

Record Linkage mit GKV-Routinedaten

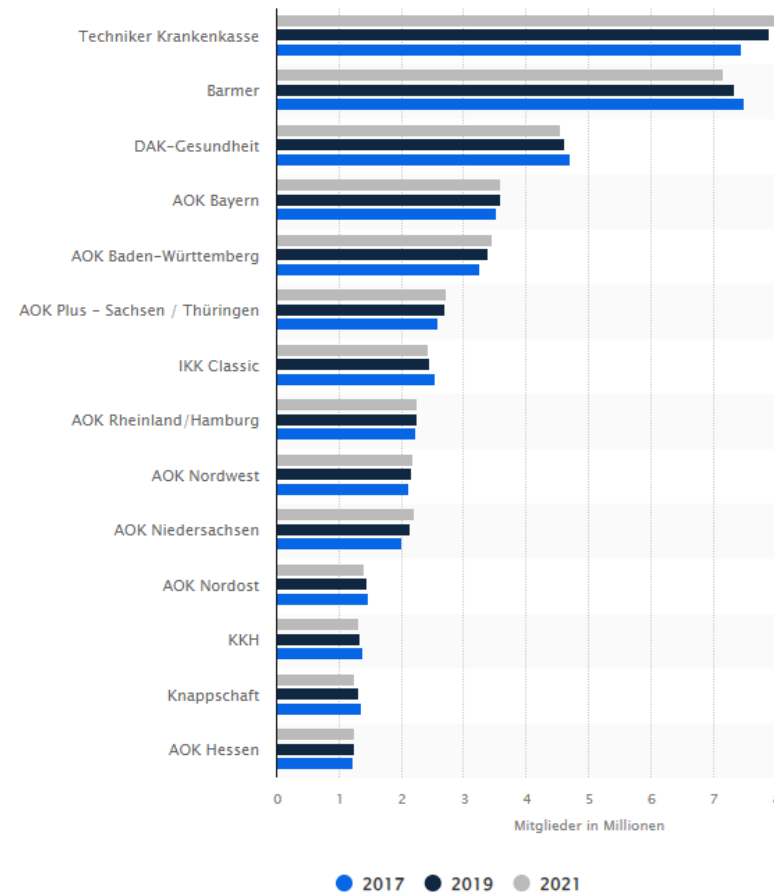
Peter Ihle

**Workshop
Record Linkage von
unterschiedlichen Datenarten
GMDS 2023, Heilbronn**

- GKV-Routinedaten
 - Übersicht
 - Datenlinkage/Krankenversicherternummer
- Projekte mit Linkage mit Routinedaten
 - VersKiK
 - RiCO (nicht präsentiert)
 - NUM-Kohorte (nicht präsentiert)
- Zusammenfassung und Fazit

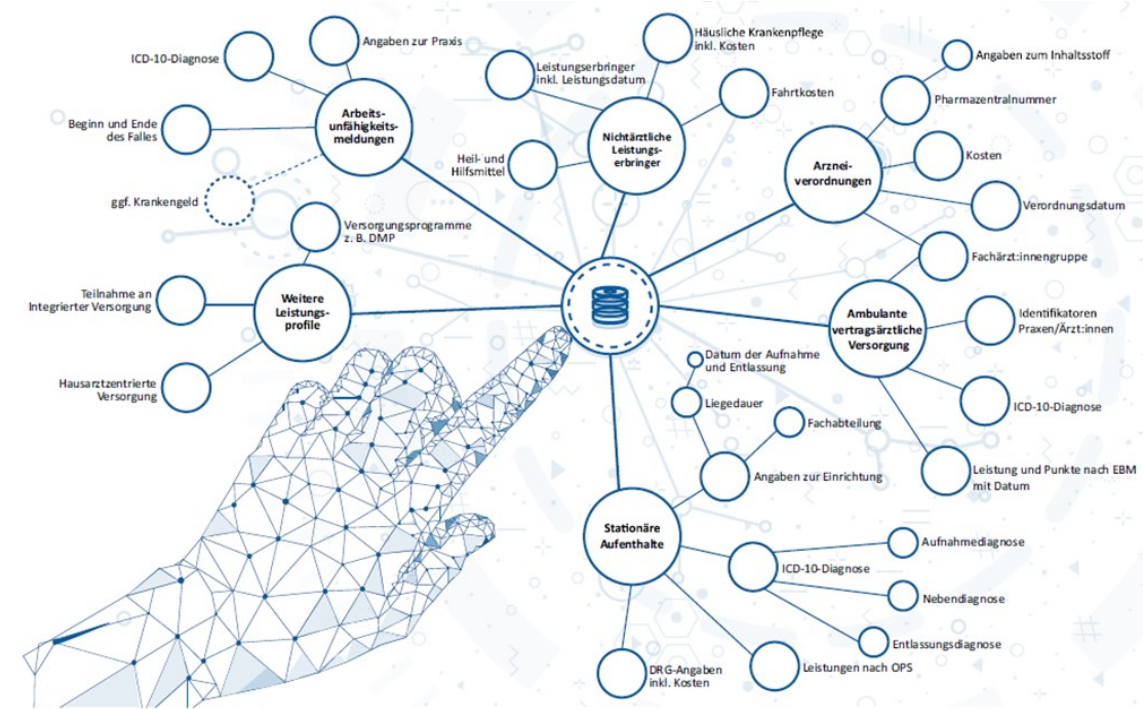
- GKV-Routinedaten sind primär Abrechnungsdaten und können sekundär für Auswertungszwecke genutzt werden (Sekundäranalyse)
- Die Daten entstehen im Rahmen der Abrechnung ohne zusätzlichen Erhebungsaufwand von Seiten der Forschung
- Der Aufwand entsteht dafür bei der Aufbereitung, Plausibilisierung und Analyse
- Nutzung der Routinedaten immer unter der selbstkritischen Frage „Sind die Daten für Beantwortung meiner Forschungsfrage geeignet?“ (LL 3 GPS - Gute Praxis Sekundärdatenanalyse)
- Spezifische Verfahren erforderlich, beispielsweise die interne Validierung zur operationalen Definition von Erkrankungen (Stichwort M2Q-Kriterium)
- Genaue Kenntnis des Kontextes der Datenentstehung

- GKV-Routinedaten liegen bei den Krankenkassen (n=96 Stand Anfang 2023)
- Kooperationsbereitschaft der Krankenkassen für Forschungsprojekte erfolgt auf freiwilliger Basis



GKV-Routinedaten

Profilgruppe	Beschreibung	Profil	Beschreibung
P0	Leistungserbringer	P0bsnr	Angaben zum Leistungserbringer
P1	Versichertendaten	P1stamm	Stammdaten des Versicherten
		P1sozio	Soziodemografische Angaben des Versicherten
		P1euro	Einkommen
		P1vers	Versicherungszeiten des Versicherten
		P1dmp	DMP-Zugehörigkeit des Versicherten
		P1hzv	HZV-Zugehörigkeit des Versicherten
P2	ambulante ärztliche Daten = Fachverfahren Ärzte	P2fall	Falldaten (Einzelfallnachweis)
		P2diag	Diagnoseangaben
		P2ebm	EBM-Leistungen
P3	Verordnungen	P3rez	Rezeptdaten
		P3pack	Packungsangaben
		P3zseg	Angaben aus dem Z-Segment für PZN-Sondernummern
		P4	Stationäre Aufenthalte inkl. Reha-Aufenthalte
		P4fall	Falldaten
P4	Stationäre Aufenthalte inkl. Reha-Aufenthalte	P4diag	Diagnoseangaben
		P4ops	OPS-Angaben
		P4drg	DRG-Angaben (Kosten)
		P4abt	Fachabteilung
		P4reha	Angaben zur stationären Rehabilitation
		P5	Arbeitsunfähigkeiten
		P5fall	Falldaten
P6	Weitere Leistungsbereiche	P5diag	Diagnoseangaben
		P6sole	Sonstige Leistungen: zur Systematik siehe Tabellenblatt
		P6soleDiag	Abrechnungsdiagnosen für die Leistungen in P6sole
P7	Pflege	P6eel	Entgelt-Ersatzleistungen
		P7fall	Falldaten (SGB XI)
		P7leist	Leistungsdaten (SGB XI)
P8	Zahnärztliche Behandlung	P8fall	Falldaten
		P8leist	Leistungsdaten
P9	Ambulantes Operieren	P9fall	Falldaten
		P9diag	Diagnoseangaben
		P9ebm	EBM-Leistungen
		P9ops	OPS-Angaben



Quelle: Cluster Medizin NRW 03/2023

- Kenntnis der Grundgesamtheit
- Sektorübergreifend
- Vollständig (ohne Untersuchungsbias, ohne Erinnerungsbias, ohne Interviewerbias)
- Lange Beobachtungszeiträume möglich
- Ziehung von Kontrollgruppen
- ...

Nachteile

- Keine Lifestyleangaben (Rauchen, Alkoholkonsum, Übergewicht, Sportliche Betätigung)
- Keine Untersuchungsergebnisse dokumentiert
- Keine oder nur schwer/unzureichend operationalisierbare Schweregrade
- Keine TNM-Dokumentation
- Arzneimitteltherapie im Krankenhaus unzureichend abgebildet
- ...

Daher inzwischen Goldstandard: **Datenlinkage mit anderen Datenquellen!**

- Krebsregister (Landeskrebsregister und Kinderkrebsregister)
- Andere Sozialversicherungsträger
z.B. Rentenversicherung oder Unfallversicherung
- (klinische) Primärstudien
- ...

Voraussetzung für das Personenbezogene Datenlinkage sind sogenannte Schlüsselvariablen:

- Personenidentifizierende Daten
wie Nachname, Vorname, Geburtsdatum, Wohnort
- Krankenversichertennummer (als unique identifier)

Krankenversichertennummer (KVNR)

(Regelung in §290 SGB V) seit 2003

- unveränderbaren Teil zur Identifikation des Versicherten
1 Buchstabe, 8 Ziffern, 1 Prüfziffer
- Vergabe/Bildung der KVNR durch die Vertrauensstelle
Krankenversichertennummer (VST) bei der ITSG
- Die KVNR wird mit Hilfe eines (geheimen) Algorithmus aus der
Rentenversicherungsnummer (RVNR) gebildet:
RVNR 13-stellig, 2-stellige Bereichsnummer, 6-stelliges Geburtsdatum,
Anfangsbuchstabe des Geburtsnamens, 3-stellige Seriennummer, 1-stellige
Prüfziffer
- Nutzung der KVNR auch durch Private Krankenversicherungen zulässig
§ 362 SGB V: Nutzung von elektronischen Gesundheitskarten oder digitalen Identitäten für Versicherte von
Unternehmen der privaten Krankenversicherung, der Postbeamtenkrankenkasse, der Krankenversorgung der
Bundesbahnbeamten, für Polizeivollzugsbeamte der Bundespolizei oder für Soldaten der Bundeswehr

Vertrauensstelle Krankenversichertennummer (VST)

- Führt ein Verzeichnis der Krankenversicherungsnummern jedoch ausschließlich zum Zweck der eindeutigen Vergabe der KVNR
- Koordiniert bei Neuvergabe einer KVNR auch die Generierung der Renten- (RVNR) bzw. Sozialversicherungsnummer (SVNR) bei der zuständigen Stelle der Rentenversicherung (in einem „vollautomatischen maschinellen Meldeverfahren“)
- Aufgabe ist die eindeutige Vergabe der KVNR

Vertrauensstelle Krankenversicherturnummer (VST)

- Regelfall: Die KVNR ist eindeutig und lebenslang
- Seltene Ausnahmefälle
 - RVNR enthält fehlerhaftes Geburtsdatum
 - > Neuvergabe der RVNR und damit auch der KVNR
 - Eine Person erhält mehr als eine RVNR
 - > Sperrung der zuviel vergebenen RVNR, Änderung der KVNR möglich
 - Eine RVNR wurde an mehrere Personen vergeben
 - > Neuvergabe der RVNR und Änderung der KVNR
- Pro Jahr ergeben sich ca. 4000 bis 6000 Änderungen
(persönliche Mitteilung, nicht veröffentlicht, keine Historie)

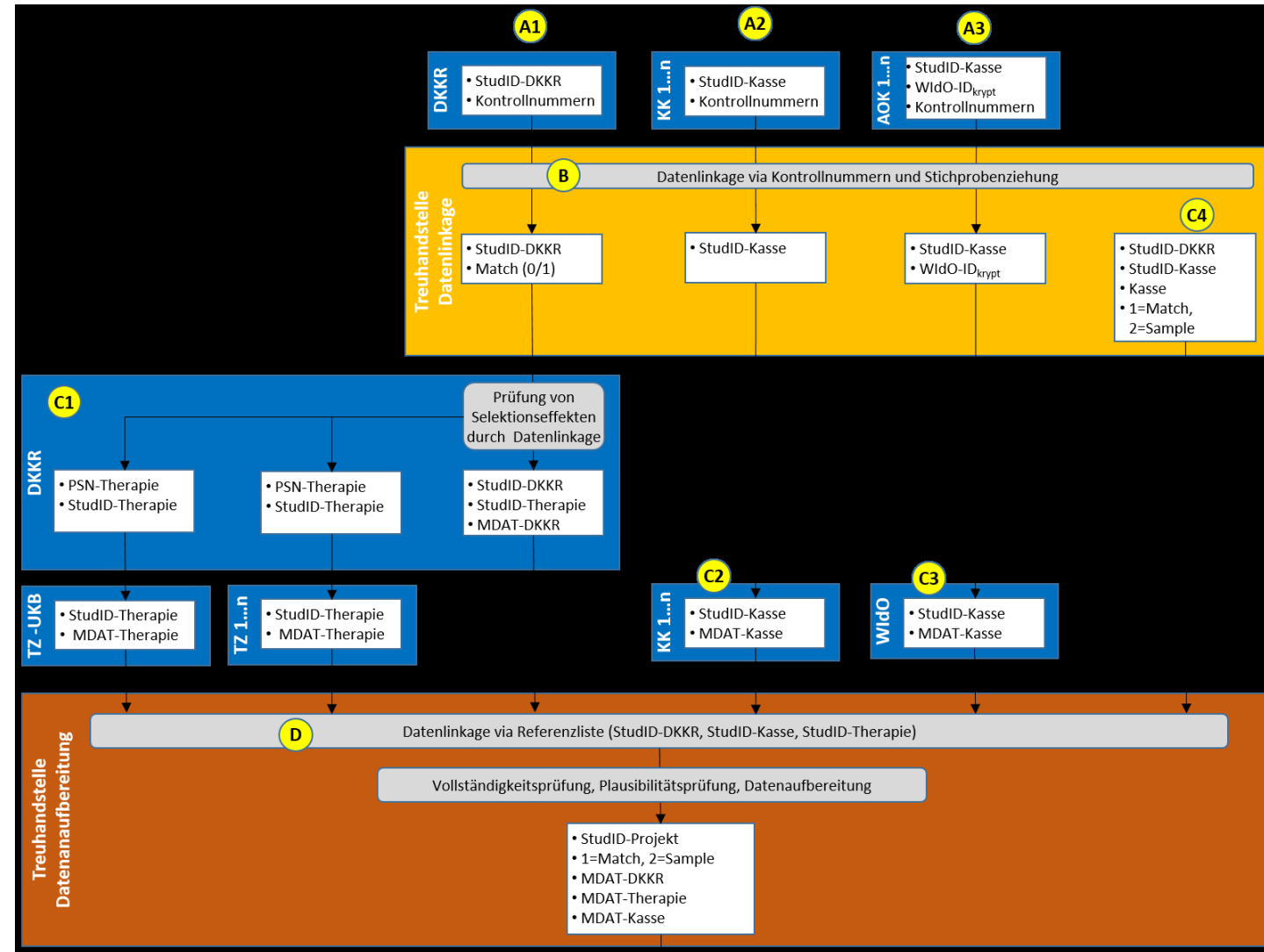
Quelle: https://www.gkv-spitzenverband.de/media/dokumente/krankenversicherung_1/telematik/telematik_3/10-2015_Information_GKV-SV_Eineutigkeit_Gueltigkeit_KVNRvo6.pdf

Fazit

- KVNR ist i.d.R. eindeutig und lebenslang
- Die zur Bildung der KVNR benötigte RVNR enthält das Geburtsdatum, daher ist das Geburtsdatum eine gut gepflegte Angabe
- Es gibt kein zentrales KVNR-Verzeichnis, bei dem die Zuordnung einer Person zu einer KVNR abgefragt werden darf
- Es gibt kein zentrales Versicherten-Verzeichnis, bei dem die Zuordnung einer Person zu einer Krankenkasse abgefragt werden kann

- VersKiK - Versorgung, Versorgungsbedarf und -bedürfnisse von Personen nach einer Krebserkrankung im Kindes- oder Jugendalter
- Innovationsfondsprojekt bis 2024
- Verlinkung von Daten des Kinderkrebsregister (Mainz) mit den Routinedaten von 13 Krankenkassen
- Linkagestelle: OFFIS, Oldenburg
- Vertrauensstelle mit Datenplausibilisierung: PMV, Köln
- Auswertungsstelle: ISMG, Magdeburg
- Antrag nach §75 SGB X bei den Aufsichtsbehörden, nicht Einwilligungsbasiert

- Datenfluss von datenbereitstellenden Kassen und dem DKKR bis zur Datenauswertung
- Teilnehmende Stellen:
 - DKKR (A₁),
 - Krankenkassen (A₂-A₃),
 - OFFIS (B),
 - Therapiezentren (C₁),
 - PMV (D),
 - ISMG (E)
- Alle Datensätze werden sicher übertragen
- Zu keinem Zeitpunkt liegen personenidentifizierende Daten im Klartext vor, außer bei den Kassen und dem DKKR



- Kontrollnummernverfahren mit Normierungsschritte für Klartextangaben
- Wichtige Schlüsselfelder:
Geburtsdatum und aktueller Wohnort, die sowohl in den Krankenkassen als auch im DKKR sehr gut gepflegt sind
- Pseudonyme auf Basis der Krankenversicherungsnummer wurden in der Projektplanungsphase verworfen, da die KVNR im DKKR nicht erfasst ist und damit als Schlüsselvariable für das Datenlinkage nicht benötigt
- Im Nachhinein für Plausibilisierungsvariable jedoch geeignet, um Matches, die in mehr als einer Krankenkasse gefunden wurden, eindeutig einer Person=KVNR zuzuordnen.

- **VersKiK**

Wenn ein zentrales Versichertenverzeichnis existiert, könnten bestehende Datenquellen, hier das Kinderkrebsregister durch die KVNR ergänzt werden. Das aufwändige (und fehlerbehaftete?) Linkage hätte damit deutlich schlanker erfolgen können, insbesondere ohne Erzeugung von Kontrollnummern aus den Klartextangaben

- Zentrales (Forschungs-)Versichertenverzeichnis
 - insbesondere für die Nacherfassung der KVNR in bestehenden Datenquellen
 - Prüfung von Zweifelsfällen (ähnlich bei Landeskrebsregister)
 - regelmäßiger Abgleich, beispielsweise auch mit Melderegisterdaten
 - Anschreiben von Personen mit der Möglichkeit, Daten bei Betroffenen zu prüfen, zu korrigieren und/oder nachzuerheben
- Geeigneter unique identifier
 - KVNR / Steuer-ID / Bürger-ID , neu zu schaffende Forschungs-ID
- Dokumentation von gelöschten/gesperren Löschung/Sperrung von RVNR/KVNR
- Auflage: Neue Datenquellen, auch solche die wie in pandemischen Zeiten auch kurzfristig aufgebaut werden, müssen das Datenlinkage als Regelfall mitdenken!