Datenschutzgerechte Erfassung von klinischen Forschungsdaten im Behandlungszusammenhang

Müller TH
Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie, Universität München, Deutschland
mueller@ibe.med.uni-muenchen.de

Einleitung und Fragestellung Die kontinuierliche wissenschaftliche Beobachtung von Patientengruppen ist bei Erkrankungen mit langen Verläufen von besonderem Interesse. Gerade diese Patienten werden oft von mehreren Ärzten parallel behandelt. Deren Kommunikation untereinander im Rahmen des Behandlungsprozesses kann durch Online-Medien wie das World Wide Web erheblich verbessert werden. Die Vernetzung von behandelnden Ärzten und Forschergruppen dient also zugleich der individuellen Behandlung und der Forschung. Jedoch muss der Persönlichkeitsschutz der Patienten angemessen gewahrt werden. Hierfür sind generische konzeptuelle Lösungen erarbeitet worden [1,2,3]. Im Folgenden wird eine Realisierung vorgestellt, die sich durch eine enge Verbindung von Datenkommunikation zu Behandlungszwecken und Datenerhebung zu Forschungszwecken auszeichnet.

Material und Methoden .Beim Design der Systemarchitektur muss die IT Infrastruktur, die in medizinischen Einrichtungen in der Regel zur Verfügung steht, und die dort üblichen Beschränkungen des Internetzugangs berücksichtigt werden. Daraus ergeben sich folgende technische Anforderungen:

- Die Kommunikation erfolgt ausschließlich über das Standardprotokoll http des World Wide Web.
- Die Sicherheit der Übertragung wird durch Verschlüsselung (SSL) gewährleistet.
- Die Installation oder das Herunterladen von ausführbaren Dateien auf die Clientsysteme ist ausgeschlossen. Ein Webbrowser mit JavaScript wird vorausgesetzt.

Die Realisierung des Prototyps in der Programmiersprache Perl beruht auf bewährten Open Source Softwarekomponenten: Linux Betriebssystem, Apache Webserver, PostgreSQL Datenbank.

Ergebnisse Ein zentraler Grundsatz des Datenschutzes ist die Datensparsamkeit. Diese bezieht sich nicht nur auf die Auswahl der Daten, die gespeichert werden, sondern auch auf die Art, wie diese Daten strukturiert und verteilt sind. Umfassende Datensammlungen zu einzelnen Personen sollen vermieden werden, da hierdurch der Schutz der Persönlichkeit in besonderem Maße gefährdet wird. Für die Erforschung von Krankheiten mit langen Verläufen sind jedoch oft umfassende und breite Datensammlungen wichtig und notwendig.

Die generischen Datenschutzkonzepte tragen diesem Zielkonflikt dadurch Rechnung, dass sie die medizinischen Daten von den Daten, die den Patienten identifizieren, trennen. Beide Datenbestände sollen sowohl physikalisch als auch organisatorisch getrennt gehalten werden. Die Trennung kann zum Beispiel durch einen bestimmten Ablauf beim Übergang der medizinischen Daten aus dem Behandlungszusammenhang in den Forschungsbestand erfolgen. Dabei werden die identifizierenden Daten durch ein Pseudonym ersetzt. Die Rückidentifizierung der Patienten durch Auflösung der Pseudonyme kann nur durch einen Treuhänder, der selbst keinen Zugang zu den medizinischen Daten hat, erfolgen. Wenn die zentral gespeicherten medizinischen Daten noch im Behandlungszusammenhang durch mehrere kooperierende Behandler genutzt werden sollen, ist ein offline Pseudonymisierungsverfahren nicht praktikabel. An seine Stelle können zwei getrennte Datenbanken treten, die im Folgenden als „Patientenliste“ und „Medizinische Datenbank“ bezeichnet werden.

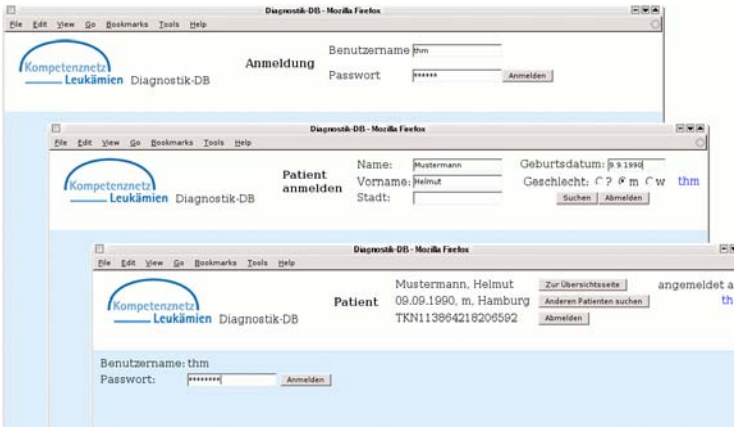

Abb. 1 Anmeldesequenz

Die Patientenliste enthält jene Daten, mit denen Patienten im Behandlungskontext üblicherweise identifiziert werden. Die Medizinische Datenbank enthält ausschließlich medizinisch relevante Daten. So ist z.B. in der Patientenliste das Geburtsdatum zur Identifizierung gespeichert, in der Medizinischen Datenbank jedoch nur das Geburtsjahr, da für die medizinische Betrachtung die Angabe des Alters in Jahren fast immer ausreicht.

Beide Datenstrukturen der beiden Datenbanken sind über einen gemeinsamen Schlüssel verbunden, der im Gegensatz zum o.g. Pseudonym geheim ist, d.h. nicht an Behandler oder Forscher weitergegeben wird. Der Zugriff auf die medizinischen Daten im Behandlungszusammenhang erfolgt über temporäre Zugriffsschlüssel, die zwischen Patientenliste und Medizinischer Datenbank automatisch vereinbart, und an den Zugriffsberechtigten weitergegeben werden. Der temporäre Schlüssel verliert seine Gültigkeit nach einer festgelegten Zeitspanne oder zum Ende der aktuellen Sitzung.

Die nebenstehende Abb. 1 zeigt den Anmeldevorgang an beiden Datenbanken im Behandlungszusammenhang aus Sicht des Clientsystems. Zur Veranschaulichung der tatsächlichen Trennung ist das Browserfenster in einen weißen Bereich für die Patientenliste und in einen hellblauen Bereich für die medizinische Datenbank unterteilt.

Der erste Schritt besteht in der Anmeldung an der Patientenliste mit Benutzername und Passwort. Danach erfolgt die Auswahl des Patienten. Hierfür stehen ausschließlich die in der Patientenliste abgelegten Informationen zur Verfügung. Eine Auswahl nachmedizinischen Daten ist an dieser Stelle nicht notwendig, da der Online-Zugang nicht für die Analyse zu Forschungszwecken verwendet werden soll. Ist der Patient identifiziert, wird ein temporärer Zugangsschlüssel (im Beispiel: TKN113864218206592) automatisch vereinbart und der Benutzer kann sich an der Medizinischen Datenbank anmelden. Dabei wird der gleiche Benutzername, jedoch mit einem anderen Passwort verwendet. Die doppelte Anmeldung ist erforderlich, weil der jeweilige Betreiber der Patientenliste und der Medizinischen Datenbank unabhängig voneinander für die Zugangskontrolle verantwortlich ist. Nach der Anmeldung erscheinen die medizinischen Daten des Patienten im Anzeigebereich der Medizinischen Datenbank (nicht in der Abb. dargestellt). Der temporäre Zugangsschlüssel wird zur visuellen Kontrolle ebenfalls mitgeführt. Auf diese Weise können Daten zu einzelnen Patienten zwischen berechtigten Behandlern ausgetauscht werden. Die Analyse der Daten über Patientengruppen zu Forschungszwecken erfolgt offline und ohne die identifizierenden Daten aus der Patientenliste.

Diskussion Das hier vorgestellte System realisiert die Trennung von identifizierenden und medizinischen Daten, so wie es das generische Datenschutzkonzept der Telematikplattform für Medizinische Forschungsnetze (TMF; www.tmf-ev.de) vorsieht, in einem Online-System das gleichzeitig die Kommunikation im Behandlungsprozess und die Erhebung von Forschungsdaten unterstützt [4]. Auch der Aufbau von Biomaterialsammlungen, die regelmäßig eines besonderen Schutzes bedürfen [5], wird berücksichtigt. Insgesamt werden durch das hier vorgestellte Verfahren einrichtungsübergreifende Erhebungen im Rahmen des normalen Behandlungsprozesses unter Wahrung des Datenschutzes ermöglicht. Das System wird nunmehr in zwei Forschungsverbünden (Kompetenznetz Hepatitis und Kompetenznetz „Akute und chronische Leukämien“) unter unterschiedlichen organisatorischen Bedingungen und mit verschiedenen medizinisch-wissenschaftlichen Zielsetzungen erprobt, um insbesondere Erkenntnisse hinsichtlich der Praktikabilität des Verfahrens zu gewinnen.

Danksagung Gefördert durch das Bundesministerium für Bildung und Forschung (Kompetenznetz Hepatitis und Kompetenznetz „Akute und chronische Leukämien“, Förderkennzeichen 01KI0403 bzw. 01G10480)

Literatur

[1] Reng CM, Debold P, Adelhard K, Pommerening K. Akzeptiertes Datenschutzkonzept Dtsch. Ärztebl. 2003; 100:A2134-2137.

[2] Semler SC, Lux A, Dolle W, Reng M, Pommerening K. Pseudonymisierung für Forschungsdatenbanken und Register. In Jäckel (Hrsg.) Telemedizinführer Deutschland, Ober-Mörlen, 2004: 198-203

[3] Pommerening K, Reng M: Secondary use of the EHR via pseudonymisation. Stud Health Technol Inform. 2004;103:441-6.

[4] Müller TH: Verbindung von Telemedizin und Forschung - Datenschutzgerechte Lösungen. in: Jäckel (ed): Telemedizinführer Deutschland Ober-Mörlen, 2005.

[5] Wellbrock R: Biobanken für die Forschung. Datenschutz und Datensicherheit. 2003 Sep;38(3):543-544.