ATG

Gesellschaft für
Versicherungswissenschaft
und -gestaltung e.V.

GVG

Management-Papier *„Pseudonymisierung / Anonymisierung“*

Aktionsforum Telematik im Gesundheitswesen

Köln, *Mai/2002*

13.06.2002

Kontakt:
Jürgen Dolle (Koordinierung), GVG
mailto: j.dolle@gvg-koeln.de

Autoren-Team:

Albert, Jürgen	VdAK /AEV
David, Dr. Dagmar M.	BÄK/ÄKNo
Junge, Medard	IKK-BV
Keil, Werner	ABDA/WuV
Rey, Heinz-Theo	KBV
Schoener, Claus	BKK-BV
Siebert, Irmgard	KZBV
Strobel, Wolfgang	AOK-BV

Inhaltsverzeichnis

1 Zusammenfassung (Management Summary)

2 Einleitung

Die Beteiligten im Gesundheitswesen verfügen auf Grund unterschiedlicher Rechtsnormen über Versorgungsdaten, die sie nach den für ihre Aufgaben entsprechenden Erfordernissen unter Berücksichtigung der geltenden datenschutzrechtlichen Bestimmungen verarbeiten. Dazu zählt auch Datentransport, der z. T. pseudonymisiert bzw. anonymisiert erfolgt.

In den von den ATG-Teams in der Vergangenheit erarbeiteten Managementpapieren zu den Themen Sicherheitsinfrastruktur, Elektronisches Rezept, Elektronischer Arztbrief und Europäische und internationale Perspektiven von Telematik im Gesundheitswesen wurden zum einen Aspekte der Pseudonymisierung / Anonymisierung bewusst ausgeklammert, zum anderen eine zeit- und sektorübergreifende Gesundheitsberichterstattung gefordert. Hinzu kommt dass auf europäischer Ebene eine umfassende Gesundheitsberichterstattung der Mitgliedsstaaten gefordert wird, um eine Vergleichbarkeit der Systeme und ihrer Versorgung (Public Health) zu ermöglichen. Auch aus diesem Grund hat die Bundesregierung Überlegungen zur Einführung eines Datentransparenzgesetzes angestellt.

Das ATG-Team Pseudonymisierung / Anonymisierung will sich mit den ordnungspolitischen Aspekten der Verfahren Pseudonymisierung und Anonymisierung unter den möglichen Szenarien zur Datenaggregation im Gesundheitswesen beschäftigen. Dabei soll ein gemeinsames Verständnis der Problemlage und der möglichen Lösungsansätze bei den Beteiligten herbeigeführt, sowie eine gemeinsame Vorgehensweise erarbeitet werden. Das Team will keine technischen Details der Verfahren beschreiben oder über den Einsatz spezieller Verfahren bei einzelnen Datensätzen entscheiden. Es will mögliche Strukturen aufzeigen und den dazu erforderlichen Handlungsbedarf beschreiben.

3 Definitionen und Abgrenzung des Themas

3.1 Zielsetzung/Abgrenzung des ATG-Teams

Wie bereits in der Einleitung angesprochen, sieht das Team "Pseudonymisierung/Anonymisierung" seinen Arbeitsauftrag nicht in der Beschreibung von Verfahren zur Erreichung von Datentransparenz.

Mit dem vorliegenden Management-Papier sollen Empfehlungen gegeben werden, wie durch technische und organisatorische Maßnahmen der Schutz der Patientendaten bei der Umsetzung von Datentransparenzverfahren gewahrt werden kann. Dabei sollen nicht dezidierte Verfahren zur Pseudonymisierung bzw. Anonymisierung beschrieben werden, sondern Ziel ist, Richtlinien für globale Verfahrensweisen bei der Pseudonymisierung/Anonymisierung von Daten aufzuzeigen (z. B. Zeitpunkt der Pseudonymisierung, zentrale/dezentrale Pseudonymisierung)

Das Team ist der Auffassung, dass ausgehend von differenzierten Anforderungen konzipiert werden muss, welche Daten zur Erfüllung der speziellen Zwecke notwendig sind, wo diese Daten vorliegen und auf welche Weise sie bereitgestellt werden können. Dabei ist die Datenhaltung zur Erfüllung der Anforderungen zu bewerten.

Das Team gibt mit diesem Papier außerdem Empfehlungen, wie Konzepte für Datentransparenz aus Sicht der Autoren ausgelegt sein sollten.

3.2 Definitionen

3.2.1 Pseudonymisierung

Mittels Pseudonymisierungsverfahren werden personenbezogene Daten, durch eine Regel derart ein-eindeutig verändert, dass sie sich weder auf eine bestimmte Person beziehen noch eine solche erkennen lassen. Durch eine inverse Rechenvorschrift oder eine Zuordnungstabelle ist eine Re-Pseudonymisierung möglich.

Pseudonymisierung ist z. B. erforderlich, wenn ein individueller Verlauf beobachtet werden soll, oder die Zusammenführung unterschiedlicher Daten zu einer Person erfolgen soll, ohne dass die Personenidentität bekannt sein muss.

3.2.2 Anonymisierung

Mit Hilfe von Anonymisierungsverfahren werden personenbezogene Daten, derart verändert, dass sie sich weder auf eine bestimmte Person beziehen noch eine solche erkennen lassen (z.B. durch Weglassen oder Änderung von personenidentifizierenden Merkmalen). Ein Wiederherstellen des ursprünglichen Personenbezuges ist nicht möglich.

Anonymisierung ist z. B. erforderlich, wenn die Personenidentitäten oder die Herkunft der Daten für die zu ermittelnden Ergebnisse nicht relevant sind oder sein sollen.

3.2.3 Datenaggregation

Unter Datenaggregation versteht man Verfahren, bei der durch Bildung von (entsprechend großen) Gruppen der Personenbezug einer Datenmenge entfernt wird. Ein Wiederherstellen des ursprünglichen Personenbezuges ist grundsätzlich nicht möglich.

4 Ist-Situation

4.1 Beschreibung realer Verfahren

4.1.1 Erzeugung von Pseudonymen am Beispiel des Deutschen Arzneimittelprüfinstituts (DAPI)

Für die apothekeneigene Statistik des DAPI sind individuelle Merkmale und Verhaltensweisen im Zusammenhang mit Medikationen von großem Interesse.

Hierzu zählen unter anderem die Informationen, ob spezielle Medikamente durch andere ersetzt worden sind und welche Co-Medikationen bei bestimmten Präparaten stattgefunden haben.

Die benötigten Arzneimitteldaten von GKV-Patienten werden von den sechs Rechenzentren, die sich zur Gesellschaft für zentrales Datenmanagement und Statistik im Gesundheitswesen (GDSG) zusammengeschlossen haben, der Werbe- und Vertriebsgesellschaft Deutscher Apotheker mbH (WuV) verschlüsselt zur Verfügung gestellt.

Zuerst werden die Verordnungsdaten mit einem nur den Mitgliedern des GDSG bekannten Schlüssel, der für alle Rechenzentren gleich ist, mittels einer Einweg-Funktion pseudonymisiert. Anschließend werden diese pseudonymisierten Daten von den Rechenzentren direkt an die WuV gesendet.

Im zweiten Schritt verschlüsselt die WuV die zuvor erzeugten Pseudonyme mit einem nur ihr bekannten Schlüssel, bevor die Daten zur weiteren Verarbeitung an das NARZ Bremen versendet werden, das anschließend die Berechnung der Statistik durchführt.

Aufgrund der Doppelrolle des NARZ Bremen als pseudonymisierende Stelle und als erstellende Institution der Statistik wurde dieses zweistufige Verfahren gewählt. Es ist mit dem Datenschutz abgestimmt.

Bedingt durch die zweistufige Prozedur ist das NARZ Bremen nicht in der Lage, auch nicht durch einen Vergleich, mit den selbst erzeugten Pseudonymen, auf die Identität der Person zu schließen, da die vorliegenden Pseudonyme mit einem dem NARZ Bremen nicht bekannten Schlüssel verschlüsselt wurden.

Auf der anderen Seite ist es der Werbe- und Vertriebsgesellschaft Deutscher Apotheker mbH ebenfalls nicht möglich, die Originaldaten zu ermitteln, da für die Pseudonymisierung eine nicht umkehrbare Einweg-Funktion verwendet wurde.

4.1.2 GKV Arzneimittel Schnellinformation (GAmSi)

Die Spitzenverbände der gesetzlichen Krankenversicherung haben sich darauf verständigt, die Verordnungsdaten zeitnah für gemeinsame Zwecke zusammenzuführen.

Dazu stellen die Spitzenverbände die Bereitstellung der Arzneimitteldaten ihrer Kassenart sicher. Die einzelnen Datenstellen bereiten die Daten mit einem einheitlichen Transformationsprogramm auf. Dabei werden die personenbezogenen Daten anonymisiert. Die so entstandenen Datenpakete werden verschlüsselt (Grundlage ist

Security-Schnittstelle für das Gesundheitswesen) und mittels DFÜ an die Relaisstelle (Informationstechnische Servicestelle der GKV – ITSG) übermittelt.

Die Relaisstelle nimmt die Anonymisierung der Absender vor und leitet die Datenpakete an die verarbeitende Stelle weiter. Die Relaisstelle betreibt keine Datenhaltung.

4.2 Rahmenbedingungen

4.2.1 Datentransparenz

Durch die Gesundheitsreform 1993 wurde der Datenaustausch mit Leistungserbringern (§§ 294 ff SGB V) gesetzlich normiert. Ziel war es unter anderem für mehr Transparenz im Vertrags- und Leistungsgeschehen im Gesundheitswesen zu sorgen. Die Leistungs- und Abrechnungsdaten liegen derzeit personenbezogen und verteilt bei verschiedenen Stellen des Gesundheitswesens vor.

Übergreifende Auswertungen, fundierte Gesamtbetrachtungen zum Versorgungsgeschehen (beispielsweise bei bestimmten Krankheitsbildern, Disease-Management-Programme), die Entwicklung des morbiditätsorientierten Risikostrukturausgleichs oder die Abbildung der medizinischen Versorgung in Städten und Regionen bzw. durch einzelne Leistungserbringer sind derzeit nur unter größtem Aufwand und nicht systematisch möglich.

Die Verfolgung dieser Ziele erfordert aus datenschutzrechtlichen Gründen ein gesichertes Pseudonymisierungs- und Anonymisierungsverfahren, das die Identifizierung eines Versicherten unmöglich macht.

4.2.2 Datenschutz

Im Bereich öffentlicher Stellen ist die Datenerhebung, -verarbeitung und –nutzung nur zulässig, wenn eine Rechtsnorm dies ausdrücklich erlaubt/vorsieht, oder die schriftliche Einwilligung des Betroffenen vorliegt (§§ 4 u. 4a BDSG). Hinzu kommt das Prinzip der Datenvermeidung und Datensparsamkeit (§ 3a BDSG). Besondere Schutzwürdigkeit wird u. a. personenbezogenen Daten mit Informationen zur Gesundheit zugeschrieben (§ 3 Abs. 9 BDSG).

Als zentraler Paragraf für den Umgang mit personenbezogenen Daten im Gesundheitswesen unter dem besonderen Aspekt der Aggregation und Auswertung wird § 3a BDSG Datenvermeidung und Datensparsamkeit angesehen.

Gestaltung und Auswahl von Datenverarbeitungssystemen haben sich danach an dem Ziel auszurichten, keine oder so wenig personenbezogene Daten wie möglich zu erheben, zu verarbeiten oder zu nutzen. Insbesondere ist von den Möglichkeiten der Anonymisierung und Pseudonymisierung Gebrauch zu machen, soweit dies möglich ist und der Aufwand in einem angemessenen Verhältnis zu dem angestrebten Schutzzweck steht.

4.3 Bewertung

Aus den gesetzlichen Vorgaben des Bundesdatenschutzgesetzes (Datensparsamkeit) und dem Wunsch Transparenz zu schaffen, ergibt sich die Notwendigkeit geeignete Verfahren der Anonymisierung/Pseudonymisierung bzw. Aggregation zu definieren und einzusetzen. Welche Verfahren das sein können, leitet sich aus den einzelnen Analyseanforderungen ab, wie die unter 4.1.1 und 4.1.2 beschriebenen Projekte beispielhaft zeigen.

Aus diesem Grund werden nachfolgend modellhafte Lösungen der Pseudonymisierung/Anonymisierung dargestellt.

(II. Meilenstein)

5 Lösungsansätze und –vorschläge

5.1 Grundsätzliche Überlegungen

5.2 Pseudonymisierung/Anonymisierung bei zentralen Datenhaltungsmodellen

5.2.1 Einheitliche oder zweckgebundene Pseudonymisierung/Anonymisierung

5.2.2 Zeitpunkt der Pseudonymisierung/Anonymisierung

5.3 Pseudonymisierung/Anonymisierung bei dezentralen Datenhaltungsmodellen

6 Empfehlungen, Maßnahmenvorschläge

Die das ATG tragenden Organisationen in alphabetischer Reihenfolge:

- **ABDA - Bundesvereinigung Deutscher Apothekerverbände**
- **Bundesärztekammer**
- **Bundesknappschaft**
- **Bundesverband der Allgemeinen Ortskrankenkassen**
- **Bundesverband der Betriebskrankenkassen**
- **Bundesverband der Innungskrankenkassen**
- **Bundesverband der landwirtschaftlichen Berufsgenossenschaften**
- **Bundesverband der landwirtschaftlichen Krankenkassen**
- **Bundesversicherungsanstalt für Angestellte**
- **Bundeszahnärztekammer**
- **Deutsche Krankenhausgesellschaft**
- **Hauptverband der gewerblichen Berufsgenossenschaften e.V.**
- **Kassenärztliche Bundesvereinigung**
- **Kassenzahnärztliche Bundesvereinigung**
- **Verband der Angestelltenkrankenkassen**
- **Verband der privaten Krankenversicherung e.V.**
- **Zentralverband der Krankengymnasten und Physiotherapeuten e.V.**