German Biobank Alliance (GBA)

Datenschutzkonzept

Projekt	German Biobank Alliance (GBA) zusammen mit German Biobank Node (GBN)
Autoren	Dr. Martin Lablans, Dr. Esther Schmidt, Petra Duhm-Harbeck, Prof. Dr. Hans-Ulrich Prokosch, Prof. Dr. Michael Hummel
Träger	Charité – Universitätsmedizin Berlin (German Biobank Node)
Version	25. Januar 2018

Inhalt

1. Art und Ziele des Projekts

Vor dem Hintergrund des Wissens um das komplexe Zusammenwirken individueller genetischer Disposition, des Lebensstils, der Umweltfaktoren und molekularer Alterationen bei der Entstehung und dem Verlauf von Erkrankungen benötigt die heutige biomedizinische Forschung neben den verschiedensten umfangreichen molekularen Daten die Beobachtung von therapeutischem Ansprechen, Krankheitsverläufen, individuellen Lebensgewohnheiten, und Umweltbedingungen über lange Zeiträume hinweg. Eine Schlüsselrolle nehmen dabei Biobanken ein, die das für diese Forschung benötigte Biomaterial der Spender[1] systematisch sammeln und mit den assoziierten klinischen Daten verknüpfen. Nur durch die Forschung in standortübergreifenden Verbünden mit ausreichenden Fallzahlen können Erkrankungen als Folge individueller genetischer Dispositionen oder immer kleinere werdende molekulare Krankheitsuntergruppen ausreichend erforscht werden.

Der German Biobank Node (GBN) hat es sich in Zusammenarbeit mit der German Biobank Alliance (GBA) zur Aufgabe gemacht, für deutsche Biobanken modulare und interoperable IT-Komponenten zu konzipieren, zu entwickeln und zu implementieren, die ein standortübergreifendes effizientes Biobanking – über die Möglichkeiten isolierter kommerzieller Biobank-Management-Systeme hinaus – möglich machen. Eine standortübergreifende Vernetzung der Biobanken ermöglicht ein schnelles Auffinden und Zusammenstellen von Biomaterial und Daten für zukünftige nationale und translationale Forschungsprojekte. Diese Vernetzung erlaubt auch eine Anbindung der deutschen Biobanken an das europäische Biobanken-Netzwerk BBMRI-ERIC[2]. Im Rahmen einer BMBF-Förderung haben sich zunächst elf deutsche Biobankstandorte zur GBA unter der Koordination von GBN zusammengeschlossen. Die zu entwickelnden IT-Komponenten und die geplante IT-Vernetzungsstruktur soll aber schrittweise auch weiteren deutschen Biobanken zur Verfügung gestellt werden, die sich mit GBA assoziieren.

Das GBA IT Core Team (ITC) entwickelt und implementiert hierfür eine IT-Infrastruktur, die es Forschern (1) zunächst ermöglicht, zu definierten Erkrankungen verfügbares Biomaterial und zugehörige klinische Daten standortübergreifend zu ermitteln, um dann (2) in Abstimmung mit den jeweiligen Biobanken diese der Forschung zugänglich zu machen. Dazu werden an den beteiligten Standorten Daten aus der klinischen Dokumentation oder aber im Kontext von Forschungsprojekten erhobene Daten mit dort vorhandenen Biomaterialproben verknüpft und über ein **lokales** Data Warehouse als Bestandteil eines "Brückenkopfes" für verteilte Suchanfragen zur Verfügung gestellt.

Um diese **lokalen** Daten einer Abfrage zugänglich zu machen, stellt das ITC eine zentrale Suchschnittstelle bereit, die es Forschern ermöglicht

- abzuschätzen, ob für ein Forschungsvorhaben innerhalb der an GBA angeschlossenen Einrichtungen Spenderproben mit den gesuchten Eigenschaften und Daten in dem benötigten Umfang **potenziell** vorhanden sind, um gegebenenfalls anschließend
- Anfragen zur Nutzung von medizinischen Daten und Biomaterialproben dieser Spender für Forschungsvorhaben zu stellen.

Dies ermöglicht, dass auch klinische Bestandsdaten, die vor der Einrichtung der GBA im Behandlungs- oder Forschungskontext erhoben wurden, für neue Kooperationsprojekte in der medizinischen Forschung nutzbar gemacht werden können. So kann durch die Verwendung bereits erhobener Datenbestände die sonst nötige Wartezeit zwi-

[1] Der Einfachheit halber wird in diesem Datenschutzkonzept die maskuline Form für beide Geschlechter verwendet. Auf eine gendergerechte Formulierung wurde aus Gründen der Lesbarkeit verzichtet. Die Bezeichnung "Spender" bezeichnet Proben- und Datenspender und trifft sowohl für Personen im Behandlungskontext als auch Probanden im Kontext von Studien zu.

[2] http://www.bbmri-eric.eu/

schen Beginn der Datenerhebung und dem Erreichen eines für Forschungszwecke ausreichend langen Beobachtungszeitraums erheblich reduziert werden. Die Nutzung von Bestandsdaten setzt eine entsprechende Patienteneinwilligung voraus. Ziel dieses Datenschutzkonzeptes ist es, darzustellen, wie die Anforderungen hinsichtlich Datenschutz, Datensicherheit, ethisch-sozialer Belange und den Vorgaben der neuen EU Datenschutz-Grundverordnung unter Berücksichtigung der Nutzbarmachung der Daten für die medizinische Forschung in Einklang mit der Wahrung der Persönlichkeitsrechte der Spender erfüllt werden.

Dieses Datenschutzkonzept beschreibt ausschließlich die Prozesse zur übergreifenden Vernetzung der teilnehmenden Biomaterialbanken im Rahmen der German Biobank Alliance. Lokale Prozesse, die in den Biobanken bereits etabliert sind, werden dadurch nicht berührt und sind nicht Bestandteil dieses Datenschutzkonzepts.

2. Organisatorische Struktur, Kooperationspartner, Verantwortlichkeiten

Die Gesamtverantwortung als Träger der GBA wird der Charité in Berlin obliegen. Hierzu werden derzeit Verträge zwischen dem Deutschen Krebsforschungszentrum (DKFZ), Heidelberg, und der Charité, Berlin, erarbeitet.

Organisatorische Rahmenbedingungen

Die datenverarbeitenden Personen und Institutionen sowie Datenempfänger in der GBA verteilen sich auf die Betreiber der zentralen Komponenten sowie die an der GBA teilnehmenden Standorte.

Betrieb der Komponenten

Der Betrieb der Brückenköpfe erfolgt durch die im GBA vertretenen Partnerstandorte in deren eigener Verantwortung:

- Aachen: Zentralisierte Biomaterialbank der RWTH Aachen University (RWTH cBMB)
- Frankfurt: Interdisziplinäre Biomaterial und Datenbank Frankfurt (iBDF) am Universitätsklinikum Frankfurt/Main
- Göttingen: Zentrale Serviceeinrichtung UMG Biobank (Universitätsmedizin Göttingen)
- Greifswald: Integrated Research Biobank (IRB) der Universitätsmedizin Greifswald (UMG)
- Hannover: Hannover Unified Biobank (HUB) der Medizinischen Hochschule Hannover (MHH)
- Heidelberg: BioMaterialBank (BMBH) (unter der Schirmherrschaft der Medizinischen Fakultät und dem Nationalen Zentrum für Tumorerkrankungen (NCT) Heidelberg)
- Jena: Integrierte Biobank Jena (IBBJ)
- Leipzig: LMB (Partner: Leipziger Forschungszentrum für Zivilisationserkrankungen (LIFE), das Institut für Anatomie und das University Cancer Centre Leipzig (UCCL))
- Lübeck: Interdisziplinäres Centrum für Biobanking-Lübeck (ICB-L) der Universität zu Lübeck
- München: Joint Biobank Munich
    - Helmholtz-Zentrum München
    - Ludwig-Maximilians-Universität München (LMU)
    - Technische Universität München (TUM)
- Würzburg: Interdisciplinary Bank of Biomaterials and Data (idbw) des Universitätsklinikums und der Medizinischen Fakultät

Der Betrieb der Server für zentrale Komponenten (Produktivinstanz des Metadata Repository, Authentifizierungsdienst, Suchbroker für die dezentrale Suche und Wartungsdienste für Brückenköpfe) erfolgt durch die Arbeitsgruppe Verbundforschung, Abteilung Medizinische Informatik in der Translationalen Onkologie, Deutsches Krebsforschungszentrum Heidelberg. Administrativen Zugriff auf die Systeme kann den Mitgliedern des GBA ITCs an den o.g. Einrichtungen in Göttingen, Hannover, Lübeck, Würzburg sowie am Lehrstuhl für Medizinische Informatik der

Friedrich-Alexander-Universität Erlangen-Nürnberg und der o.g. Arbeitsgruppe Verbundforschung gewährt werden.

Träger

Träger dieses Vorhabens ist die Charité – Universitätsmedizin Berlin (German Biobank Node, GBN).

3. IT- und Netzarchitektur, Daten, Prozesse und Kommunikationswege

3.1 IT- und Netzarchitektur

An jedem Standort wird ein sogenannter GBA-Brückenkopf etabliert. Er besteht aus Hard- und Software-Komponenten, die unter der Hoheit jedes Standorts installiert und betrieben werden und dient dazu, die Daten des jeweiligen Standorts in ein GBA-kompatibles Format zu überführen und für die anderen Komponenten nutzbar zu machen. Seine Aufgaben sind:

- *Datenharmonisierung*: Daten werden im Brückenkopf so harmonisiert (und in dieser Form dupliziert gespeichert), dass sie von den übrigen Komponenten der GBA verstanden werden.
- *Sichtbarmachung für die GBA*: z.B. Ausgabe aggregierter Ergebnisse für anfragende Forscher im Rahmen der dezentralen Suche.
- *Einhaltung von Datenhoheit*: Der Brückenkopf erlaubt dem Standort eine Teilnahme an der GBA, auch ohne spenderbezogene Daten „auf Verdacht" an eine externe Stelle hochladen zu müssen, was Datenschutz und Datenhoheit fördert.

Der Brückenkopf besteht aus den folgenden, lokal in der Einrichtung der Biobank installierten Softwarekomponenten:

- *Lokales Data Warehouse*: Stellt die in den lokalen Primärsystemen prinzipiell vorliegenden, aber unterschiedlich strukturierten Datenbestände für eine Nutzung in der GBA bereit.
- *Teiler*: Leistet eine kontrollierte Freigabe aggregierter Informationen aus dem lokalen Data Warehouse zur Nutzung in Projektanfragen.
- *Lokales Identitätsmanagement*: Stellt für die Pseudonymisierung von Spendern die in 3.3 beschriebene Funktionalität bereit.

Die im Brückenkopf gespeicherten medizinischen Daten (MDAT) können Daten zu Biomaterialproben selbst und klinische, Probenspender charakterisierende Daten (vgl. GBA-Datensatz in Anhang 2) umfassen.

Diese Komponenten stehen unter Kontrolle des jeweiligen Standortes, das heißt, die in diesen Komponenten gespeicherten Daten stehen weiter unter der Hoheit der Institution, in der sie erhoben wurden. Gegebenenfalls ist der Zugriff auf die behandelnde Einheit, z.B. die Fachabteilung, einzuschränken (vgl. auch Abschnitt 5).

3.2 Daten

An den teilnehmenden Partnerstandorten wurden und werden Daten von Spendern, die dort behandelt/betreut werden, erhoben und verarbeitet. Sie werden zum großen Teil aus vorhandenen Datenverarbeitungssystemen (z.B. Klinische Arbeitsplatzsystemen, Labordateninformationssystemen oder auch Tumordokumentationssystemen) entnommen, um über die Komponenten der GBA für die Beantwortung von Forschungsanfragen in aggregierter Form zur Verfügung gestellt werden zu können.

Grundsätzlich teilen sich die erhobenen datenschutzrelevanten Daten in medizinische und identifizierende Daten auf, die im Folgenden in Anlehnung an den "TMF- Leitfaden zum Datenschutz in medizinischen Forschungsprojekten"[3] als MDAT und IDAT bezeichnet werden. Die erhobenen MDAT umfassen

- Daten zu Biomaterialproben, wie z.B. die Probenart oder auch Informationen dazu, ob bereits in anderen Projekten bestimmte Analysen mit diesem Probenmaterial durchgeführt wurde,
- Klinische Daten, wie zum Beispiel codierte Diagnosen und Tumor-Klassifikationen im Fall von Krebserkrankungen, Labordaten, Informationen zur Therapie und Therapieerfolg, etc.
- Organisatorische Informationen zum Probenspender (z.B. ob er an einer Studie teilnimmt)
- Hinreichend vergröberte demographische Daten (z.B. Geschlecht und Geburtsjahr).

Bei der Bereitstellung von Daten im MDAT-Datensatz gilt immer das Prinzip der Datensparsamkeit, so dass lediglich Daten bereitgestellt werden, die für den Zweck einer Biomaterialanfrage auch erforderlich sind. Die IDAT enthalten demographische Daten, die eine eindeutige Identifikation des Spenders erlauben. Genauere Informationen zur Verarbeitung dieser Daten finden sich in Abschnitt 3.3.

3.3 Prozesse und Kommunikationswege

Identitätsmanagement

Pseudonymisierung dient dazu, Datensätze aus verschiedenen Quellen (hier: Daten zu Bioproben und klinische Daten) zu verknüpfen und gleichzeitig ein hohes Datenschutzniveau aufrechtzuerhalten, damit der Spender vor Rückidentifizierung geschützt wird. An die Stelle seiner identifizierenden Daten (IDAT) treten Pseudonyme. Der Brückenkopf stellt ein *lokales Identitätsmanagement* für die Pseudonymisierung von Spendern bereit. Alternativ können an den Standorten eigene Pseudonymisierungstools mit gleicher Funktionalität zum Einsatz kommen. Es erfolgt auf jeden Fall an jedem einzelnen Standort eine Pseudonymisierung in bzw. vor dessen Brückenkopf, sodass in das lokale Data Warehouse im Brückenkopf nur pseudonymisierte Daten übertragen werden. Das Pseudonymisierungsverfahren unterliegt der Kontrolle des jeweiligen Standorts und damit den lokal geltenden Datenschutzrichtlinien.

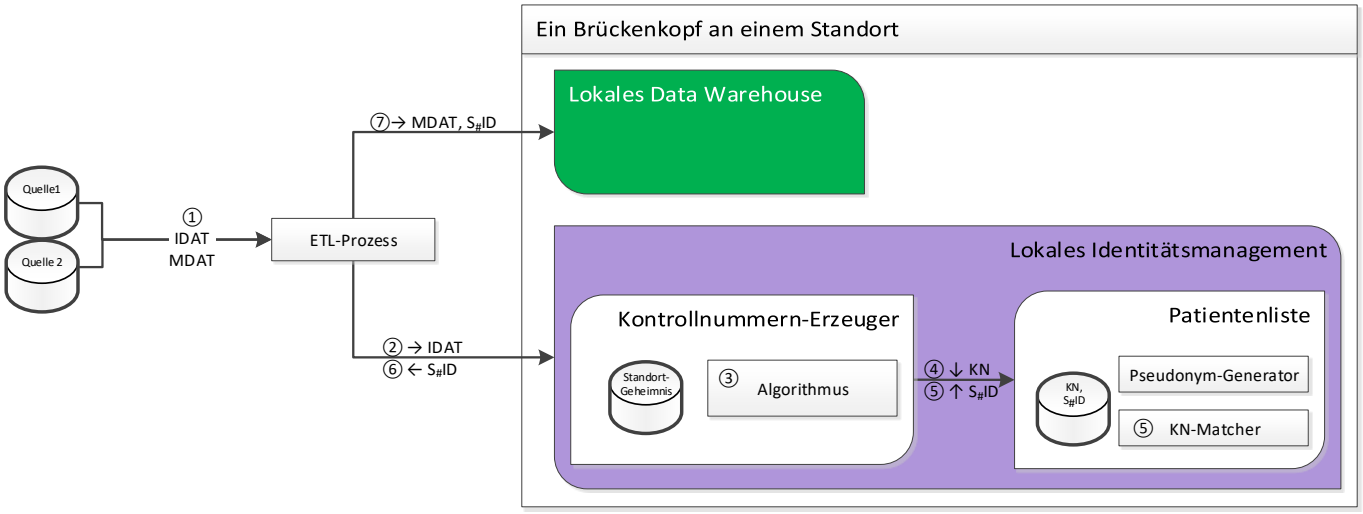

Abbildung 1 - Import von Daten in den Brückenkopf sowie Erzeugung lokaler Pseudonyme.

[3] Technologie- und Methodenplattform für die vernetzte medizinische Forschung e.V., "Leitfaden zum Datenschutz in medizinischen Forschungsprojekten" (http://www.tmf-ev.de/Publikationen/www.tmf-ev.de/datenschutz-leitfaden)

Import in Brückenkopf und Pseudonymisierung im lokalen Identitätsmanagement

Abbildung 1 zeigt, wie Daten aus den verschiedenen Quellsystemen eines Standorts in den lokalen Brückenkopf importiert werden und wie die Pseudonymisierung im Rahmen des ETL-Prozesses genutzt werden kann, um einen primären Spenderidentifikator erster Stufe zu erhalten. Es wird ein lokales Pseudonym ($S_{\#}$ID) erzeugt, d.h. eines, das nur innerhalb des Standorts zugeordnet werden kann. Man kann es etwa dazu verwenden, um Daten zur Bioprobe mit zugehörigen klinischen Daten zu verknüpfen. Es ist aber nicht möglich, einen Spender über Standortgrenzen hinweg zuzuordnen.

1. Identifizierende, medizinische und Probendaten werden aus mehreren Quellsystemen durch einen ETL-Prozess extrahiert.
2. Das lokale Identitätsmanagement erhält die Klartext-IDAT eines Spenders. Diese werden an den Kontrollnummern-Erzeuger übertragen.
3. Der Kontrollnummern-Erzeuger errechnet aus den Klartext-IDAT und seinem Geheimnis Kontrollnummern (KN) und verwirft die IDAT.
4. Der KN-Erzeuger übermittelt die Kontrollnummern an die lokale Patientenliste.
5. Die Patientenliste gleicht die erhaltenen Kontrollnummern mit den bestehenden ab (KN-Matcher). Im Falle eines Treffers wird eine bestehende ID ($S_{\#}$ID) zurückgegeben; falls kein passender Datensatz gefunden wird, wird mithilfe des Pseudonym-Generators eine neue ID erzeugt und zusammen mit den Kontrollnummern in der Patientenlisten-Datenbank gespeichert.
6. Der ETL-Prozess ordnet den Identifikator dem Datensatz zu.
7. Daten werden im lokalen Data Warehouse abgelegt.

Falls bereits ein für alle Quellen gemeinsames Pseudonym mit analogem Verfahren außerhalb des Brückenkopfes erzeugt wurde, das den Anforderungen an die Speicherung im lokalen Data Warehouse entspricht, können die Schritte 2–6 entfallen. Dann werden keine IDAT verarbeitet, sondern das extern erzeugte Pseudonym ($S_{\#}$ID) direkt mit den MDAT in das lokale Data Warehouse abgelegt.

Metadata Repository

Das Metadata Repository (MDR) speichert die Bedeutung (Semantik) sämtlicher in der GBA verwendeten (Nutz-)Datenelemente. Es bietet ein kontrolliertes Vokabular (Syntax) und kann maschinenlesbare, strukturierte Aussagen über Datenelemente machen, bspw. konzeptuelle Domänen oder Wertebereiche. Da das MDR keine personenbezogenen Daten verarbeitet, wird innerhalb dieses Datenschutzkonzepts nicht weiter darauf eingegangen.

Dezentrale Suche

Die dezentrale Suche dient dem Auffinden von geeigneten Bioproben und zugehörigen klinischen Daten für Forschungsvorhaben in aggregierter Form. Der Suchbroker für die dezentrale Suche stellt eine Schnittstelle zur Formulierung von Anfragen zur Verfügung und verwaltet diese Anfragen. Er verarbeitet keine personenbezogenen Daten von Spendern. Personenbezogene Daten von zugreifenden Benutzern hingegen können im Rahmen der Protokollierung (vgl. Abschnitt 5) gespeichert werden. Darauf werden Nutzer hingewiesen.

Das Suchformular der dezentralen Suche wird vom Suchbroker bereitgestellt. Sie erlaubt die Suche nach allen im Metadata Repository abgelegten Begriffen; zusätzlich sind Ergänzungen im Freitext möglich. Der anzufragende Datensatz ist hier also prinzipiell unbeschränkt. Die Anfrage wird zunächst im Suchbroker gespeichert. Die Brückenköpfe der Standorte rufen in regelmäßigen Abständen neu hinzugekommene Anfragen vom Suchbroker ab und ermitteln, welche Datensätze im lokalen Data Warehouse – und folglich welches Probenmaterial in der Biobank – den Suchkriterien entsprechen. Der Inhalt der Anfrage sowie die gefundenen Datensätze können an jedem Standort von einer dazu berechtigten Person eingesehen werden.

Im Auslieferungszustand senden Brückenköpfe dem Suchbroker lediglich eine Übermittlungsbestätigung und informieren den Biobankverantwortlichen über die eingegangene Anfrage. Durch die Rückgabe einer aggregierten Anzahl gefundener Datensätze, die den Suchkriterien entsprechen, kann der Anfragende eine vorläufige Übersicht gewinnen. Um dies zu ermöglichen, kann der lokale Biobank-Verantwortliche die Suche so konfigurieren, dass (bestimmte) Anfragen automatisch beantwortet werden. In der lokalen Instanz der dezentralen Suche ist ein flexibles Regelsystem implementiert, mit dem die Rückmeldung einer Antwort u.a. von der jeweils anfragenden Person, deren Institution und dem Projektkontext abhängig gemacht werden kann. Damit soll es ermöglicht werden, einen lokalen Prüfschritt durch den Biobank-Verantwortlichen vorzuschalten. Die entsprechende Konfiguration dieses Regelsystems obliegt dem jeweiligen lokalen Biobank-Verantwortlichen.

Eine mögliche Vereinbarung über die konkrete Weitergabe von Daten und/oder Proben erfolgt dann im direkten Verhältnis zwischen dem Anfragenden und der/den Biobank(en). Damit verbundene datenschutzrechtliche Aspekte müssen im Einzelfall von den beteiligten Personen geklärt werden.

4. Organisatorische Maßnahmen

Zugriff durch Systemadministratoren

Die in den lokalen Data Warehouses der GBA-Brückenköpfe gespeicherten Daten können prinzipiell von den Administratoren der verwendeten IT-Infrastruktur eingesehen werden. Zugriffe auf die Daten durch Administratoren dürfen nur erfolgen, wenn dies zur Erfüllung ihrer Aufgaben zwingend erforderlich ist. Alle Administratoren sind auf diesen Grundsatz und auf ihre Pflicht zur Verschwiegenheit hinzuweisen. Dies sollte in der Regel im Rahmen des Arbeitsverhältnisses an der zuständigen Institution ohnehin geschehen sein.

Authentifizierung von Benutzern

Die Authentifizierung von Benutzern an zentralen Komponenten der GBA erfolgt mithilfe eines zentralen Authentifizierungsdienstes, der eine einheitliche, einmalige Anmeldung unter Nutzung derselben Zugangsdaten ("Single Sign-On") bereitstellt. An Standorten, die an der DFN-AAI-Föderation des DFN e.V.[4] teilnehmen, wird die Prüfung von Nutzernamen und Passwort an die Identity Provider der jeweiligen Standorte delegiert; Benutzer können sich also auf sichere Weise mit den Zugängen ihrer Arbeitgeber anmelden. Zusätzlich erfolgt eine Prüfung von Identität und Berechtigung von Benutzern durch die Geschäftsstelle des German Biobank Node in Berlin im Rahmen der Accountfreischaltung in Form einer E-Mail-Adress-Verifikation.

Wie alle zentralen Komponenten wird der Authentifizierungsdienst auf Servern des DKFZ betrieben (vgl. Abschnitt 2); die Verfahrensverantwortung verbleibt jedoch beim Träger.

Die Authentifizierung von Benutzern für die Nutzung von lokalen Komponenten obliegt den Standorten selbst.

Authentifizierung von Komponenten

Zugriffe einer GBA-Komponente auf eine andere über das Internet erfolgen nur nach erfolgreicher Authentifizierung, d.h. nicht nur die Berechtigung (Autorisierung), sondern auch die Identität der zugreifenden Komponente wird geprüft.

[4] Die DFN-AAI-Föderation (https://www.aai.dfn.de) ist ein Dienst des DFN-Vereins für wissenschaftliche Einrichtungen (Universitäten, Institute) und Anbieter (kommerziell und nicht kommerziell). Sie schafft das notwendige Vertrauensverhältnis sowie einen organisatorischen und technischen Rahmen für den Austausch von Benutzerinformationen zwischen Einrichtungen und Anbietern.

5. Technische Maßnahmen

Sicherheit der gespeicherten Daten

Die Sicherheit der gespeicherten Daten obliegt dem jeweiligen Standort und wird im Standort-internen Datenschutzkonzept der jeweiligen Biobank dargelegt. Ebenso unterliegen die ETL-Prozesse zur Befüllung des lokalen Brückenkopfs aus den Quellsystemen den lokalen Datenschutzrichtlinien und sind nicht Bestandteil dieses Datenschutzkonzepts.

Sicherheit der Kommunikation

Die Komponenten der GBA werden prinzipbedingt verteilt betrieben und kommunizieren über das öffentliche Internet. Die Vertraulichkeit ihrer Kommunikation wird durch folgende Maßnahmen sichergestellt:

- Die Kommunikation zwischen den einzelnen Komponenten erfolgt grundsätzlich über verschlüsselte Verbindungen (HTTPS). Die dafür eingesetzten Schlüssel und Zertifikate sind so zu erstellen, dass sie den aktuell anerkannten Anforderungen entsprechen (z.B. Schlüssellänge, Algorithmus).
- Durch Firewalls ist sichergestellt, dass die Server, auf denen die zentralen Komponenten laufen, nur über diejenigen Protokolle und Ports erreichbar sind, die für die Kommunikation mit Benutzern oder anderen Komponenten erforderlich sind (in der Regel HTTPS-Verbindungen). Der administrative Zugang ist auf das Intranet des jeweiligen Betreibers beschränkt.
- Alle Kommunikationsvorgänge zwischen den Brückenköpfen der Standorte und den zentralen Komponenten werden von den Brückenköpfen initiiert. Diese können dadurch hinter einer Firewall oder einem Proxyserver betrieben werden, ohne über eine öffentliche Adresse aus dem Internet erreichbar zu sein.

Protokollierung

Es erfolgt eine Protokollierung der Zugriffe von Forschern auf die Komponenten sowie zwischen den Komponenten untereinander. Der Nutzer wird bei Erstzugriff durch den Authentifizierungsdienst in Form eines "Terms of use" darüber informiert und seine Zustimmung eingeholt. Das Protokoll enthält mindestens:

- Die Identität der zugreifenden Person oder Komponente.
- Datum und Uhrzeit des Zugriffs.
- Den Inhalt des Zugriffs (die übermittelten Daten, ggfls. aggregiert) oder Informationen, aus denen dieser rekonstruiert werden kann (z.B. Verweis auf einen Datenbankeintrag o.ä.).

Das Protokoll wird zusammen mit den Nutzdaten des entsprechenden Servers auf diesem Server gespeichert. Nach zwölf Monaten werden entsprechende Protokolldateien gelöscht. Die aufgezeichneten Daten werden nur für folgende Zwecke verarbeitet und eingesehen:

- Im Rahmen der technischen Administration (insbesondere zur Fehlersuche).
- Zur Aufdeckung möglicher Missbrauchsfälle.
- Zur Erstellung anonymisierter Nutzungsstatistiken.

6. Wahrung von Betroffenenrechten

Rechtsgrundlage

Daten von Spendern verbleiben unter der Hoheit des Standorts. In Hinblick auf die Rechtsgrundlage sowie die Institutions-internen Prozesse der Datenverarbeitung wird auf die Datenschutzkonzepte der teilnehmenden Standorte verwiesen. Diese sehen vor, dass die Spender eine Einwilligung gegeben haben, die eine Verwendung ihrer Biomaterialproben und -daten im Rahmen der hier beschriebenen standortübergreifenden Prozesse erlaubt.

Lokal (d.h. im Brückenkopf) gespeicherte Datensätze können über die dezentrale Suche der GBA angefragt werden. Dabei verlassen je nach Konfiguration entweder keine Daten oder nur aggregierte Fallzahlen den Standort (vgl. Abschnitt 3.3). In beiden Fällen erfolgt die Beantwortung gemäß den am Standort anwendbaren landes- und bundesrechtlichen Datenschutzbestimmungen. Die beschriebenen Daten werden durch nachvollziehbare Zugriffsrechte geschützt. Die Beschreibung der dafür notwendigen Komponenten und Prozesse wird in Abschnitt 4 gegeben.

Datenschutzrechtliche Abgrenzung der GBA zu den einzelnen Standorten

Die Erhebung und Speicherung der Daten von Spendern erfolgt nur in der jeweiligen Institution. Dort ist zunächst zu prüfen, ob eine lokale Einwilligung in die Verwendung und Weitergabe von klinischen Daten und/oder Biomaterialen als Rechtsgrundlage vorliegt. Falls dies nicht zutrifft, sind die jeweiligen landesrechtlichen Regelungen mit den entsprechenden Ausnahmetatbeständen zu prüfen. Sollten diese im Einzelfall eine Verwendung von Bestandsdaten aus dem Behandlungskontext für die medizinische Forschung zulassen, können diese auch ohne Einwilligung verwendet werden. Ebenso ist die Speicherung von Daten im Brückenkopf unbedenklich, sofern sichergestellt ist, dass diese Daten auch nach der Speicherung im Brückenkopf nur von der jeweiligen Institution eingesehen werden können (vgl. Abschnitt 7).

Die MDAT der Spender werden in einer lokalen Komponente, dem lokalen Data Warehouse im Brückenkopf, in pseudonymisierter Form gespeichert. Der Brückenkopf steht unter lokaler Kontrolle des jeweiligen Standorts, und auch nur dort kann (mit erheblichem technischen Aufwand) mithilfe des Pseudonyms auf die Identität des Spenders geschlossen werden.

Die IDAT werden durch ein geeignetes Verfahren so verschlüsselt, dass eine Re-Identifizierung des Spenders praktisch unmöglich ist. Dies geschieht je nach Konfiguration entweder außerhalb des Brückenkopfes, unter lokaler Kontrolle des Standorts, oder im lokalen Identitätsmanagement des Brückenkopfes (vgl. Abschnitt 3.3).

Falls es für einen der beteiligten Standorte keine spezialgesetzlichen Ermächtigungsregelungen für die Verwendbarkeit der Daten aus dem Behandlungskontext zu Forschungszwecken (z.B. Landeskrankenhausregelungen) gibt, so kann ggf. eine Erhebung der Daten auf Basis der Forschungsklauseln des jeweiligen Landesdatenschutzrechts (für öffentliche Stellen der Länder) oder des Bundesdatenschutzgesetzes (für Stellen in privater Trägerschaft) möglich sein.

Aufklärung und Einwilligung

Eine informierte Einwilligung, in der der Spender auch über sein Recht auf Auskunft und Widerruf aufgeklärt wird, ist Rechtsgrundlage der Datenverarbeitung sowie der Nutzung und Weitergabe von Biomaterialproben. Ein vom Arbeitskreis Medizinischer Ethik-Kommissionen in der Bundesrepublik Deutschland e.V. erstellter Mustertext findet sich im Anhang. Standorte, die eine elektronische Erfassung des Einwilligungsstatus für ihre Spender implementiert haben, können diesen Einwilligungsstatus im ETL-Prozess bei der Befüllung des Brückenkopfs übergeben.

Auskunft über gespeicherte Daten

Alle Spender, deren Daten in den technischen Komponenten verwendet werden, haben das Recht, Auskunft über die über sie gespeicherten Daten zu erhalten. Der Antrag auf Auskunft ist schriftlich an die behandelnde Klinik/verantwortliche Institution zu stellen.

Widerruf, Löschung

Sämtliche Spender, deren Daten in den technischen Komponenten der GBA gespeichert und verwendet werden, haben jederzeit das Recht, die Einwilligung in die Verarbeitung der Daten in der GBA zu widerrufen. Infolgedessen

wird eine Löschung der Daten im Brückenkopf vorgenommen. Der Widerruf ist an den Standort zu richten, der die Bioproben und deren Daten hält.

Diese Löschung der Daten im Brückenkopf ist von den zuständigen Betreibern am jeweiligen Standort in einem angemessenen Zeitraum vorzunehmen und der Spender vom Vollzug zu unterrichten. Die meist impraktikable Löschung in Datensicherungen ist verzichtbar, sofern die Sicherungen nur durch den zuständigen Systemadministrator eingesehen werden können und alte Sicherungen regelmäßig gelöscht werden. Für den Sonderfall einer eventuell notwendig werdenden Datenwiederherstellung aus einer Datensicherung wird an jedem Standort Sorge getragen, dass Daten von Patienten, die ihre Einwilligung widerrufen haben, nicht wieder in den Brückenkopf zurückgespielt werden.

Dauer der Speicherung

Die erhobenen Daten bleiben in den lokalen Komponenten der GBA gespeichert, wie es im Rahmen der Patienteneinwilligung vorgesehen ist. Falls die Daten nicht mehr in der vorgesehenen Form genutzt werden können (z.B. falls der Betrieb der GBA nicht fortgeführt wird), ist von jedem Standort individuell eine Entscheidung über den weiteren Umgang mit den Daten zu treffen, da diese möglicherweise weiter für eigene Forschungsvorhaben genutzt werden können.

7. Vergleich mit dem TMF-Datenschutzleitfaden

Das Konzept folgt dem Klinischen Modul und dem Forschungsmodul des TMF-Datenschutzleitfadens.

8. Lokale Umgebung

An dieser Stelle können GBA-Biobanken Besonderheiten ihres jeweiligen Standorts beschreiben. Nach Möglichkeit sollte hierfür auf das am Standort bestehende Datenschutzkonzept verwiesen werden.

Anhang

Patienteneinwilligung

Der angehängte „Mustertext zur Spende, Einlagerung und Nutzung von Biomaterialien sowie zur Erhebung, Verarbeitung und Nutzung von Daten in Biobanken“ wurde vom Arbeitskreis Medizinischer Ethik-Kommissionen in der Bundesrepublik Deutschland e.V.[5] erstellt und dient als Vorlage, die von jedem Standort an lokale Erfordernisse angepasst wird. Die tatsächlich eingesetzte Einwilligungserklärung erhalten Sie beim jeweiligen GBA-Standortvertreter.

GBA-Datensatz

Der in der GBA zur Verwendung kommende Datensatz ist auf SIMPLIFIER.net abgelegt.

Votum der AG Datenschutz der TMF

Dieses Datenschutzkonzept wurde der Arbeitsgruppe Datenschutz der Technologie- und Methodenplattform für die vernetzte medizinische Forschung (TMF e.V.) zur Erteilung eines Votums vorgelegt. Von Seiten der TMF AG wurde beschieden, dass keine Bedenken bzgl. der Umsetzung des vorgelegten GBA-Datenschutzkonzeptes bestehen (Votum).

Datenschutz-Folgeabschätzung

Eine Datenschutz-Folgeabschätzung (Autorin: Petra Duhm-Harbeck) zu lokalen Bearbeitung wurde ebenfalls erstellt.

Danksagung

Dieses Datenschutzkonzept beruht auf dem generischen Datenschutzleitfaden der TMF und dem Datenschutzkonzept der Clinical Communication Platform (CCP-IT) des Deutschen Konsortiums für Translationale Krebsforschung (DKTK), welches unter maßgeblicher Mitwirkung von Andreas Borg (Universitätsmedizin Mainz) entstand. Es kann unter https://dktk.dkfz.de eingesehen werden.

[5] http://www.ak-med-ethik-komm.de/index.php?option=com_content&view=article&id=145&Itemid=154&lang=de