

BEISPIELE FÜR DIE VERKNÜPFUNG VERSCHIEDENER DATENQUELLEN MIT DEN DATEN GESETZLICHER KRANKENKASSEN

Iris Pigeot, Bianca Kollhorst
Leibniz-Institut für Präventionsforschung und Epidemiologie – BIPS

19.07.2018, Bamberg

Inhalt

- Pharmakoepidemiologische Forschung
- Die Forschungsdatenbank GePaRD
- Bisherige Linkage-Erfahrungen an fünf Beispielen
- Fazit

- Untersucht Anwendung, Wirkungen und Risiken von Arzneimitteln nach Marktzulassung
- **Limitationen der klinischen Studien vor Zulassung**
 - Zu geringe Größe (i.d.R. 3000-5000 Patienten)
 - Zu geringe Studiendauer
 - Vulnerable Patientengruppen häufig ausgeschlossen
- **Risiken von nicht bestimmungsgemäßigem Gebrauch nach Marktzulassung**
 - Nicht zugelassene Indikation, Unter- oder Überdosierung
 - Nichteinhaltung von Kontraindikationen
 - Verbotene Begleitmedikation

Die Forschungsdatenbank GePaRD



Aufbau

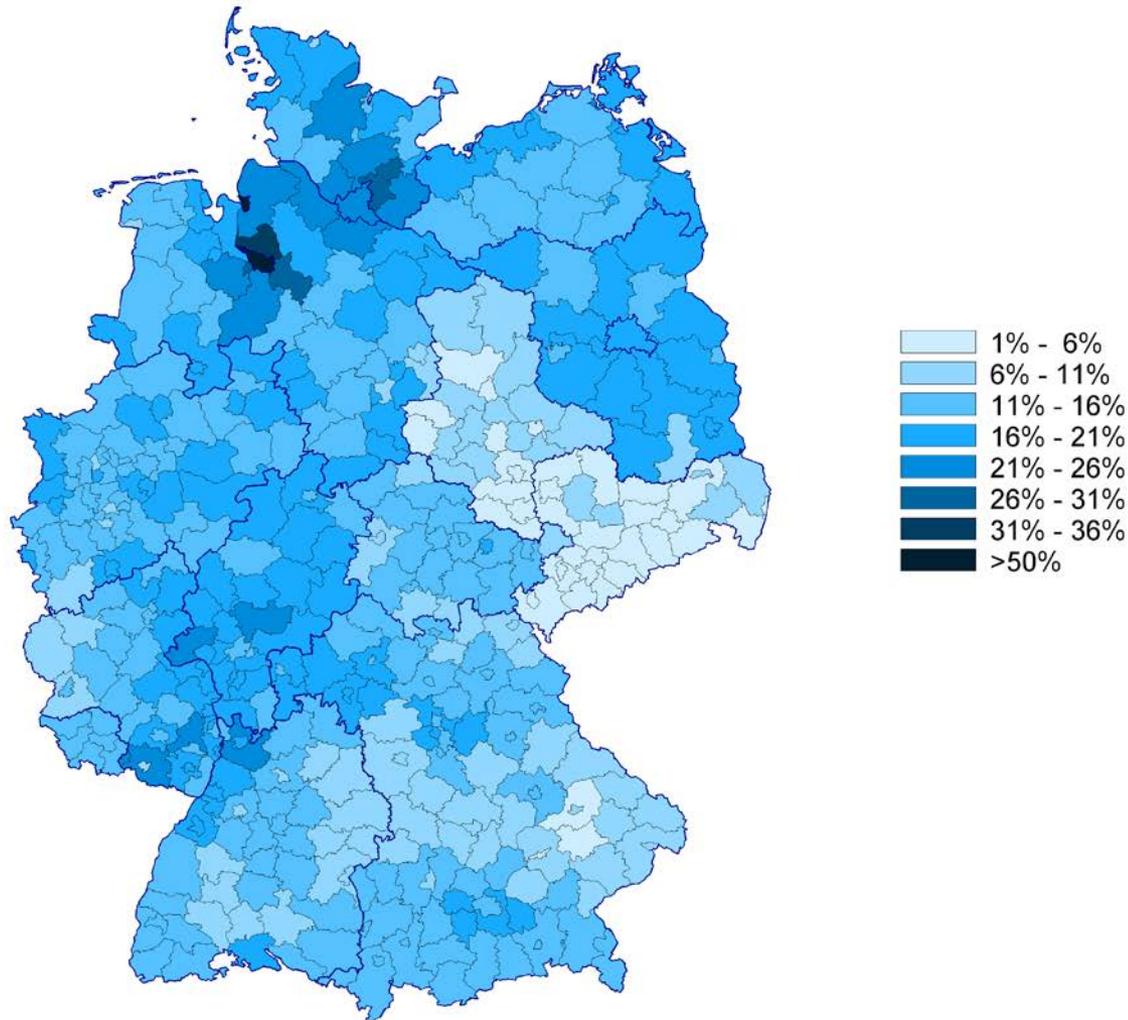
4

- Aufbau und Pflege **seit 2004**
- Basiert auf Abrechnungsdaten **4 gesetzlicher Krankenversicherungen** von **zurzeit 24 Mio. Versicherten**
- Enthält Daten von ungefähr **17% der deutschen Allgemeinbevölkerung** und deckt alle geographischen Regionen der Bundesrepublik ab

Die Forschungsdatenbank GePaRD

Verteilung der Versicherten über Deutschland

5



Die Forschungsdatenbank GePaRD

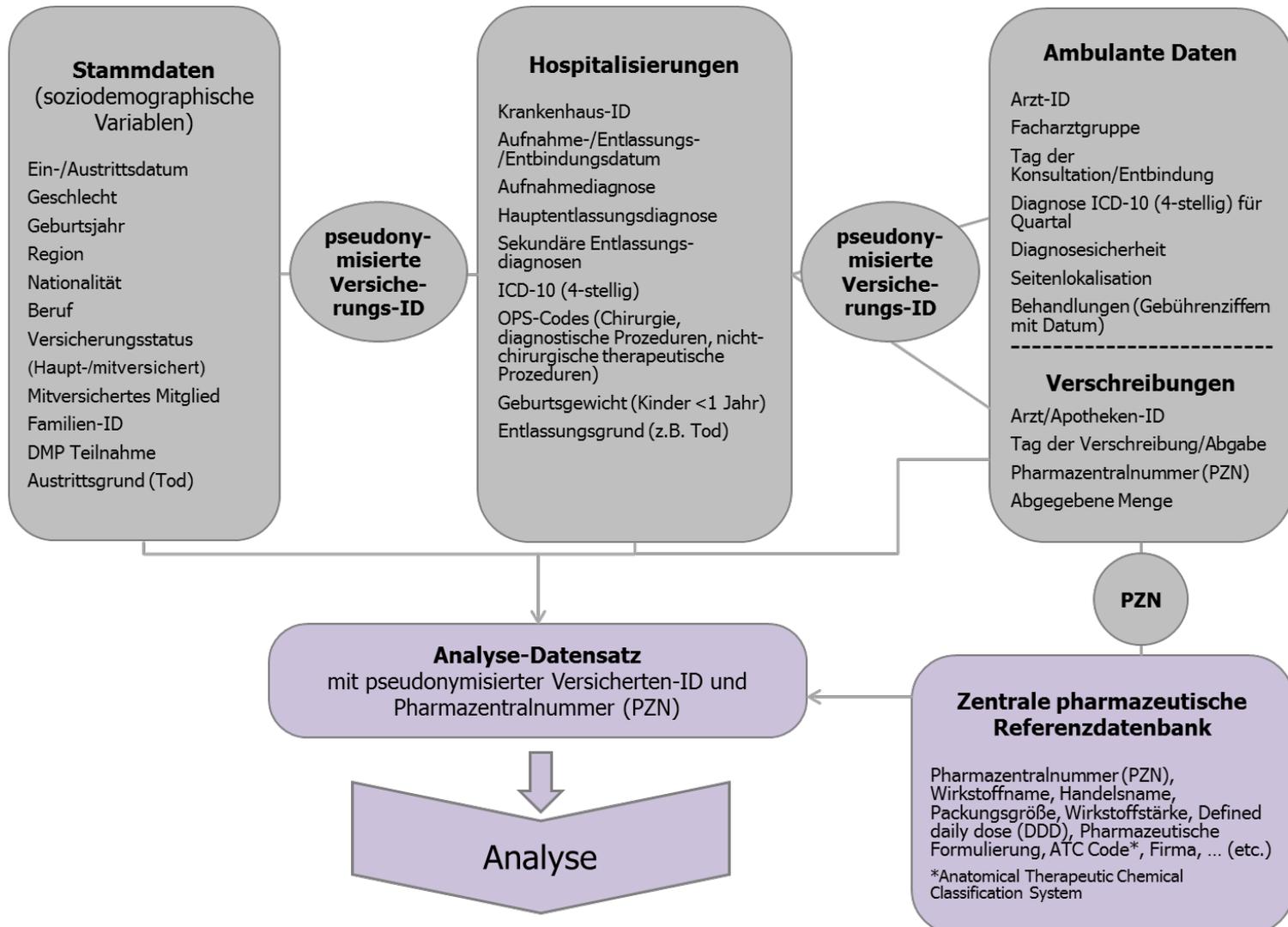


Aufbau

6

- Aufbau und Pflege **seit 2004**
- Basiert auf Abrechnungsdaten **4 gesetzlicher Krankenversicherungen** von **zurzeit 24 Mio. Versicherten**
- Enthält Daten von ungefähr **17% der deutschen Allgemeinbevölkerung** und deckt alle geographischen Regionen der Bundesrepublik ab
- Enthält **demographische Angaben**, Informationen zu **Krankenhausaufenthalten, ambulanten Arztbesuchen und Arzneimittelverschreibungen**

Die Forschungsdatenbank GePaRD



- Monitoring der Anwendung von Arzneimitteln und Impfstoffen nach Zulassung
- Untersuchung von
 - Arzneimittel- und Impfstoffanwendung, Fehlversorgung bei bestimmten Indikationen
 - (Seltenen) Arzneimittel- und Impfstoffrisiken
 - Arzneimittel-/Impfstoffrisiken mit einer langen Latenzzeit
 - Arzneimittelinteraktionen

- Studien vergleichsweise **schnell** und **kostengünstig** durchführbar
- **Retrospektive Kohortenstudien** und **Fall-Kontroll-Studien** möglich
- **Große Patientenzahl** und **viele Datenjahre**
- Untersuchung **seltener**, **schwerwiegender** und **sich spät manifestierender Ereignisse** möglich
- Berücksichtigung **besonderer Patientengruppen**
(z.B. Ältere, Schwangere, Kinder)
- Spezifische Angaben zu **Leistungen**, **Diagnosen** und **Arzneimitteln**
- **Keine studienbedingten Verzerrungen**
(Erinnerungsfehler etc.)
- **Hohe Repräsentativität**
(keine Einwilligung nötig, kaum Loss-to-Follow-up)

Daten wurden zu anderem Zweck erhoben:

Großes Wissen über Möglichkeiten und Limitationen unerlässlich!!

Pharmacoepidemiology and Drug Safety [Explore this journal >](#)

Original Report

Linkage of mother-baby pairs in the German Pharmacoepidemiological Research Database

Edeltraut Garbe [✉](#), Marc Suling, Sebastian Kloss, Christina Lindemann, Ulrike Schmid

First published: 11 November 2010 [Full publication history](#)

Original Report

Evaluation of pregnancy outcome records in the German Pharmacoepidemiological Research Database (GePaRD)[†]

RT Mikolajczyk [✉](#), AA Kraut, E Garbe

First published: 3 June 2013 [Full publication history](#)

DOI: 10.1002/pds.3467 [View/save citation](#)

Behr S, Kollhorst B, Dippel F-W, Garbe E. **Identification of patients with diabetes mellitus type 2 in insurance claims data - Are problems caused by ambiguous ICD-10 codes?** 28th International Conference on Pharmacoepidemiology & Therapeutic Risk Management (ICPE), 23-26 August 2012, Barcelona, Spain. (Abstract published in: Pharmacoepidemiology & Drug Safety. 2012;21(Suppl.3):396)

- Potential für Über-, Unter- und Fehlkodierungen durch Verwendung von Abrechnungsdaten
 - Fehlende Zeiträume vor Einrichtung der Datenbank
 - **Fehlende Angaben zu Arzneimitteln**
(tatsächliche Arzneimittelleinnahme, verschriebene Tagesdosis, Dauer der Therapie, OTC-Arzneimittel, Medikation im Krankenhaus)
 - **Keine Daten zu klinischen Parametern**
 - **Keine Daten zu Lebensstilfaktoren**
(z.B. Bewegung, Gewicht, Rauchen)
- **Mögliche Lösung: Linkage mit anderen Datenquellen**

Bisherige Linkage-Erfahrungen

- Daten eines **Krankenhausinformationssystems (KIS)**
- Evaluierung eines **indirekten Linkage-Ansatzes** im Vergleich zu direktem Linkage (über pseudonymisierte KVNR) als Goldstandard
- Indirekt: Übereinstimmung von **Aufnahme- und Entlassungsdiagnose** sowie definiertem Anteil **zusätzlicher Diagnosen**
- Spezifität durchgängig hoch
- Sensitivität abhängig vom Grad der Übereinstimmung der Diagnosen
 - 30%-70%: 86,7% vs. 100%: 41,7%
- Hohe datenschutzrechtliche und damit große zeitliche/finanzielle Anforderungen
- Viele **fehlende Angaben in KIS**
- **Aufwand gerechtfertigt?**

Gesundheitswesen 2015; 77(02): e8-e14
DOI: 10.1055/s-0034-1395644

Originalarbeit

© Georg Thieme Verlag KG Stuttgart · New York

Verknüpfung von Routinedaten der Gesetzlichen Krankenversicherung mit Daten eines Krankenhausinformationssystems: Machbar, aber auch „nützlich“?

Linkage of Statutory Health Insurance Data with those of a Hospital Information System: Feasible, but also "Useful"?

C. Ohlmeier¹, F. Hoffmann²⁻⁷, K. Giersiepen²⁻⁷, H. Rothgang²⁻⁷, R. Mikolajczyk³⁻⁴, H.-J. Apperlat⁵, A. Elsässer⁶, E. Garbe²⁻⁷

Bisherige Linkage-Erfahrungen

Beispielprojekt II: Mortalitätsregisterdaten

13

- Daten des **Bremer Mortalitätsindex (BreMI)**
- Deterministisches Rekord-Linkage anhand der pseudonymisierten KVNR nicht möglich, da nicht im BreMI enthalten
- **Probabilistisches Linkage** auf Basis personenbezogener Daten (Vorname, 2. Vorname, Nachname, Geburtsname, Geschlecht, Geburtsdatum, Adresse des letzten Wohnorts)
- **Linkage möglich für 83,7% aller Todesfälle** aus GePaRD
- **Todesdatum in 97,1% korrekt**
- **Keine Unterschiede** bzgl. Geschlecht, Alter oder Versichertenstatus

Pharmacoepidemiology and Drug Safety [Explore this journal >](#)

Original Report

Validating mortality in the German Pharmacoepidemiological Research Database (GePaRD) against a mortality registry

Christoph Ohlmeier , Ingo Langner, Edeltraut Garbe, Oliver Riedel

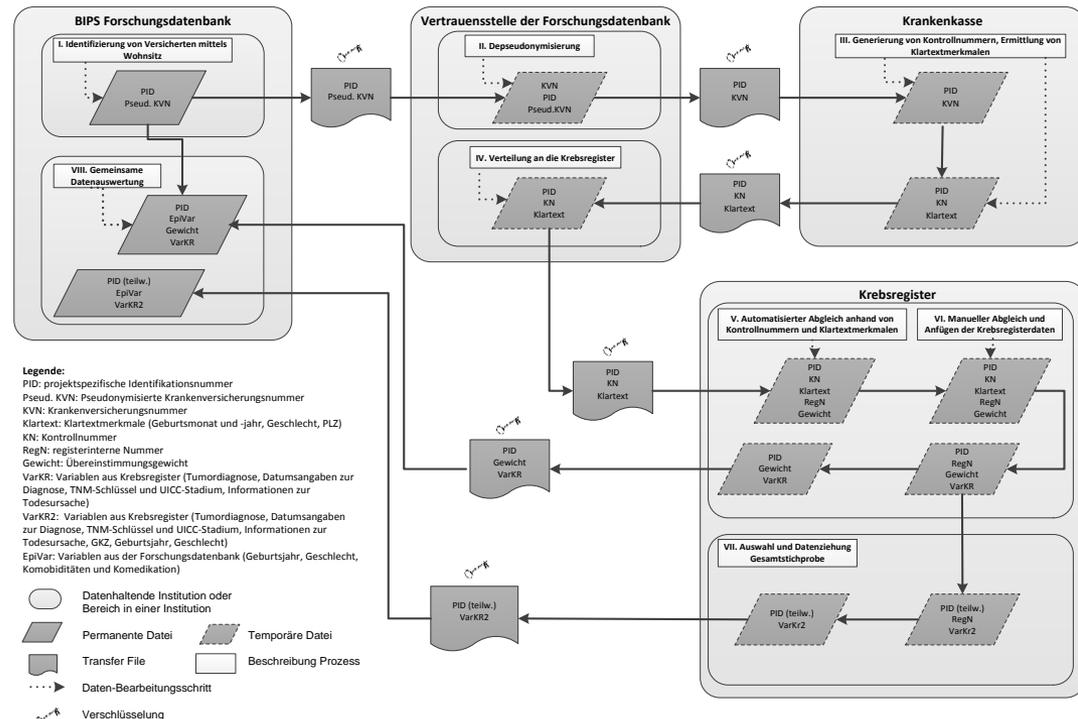
First published: 7 April 2016 [Full publication history](#)

Bisherige Linkage-Erfahrungen

Beispielprojekt III: Krebsregisterdaten

14

- Daten von sechs epidemiologischen Krebsregistern
- **Ziel:** Evaluierung eines indirekten Linkage-Ansatzes im Vergleich zu probabilistischen Linkage (personenbezogene Daten)
- **Indirekt:** Geburtsjahr, Geschlecht, GKZ, Datum der Tumordiagnose
- Hohe datenschutzrechtliche und damit große zeitliche/finanzielle Anforderungen



Bisherige Linkage-Erfahrungen

- Befragungsdaten (Einverständniserklärung erforderlich)
- Aufwendiges Datenschutzkonzept
- Two-Phase Design, d.h. Teilstichprobe mit Befragungsdaten
- Geringe Response: 15.4%
(4000 angeschrieben → 3280 Briefe zugestellt → 505 Interviews durchgeführt)

Pharmacoepidemiology and Drug Safety [Explore this journal >](#)

Original Report

Does additional confounder information alter the estimated risk of bleeding associated with phenprocoumon use—results of a two-phase study

[Sigrid Behr](#) , [Walter Schill](#), [Iris Pigeot](#)

First published: 2 February 2012 [Full publication history](#)

Bisherige Linkage-Erfahrungen

- NAKO Gesundheitsstudie: prospektive Kohortenstudie
- Kompetenznetz Sekundär- und Registerdaten
- **Aufgabe:** Beschaffung und Aufbereitung von GKV-Daten, Beratung von NutzerInnen
- Sehr hohe datenschutzrechtliche Anforderungen
- Einzigartige Datenbasis durch Kombination verschiedener Datenquellen für Forschung nutzbar
- Bisher **hoher Anteil** ($\geq 90\%$) an **Zustimmung** zur Verlinkung der Primärdaten mit verschiedenen Sekundärdatenquellen



- Wichtige Informationen **ohne zusätzliche zeitliche Belastung** der Teilnehmer (z.B. Berufsbiographie)
- **Keine Erinnerungslücken** und **hohes Maß an Spezifität**
 - kann bei Befragungen nicht erreicht werden (z.B. Medikamenteneinnahme)
- Erfassung **vergängerer und zukünftiger Ereignisse** (z.B. Erkrankungen)
 - Auch wenn TeilnehmerIn (z.B. aus gesundheitlichen Gründen) **nicht ins Studienzentrum kommen kann**
- **Validierung/Ergänzung** der Primärdaten
- **Vermeidung von Fehlklassifikation**
- Unterstützung des **Follow-up**
 - Kontinuierliche detaillierte Informationen zu Gesundheit, Leistungsanspruchnahme, Rentenverlauf, Berufsbiografie

- Linkage auf Basis personenidentifizierender Daten
 - Hohe datenschutzrechtliche Anforderungen und hoher administrativer Aufwand
 - Genehmigung durch Krankenkassen und Aufsichtsbehörden
 - Datenschutzkonzept
 - (De-)Pseudonymisierung über Vertrauensstelle(n)
 - ...
- Direktes Linkage
 - Teilweise aufgrund der Heterogenität der Daten zwischen einzelnen Leistungserbringern und Datenquellen nicht möglich
- Indirektes Linkage
 - Teilweise Einschränkungen in der Qualität der Ergebnisse

Bisherige Linkage-Erfahrungen

Probleme:

- Löst Problem fehlender Informationen zur Beantwortung von Arzneimittelsicherheitsfragen nicht vollständig
- Aufbereitung der uneinheitlichen Daten und Datenstrukturen
- Datenqualität

Vielen Dank für Ihre Aufmerksamkeit!

www.leibniz-bips.de

Kontakt

Prof. Dr. Iris Pigeot, Dr. Bianca Kollhorst
Leibniz-Institut für Präventionsforschung
und Epidemiologie – BIPS
Achterstraße 30
28359 Bremen

pigeot@leibniz-bips.de
kollhorst@leibniz-bips.de

