

Aufbau einer krankenkassenübergreifenden GKV-Routinedatenbank – Lessons Learned

Nadine Scholten*, Peter Ihle**

* IMVR, Universität zu Köln, ** PMV forschungsgruppe, Uniklinik Köln

Hintergrund: CoRe-Net ist ein BMBF gefördertes Kompetenznetzwerk (FKZ: 01GY1606) mit dem Ziel die Forschung und Versorgungspraxis in Köln näher zusammen zu bringen. Hierzu werden Strukturen geschaffen, die den Austausch untereinander fördern und darüber neue, innovative Versorgungs(forschungs)projekte für den Raum Köln ermöglichen. CoRe-Net ist inhaltlich angelehnt an das value-based-healthcare Konzept nach Porter und somit auf die Umgestaltung von Versorgungsprozessen und -strukturen hin zu mehr Patientenzentrierung bei gleichzeitigem Kostenbewusstsein fokussiert. Im Rahmen dieses Projektes wird Köln zu einer Modellregion für die Versorgungsforschung. Ein Pfeiler dieser neu zu schaffenden Struktur ist die einzigartige Datenbank für Köln die aufgebaut wird. Diese beinhaltet u.a. krankenkassenübergreifende GKV Routinedaten, Sozialdaten und Primärdaten.

Ziel: Dieser Datensatz soll GKV-Daten 5 Jahre retrospektiv und über die gesamte Projektlaufzeit (geplant bis zu 15 Jahren) beinhalten. Hierzu liefern die kooperierenden Krankenkassen jährlich Daten für die in Köln gemeldeten Versicherten. Anhand dieser Datenbank, die krankenkassenübergreifend Daten, auch über die internen Löschrufen der einzelnen Krankenkassen hinaus, beinhaltet, wird es möglich sein, die Versorgung in Köln über längere Zeiträume zu analysieren und die langfristigen Auswirkungen von regionalen Interventionen zu evaluieren.

Methodisches Kernproblem: Um einen Datenbestand über die Löschrufen der einzelnen Kassen hinaus aufzubauen, ist es notwendig den in der Datenbank befindlichen alten Datenbestand jährlich um die aktuelle Datenlieferung der Kassen zu erweitern. Hierzu müssen die Daten von den Krankenkassen über die Zeit mit dem krankenkassenspezifischen identischen Pseudonym übermittelt werden. Die Übermittlung dieser Sozialdaten basiert auf dem §75 SGB X. Hierin wird eine Übermittlung ohne Einwilligung des Betroffenen ausgeschlossen, soweit es zumutbar ist, die Einwilligung des Betroffenen nach § 67b einzuholen. Die Auslegung dieses Paragraphen durch die Ethikkommission der Uniklinik Köln soll im Folgenden diskutiert werden.

Lösungsansätze: Der Aufbau der Routinedatenbank erfolgt auf Basis versichertenbezogener Daten, die im Lauf der Übermittlung von den Krankenkassen an CoRe-Net durch mehrere kaskadierende Bearbeitungsschritte für die CoRe-Net-Subprojekte anonymisiert werden. Hierfür werden bereits in den Krankenkassen die direkt versichertenidentifizierenden Angaben durch ein nach krankenkasseninternen Vorgaben erzeugtes Pseudonym ersetzt und an die CoRe-

Net- Vertrauensstelle, transportverschlüsselt übermittelt. Durch den Pseudonymisierungsdienst (CoRe-Net extern) werden die krankenkassenspezifischen Pseudonyme durch ein Einweghashverfahren (SHA-3) unter Verwendung eines über die Laufzeit des CoRe-Net-Projekts eindeutigen Salts zweitpseudonymisiert und an die CoRe-Net-Vertrauensstelle (intern), wieder kryptographisch verschlüsselt, übermittelt. Die Vertrauensstelle generiert aus dem Hashpseudonym einen CoRe-Net spezifischen Studienidentifikator (Dritt pseudonym), der über die Laufzeit von CoRe-Net identisch bleibt. Die Vertrauensstelle plausibilisiert die GKV-Routinedaten und verlinkt die jahresbezogenen Lieferungen über das Dritt pseudonym. Die Bereitstellung der GKV-Routinedaten für die einzelnen CoRe-Net-Projekte erfolgt projektspezifisch. Nach einem positiven Votum eines für die GKV-Routinedaten zuständigen CoRe-Net-Beirats, der unter anderem aus Vertretern der Daten liefernden Krankenkassen besteht und einstimmig entscheidet, sowie einem positiven Votum der Ethikkommission des Universitätsklinikums Köln, werden die für die Beantwortung der Fragestellung notwendigen Daten aus dem GKV-Pool selektiert. Hierbei sollen die Daten soweit aggregiert werden, dass Zellbesetzungen von kleiner als fünf nicht auftreten. Diese Prüf- und Aufbereitungsschritte können nicht einmalig und statisch durchgeführt werden sondern müssen jeweils projektbezogen erfolgen, da mit zunehmender Beobachtungsdauer auch die pro Versichertem gespeicherten Ereignisse sowie die Detailtiefe zunimmt. Aus diesem Grund sind diese Aufbereitungsschritte in der Vertrauensstelle des CoRe-Net verortet, in der die Sachkenntnis der GKV-Routinedaten vorhanden ist und die Daten jeweils projektspezifisch zugeschnitten und aggregiert werden können. Hier erfolgt auch die projektindividuelle Viert pseudonymisierung der Daten, so dass ausgeschlossen ist, dass Daten aus den einzelnen Projekten über das Viert pseudonym wieder verlinkt werden können.

Diskussion: Grundsätzlich diskutiert werden kann der über die einzelnen Stellen hinweg fließende Übergang von pseudonymen Daten zu anonymen Daten und den spezifischen Umgang mit diesen sensiblen Daten. Die Auslegung der Ethikkommission und der sich hieraus ergebende Lösungsansatz soll vorgestellt und diskutiert werden.

Schlussfolgerungen: Der Umgang mit GKV-Routinedaten, hinsichtlich des Datenschutzes ist ein Thema, welches im Rahmen der zunehmenden Verwendung von Routinedaten in der Versorgungsforschung an Relevanz gewinnt. Die Unterschiedlichen Auslegungen der Gesetzeslage verhindern aktuell ein einheitliches Vorgehen in Bezug auf die Wahrung des Datenschutzes.

Korrespondenzadresse:

Dr. Nadine Scholten
IMVR – Universität zu Köln
Eupener Strasse 129
50933 Köln
nadine.scholten@uk-koeln.de