

SEKUNDÄRDATEN IN DER PHARMAKOEPIDEMIOLOGISCHEN FORSCHUNG:

KÖNNEN MODERNE STATISTISCHE VERFAHREN LIMITATIONEN DER DATEN AUSGLEICHEN?

Iris Pigeot, Vanessa Didelez, Dirk Enders, Bianca Kollhorst

Leibniz-Institut für Präventionsforschung und Epidemiologie – BIPS

2. März 2018

2. DVMD-Frühjahrssymposium: „Klassifikationen und Gesundheitsstatistik“
Duisburg

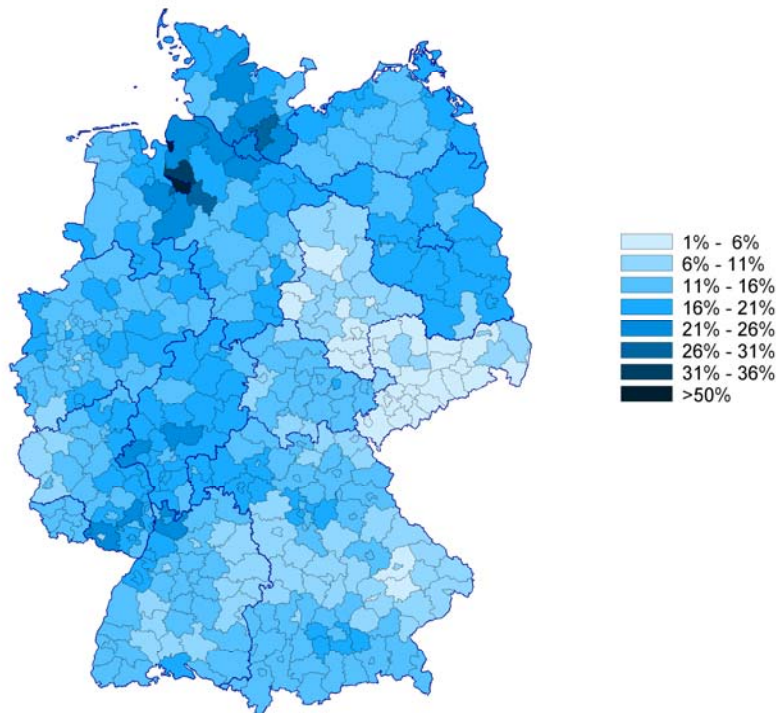
Inhalt

-
- Pharmakoepidemiologische Forschung
 - Die Forschungsdatenbank GePaRD
 - Statistische Ansätze zur Berücksichtigung von Confounding
 - Bisherige Linkage-Erfahrungen

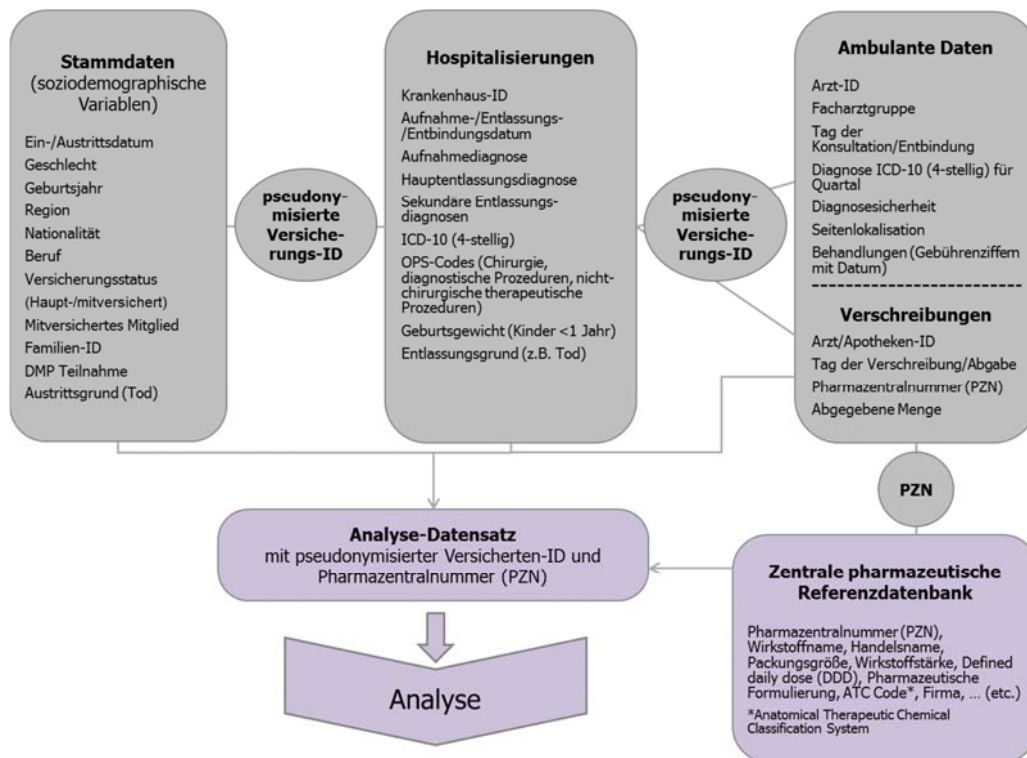
- Untersucht Anwendung, Wirkungen und Risiken von Arzneimitteln nach Marktzulassung
- **Limitationen der klinischen Studien vor Zulassung**
 - Zu geringe Größe (i.d.R. 3000-5000 Patienten)
 - Zu geringe Studiendauer
 - Vulnerable Patientengruppen häufig ausgeschlossen
- **Risiken von nicht bestimmungsgemäßigem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

Die Forschungsdatenbank GePaRD

- Aufbau 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



- 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**



- **Monitoring der Anwendung von Arzneimitteln und Impfstoffen nach Zulassung**
- **Untersuchung von**
 - Arzneimittel- und Impfstoffanwendung, Fehlversorgung bei bestimmten Indikationen
 - (Seltene) 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. Alte, Pflegebedürftige, Schwerkranke, Verstorbene)
- Spezifische Angaben zu erhaltenen **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)[†]

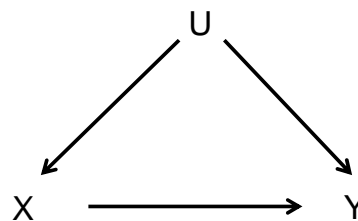
RT Mikolajczyk [✉](#), AA Kraut, E Garbe

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 Arzneimittel-Angaben**
(z.B. tatsächliche Arzneimitteleinnahme, Tagesdosis, Therapiedauer, OTC-Präparate)
- **Fehlende Laborparameter**
- **Fehlende Angaben zu Lebensstilfaktoren**
(z.B. Bewegung, Gewicht, Rauchen)
- Ungemessenes Confounding



Mögliche Lösungsansätze

- **Statistische Ansätze**
 - Propensity Scores
 - Instrumental Variables
 - Two-Phase Designs
 -
- **Datenanreicherung durch Linkage mit anderen Datenquellen**
 - Primärdaten
 - Registerdaten
 - Sekundärdaten
 - ...

High-dimensional Propensity Scores I



Methode: Idee

13

- **Propensity Score (PS):** W-keit für Behandlung, bedingt auf **manuell** definierten Kovariablen (X_m)
- **Problem:** Viel Information der Datenbank bleibt ungenutzt
- **High-dimensional PS (HDPS) Algorithmus:** Erweiterung des PS um **automatisch** selektierte Kovariablen (X_a) mit größtem Potential zur Confounderkontrolle, d.h.
 1. Bildung von Kovariablen zu häufigsten Codes (ICD, ATC, OPS, EBM) in Datenbank für jeweilige Studienpopulation
 2. Sortierung nach Relevanz, X_a = top n Kovariablen (oft mehrere hundert)
 3. Anpassung eines logistischen Modells: $\text{logit}(HDPS) = \beta_0 + \beta'_m X_m + \beta'_a X_a$

High-dimensional Propensity Scores II



Anwendungsbeispiel: Studiendesign

14

- **Ziel:**

Vergleich des Mortalitätsrisikos zwischen ambulant und stationär durchgeführten perkutanen koronaren Interventionen (PCIs) zur Behandlung der koronaren Herzkrankheit (KHK)
- **Design:**

Kohortenstudie basierend auf Daten von 3 Krankenkassen (2004-2009)
- **Confounding durch Krankheitsschwere:**

Sektor der PCI abhängig von allg. Gesundheitszustand des Patienten und Schwere der KHK
- **Statistische Analysen zur Kontrolle des Confounding:**
 - Primäre Analyse: Multivariable Cox-Regression, adjustiert für kardiovaskuläre Komorbiditäten und Prädiktoren von Krankenhausmortalität
 - Sensitivitätsanalysen basierend auf PS (z.B. Matching)

High-dimensional Propensity Scores III



Anwendungsbeispiel: Ergebnisse; Teil 1

15

- **Ergebnis:**
Signifikant reduzierte Mortalität um 40% für ambulante PCIs in allen Analysen
- **Kontraintuitiv:**
Leicht erhöhte Mortalität erwartet (Standardisierung und bessere Überwachung im Krankenhaus)
- **Möglicherweise verursacht durch ungemessenes Confounding:**
Schwere der KHK nicht erfasst in GePaRD
- **Lösungsansatz: High-dimensional PS (HDPS)**
 - Automatische Selektion von hunderten empirischen Confoundern aus GePaRD
 - „Hoffnung“: Gesamtheit an empirischen Confoundern ist Proxy für Schwere der KHK

High-dimensional Propensity Scores IV



Anwendungsbeispiel: Ergebnisse; Teil 2

16

- Erneute Analyse der Kohorte basierend auf HDPS
- **Ergebnis:**
Erhöhte Mortalität um 20% für ambulante PCIs in HDPS-Analysen (nicht signifikant)
- **Plausible Ergebnisse** der HDPS-Analysen
- **Mit Vorsicht zu interpretieren:**
 - Generalisierbarkeit eingeschränkt: Gesundere und wesentlich kleinere Studienpopulation als in der ursprünglichen Kohorte
 - Ungemessenes Confounding trotz umfangreicher Adjustierung nie auszuschließen



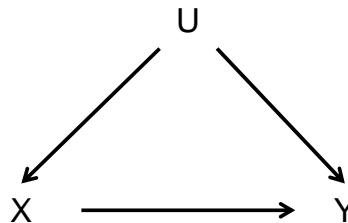
© Health Research and Educational Trust
DOI: 10.1111/1473-6773.12653
RESEARCH ARTICLE

The Potential of High-Dimensional Propensity Scores in Health Services Research: An Exemplary Study on the Quality of Care for Elective Percutaneous Coronary Interventions

Dirk Enders, Christoph Ohlmeier, and Edeltraut Garbe

Interesse:

Effekt einer Exposition X auf den Gesundheitsstatus Y



Problem:

Exposition und Gesundheitsstatus abhängig von weiteren, insb. **unbeobachteten** Einflussfaktoren U

➤ **Confounding** → verzerrte Schätzung des Expositionseffekts

Instrumentelle Variablen II

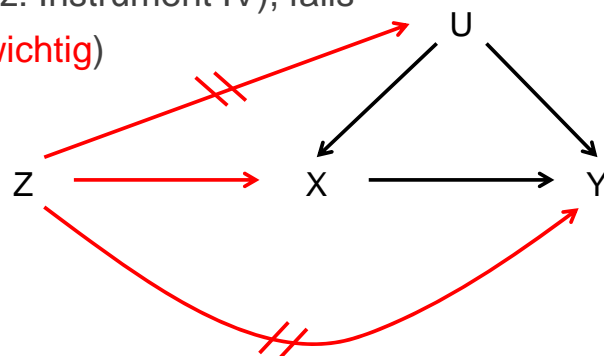
Idee: Finde beobachtbare Variable Z, die zufällig Exposition voraussagt, aber keinen direkten Einfluss auf Gesundheitsstatus hat

➤ **Z instrumentelle Variable** (kurz: Instrument IV), falls

1. $Z \not\perp\!\!\!\perp X$ (**starke Assoziation wichtig**)

2. $Z \perp\!\!\!\perp U$
Independence Assumption

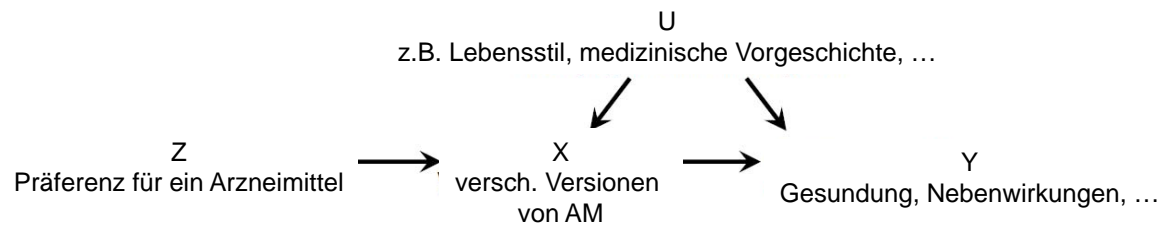
3. $Z \perp\!\!\!\perp Y | (X,U)$
Exclusion Restriction



➤ **Problem:** 2. und 3. nicht empirisch überprüfbar → **Hintergrundwissen!**

$A \perp\!\!\!\perp B | C$: A bedingt unabhängig von B, gegeben C

Beliebtes IV: Arztpräferenz



Beispiel:

- Behandlung X = Coxibe vs. NSAIDs (Brookhart et al. 2006)
- „Confounding by Indication“ großes Problem
- **Vorschlag:** Arztpräferenz für bestimmtes Arzneimittel als IV

Instrumentelle Variablen IV

- **Mögliche Definitionen, um die Arztpräferenz zu schätzen:**
 - Behandlung, die dem **vorherigen Patienten** verschrieben wurde
 - Anteil aller **vorheriger Patienten**, denen COX-2 Hemmer verschrieben wurden
 - Nur moderate Stärke der Assoziation zwischen Behandlung und IV
- **Valides Instrument in einem Gesundheitssystem nicht zwangsläufig übertragbar auf andere Gesundheitssysteme**

Instrumentelle Variablen V



Anwendungsstudie: Ergebnisse

21

	Konventionelle Analyse RD (95% KI)	IV Analyse RD (95% KI)
Gesamte Kohorte	-0,01 (-0,15; 0,12)	-0,02 (-0,80; 0,76)

RD: Risikodifferenz; KI: Konfidenzintervall

- IV Schätzer weiteres Konfidenzintervall als konventioneller Schätzer, bedingt durch moderate Assoziation zwischen Behandlung und Instrument
- Möglicherweise keine Beeinträchtigung der konventionellen Analyse durch ungemessenes Confounding

Instrumentelle Variablen V



Anwendungsstudie: Ergebnisse

22

	Konventionelle Analyse RD (95% KI)	IV Analyse RD (95% KI)
Gesamte Kohorte	-0,01 (-0,15; 0,12)	-0,02 (-0,80; 0,76)

RD: Risikodifferenz; KI: Konfidenzintervall

- IV Schätzer weiteres Konfidenzintervall als konventioneller Schätzer, bedingt durch moderate Assoziation zwischen Behandlung und Instrument
- Möglicherweise keine Beeinträchtigung der konventionellen Analyse durch ungemessenes Confounding



Journal of Clinical Epidemiology 69 (2016) 96–106



The proportion of all previous patients was a potential instrument for patients' actual prescriptions of nonsteroidal anti-inflammatory drugs

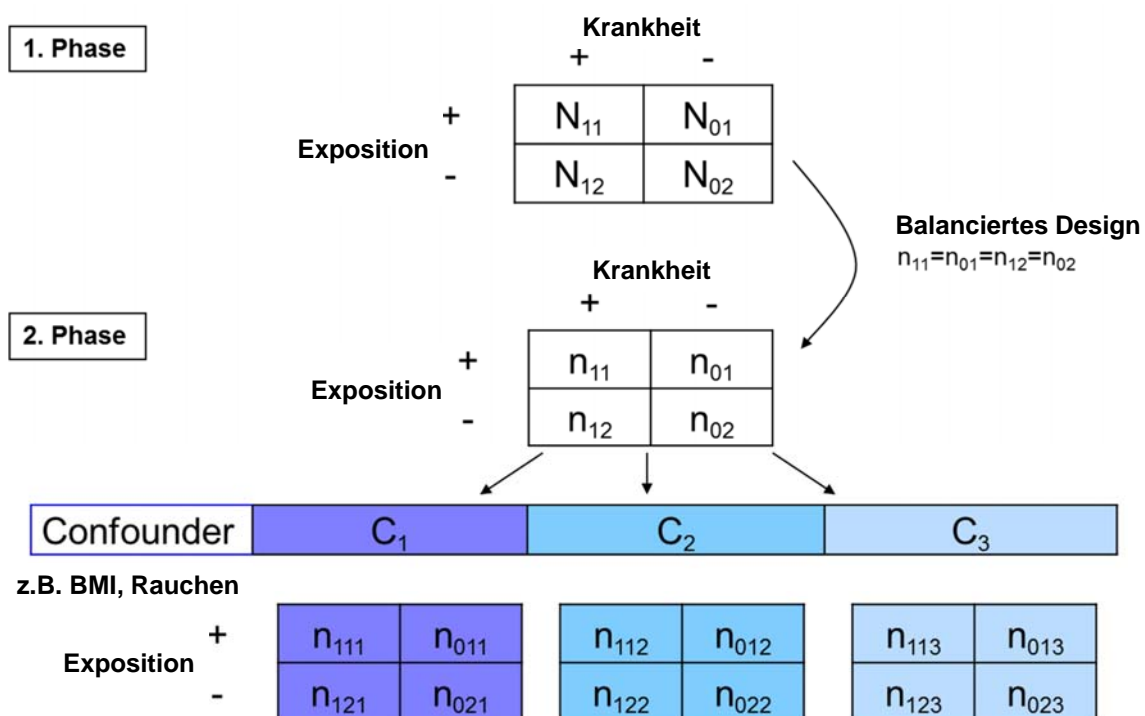
Bianca Kollhorst^{a,b}, Michal Abrahamowicz^{b,c}, Iris Pigeot^{a,d}

^a Leibniz Institute for Prevention Research and Epidemiology—BIPS, Department Biometry and Data Management, Achterstr. 30, Bremen 28359, Germany
^b Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, 3820 Pine Avenue West, Montreal H3A 1A2, Quebec, Canada
^c Division of Epidemiology, McGill University Health Centre Royal Victoria Hospital, 687 Pine Avenue West, Montreal H3A 1A1, Quebec, Canada
^d Faculty of Mathematics and Computer Science, University of Bremen, Bibliothekstraße 1, Bremen 28359, Germany

Accepted 24 August 2015; Published online 2 September 2015

1. **Studiendesign:** Fall-Kontroll-Studie
2. **Zwei Datenquellen:**
 - Daten für alle Patienten (**Phase 1**): Fall-Kontroll-Status, Exposition, Confounder
 - Zusätzliche Confounder für Teil der Patienten (**Phase 2**)
3. **Stratifizierung aus Exposition und Auswahl an Confoundern in Phase 1**
 - Teilnahme an Phase 2 hängt nur von Stratifizierung und Fall-Kontroll-Status ab
 - Keine leeren Schichten in Phase 1 und Phase 2 für Fälle bzw. Kontrollen
4. **Analysemethoden:**
Weighted Likelihood, Pseudo Likelihood, **Maximum Likelihood**

Two-Phase Designs II



▪ **Ziel:**

Vergleich Therapie Sulfonylharnstoffe vs. DPP4-Inhibitoren hinsichtlich des kombinierten Risikos aus Myokardinfarkt, Schlaganfall und Herzinsuffizienz in Typ 2-Diabetikern mit 1st-Line Therapie Metformin

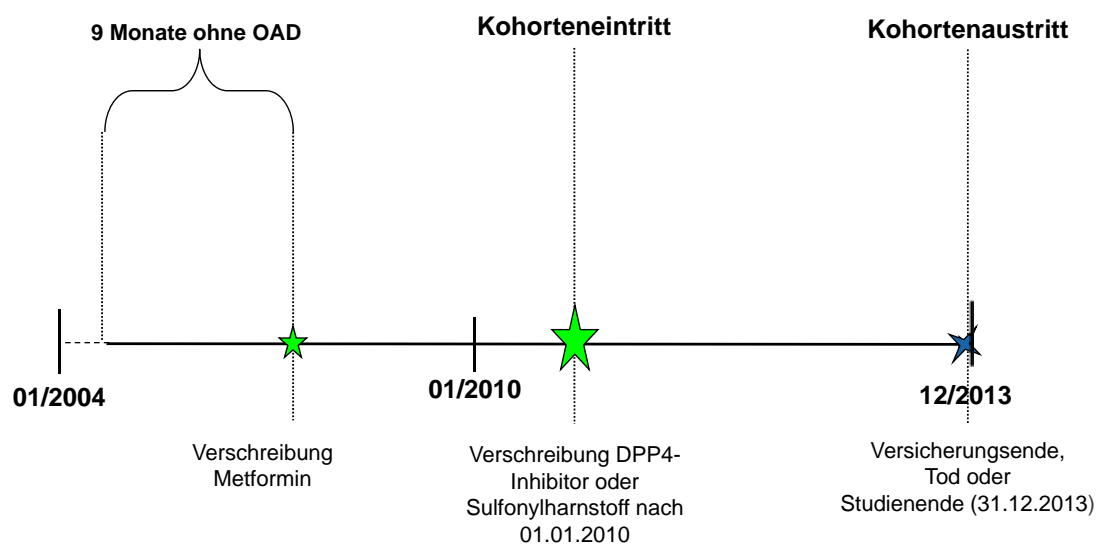
▪ **Studiendesign:**

Eingebettete Zwei-Phasen Fall-Kontroll-Studie

▪ **Datenquellen:**

- **Phase 1:** Pharmakoepidemiologische Forschungsdatenbank (GePaRD)
- **Phase 2:** Disease Management Programm Diabetes mellitus Typ 2

Two-Phase Designs IV



Two-Phase Designs V

- **Fälle:**
Patienten mit Myokardinfarkt, Schlaganfall oder Herzinsuffizienz innerhalb des Follow-up
 - **Kontrollen:** gematcht im Verhältnis 1:10
 - **Exposition (2 Kategorien):**
 - Metformin + Sulfonylharnstoffe
 - Metformin + DPP4-Inhibitoren
- **Teilnahme an Phase 2:**
Untersuchung im DMP vor Kohorteneintritt bis 28 Tage danach

Two-Phase Designs VI

Phase 1

GePaRD 

- Alter, Geschlecht
- Kardiovaskuläre Komorbiditäten
- Schwere des Diabetes
- Sonstige Komorbiditäten
- Komedikationen

Phase 2



- Rauchverhalten
- BMI
- HbA1c-Wert
- Schulungen
- Diabetes-bezogene Überweisungen

1. Phase 1 Analyse

- Logistische Regression mit allen Patienten
- Nur Confounder aus GePaRD

2. Complete-Case Analyse

- Gewichtete logistische Regression mit Patienten aus Phase 2
- Confounder aus GePaRD und DMP
- Gewicht=Inverse der Selektionswahrscheinlichkeit für Teilnahme an Phase 2 (geschätzt)

3. 2-Phasen Analyse

- 2-Phasen logistische Regression
- Confounder aus GePaRD und DMP
- Stratifizierung nach Exposition, Alter & Geschlecht
- Maximum Likelihood Methode

Two-Phase Designs VIII

	Phase 1 (N=3.179)		Phase 2 (N=1.968)	
	Fälle N=289	Kontrollen N=2.890	Fälle N=168	Kontrollen N=1.800
Mittleres Alter in Jahren (StD)	66,1 (10,03)	59,0 (10,59)	65,6 (9,86)	59,0 (10,43)
Männliches Geschlecht	229 (79,2%)	2.072 (71,7%)	136 (81,0%)	1.253 (69,6%)
Exposition				
Metformin + Sulfonylharnstoffe	76 (26.3%)	581 (20.1%)	46 (27.4%)	365 (20.3%)
Metformin + DPP4-Inhibitoren	213 (73.7%)	2.309 (79.9%)	122 (72.6%)	1.435 (79.7%)


Odds Ratio (95%-Konfidenzintervall)

	Phase 1 Analyse	Complete-Case Analyse	2-Phasen Analyse
DPP4-Inhibitoren vs. Sulfonylharnstoffe	0.81 (0.60–1.09)	0.76 (0.52–1.12)	0.83 (0.61–1.13)
Geschlecht (Referenz: weiblich)	1.43 (1.05–1.96)	1.52 (1.01–2.29)	1.37 (0.99–1.89)
Alter (Referenz: <55 Jahre)			
55- < 65 Jahre	1.60 (1.08–2.39)	1.78 (1.05–3.04)	1.62 (1.08–2.45)
>= 65 Jahre	2.72 (1.87–3.97)	2.78 (1.65–4.68)	3.01 (1.98–4.60)
Vorgeschichte ischämische Herzkrankheit (Referenz: nein)	1.91 (1.43–2.55)	1.76 (1.18–2.63)	1.78 (1.22–2.59)
Vorgeschichte Schlaganfall (Referenz: nein)	3.19 (2.00–5.10)	2.81 (1.38–5.71)	2.64 (1.31–5.29)
Vorgeschichte Herzinsuffizienz (Referenz: nein)	2.29 (1.58–3.31)	2.09 (1.27–3.44)	2.18 (1.32–3.58)
HbA1c >=7.5% (Referenz: nein)	-	1.23 (0.86–1.75)	1.22 (0.86–1.72)
BMI (Referenz: >=30)	-	-	-
<25	-	1.17 (0.66–2.09)	1.17 (0.66–2.09)
25-<30	-	1.09 (0.76–1.56)	1.09 (0.76–1.56)
Raucher (Referenz: nein)	-	1.38 (0.83–2.30)	1.38 (0.83–2.30)

Two-Phase Designs X

- Odds Ratio für Behandlung **nur unwesentlich** durch Informationen des DMP **beeinflusst**
- Mögliche Erklärungen:
 - **Adäquate Berücksichtigung wesentlicher Confounder** durch Information in GePaRD
 - **Geringe Qualität der DMP-Daten**
- **2-Phasen Analyse** im Vergleich zu Complete-Case Analyse **präziser**


Journal of Diabetes and Its Complications 30 (2016) 1339–1346



Contents lists available at ScienceDirect

Journal of Diabetes and Its Complications

journal homepage: WWW.JDCJOURNAL.COM



Comparative risk for cardiovascular diseases of dipeptidyl peptidase-4 inhibitors vs. sulfonylureas in combination with metformin: Results of a two-phase study



Dirk Enders ^a, Bianca Kollhorst ^a, Susanne Engel ^b, Roland Linder ^b, Frank Verheyen ^b, Iris Pigeot ^{a,*}

^a Leibniz Institute for Prevention Research and Epidemiology—BIPS, Bremen, Germany
^b Scientific Institute of TK for Benefit and Efficiency in Health Care, Hamburg, Germany

Bisherige Linkage-Erfahrungen I

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



Bisherige Linkage-Erfahrungen II

- Daten des **Bremer Mortalitätsindex (BreMI)**
- Deterministisches Rekord-Linkage anhand von pseudonymisierter Versichertennr. nicht möglich, da nicht vorhanden im BreMI
- **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

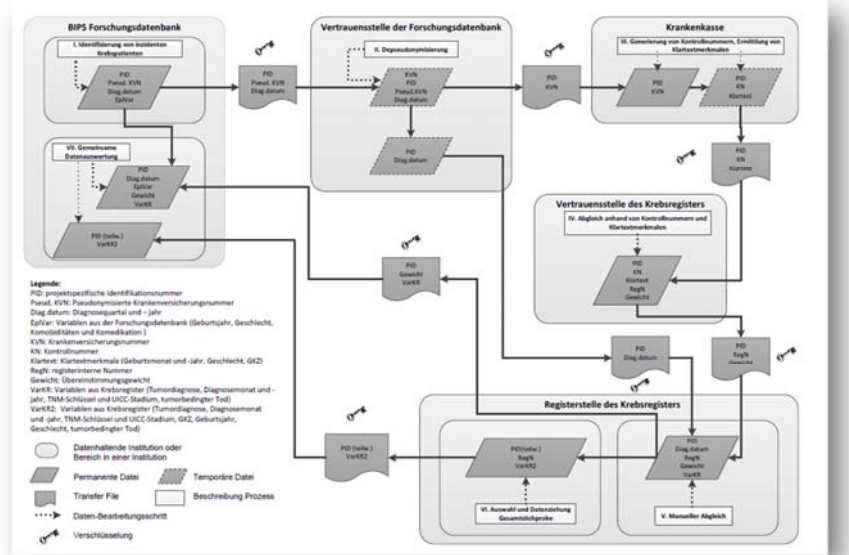
Validating mortality in the German Pharmacoepidemiological Research Database (GePaRD) against a mortality registry

Christoph Ohlmeier , Ingo Langner, Edeltraut Garbe, Oliver Riedel

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

Bisherige Linkage-Erfahrungen III

- Daten epidemiologischer Krebsregister
- DFG-Projekt: Beginn 01.12.2016
- **Ziel:** Evaluierung eines **indirekten Linkage-Ansatzes** im Vergleich zu direktem Linkage → kanzerogenes Potential von Antidiabetika
- Hohe datenschutzrechtliche und damit große zeitliche/finanzielle Anforderungen



Bisherige Linkage-Erfahrungen IV

- 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](#)

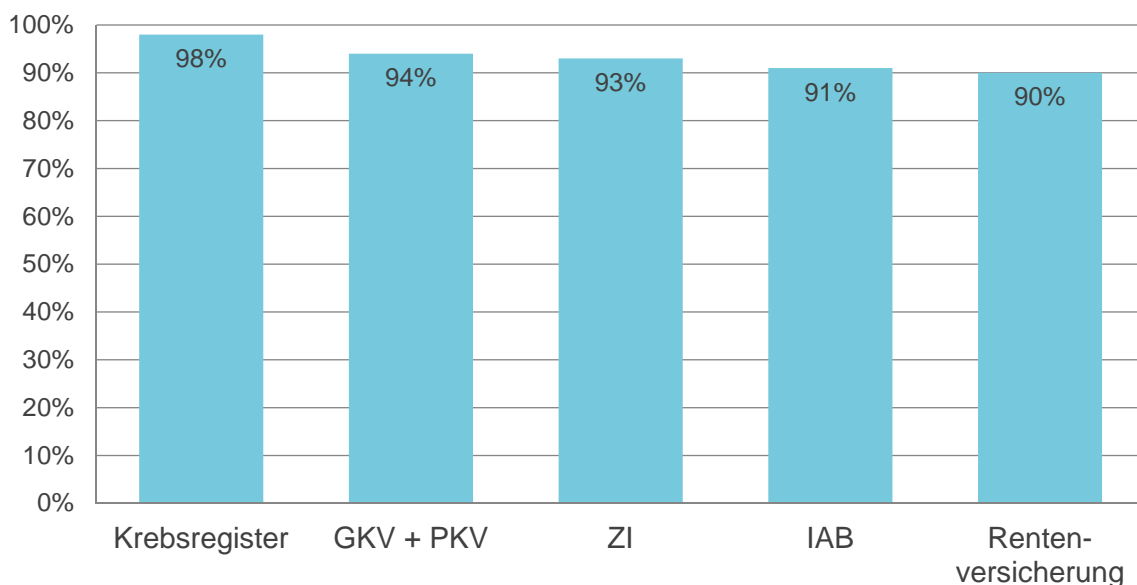
Bisherige Linkage-Erfahrungen V

- NAKO Gesundheitsstudie: prospektive Kohortenstudie (2013-2022; geplant für weitere 20-30+ Jahre) mit 200.000 Personen, 20-69 Jahre, 18 Studienzentren
- Kompetenznetz Sekundär- und Registerdaten
 - **Aufgabe:** Beschaffung und Aufbereitung von GKV-Daten, Beratung von Nutzern
 - Sehr hohe datenschutzrechtliche Anforderungen
 - Sobald Daten vorliegen, niedrighschwellige Nutzung für Forschung möglich



NAKO Gesundheitsstudie

Anteil der Zustimmungen bis 31.12.2017: n = 131.619



- Wichtige Informationen **ohne zusätzliche zeitliche Belastung** der Teilnehmer (z.B. Berufsbiographie)
- **Keine Erinnerungslücken** und **hohes Maß an Spezifität**, die bei Befragungen nicht erreicht werden kann (z.B. Medikamenteneinnahme)
- **Einzigartige Datenbasis** durch Kombination verschiedener Datenquellen
- **Validierung/Ergänzung** der Primärdaten
- **Vermeidung von Fehlklassifikation**
- Unterstützung des **Follow-up**: Kontinuierliche detaillierte Informationen zu Gesundheit, Leistungsanspruchnahme, Rentenverlauf und Berufsbiografie

Bisherige Linkage-Erfahrungen

- **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

Vielen Dank für Ihre Aufmerksamkeit!

www.leibniz-bips.de

Kontakt

Prof. Dr. Iris Pigeot
Leibniz-Institut für Präventionsforschung
und Epidemiologie – BIPS
Achterstraße 30
28359 Bremen
pigeot@leibniz-bips.de



Literatur



42

Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. (2006) Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology* 17:268-75

Schneeweiss S (2008) Confounding. In: *Pharmacoepidemiology and Therapeutic Risk Management*. Ed. by AG Hartzema, HH Tilson, KA Chan. Harvey Whitney Books: Cincinnati, pp. 273-99

Schneeweiss S, Suissa S (2012) Advanced approaches to controlling confounding in pharmacoepidemiologic studies. In: *Pharmacoepidemiology*. Ed. by BL Strom, SE Kimmel, S Hennessy. 5th edn. John Wiley and Sons: Chichester, pp. 868-91

Staiger D, Stock JH (1997) Instrumental variables regression with weak instruments. *Econometrica* 65:557-86

Stock JH, Wright JH, Motohiro Y (2002) Survey of weak instruments and weak identification in generalized method of moments. *J Bus Econ Stat* 20:518-29

Uddin MJ, Groenwold RHH, Sanni Ali M, de Boer A, Roes KCB, Chowdhury MAB, Klungel OH (2016) Methods to control for unmeasured confounding in pharmacoepidemiology: An overview. *Int J Clin Pharm* 38: 714-23