

Personenbezogene Verknüpfung von Befragungs-, Versorgungs- und GKV-Routinedaten zur Evaluation einer komplexen Intervention: Zum Konzept der Datenflüsse

Sandra Salm, Nadine Scholten, Holger Pfaff & Antje Dresen
Institut für Medizinsoziologie, Versorgungsforschung und Rehabilitationswissenschaft, Universität zu Köln

Hintergrund & Ziel

Im Projekt isPO (Integrierte, sektorenübergreifende Psychoonkologie) wird ein Versorgungsprogramm entwickelt, umgesetzt und evaluiert. Hierzu erfolgt ein IT-gestütztes, gestuftes Vorgehen, das zusätzlich Maßnahmen der Qualitätssicherung einschließt.

Je nach individuellem Bedarf werden die Krebspatient*innen inklusive eines Informationsangebots der Krebsselfhilfe über ein Jahr ab Erstdiagnose psychosozial und/oder psychotherapeutisch begleitet.

Um die Wirksamkeit und Qualität des isPO-Programms auf Patient*innenebene zu überprüfen, sollen Befragungsdaten mit Versorgungs- sowie GKV-Routinedaten von drei Kassen verknüpft werden.

Methodisches Kernproblem

Neben der Analyse diverser Outcome-Parameter liegt der Vorteil des Datenlinkages in der Möglichkeit der externen Validierung der GKV-Daten mithilfe eines Goldstandards (Hoffmann et al., 2008). Die besondere Herausforderung ist in isPO durch die Anzahl der Datenquellen sowie der mehrmaligen Datenerhebung bzw. -lieferung gegeben. Die Versorgungsdaten werden dem isPO-Data Warehouse (isPO-DW) vierteljährlich übermittelt. Die Befragungen erfolgen 4 und 12 Monate nach der Erstdiagnose. Zudem ist die Zusammenführung der Daten der drei GKVen erforderlich, die mitunter die gleichen Informationen in unterschiedlicher Datenarchitektur speichern (March et al., 2012). Des Weiteren stellt das personenbezogene Datenlinkage zusätzliche Ansprüche an den Datenschutz.

Lösungsansätze

Um das Linkage der verschiedenen Daten und die Wahrung des Datenschutzes zu gewährleisten, ist ein Konzept erstellt worden, das die Datenflüsse und Pseudonymzuordnungen transparent macht (Abb. 1).

Die Versorgungseinrichtungen geben die Versorgungsdaten pseudonymisiert an das isPO-DW weiter (Ps. 1).

Lösungsansätze (Fortsetzung)

Neben den Versorgenden kennt nur das eingerichtete Trust Center (TC) die Zuordnung von Pseudonym und versorgter Person. Das TC generiert zwei weitere Pseudonyme; eines, das auf den Fragebögen enthalten ist (Ps. 2) und ein weiteres für die Datenanfrage bei den GKVen (Ps. 3). Mithilfe von identifizierenden Daten (Name, Versichertenr., Geburtsdatum) können die Kassen das Ps. 3 ihren Daten zuordnen und nach Entfernung der identifizierenden Variablen einen pseudonymisierten Datensatz an das TC übermitteln. Um die Zusammenführung der verschiedenen GKV-Daten im TC zu gewährleisten, erfolgt eine Synchronisierung der Datensätze durch die Kassen selbst im Zuge einer gegründeten AG Datenmanagement.

Im TC werden in den Befragungs- und GKV-Daten die eigens generierten Pseudonyme durch das Ps. 1 ersetzt und an das isPO-DW übermittelt und dort mit den Versorgungsdaten verlinkt.

Diskussion

Durch die Generierung mehrerer Pseudonyme und der alleinigen Kenntnis des Pseudonym-Triples durch das TC soll der Rückschluss auf einzelne Personen ausgeschlossen werden. Allerdings können vor allem aufgrund der mehrfachen Datenerhebungen bzw. -lieferungen Fehler entstehen, wenn das Pseudonymisierungsverfahren nicht gleichgehalten wird, was wiederum das Datenlinkage erschweren könnte.

Außerdem steht zur Diskussion, wie sich die Bereitschaft zur Datennutzung sowie zum Datenlinkage und damit die Teilnahme am isPO-Programm darstellen. Denn in diesem besonderen Setting geht es um Menschen, die erst kürzlich die Diagnose Krebs erhalten haben.

Schlussfolgerung

Die Verknüpfung von Befragungs-, Versorgungs- und GKV-Daten stellt eine Bereicherung zur Outcome-Evaluation einer komplexen Intervention unter Alltagsbedingungen dar. Allerdings gilt es, die hohen Anforderungen an den Datenschutz umzusetzen und gleichbleibende Verfahren zur (Um-)Pseudonymisierung sicherzustellen.



isPO
Integrierte, sektorenübergreifende
Psychoonkologie

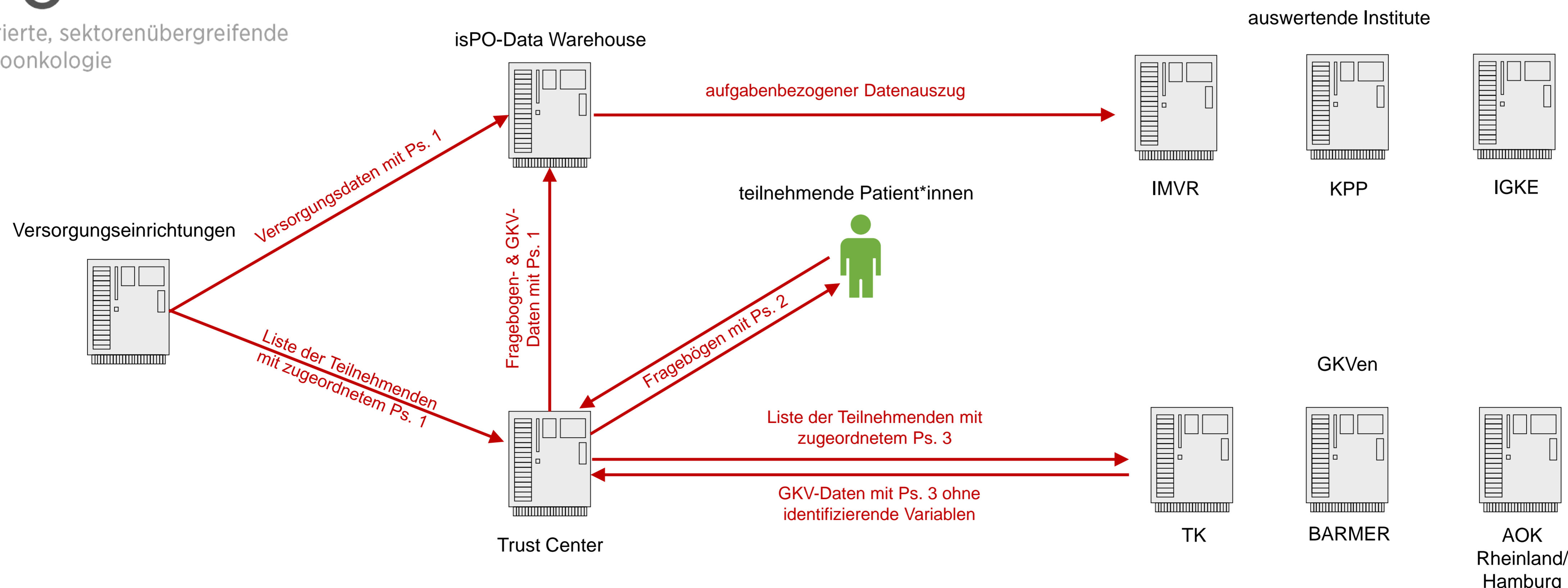


Abb. 1. Schematische Darstellung der Datenflüsse und Pseudonymisierungen im isPO-Projekt

Kontakt

Sandra Salm, M.Sc.

IMVR
Eupener Straße 129
D-50933 Köln

Telefon +49 (0)221 478 97131
sandra.salm@uk-koeln.de
www.imvr.de

gefördert durch



**Gemeinsamer
Bundesausschuss
Innovationsausschuss**

Literatur

Hoffmann, F., Andersohn, F., Giersiepen, K., Scharnetzky, E., & Garbe, E. (2008). Validierung von Sekundärdaten: Grenzen und Möglichkeiten. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, 51, 1118-1126. doi: 10.1007/s00103-008-0646-y

March, S., Rauch, A., Thomas, D., Bender, S., & Swart, E. (2012). Datenschutzrechtliche Vorgehensweise bei der Verknüpfung von Primär- und Sekundärdaten in einer Kohortenstudie: die lidA-Studie. *Das Gesundheitswesen*, 74, e22-e129. doi: 10.1055/s-0031-1301276