZ Gesundheit	Voraussichtlich: 17. November 2026	

Die aufgezeichneten Vorträge, inklusive Folien, finden Sie in unserer [Mediathek](#).

[Link zu unseren Veranstaltungen](#)

Vielen Dank für Ihre Aufmerksamkeit!



Kontakt

Bundesinstitut für Arzneimittel und Medizinprodukte
Forschungsdatenzentrum
Kurt-Georg-Kiesinger-Allee 3
53175 Bonn

Ansprechpartner
Alisa Hauser
Alisa.hauser@bfarm.de
www.bfarm.de
www.forschungsdatenzentrum-gesundheit.de

